# Enhancing the sensitivity of the envelope-following response for cochlear synaptopathy screening in humans: the role of stimulus envelope

Viacheslav Vasilkov[a], Markus Garrett[b], Manfred Mauermann[b], Sarah Verhulst[a,*]

[a]*Hearing Technology @ WAVES, Department of Information Technology, Ghent University, Technologiepark 15, 9052 Zwijnaarde, Belgium*
[b]*Medizinische Physik and Cluster of Excellence "Hearing4all", Department of Medical Physics and Acoustics, Carl von Ossietzky University of Oldenburg, Carl-von-Ossietzky-Straße 9-11, 26129, Oldenburg, Germany*

## Abstract

Auditory de-afferentation, a permanent reduction in the number of inner-hair-cells and auditory-nerve synapses due to cochlear synaptopathy or damage, can reliably be quantified using temporal bone histology and immunostaining. There is, however, an urgent need for non-invasive markers of synaptopathy to study its perceptual consequences in live humans and to develop effective therapeutic interventions. While animal studies have identified candidate auditory-evoked-potential (AEP) based markers for synaptopathy, their interpretation in humans has suffered from translational issues related to neural generator differences, unknown hearing-damage histopathologies or measurement sensitivity. To render AEP-based markers of synaptopathy more robust and differential to the synaptopathy aspect of sensorineural hearing loss, we followed a combined computational and experimental approach. Starting from the known characteristics of auditory-nerve physiology, we op-



*Corresponding author
*Email address:* s.verhulst@ugent.be (Sarah Verhulst)

timized the stimulus envelope for envelope-following-responses (EFRs) to optimally and synchronously stimulate the available auditory-nerve population and consequently generate a strong AEP. We additionally used model simulations to explore which stimuli evoked a response which was sensitive to synaptopathy, while being insensitive to possible co-existing outer-hair-cell pathologies. We compared the model-predicted trends to AEPs recorded in younger and older listeners (N=44, 24f) who either had normal or impaired audiograms. We conclude that optimal stimulation paradigms for EFR-based quantification of synaptopathy should have sharply rising envelope shapes, a minimal plateau duration of 1.7-2.1 ms for a 120 Hz modulation rate, and inter-peak intervals which contain near-zero amplitudes. From our recorded conditions, the optimal EFR-evoking stimulus had a rectangular envelope shape with a 25% duty cycle and a 95% modulation depth.

## Significance Statement

Even though cochlear synaptopathy is since 2009 identified as a form of sensorineural hearing loss (SNHL) which also affect primates and humans, clinical practice does not routinely screen for it and the role of synaptopathy for sound and speech perception is presently unclear. Consequently, cochlear synaptopathy may be underdiagnosed in the ageing population with self-reported hearing difficulties and its perceptual impact underestimated. To enable a differential EEG-based diagnosis of synaptopathy in humans, it is crucial to adopt a stimulation and analysis method which yields a robust response which shows large inter-individual differences which are sensitive to synaptopathy but not affected by other SNHL aspects. Our study uniquely combines computational modeling with experiments in normal and hearing-impaired listeners to design a EFR stimulus which can be used for the differential diagnosis of synaptopathy in humans.

*Abbreviations*

ABR - auditory brainstem response; AEP - auditory evoked potentials; AM - amplitude modulation; ANF - auditory-nerve fiber; BB - broadband; BM - basilar membrane; CF - characteristic frequency; CN - cochlear nucleus; EFR - envelope following response; H/M/LSR - high/medium/low spontaneous rate; HI - hearing-impaired; IC - inferior colliculus; IHC - inner-hair-cell; MD - modulation depth; NH - normal-hearing; OAE - otoacoustic emission; OHC - outer hair cell; peSPL - peak-equivalent sound pressure level;

RAM - rectangular-wave amplitude-modulated; RMS - root mean square; SAM - sinusoidal amplitude-modulated; SNHL - sensorineural hearing loss;

## Introduction

Noise overexposure, ototoxicity and aging can cause primary cochlear de-afferentation, i.e. progressive and irreversible damage to the afferent neuronal structures in the auditory periphery. One form of auditory de-afferentation is cochlear synaptopathy and refers to damaged synapses between the inner-hair-cells (IHCs) and auditory-nerve fibers (ANFs). This type of sensorineural hearing loss (SNHL) was first discovered in mouse models (Kujawa & Liberman, 2009) and has since been shown to exist in macaques and humans as well (Wu et al., 2018; Viana et al., 2015; Makary et al., 2011; Valero et al., 2017). Cochlear synaptopathy specifically degenerates the synaptic terminals of the spiral ganglion cells and precedes hair cell damage in the ageing process (Wu et al., 2018; Fernandez et al., 2015; Sergeyenko et al., 2013). Also Ouabain and Kanic-acid treatment (Bourien et al., 2014; Shaheen et al., 2015; Chambers et al., 2016; Sheets, 2017) or noise-induced insults which only cause a temporary threshold shift have shown to result in cochlear synaptopathy (Kujawa & Liberman, 2009; Furman et al., 2013). The compelling histopathological evidence, along with outcomes from animal behavior studies of auditory de-afferentation (Schuknecht & Woellner, 1955; Lobarinas et al., 2013), have shown that cochlear synaptopathy has little effect on hearing sensitivity assessed through the behavioral audiogram or

physiological threshold measures (e.g., distortion-product otoacoustic emissions; DPOAEs or auditory brainstem responses; ABRs). Cochlear synaptopathy might hence remain hidden during routine clinical hearing screening (Schaette & McAlpine, 2011), which typically assesses hearing sensitivity using the audiogram. We might hence overlook a large population of listeners who have accrued synaptopathy while their audiograms reflect normal hearing. Additionally, in listeners with age-related audiometric declines (ISO 7029), we hitherto only diagnosed and treated the hair-cell-damage aspect of SNHL, while disregarding the possible co-existing synaptopathy aspect. To study the prevalence of synaptopathy, and its consequences for sound perception in humans, it is hence crucial to develop a non-invasive differential diagnostic test for synaptopathy which offers a pathway to effective therapeutic interventions.

The search for candidate non-invasive markers of synaptopathy has been ongoing since its discovery and has shown a promising role for auditory-evoked potentials (AEPs). Specifically, a reduction of the supra-threshold auditory brainstem response (ABR) amplitude was directly associated with histologically-verified cochlear synaptopathy (Kujawa & Liberman, 2009; Furman et al., 2013; Sergeyenko et al., 2013; Bourien et al., 2014; Möhrle et al., 2016a). Particularly, the ABR wave-I amplitude is currently considered as the most direct metric of cochlear synaptopathy (Kujawa & Liberman, 2009; Schaette & McAlpine, 2011; Lin et al., 2011; Furman et al., 2013; Prendergast et al., 2017a, 2018; Plack et al., 2016; Bramhall et al., 2019). The second

measure proposed from cochlear synaptopathy studies in animal models is the envelope-following response (EFR), an AEP-type which is of predominant subcortical origin when the amplitude modulation (AM) rate of the sustained stimulus is above 80 Hz (Purcell et al., 2004). EFRs offer a more robust metric of cochlear synaptopathy than ABRs, as synaptopathy-induced EFR changes are greater than ABR amplitude reductions in the same animal (Shaheen et al., 2015; Parthasarathy et al., 2018).

Despite the compelling experimental evidence of studies which combined AEP recordings with direct post-mortem synapse counts, a direct translation toward a differential synaptopathy diagnosis in humans has proven difficult (Plack et al., 2016; Guest et al., 2017, 2018; Prendergast et al., 2017a, 2018; Bramhall et al., 2019; Garrett & Verhulst, 2019; Bharadwaj et al., 2019). The human data is not unambiguous in demonstrating reduced AEP metrics in listener groups with suspected synaptopathy (e.g., as induced through accumulated noise-exposure or age), and a number of studies report subtle or non-significant correlations between different electrophysiological markers which are sensitive to synaptopathy in animals (e.g. ABR and EFR amplitudes or slope changes, middle-ear-muscle reflex strength; Prendergast et al., 2017a; Guest et al., 2019; Garrett & Verhulst, 2019). Also individual differences in psychoacoustic tasks thought to be sensitive to cochlear synaptopathy (e.g. speech perception in noise, frequency discrimination, amplitude-modulation detection) do in some studies correlate with physiological markers of synaptopathy (e.g.; Bharadwaj et al., 2015; Mehraei et al., 2016; Liberman et al.,

2016; Verhulst et al., 2018b), whereas in others they do not (e.g.; Schoof & Rosen, 2016; Guest et al., 2018; Prendergast et al., 2017b; Plack et al., 2014; Johannesen et al., 2019).

There are several aspects which contribute to these translational issues: the adopted physiological markers may be affected by species-specific biophysical processes (e.g. humans may be less vulnerable to noise damage than other species; Dobie & Humes, 2017; Valero et al., 2017; Hickox et al., 2017). Secondly, the markers may be differently impacted by the functionality of SNHL aspects (e.g., OHC nonlinearity vs ANF dynamic range coding; Bramhall et al., 2019; Garrett & Verhulst, 2019), which may complicate their interpretation in terms of synaptopathy. Another aspect relates to the limited extent by which cochlear synaptopathy might affect the considered the perceptual tasks (e.g. 50% ANF loss might be required to see a perceptual effect; Oxenham, 2016), and different ANF types may contribute differently to the considered electrophysiological markers. E.g., low-spontaneous rate ANFs do not contribute strongly to the transient ABR wave-I (Bourien et al., 2014), but may contribute strongly to the EFR to low-modulation depth stimulation (Bharadwaj et al., 2014). Lastly, it is possible that electrophysiological markers of synaptopathy simply have limited test-retest reliability for human use (D'haenens et al., 2008; Prendergast et al., 2018). Before resorting to non-AEP based diagnostic markers, it is hence worthwhile to optimize existing AEP stimulation paradigms and analysis approaches to enhance the signal-to-noise ratio of the AEP, and as consequence, improve test reliability.

This route may yield a robust and sensitive diagnostic marker for auditory deafferentation in humans, and help resolve the role of cochlear synaptopathy for sound perception.

To address the above translational issues, this study focusses on optimizing stimulation and analysis paradigms to yield a reliable EFR-based cochlear synaptopathy diagnosis in humans. We focus on the EFR, a particular AEP type evoked by sustained periodic stimuli with constant carrier and modulation frequencies. In other clinical studies, responses to this type of stimulation are also referred to as auditory steady-state responses (ASSR). However, to remain consistent with the nomenclature adopted in cochlear synaptopathy studies, we will use the more general EFR term. To develop more robust EFR markers, we draw from functional IHC-AN and peripheral auditory processing properties to develop stimulation paradigms which better reflect the available ANF population. This is important because a strong baseline EFR response will be more sensitive to changes induced by alterations in the ANF population. At the same time, we adopt an optimized analysis method which extracts all the relevant envelope-following components from the raw EEG recordings (Vasilkov & Verhulst, 2019). Furthermore, we studied how OHC functionality affects the EFR generators to different stimulation paradigms to evaluate which EFR markers are differentially sensitive to synaptopathy, even when OHC damage is simultaneously present. We tested our biophysically-inspired stimulation paradigms in a computational model of the human auditory periphery (Verhulst et al., 2018a; Vecchi & Verhulst,

2019, v1.2) which can simulate EFRs for several frequency-specific SNHL profiles with different combinations of OHC and synaptopathy damage. The model simulations were compared against reference data recorded from three groups of study participants: young and older listeners with normal or elevated audiometric thresholds. The latter two groups show age-related SNHL damage with less or more OHC damage, whereas the first two groups might show differences in age-related synaptopathy. Both model simulations and experiments confirmed that our proposed modifications to the stimulation and analysis paradigms increased the EFR magnitude and its sensitivity to selectively detecting changes in the ANF population, to yield a more-robust non-invasive marker of cochlear synaptopathy for use in humans.

## Materials and Methods

*Stimuli*

All acoustic AEP stimuli were generated in MATLAB R2015b (The MathWorks Inc., 2015) and had a sampling rate of 48 kHz for the recording sessions and of 100 kHz for the model simulations. Specifically, we designed our stimuli on the basis of known observations in psychoacoustic and physiological studies of AM (e.g., van de Par & Kohlrausch, 1997; John et al., 2002, 2003; Bernstein & Trahiotis, 2002, 2009; Stürzebecher et al., 2003; Griffin et al., 2005; Dreyer & Delgutte, 2006; Laback et al., 2011; Klein-Hennig et al., 2011; Greenberg et al., 2017; Van Canneyt et al., 2019), and hypothesize

that overall stronger EFRs might render individual EFR differences more robust. From the analysis of the named studies, we suspect that a combination of increased silence gaps between the stimulus peaks and shorter stimulus duty-cycles might cause more synchronized ANF activity, which, in turn might result in stronger EFRs. To test this hypothesis, we designed seven stimulus conditions which had the same amplitude modulated (120 Hz) 400-ms tonal/noise stimuli, but had different stimulus levels or envelope shapes. To validate our predictions and study the relative contribution of synaptopathy and OHC deficits, we simulated single-unit ANF responses as well as EFRs, which we also recorded EFRs in 44 participants. We used the widely-adopted sinusoidal amplitude-modulation (SAM) as the reference condition:

$$m\,(t) = \frac{\mathrm{A}_m}{2} \sin\,(2\pi \mathrm{f}_m t), \tag{1}$$

where $\mathrm{A}_m$ corresponds to the peak-to-peak amplitude, $\mathrm{f}_m$ is the modulation frequency, $t$ is the time vector. Two carrier types were considered: a 4-kHz pure tone (PT) and a white noise carrier with a 50-16000 Hz bandwidth (BB) which stimulates a tonotopically broader cochlear region. Amplitude modulation was implemented by multiplying the carrier $c(t)$ with peak-to-peak amplitude $\mathrm{A}_c$ with $[1 + md * m(t)/(\mathrm{A}_m)]$, where $md = \mathrm{A}_m/\mathrm{A}_c$. Figure 1a represents two cycles of the reference SAM stimulus with the 4-kHz PT carrier presented in sine phase. EFRs to the reference SAM stimuli were compared

to EFRs evoked by five AM stimuli which had the same 4-kHz pure tone carrier, but had different modulators or sound levels.

The second, non-sinusoidal periodic modulator was a rectangular waveform with a period of $2\pi$ and a 25% duty-cycle (RAM25, Fig. 1b) which was generated using the *square(t, 100d)* function (MATLAB R2015b). The RAM25 modulator can be described as the Fourier series expansion:

$$m\left(t\right) = \mathrm{A}_m \left( \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin\left(\pi n d\right) \cos\left(2\pi n \mathrm{f}_m t\right)}{n} - \frac{1}{2} + d \right), \tag{2}$$

where d = 0.25 denotes the duty-cycle, which represented a ratio between the pulse width and the total period of the modulator, n is the harmonic number of the series.

The third modulator was a rectangular waveform with a 50% duty-cycle (RAM50; Fig. 1c) which was generated using the *square(t, 100d)* function with d = 0.5, and can be described as:

$$m\left(t\right) = \frac{2\mathrm{A}_m}{\pi} \sum_{n=1}^{\infty} \frac{\sin\left(2\pi\left(2n-1\right)\mathrm{f}_m t\right)}{\left(2n-1\right)}. \tag{3}$$

The fourth modulator was a ten-harmonic complex (H10AM, Fig. 1d) presented in cosine phase and is defined as:

$$m\left(t\right) = \frac{\mathrm{A}_m}{2} \sum_{n=1}^{10} \cos\left(2\pi n \mathrm{f}_m t\right). \tag{4}$$

The different AM stimuli were calibrated to have the same root-mean-square (RMS) sound pressure of 70 dB SPL. To study whether there was an effect of RMS versus peak-to-peak sound calibration, two stimuli were also calibrated to have the same peak-to-peak amplitude as the reference SAM tone (i.e. $RAM25_{ptp}$ and $RAM50_{ptp}$; Fig. 1b and Fig. 1c, cyan). They were presented at 68.18 dB SPL and 71.18 dB SPL, respectively. All stimuli were 95% amplitude modulated (e.g., -0.45 dB re 100% modulation) with a starting phase shift of $3\pi/2$ (except for the H10 complex modulator which had a 0 starting phase shift). Each stimulus was ramped using a 2.5% tapered-cosine (Tukey) window and was presented 1000 times using 500 repetitions per polarity. Stimuli were presented monaurally and a uniformly distributed inter-stimulus silence interval of 100 $\pm$ 10 ms was applied.

**ABR**s were recorded to 3000 repetitions of a 80-$\mu$s click presented monaurally with alternating (condensation/rarefaction) polarity at a mean rate of 10 Hz (including the uniformly distributed 10% silence jitter). Three stimulus levels were tested (70, 85, and 100 dB peSPL) and we only considered the 70 and 100 dB peSPL conditions conditions for this study. ABR and EFR stimuli were calibrated using an oscilloscope, ear simulator (Brüel & Kjær, type 4157) and B&K 2610 sound level meter.

*Model of the auditory periphery*

The auditory periphery model we adopted for this study (Verhulst et al., 2018a; Vecchi & Verhulst, 2019, model implementation v1.2) simulates audi-

tory processing along the ascending pathways (Fig. 3) and includes middle-ear filtering, a nonlinear transmission-line representation of human cochlear mechanics (Verhulst et al., 2012; Altoè et al., 2014), a biophysical inspired model of the IHC-AN complex (Altoè et al., 2018), and a phenomenological description of ventral cochlear nucleus (CN) and inferior colliculus (IC) neurons (Nelson & Carney, 2004). The model reasonably captures properties of AN fiber types with different spontaneous rates, level-dependent ABR/EFR characteristics, and furthermore can mimic frequency-specific hearing impairments related to OHC damage and cochlear deafferentation or synaptopathy (Verhulst et al., 2015, 2018a).

Cochlear synaptopathy was modeled by reducing the number of IHC-AN synapses and functional ANFs of different types at each simulated tonotopic location. The normal-hearing (NH) model had 19 fibers with three spontaneous rate (SR) types synapsing onto each IHC (Verhulst et al., 2018a): 3 low (LSR), 3 medium (MSR) and 13 high (HSR), following the ratio observed in cats (Liberman, 1978). Three synaptopathy profiles were implemented by removing the following fiber types across the tonotopic axis: (i) all LSR and MSR fibers ($\mathrm{HI_{CS:0L,0M,13H}}$), (ii) all LSR, MSR and 50 % of the HSR fibers ($\mathrm{HI_{CS:0L,0M,07H}}$), and (iii) all LSR, MSR and 80 % of the HSR fibers ($\mathrm{HI_{CS:0L,0M,03H}}$). We limited our simulations to uniform, CF-independent synaptopathy profiles. No IHC-specific dysfunctions were simulated in the current study as synaptopathy was suggested to occur without destroying the sensory cells (Kujawa & Liberman, 2009; Lin et al., 2011; Furman et al.,

2013; Shaheen et al., 2015). However, IHC loss can be simulated in this framework by introducing a complete synaptopathy (0 LSR, 0 MSR, 0 HSR fibers).

Simulation of OHC dysfunction caused by damaged mechano-receptors or presbycusis is possible by adjusting the parameters of the simulated cochlear filters to yield frequency-specific gain loss profiles. Figure 2a shows mean audiometric thresholds of the study participants along with corresponding simulated cochlear gain loss profiles (dashed and solid lines, respectively). These gain loss profiles (in dB HL) were used to determine the parameters of the cochlear filter gain relative to the normal-hearing cochlear filter gain at CFs corresponding to the audiometric testing frequencies (see Fig.2 in Verhulst et al., 2016, for the relationship between filter gain and the value of double-pole of the basilar-membrane (BM) admittance in the model). Even though the model can simulate individual human audiograms in great detail, we limited our simulations to a range of sloping high-frequency audiograms approximating the average audiograms of each participant group (yNH, oNH, oHI; see section "Participants").

AEPs were simulated by adding up instantaneous firing rates across a tonotopic array of 401 IHC-AN/CN/IC units (Verhulst et al., 2018a) positioned along the cochlea according to the frequency-position map (Greenwood, 1990). The responses from 19 AN fibers of three SR types (or all available ANFs for HI profiles) which synapse to a single IHC at each CF were summed and projected to a single CN unit of the same CF. The instantaneous firing rate

of a single CN unit served as input to a single IC unit. A same-frequency inhibition and excitation model for the CN and IC units (which captures the modulation filtering and onset enhancement characteristics of auditory brainstem and midbrain neurons; Nelson & Carney, 2004) was adopted. Population responses were obtained by adding up instantaneous firing rates across all simulated CFs for three processing stages: (i) the AN, after summing up 19 ANF responses across each IHC with different CFs, which yields the W-I response in Fig. 3; (ii) the CN and (iii) IC model stages yielding the W-III and W-V response respectively. For EFR simulations, the summed population responses from the AN, CN and IC processing stages were added up (Verhulst et al., 2018a) to realistically capture the different subcortical sources that contribute to EFRs (Dolphin & Mountain, 1992; Kuwada et al., 2002).

*Participants*

A total of 44 participants were recruited into three groups based on the combination of two criteria: age and audiometric profile. The young normal-hearing (yNH; Fig. 2) group consisted of 15 participants (8 females) with ages between 20-30 years (24.5 $\pm$ 2.2 y/o) and pure-tone hearing thresholds below 25 dB HL across the standard audiometric frequency range. The older normal-hearing (oNH; Fig. 2) group comprised 16 participants (8 females) with ages between 60-70 years (64.3 $\pm$ 1.8 y/o) and normal hearing thresholds (below 25 dB HL; Table 2) across the audiometric frequencies 0.125-4

kHz (where 4-kHz corresponds to the pure-tone carrier frequency of the AM signal). The older hearing-impaired (oHI; Fig. 2) group consisted of 13 participants (8 females) with ages 60-70 years (65.2 ± 1.8 y/o) and sloping high-frequency pure-tone audiograms that exceeded 25 dB HL at 4-kHz. An otoscopic inspection was performed to ensure that participants had no visible pathologies or obstructions. The audiometric thresholds, gender and ages of all participating individuals are listed in Table 2. Participants were informed about the experimental procedures and the experiments were approved by the ethical commission of the University of Oldenburg. Participants gave a written informed consent and were paid for their participation.

Audiograms were measured for standard frequencies between 0.125-8 kHz using a clinical audiometer (Auritec AT 900) and over-ear audiometric headphones (Sennheiser HDA 200). Individual hearing thresholds of the audiometrically better ear (which was used for the auditory stimulation) are depicted in Fig. 2A. Figure 2B shows individual hearing thresholds at 4-kHz (which corresponds to the PT carrier frequency of the EFR stimuli) for yNH (3.3 ± 3.5 dB HL), oNH (11.6 ± 3.8 dB HL) and oHI (37.7 ± 6.4) groups. Additionally, we recorded distortion-product otoacoustic emissions (DPOAEs) to isolate the OHC-related aspect of SNHL. To this end, ER-2 insert earphones were coupled to the ER-10B+ microphone system (Etymotic Research) and we used a custom-made MATLAB software (Mauermann, 2013) for DPOAE recording and analysis. Primary tone pairs were simultaneously presented with a fixed $f_2/f_1$ ratio of 1.2 using a continuously sweeping DPOAE

paradigm (Long et al., 2008). Primary-frequencies were exponentially swept (2 s/octave) during stimulus presentation over a 1/3 octave range around the geometric mean. $f_2$ ranged from 1000 to 4000 Hz using octave steps. Levels were set according to the "scissors" level paradigm ($L_1 = 0.4L_2 + 39$ dB; Kummer et al., 1998). $L_2$ levels ranged in 6 dB steps between 30-60 dB SPL for both yNH and oNH participants, and 30-72 dB SPL for oHI participants. DPOAE thresholds were derived from recorded DPOAE-$L_2$-level series for each mean $f_2$ frequency within the measured frequency range using a bootstrapping procedure. Extracted distortion components were bootstrapped 200 times and, for each bootstrap average, a tailored cubic growth function was fit through the data-points (Verhulst et al., 2016). DPOAE thresholds were determined as the median of the $L_2$ levels at which the cubic curve fit reached a level of -25 dB SPL (Boege & Janssen, 2002) for each bootstrap average. DPOAE thresholds at 4 kHz are depicted in Fig.2C for yNH (16.1 $\pm$ 10.0 dB HL), oNH (28.9 $\pm$ 7.5 dB HL) and oHI (48.2 $\pm$ 10.5) groups.

*Recording setup and data preprocessing*

Measurements were performed in a double-walled electrically-shielded booth while participants sat comfortably in a reclining chair and watched a silent movie. Stimuli were presented monaurally (using the audiometrically better ear) over magnetically-shielded ER-2 insert earphones (Etymotic Research) connected to a TDT-HB7 headphone driver (Tucker-Davis) and a Fireface UCX sound card (RME). EEG data were recorded using a

64-channel recording system (BioSemi) and BioSemi Active-electrodes which were spaced equidistantly in an EEG recording cap (EasyCap). A common-mode-sense active electrode was placed on the fronto-central midline and a driven-right-leg passive electrode was placed on the tip of the nose of the participant. Reference electrodes were placed on each earlobe and electrode offset voltages were kept below 25 mV for all electrodes. A 24-bit AD conversion with sampling rate of 16384 Hz was used to digitize and store the raw data (for additional setup details see Garrett et al., 2019).

The raw data were preprocessed using Python (version 2.7.10) and the MNE-Python (version 0.9.0) open-source software package (Gramfort et al., 2013, 2014). The vertex (Cz) channel potentials were re-referenced to the off-line-average of the two earlobe channel potentials to AEPs. EFR data were epoched to 400-ms windows starting from the stimulus onset and were baseline corrected using the average amplitude per epoch. ABR recordings to positive and negative polarity clicks were high-pass filtered with a cut-off frequency of 200 Hz and then low-pass filtered with a cutoff frequency of 2000 Hz using a zero-phase filter (4th order IIR Butterworth filter). ABR recordings were epoched into 20 ms windows relative to the stimulus onset. Bad epochs were identified using the joint probability criteria as implemented in EEGLAB (Brunner et al., 2013). For additional details on the ABR data preprocessing see Garrett & Verhulst (2019). To allow a fair comparison across condition and subjects, a constant number (100) of pair-averaged epochs (out of 1500) with the highest peak-to-trough amplitudes were rejected. The

peak-to-trough amplitudes which remained after artifact rejection were below 25 $\mu$V.

*EEG analysis*

Pre-processed time-domain EEG waveforms were further processed in MATLAB R2014b (The MathWorks Inc., 2014) to perform waveform averaging, bootstrapping and feature extraction. **EFR** magnitudes were derived by estimating the amplitude of a time-domain response which contained predominantly stimulus-driven energy. This signal was obtained by removing the individual electrophysiological noise floor (NF) and stimulus-irrelevant EEG components (Fig. 4). We calculated the mean EFR magnitude and corresponding standard deviation across the available epochs using a bootstrap procedure (Zhu et al., 2013). In each bootstrap run, a magnitude spectrum (in $\mu$V) was calculated by calculating the FFT of the time-domain average of 1000 randomly sampled response epochs (500 epochs per each stimulus polarity) with replacement. Epochs were ramped using a 2% tapered-cosine (Tukey) window before the frequency-domain transformation was applied. An example of an EEG magnitude spectrum for one bootstrap average and corresponding NF estimates is shown for a listener from the NH group (#7 and 70-dB-SPL RAM25$_{rms}$ stimulation) in Fig. 4A. To include all available envelope-related components, the EFR magnitude was computed from the EEG spectrum based on the energy at the frequency corresponding to the stimulus modulation rate ($f_0$=120 Hz) and harmonics of the fundamental

modulation frequency ($f_{(k-1)}$=$k$*$f_0$, $k$=[1..5] for our recordings) using the energy above the NF. The noise floor at $f_0$-$f_4$ was computed as the average magnitude across the ten bins centered around the corresponding frequency (5 bins on either side). Spectral peaks at $f_0$-$f_4$ ($F_n$) were then corrected by subtracting the respective $NF_n$ values to yield a relative peak-to-noise-floor (PtN) magnitude estimates (blue arrows; Fig. 4A). Negative PtN estimates (e.g., when spectral peaks $F_n$ were smaller than the noise-floor $NF_n$) were set to zero and energy at other frequencies were removed. The EFR waveform was obtained after performing an iFFT which included the noise-floor corrected peaks ($F_n$-$NF_n$) and their corresponding phase angle values ($\theta_n$) to yield a time-domain signal which mostly contains response energy related to the AM stimulation (Vasilkov & Verhulst, 2019). This procedure allowed us to focus on the individual NF-corrected component of the recording and uses absolute signal values (in $\mu$V) instead of SNR values (which can be affected by noise-floor level variability between NH and HI listeners). Figure 4B depicts the comparison between reconstructed and recorded time-domain signals (solid blue and thin gray traces, respectively). Note that the recordings in Fig. 4B were band-pass filtered between 117 Hz and 603 Hz to keep stimulus-related components and to remove irrelevant energy beyond the fundamental modulation frequency and its harmonics for visual clarity. Finally, the EFR magnitude was defined as half the peak-to-peak amplitude

of the reconstructed time domain-signal waveform, i.e:

$$\mathrm{EFR}_{\mathrm{PtN}} = \frac{\textit{peak-to-peak}\left(\frac{1}{N}\sum_{n=0}^{N-1}\left(\mathrm{F}_n - \mathrm{NF}_n\right)e^{i\theta_n}\right)}{2}; \qquad (5)$$

$$\text{if } n \neq \frac{k\mathrm{f}_0}{\mathrm{f}_s}N, \text{ then } \mathrm{F}_n, \mathrm{NF}_n = 0, \text{for } k = [1..5],$$

where $N$ corresponds to the length of the magnitude spectrum, and $\mathrm{f}_s$ is the sampling rate. As a result of the bootstrapping procedure, we obtained 200 reconstructed time-domain waveforms for each listener and stimulus condition, which we used to accurately estimate the $\mathrm{EFR}_{\mathrm{PtN}}$ magnitude and its standard deviation. Simulated EFR magnitudes for different SNHL profiles were directly derived from the time-domain responses, because no noise or stochastic processes were implemented in the adopted model version. The simulated EFR magnitude was defined as half of the peak-to-peak amplitude of the average one-modulation-cycle waveform across the 400-ms epoch duration.

**ABR** waveforms, variability and noise floors were estimated using a bootstrap procedure. For each condition, 1000 time-domain epochs (for positive and negative stimulus polarities) were randomly drawn with replacement and averaged 200 times. To estimate the noise-floor, epoch averaging was repeated 1000 times, but half of the total (1000) randomly drawn epochs were multiplied by -1 before averaging. Both ABR wave-I and wave-V and

their corresponding peaks were identified by visual inspection. Picked ABR peaks were corrected by noise floor estimates. ABR amplitudes [in $\mu$V] were defined as half the amplitude difference between the corresponding positive peak and maximal negative deflection before the next up-going slope (Picton, 2010).

## Results

*Simulated single-unit responses to EFR stimuli with different envelope shapes*

Simulated ANF firing rates at the CF of 4 kHz were summed across the different fiber types and shown in Fig. 1e-h for the different stimuli. Responses to two cycles of the reference 70-dB-SPL SAM tone are shown as well as responses to the other three envelope shape stimuli. Each panel depicts simulations of the summed responses of the normal-hearing model (NH: green) as well as for two models which simulated different aspects of SNHL: two degrees of simulated OHC damage ($HI_{OHC:10@4K}$: blue, $HI_{OHC:35@4K}$: red) and one synaptopathy profile ($HI_{CS:0L,0M,03H}$: black). The simulated NH responses generally follow the stimulus envelope shape, but due to the nonlinear properties of the auditory periphery, the individual responses show different degrees of response strength and distortion. In particular, the NH reference SAM response in Fig. 1e shows a distorted shape with (i) strong firing rates in response to the sloping parts of each stimulation cycle due to nonlinear cochlear responses properties. The same mechanism caused only short temporal regions where there was no firing in response to stimulus en-

velope minima, and (ii), low firing rates towards the end of each cycle (due to IHC-AN adaptation properties; Altoè et al., 2018).

Despite their similar modulation rates and SPL levels, NH responses in Fig. 1f,g had steeper attack/decay slopes with broader temporal regions with near-zero firing rates. The RAM stimuli evoked stronger responses compared to the reference SAM condition. Simulated NH ANF responses evoked by the H10AM stimulus (Fig. 1h, green) were characterized by sharp peaks and showed pronounced firing to the near-threshold stimulus fluctuations between two cycle peaks (due to OHC-related amplification for low stimulus levels). Comparison between the conditions shows that long inter-peak-intervals (IPI) are a necessary condition to yield high peak firing rates during supra-threshold stimulation with modulated stimuli of high modulation rates (Fig. 1f, green; RAM25). Longer IPIs may provide more time for the neuron to recover (e.g. replenish neurotransmitter) such that it can respond more reliably to each stimulation cycle. However, a comparison between RAM25 and H10AM, or SAM and RAM50 firing rates confirms that long IPIs are a necessary, but insufficient, condition. In fact, stimulation within the IPI can reduce the peak firing rates (e.g. see H10AM). Comparison between the RAM50 and RAM25 firing rates shows that the firing rate increases with decreasing duty-cycle of the stimulus envelope. This reflects more synchronized spiking activity to the stimulation plateau and reduced spiking in the silence windows caused by the longer RAM25 IPI.

Cochlear amplification responds differently to the different stimulus envelope

shapes, and consequently, OHC damage might influence the simulated ANF responses differently. Simulated NH ANF rates (green) and HI rates for 10 (blue) and 35 dB HL (red) loss at 4 kHz are depicted in Fig. 1e-h. For the reference SAM stimulus (Fig. 1e), increasing the degree of cochlear gain loss resulted in linearized and less distorted ANF firing rates and broader silence regions between the response cycles. Additionally, enhanced response amplitudes were observed in comparison to the NH ANF responses near the stimulus envelope maxima. Stronger peak ANF rates for HI vs NH simulations were also observed for the H10AM condition. Elevated threshold sensitivity in combination with small ANF rates between the stimulus peaks, caused strong HI ANF peak responses near the envelope maxima. Cochlear gain loss affected the RAM firing rates differently: ANF rates to all simulated cochlear gain loss profiles largely overlapped and showed only marginal differences between the peak rates of NH and HI responses. These simulations show that different degrees of OHC damage had a negligible effect on the peak RAM-ANF rates at CF.

Introducing synaptopathy reduced the firing rates to all stimuli (Fig. 1e-h, black dashed lines). This reduction was proportional to the remaining number of intact ANFs. Figures 1e-h show that the stimulus envelope shape can have an important effect on how the ANF firing patterns are affected by different aspects of SNHL. Both responses to the SAM and H10AM stimuli show that inter-peak envelope components with low sound intensities can yield stronger peak responses after OHC damage, which - to a certain degree

- can compensate for the reduced firing rates caused by synaptopathy when both SNHL aspects are present. In contrast, the ANF rates to the RAM stimuli were strongly affected by synaptopathy, irrespective of the OHC damage pattern. These ANF response simulations at CF hence suggest that AM stimuli with rectangular envelope shapes can provide a differential and enhanced sensitivity to the synaptopathy aspect of SNHL.

*Simulated and recorded EFRs: time domain comparison*

To investigate whether the EFR, as a population response across a neuronal population of different CFs, follows the on-CF ANF response trends, Figs 1i-l show simulated (open traces) and recorded (filled traces) EFR waveform averages per group. Simulated and recorded NH EFRs (Fig. 1i-l, green) generally followed the trends observed in the ANF responses by showing the strongest response maxima for the RAM25 stimulus. Comparing the H10AM to the RAM25 EFR confirms that long IPIs and short duty-cycles are important, but not sufficient to evoke a strong EFR, and that low-level stimulation between two envelope maxima can end up reducing the EFR strenght. It is noteworthy that the RAM50 recording showed two peaks per cycle whereas neither the RAM50 simulations, nor the RAM25 responses showed such a double response peak within an envelope cycle. Our simulations showed that double responses can occur when apical off-CF BM vibrations (approximately up to one octave below the 4-kHz CF) contribute to the ANF population re-

sponse. However, the CN/IC filtering properties of our model removed this second peak from our EFR simulations.

In agreement with ANF simulations, synaptopathy reduced the peak-to-peak amplitudes of the simulated EFRs considerably. The recordings showed reduced EFR amplitudes in both groups of older listeners (oNH and oHI) compared to the yNH group (Fig. 1i-l). Our simulations suggest that age-related synaptopathy was the cause of this reduction in the older participant groups. One might further speculate that the somewhat smaller oHI than oNH responses for the RAM conditions reflects a stronger degree of synaptopathy in the oHI group. This statement is supported by the predicted insensitivity of the RAM EFR to OHC damage. At the same time, it is important to note that the H10AM response difference between oNH and oHI groups disappears (Fig. 1l). This observed trend corroborates the simulations which show that the OHC damage aspect can counteract the synaptopathy-induced response reduction for the H10AM, but not RAM, stimuli.

*Group EFR magnitudes across stimulus conditions*

Fig. 5a compares recorded (filled symbols) and simulated EFR magnitudes (open symbols) across conditions and subject groups, and Table 3 reports the associated group means and variability. NH SAM-EFR magnitudes were significantly smaller than for the other conditions (paired t-tests between SAM-EFR and the other conditions showed $p<0.001$ for all tests). Despite the sensitivity of this metric to synaptopathy (model simulations

and Parthasarathy et al., 2018), the distribution of SAM-EFR data-points showed considerable overlap across the three subject groups and limits its potential for individual diagnosis in humans. NH EFR magnitudes to the other stimulus conditions were overall larger and showed a larger spread around the mean (Table 3), which emphasizes individual EFR differences and can hence improve diagnostic interpretation. Specifically, the RAM25 magnitudes showed non-overlapping interquartile ranges between the groups and demonstrates that this metric is more sensitive to capturing age-related SNHL aspects than the reference SAM-EFR. The overlapping SAM-EFR responses between yNH and oNH groups were not observed for the RAM25-EFR, consistent with the simulations which showed that synaptopathy had a greater effect on the latter metric. At the same time, the RAM-EFR was less influenced by individual OHC damage differences. Experimental support for this statement follows from the observation that the EFR group means were larger between the yNH and oNH/oHI groups (t=4.91 and 8.38 with $p<0.00005$ in both cases) than between the oNH and oHI group (t=3.70 and p=0.001), whereas the 4-kHz hearing threshold differences between the groups (Fig.2) were larger between the oNH and oHI group (t=12.48, p=0) than between the yNH and oNH group (t=6.04; p=0). Lastly, in agreement with the ANF simulations of Fig.1, the longer duty-cycle RAM50 EFRs had significantly smaller magnitudes than the RAM25 EFR (NH ptp/rms condition, t=6.87/8.65, p=0).

Lastly, it is of interest to study whether the stronger RAM vs SAM EFR

magnitudes were caused by the broader tonotopic excitation of the RAM stimulus or whether the shape of the temporal envelope was most efficient at eliciting a stronger synchronized neuronal response. To this end, we compared both EFR magnitudes those of an EFR evoked by an 70-dB-SPL SAM broadband white noise carrier (Fig. 5a; $SAM_{BB}$). $SAM_{BB}$ magnitudes were on average larger than reference SAM magnitudes (NH: t=6.61, p=0), but were smaller than the RAM25 EFRs (NH: t=7.25, p=0; see also Table 3). This observation confirms that, even for a carrier with a broad tonotopic excitation, SAM EFRs were limited by the neuronal/synaptic saturation properties in response to the SAM envelope shape. Stimulus envelopes which are optimized to enhance synchronous ANF firing can result in stronger EFRs, even for narrow-band carriers. Comparison between RAM-EFRs elicited by same-rms and/or same-peak-equivalent SPLs only showed marginal magnitude differences, in line with the model predictions (Fig. 4a, open symbols). Given the small level-differences between the two stimulation paradigms (less than 2 dB SPL), there was no observed benefit of using one over the other.

*Individual EFR differences and their relationship to different SNHL aspects*

Figure 6 depicts the relationship between individual EFR magnitudes and 4-kHz DPOAE thresholds across all subjects, and shows mean EFR magnitudes for groups separated by their 4-kHz audiometric threshold (yNH/oNH $<$ 25 dB HL $<=$ oHI, Table2) or by their age (yNH $<$ 30 y/o and oNH/oHI $>$ 60 y/o). The group data in Figure 6a shows that it was not possible to

discriminate between younger and older participants on the basis of the SAM-EFR magnitude. The same conclusion is drawn for the normal or elevated audiometric thresholds (4-kHz) groups. This suggest that the SAM-EFR is a weak response, and is affected by both age-related and OHC-damage aspects of SNHL, which corroborates the model predictions showing that both synaptopathy and OHC damage can affect the SAM-EFR. Despite overall stronger H10AM EFR magnitudes (Fig. 6b), they were similarly unable to segregate groups of young/old age or normal/elevated threshold listeners. Differently, the RAM-EFR magnitudes (Fig. 6c,d) were able to separate listeners into groups of younger or older listeners, demonstrating that this condition is more susceptible to the age-related aspect of SNHL than to the OHC damage aspect of SNHL. This latter statement is further supported by the observation that the separation into groups of younger/older listeners was better on the basis of the RAM-EFR magnitude than on the basis of the DPOAE threshold. Even though OHC and age-related SNHL deficits can coexist in older listeners, the RAM-EFR stimulus was more sensitive at isolating the age-related aspect than the other considered stimuli. These results corroborate our model simulations which show a near-differential sensitivity of the RAM-EFR to synaptopathy, and a mixed sensitivity of the SAM, H10AM-EFR to OHC damage and synaptopathy. Because at the same time, animal studies of age-related and histologically-verified synaptopathy show reduced EFRs (Sergeyenko et al., 2013; Fernandez et al., 2015; Möhrle et al., 2016b; Parthasarathy & Kujawa, 2018), we have strong circumstantial evidence that

the RAM-EFR magnitude was successful at separating listeners into groups with and without age-related auditory de-afferentation, including cochlear synaptopathy.

To further quantify the association between age-related SNHL factors (e.g. cochlear synaptopathy and OHC damage) and individual EFR magnitudes, we constructed a linear regression model of the form: $\mathrm{EFR}_{\mathrm{PtN}} = \beta_0 + \beta_1 * \mathrm{Age} + \beta_2 * \mathrm{DPOAE}_{@4\mathrm{kHz}}$, including data of all participants (N=44) and decomposed the models' $R^2$ into commonality coefficients (using the statistical package of the R programming language and environment; R Core Team, 2019; Nimon et al., 2008). The results are summarized in Table1.

Significant relationships between the EFR magnitude and predictive variables age and $\mathrm{DPOAE}_{@4\mathrm{kHz}}$ shared approximately half of the total explained variance across condition ("common"). However, OHC damage (as reflected in the $\mathrm{DPOAE}_{@4\mathrm{kHz}}$ threshold) showed the lowest contribution to the EFR magnitude for all conditions. When the variance of age-related factors are accounted for, the unique $\mathrm{DPOAE}_{@4\mathrm{kHz}}$ contribution became negligible. This commonality analysis shows that -in case of co-existence of OHC damage and synaptopathy-, the observed relationships between the EFR magnitude and $\mathrm{DPOAE}_{@4\mathrm{kHz}}$ are merely driven by the co-existing synaptopathy aspect of SNHL (which the RAM-stimuli are especially sensitive to).

Laslty, it is worthwhile mentioning that a multiple regression model which uses the commonly-adopted magnitude metric for the SAM-EFR stimulation as the explaining variable, i.e the spectral peak $F_n$ at fundamental frequency

Table 1: Results for the multiple regression model and commonality analysis

| Stimulus | $R^2$ | Adj. $R^2$ | p-value | beta | unique† | common† |
|---|---|---|---|---|---|---|
| SAM | 0.314 | 0.281 | 0.0004 | $\beta_1$=-0.0005* | 26.60 | 64.47 |
| | | | | $\beta_2$=-0.0003 | 8.93 | |
| $H10AM_{rms}$ | 0.220 | 0.182 | 0.0061 | $\beta_1$=-0.0010* | 52.89 | 47.01 |
| | | | | $\beta_2$=-0.0001 | 0.10 | |
| $RAM50_{ptp}$ | 0.606 | 0.587 | <0.0001 | $\beta_1$=-0.0013*** | 41.98 | 56.06 |
| | | | | $\beta_2$=-0.0004 | 1.96 | |
| $RAM25_{ptp}$ | 0.585 | 0.565 | <0.0001 | $\beta_1$=-0.0022*** | 38.18 | 58.64 |
| | | | | $\beta_2$=-0.0008 | 3.18 | |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
† % of $R^2$

$f_0$ (Fig. 4), did not reach significance (multiple $R^2$ = 0.073, adjusted $R^2$ = 0.028, p-value = 0.2127). The weak SAM-EFR response, the NF confounds which affect $f_N$ differently across listeners, and the overall underestimation of stimulus-envelope related energy in the EEG spectrum (i.e. due to omission of harmonics), are together responsible for this outcome and stress the need to use optimized stimuli and analysis methods.

*Comparison between EFR magnitudes and ABR amplitudes*

Figure 5b compares the SAM and RAM25 EFRs to features derived from low-rate (10 Hz) ABRs recorded in the same listeners. ABR amplitudes were defined as half the peak-to-through amplitude, to allow a fair comparison to the EFR magnitudes, which were calculated as half the reconstructed EFR waveform amplitude (Eq.5). ABR W-I and W-V amplitudes were calculated between corresponding positive peak and subsequent negative through according to the widely adopted approach (see Picton, 2010). In agreement with normative data (Picton 2010), NH ABR amplitudes followed the expected trend with (on average) smaller W-I than W-V amplitudes and larger amplitudes with increasing SPL (see also Table 4). Overall, both oNH and oHI groups showed reduced group medians in comparison to yNH W-I amplitudes. However, the interquartile ranges overlapped and only showed significant differences between yNH and oHI W-I amplitudes for the 70-dB-peSPL condition ($t = 4.48$, $p = 0.02$). Similar trends were also observed for the 100-dB-peSPL W-I amplitude (see Table4), which is a widely adopted marker for synaptopathy screening as high-level transients can evoke a synchronous ANF response across a broad tonotopic region and different fiber types (Liberman, 1978). In our recordings, the 100-dB-peSPL W-I amplitude was able to separate the yNH and oNH/oHI groups ($t=3.07$, $p=0.004$ and $t=4.31$, $p=0$), but not the oNH and oHI groups, consistent with the view that this marker was able to detect age-related synaptopathy aspects on a group level. The ABR W-V amplitudes were overall larger, but still

showed considerable overlap between the groups. The 70-dB-peSPL W-V was able to separate the yNH from the oNH/oHI groups (t=3.48, p= 0.002 and t=4.6, p=0), and the 100-dB-peSPL W-V condition was able to separate the yNH from oHI groups (t=2.76, p=0.01), but not the yNH from the oNH listeners. Comparing ABRs to EFRs recorded from the same listeners, it is clear that the latter were characterized by a reduced magnitude distribution in comparison to the ABRs (e.g., compare their interquartile ranges, especially for oNH and oHI groups). EFRs were computed based on automatic procedures, and corrected for by the individual NF, which may partly explain this observation. At the same time, the RAM25-EFRs had significantly overall larger magnitudes than the SAM-EFR (t=5.13, p=0, n=44) and comparable magnitudes to the ABRs. Despite similar RAM25-EFR and ABR amplitudes, the group means were much more separated in the EFR than ABR condition. At the same time, there is a difference in their sensitivity to different aspects of SNHL. A recent modeling study showed that both synaptopathy and OHC can reduce the generator strength of the ABR (Verhulst et al., 2016; Vasilkov & Verhulst, 2019), whereas the present study shows that the RAM-EFR is maximally sensitive to synaptopathy. Taken together, we conclude that the RAM25-EFR has a better sensitivity than the ABR in identifying the age-related (synaptopathy) aspect of SNHL.

## Discussion

We adopted a combined computational and experimental approach to investigate which EFR paradigms and analysis methods would enhance its sensitivity to isolate the synaptopathy aspect of SNHL in listeners who may have mixed OHC-damage/synaptopathy pathologies. The modeling work incorporates our latest knowledge of the physiology of hearing and hearing damage, and the predicted the outcomes of the experimental study well. Even though we did not have access to animal physiology methods in this study, our approach strongly supports the use of the RAM25 stimulus as a sensitive marker for synaptopathy in humans.

*The effect of stimulus envelope characteristics on the EFR*

The results of this study revealed that short duty-cycles can yield stronger EFR magnitudes due to the longer IPI and more synchronized ANF responses per each cycle. These observations corroborate the experimental findings of Dreyer & Delgutte (2006) which show stronger ANF responses for transposed-vs-SAM tones. Transposed tones have envelope shapes which rise faster than SAM envelopes, and even though they are not as sharp as the RAM stimuli considered here, single-unit ANF responses show the same trends as observed here. To further explore how the IPI and envelope maximum duration affects the EFR magnitude, we simulated EFRs for modulation rates of 10 Hz (Fig.7a) while changing the duty cycle from 0.2 to 25%. To study the respective effects of rate and duty-cycle, the 120-Hz modulation rate for dif-

ferent duty-cycles was used as the reference (Fig.7b), after which the IPI was changed to 100 ms off-time to simulate the effect of lowering the repetition rate to 10 Hz.

In agreement with how ABRs to click trains with low repetition rates and long IPIs (10 Hz) yield robust ABRs (Picton, 2010), simulated 10-Hz AM EFRs evoked stronger responses compared to the 120-Hz condition for all duty-cycles (Fig.7a vs b). At the same time, stimuli of both repetition rates evoked substantially reduced EFRs when the duty-cycle reduced from 25% to 0.2% (where 25% corresponds to the experimental RAM25 condition). These simulations suggest that one of the factors responsible for weak EFRs could relate to the lower amount of sound energy carried in each short duty-cycle, which might compromize a synchronized and robust ANF response. Moreover, large duty-cycles (e.g. 50% as used in RAM50) were also shown to evoke reduced EFRs when compared to the 25% duty-cycle (Fig. 1g,k vs f,j). The recordings even showed a second response peak per duty-cycle for the RAM50, which may have compromized the synchronous response to the stimulus envelope frequency. The experimental results and model simulations hence suggest that the IPI (which determines the silence interval between the stimulus peaks) was more important than the duration of the duty-cycle to yield robust EFR responses. However, too short, or too long, pulse durations can also compromise the stimulation efficiency, and hence point to a sweet-spot duty-cycle of 15-25% for most efficient EFR stimulation.

Furthermore, stimuli with short (click-like) duty-cycles, which contain limited sound energy per cycle as a result, evoked responses that were strongly influenced by the OHC aspect of SNHL. This aspect is observed in Fig.6b, which shows a greater influence of cochlear gain loss when the duty-cycle reduced to 0.2%. Figure 7a furthermore shows that the overall stimulation rate has an effect as well: the 10-Hz rate condition was more affected by cochlear gain loss than was the 120-Hz condition. ANF responses evoked by transient stimuli presented at low repetition rates might be strongly affected by the active components of the BM impulse response, and hence strongly be attenuated when OHC dysfunction occurs. Differently, for the 120-Hz AM stimuli, responses stem from-more saturated ANF responses (e.g., 0.129 $\mu$V and 0.075 $\mu$V reduction for Sq0.5 presented with 10 Hz and 120 Hz, respectively) and are hence more sensitive to the number of intact ANFs than to OHC dysfunction (Fig. 1e-l). To support the predominantly saturated ANF response origin of the 120-Hz condition, simulated EFRs (both SAM and RAM25) with duty-cycles greater than 10% showed slightly elevated EFR magnitudes for high-frequency OHC damage (Fig. 7b, NH to $\mathrm{HI}_{\mathrm{OHC:10@4K}}$). These enhanced EFR magnitudes stemmed from the linearized (basal) cochlear responses and attenuated BM input to the ANF in the IPI, which pulls the ANFs out of saturation to yield a stronger modulated response (Joris & Yin, 1992).

*Effect of the modulation frequency on EFR strength*

EFRs evoked by AM signals are known to vary in magnitude depending on the modulation rate (Purcell et al., 2004; Parthasarathy et al., 2018). There are likely several aspects underlying this variation, including cochlear processing, neuronal properties, and different neuronal generators which contribute to the population responses (Joris et al., 2004; Purcell et al., 2004; Picton, 2010). The latter aspect can be used to steer the focus of EFR screening towards specific neuronal structures. It is assumed that cortical sources dominate the EFR to AM stimuli with modulation rates below 50-100 Hz, but that more peripheral (subcortical) neurons can follow higher modulation rates (Purcell et al., 2004; Herdman et al., 2002; Bidelman, 2015, 2018). Consequently, it might be possible to render the EFR more sensitivity to central or peripheral aspects of the auditory functioning by de- or increasing the modulation frequency, respectively. However, in humans, the factor which limits the upper modulation frequency limit is the robustness of the recorded signal which substantially starts declining for modulation rates above 60-70 Hz and becomes statistically indistinguishable from the background noise for modulation rates above 250 Hz (Purcell et al., 2004; Picton, 2010; Garrett & Verhulst, 2019).

To further explore the relationship between EFR strength and the stimulus modulation rate, we simulated the envelope-locking limit for the RAM25 condition which had the largest median response magnitude (Fig. 4a). Figure 8 shows simulated EFR magnitudes for modulation frequencies between

120 and 600 Hz. The two filled symbols correspond to the median NH EFR magnitudes for the SAM and RAM25 stimuli. The simulations show that the insensitivity of the RAM25 to possibly co-occurring OHC deficits can be further increased by increasing the modulation frequency to higher modulation rates, e.g. at the 200 Hz modulation rate, there was no influence of OHC damage, while the RAM25 stimulus still evoked a 1.8 times larger response than for the reference SAM stimulus. However, this increased differential sensitivity to synaptopathy happens at the cost of overall reduced RAM25 magnitudes which might compromise the sensitivity of the metric toward detecting individual synaptopathy differences. A compromise which takes both of these aspects into account might thus be ideal. If we take the NH SAM EFR as an acceptable reference magnitude, RAM EFRs with modulation frequencies up to 240 Hz can be adopted to yield the same response sensitivity.

*Implications for diagnostic applications*

Finding an AEP-based metric which is differentially sensitive to the synaptopathy aspect of SNHL, even when OHC deficits are present, is an important pursuit which requires a multi-center and interdisciplinary approach. On the one hand, there is compelling evidence from animal studies that the ABR wave-I and SAM EFRs are compromised after histologically-verified synaptopathy (Kujawa & Liberman, 2009; Bourien et al., 2014; Sergeyenko et al., 2013; Shaheen et al., 2015; Möhrle et al., 2016a; Chambers et al.,

2016; Parthasarathy et al., 2018), but on the other, little is known about the respective roles of OHC and synaptopathy aspects in this degradation. Animal studies of synaptopathy often focus on individual ABR/EFR markers after controlled synaptopathy-induction (e.g. quietly-raised animals with age effects, ototoxic-induced) and most human studies focus on clinically normal-hearing subjects (e.g., Kujawa & Liberman, 2009; Mehraei et al., 2016; Prendergast et al., 2017a; Guest et al., 2017). Lastly, human AEP studies on listeners with impaired audiograms have the drawback that it is presently not possible to connect the individual histopathology to AEP alterations. To bridge this translational gap, model-based approaches can have a pivotal role. Even though model studies are limited by the quality of the used model, they can be effective in narrowing down the parameter space of potentially sensitive AEP markers. Promising candidate markers which can afterwards be tested more efficiently in experiments with human and animal models. Despite the known limitations of model approaches, there is a present absence of controlled experimental approaches which vary the degree of OHC and synaptopathy damage in a controlled way. In the meantime, models are the only available tools which can be used to study the relative contribution of OHC damage and synaptopathy on the source generators of human ABRs and EFRs.

Despite the theoretical starting-point we took in the design of our stimulus set, the model was able to validate our single-unit predictions, provided confidence that the associated population EFR responses would follow these

trends, while showing differential sensitivity to either OHC or synaptopathy aspects. Even though the model simulates the functional signal representation along the auditory pathway, and does not perfectly incorporate all different brainstem neuron types, it has previously shown its merit at reasonably and collectively simulating the level-dependence of human OAEs, ABRs, EFRs while accounting for the level-dependence and adaptation properties of single unit-ANF fibers (Altoè et al., 2018; Verhulst et al., 2012, 2015, 2018a; Keshishzadeh et al., 2020). It is hence not incidental that the predicted changes in EFR strength due to stimulus envelope changes or their sensitivity to different aspects of SNHL were confirmed experimentally.

At the same time, our study showed that RAM25 EFRs were more effective at identifying age-related SNHL differences than either the SAM EFR or the ABR waves, which can advance the field in a couple of ways. First, overall stronger EFRs improve their application range towards listeners with more severe SNHL pathologies. Whereas the SAM EFRs did not show group differences between the oNH and oHI listeners, the RAM EFR was able to do this, and can hence offer a more fine-grained estimate of the degree of synaptopathy in listeners with normal or impaired audiograms. Secondly, both the model simulations and communality analysis showed that the RAM EFR is more sensitive to (age-related) synaptopathy and less sensitive possibly co-existing OHC damage aspects. Even though this finding should ideally be confirmed in animal histopathology studies of synaptopathy and OHC damage, it is clear that when confirmed, the RAM25 EFR might help

therapeutic interventions or studies which aim to study the perceptual consequences of synaptopathy. Aside from the Bharadwaj et al. (2015); Guest et al. (2017) studies which use transposed-tone EFRs with sharp envelopes, most human studies use the smaller ABR wave-I amplitude or SAM-EFR marker to study how individual physiological responses relate to sound perception. This means that the ABR wave-I and SAM-EFRs in those studies could have been troubled by response analysis (i.e. $f_0$ vs $f_0$+harmonics vs noise floor correction) and confounding OHC damage factors (see also Verhulst et al., 2016), which introduced an inherent variability in the presumed physiological markers of synaptopathy. It would be worthwhile to re-analyse the SAM-EFRs of those studies using the proposed analysis method, and adopt RAM stimuli for future human synaptopathy studies in which co-existing OHC damage is possible. Using a sensitive and differential marker of synaptopathy is essential to enable a causal relationship between the origin of the marker reduction (synaptopathy) and its impact on sound perception. It is known that age can result both in OHC damage (ISO 7029) and synaptopathy (Sergeyenko et al., 2013; Parthasarathy et al., 2018), and hence people with OHC damage may also suffer from synaptopathy. This means that when audiometric thresholds predict outcomes on psychoacoustic tasks in ageing studies, we should leave the possibility open that the co-existing synaptopathy aspect of SNHL was responsible for driving the reduction in task performance.

## Conclusion

We adopted a combined theoretical and human experimental approach to develop AEP-based stimuli which showed enhanced sensitivity to the (age-related) synaptopathy aspect of SNHL. We conclude that supra-threshold RAM stimuli with duty cycles between 20-25% for a 120-Hz modulation rate, are maximally efficient to both yield a strong response magnitude and a differential sensitivity to the synaptopathy aspect of SNHL. RAM25 amplitudes were considerably larger than commonly-used SAM-EFR markers of synaptopathy, and showed more pronounced age-related differences than ABR markers. Improving the analysis method to include the harmonics and perform a noise-floor correction further improved the robustness of the RAM-EFR. Taken together, we hope that the outcomes of this theoretical-experimental study will improve the interpretation possibilities of future studies aimed at studying the role of synaptopathy/deafferentation for sound perception and will yield a set of robust and sensitive markers of cochlear synaptopathy for use in animal and human studies.

## Author Contributions

VV: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing: Original Draft, Visualization; MG: Methodology, Software, Investigation; MM: Methodology, Software, Investigation; SV: Conceptualization, Methodology, Resources, Writing: Orig-

inal Draft, Writing: Review & Editing, Supervision, Project administration, Funding acquisition.

## Acknowledgements

This work was supported by the German Research Foundation (PP1608, VE924/1-1; VV, SV), the European Research Council (ERC-StG-678120, RobSpear; VV, SV) and DFG Cluster of Excellence (EXC 1077/1, "Hearing4all"; MG, MM). The authors would like to thank the study participants as well as the Oldenburg Hörzentrum for helping with their recruitment.

## Conflict of interest statement

Ghent University filed a patent application (PCTEP2020053192) which covers some of the ideas presented in this paper. Sarah Verhulst and Viacheslav Vasilkov are inventors.

## Data availability

The model code used for the simulations is available via 10.5281/zenodo.3717800 or github.com/HearingTechnology/Verhulstetal2018Model. Software to extract the EFR magnitudes from the raw-EEG recordings as well as code to generate the RAM stimuli will be made available on github.com/HearingTechnology upon the publication of the manuscript.

Altoè, A., Pulkki, V., & Verhulst, S. (2014). Transmission line cochlear models: improved accuracy and efficiency. *The Journal of the Acoustical Society of America*, *136*, EL302–EL308.

Altoè, A., Pulkki, V., & Verhulst, S. (2018). The effects of the activation of the inner-hair-cell basolateral k+ channels on auditory nerve responses. *Hearing research*, *364*, 68–80.

Bernstein, L. R., & Trahiotis, C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli". *The Journal of the Acoustical Society of America*, *112*, 1026–1036.

Bernstein, L. R., & Trahiotis, C. (2009). How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli. *The Journal of the Acoustical Society of America*, *125*, 3234–3242.

Bharadwaj, H. M., Mai, A. R., Simpson, J. M., Choi, I., Heinz, M. G., & Shinn-Cunningham, B. G. (2019). Non-invasive assays of cochlear synaptopathy–candidates and considerations. *Neuroscience*, *407*, 53–66.

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *Journal of Neuroscience*, *35*, 2161–2172.

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., & Shinn-

Cunningham, B. G. (2014). Cochlear neuropathy and the coding of supra-threshold sound. *Frontiers in systems neuroscience*, *8*, 26.

Bidelman, G. M. (2015). Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient abr. *Hearing research*, *323*, 68–80.

Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage*, *175*, 56–69.

Boege, P., & Janssen, T. (2002). Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission I/O-functions in normal and cochlear hearing loss ears. *The Journal of the Acoustical Society of America*, *111*, 1810–1818. doi:10.1121/1.1460923.

Bourien, J., Tang, Y., Batrel, C., Huet, A., Lenoir, M., Ladrech, S., Desmadryl, G., Nouvian, R., Puel, J.-L., & Wang, J. (2014). Contribution of auditory nerve fibers to compound action potential of the auditory nerve. *Journal of neurophysiology*, *112*, 1025–1039.

Bramhall, N., Beach, E., Epp, B., LePrell, C. G., Lopez-Poveda, E. A., Plack, C., Schaette, R., Verhulst, S., & Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing research*, .

Brunner, C., Delorme, A., & Makeig, S. (2013). Eeglab–an open source

matlab toolbox for electrophysiological research. *Biomedical Engineering/Biomedizinische Technik*, *58*.

Chambers, A. R., Resnik, J., Yuan, Y., Whitton, J. P., Edge, A. S., Liberman, M. C., & Polley, D. B. (2016). Central gain restores auditory processing following near-complete cochlear denervation. *Neuron*, *89*, 867–879.

D'haenens, W., Vinck, B. M., De Vel, E., Maes, L., Bockstael, A., Keppler, H., Philips, B., Swinnen, F., & Dhooge, I. (2008). Auditory steady-state responses in normal hearing adults: a test-retest reliability study. *International journal of audiology*, *47*, 489–498.

Dobie, R. A., & Humes, L. E. (2017). Commentary on the regulatory implications of noise-induced cochlear neuropathy. *International journal of audiology*, *56*, 74–78.

Dolphin, W., & Mountain, D. (1992). The envelope following response: scalp potentials elicited in the mongolian gerbil using sinusoidally am acoustic signals. *Hearing research*, *58*, 70–78.

Dreyer, A., & Delgutte, B. (2006). Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: implications for sound localization. *Journal of neurophysiology*, *96*, 2327–2341.

Fernandez, K. A., Jeffers, P. W., Lall, K., Liberman, M. C., & Kujawa, S. G. (2015). Aging after noise exposure: acceleration of cochlear synaptopathy in "recovered" ears. *Journal of Neuroscience*, *35*, 7509–7520.

Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of neurophysiology*, *110*, 577–586.

Garrett, M., Debener, S., & Verhulst, S. (2019). Acquisition of subcortical auditory potentials with around-the-ear ceegrid technology in normal and hearing impaired listeners. *Frontiers in neuroscience*, *13*, 730.

Garrett, M., & Verhulst, S. (2019). Applicability of subcortical eeg metrics of synaptopathy to older listeners with impaired audiograms. *Hearing research*, *380*, 150–165.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L. et al. (2013). Meg and eeg data analysis with mne-python. *Frontiers in neuroscience*, *7*, 267.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). Mne software for processing meg and eeg data. *Neuroimage*, *86*, 446–460.

Greenberg, D., Monaghan, J. J., Dietz, M., Marquardt, T., & McAlpine, D. (2017). Influence of envelope waveform on itd sensitivity of neurons in the auditory midbrain. *Journal of neurophysiology*, *118*, 2358–2370.

Greenwood, D. D. (1990). A cochlear frequency-position function for several

species—29 years later. *The Journal of the Acoustical Society of America*, *87*, 2592–2605.

Griffin, S. J., Bernstein, L. R., Ingham, N. J., & McAlpine, D. (2005). Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig. *Journal of neurophysiology*, *93*, 3463–3478.

Guest, H., Munro, K. J., & Plack, C. J. (2019). Acoustic middle-ear-muscle-reflex thresholds in humans with normal audiograms: No relations to tinnitus, speech perception in noise, or noise exposure. *Neuroscience*, *407*, 75–82.

Guest, H., Munro, K. J., Prendergast, G., Howe, S., & Plack, C. J. (2017). Tinnitus with a normal audiogram: Relation to noise exposure but no evidence for cochlear synaptopathy. *Hearing Research*, *344*, 265–274.

Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018). Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing research*, *364*, 142–151.

Herdman, A. T., Lins, O., Van Roon, P., Stapells, D. R., Scherg, M., & Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain topography*, *15*, 69–86.

Hickox, A. E., Larsen, E., Heinz, M. G., Shinobu, L., & Whitton, J. P.

(2017). Translational issues in cochlear synaptopathy. *Hearing research*, *349*, 164–171.

Johannesen, P. T., Buzo, B. C., & Lopez-Poveda, E. A. (2019). Evidence for age-related cochlear synaptopathy in humans unconnected to speech-in-noise intelligibility deficits. *Hearing research*, *374*, 35–48.

John, M., Dimitrijevic, A., & Picton, T. (2003). Efficient stimuli for evoking auditory steady-state responses. *Ear and hearing*, *24*, 406–423.

John, M. S., Dimitrijevic, A., & Picton, T. W. (2002). Auditory steady-state responses to exponential modulation envelopes. *Ear & Hearing*, *23*, 106–117.

Joris, P., Schreiner, C., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological reviews*, *84*, 541–577.

Joris, P. X., & Yin, T. C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America*, *91*, 215–232.

Keshishzadeh, S., Garrett, M., Vasilkov, V., & Verhulst, S. (2020). The derived-band envelope following response and its sensitivity to sensorineural hearing deficits. *Hearing Research*, (p. 107979).

Klein-Hennig, M., Dietz, M., Hohmann, V., & Ewert, S. D. (2011). The influence of different segments of the ongoing envelope on sensitivity to

interaural time delays. *The Journal of the Acoustical Society of America*, *129*, 3856–3872.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. *Journal of Neuroscience*, *29*, 14077–14085.

Kummer, P., Janssen, T., & Arnold, W. (1998). The level and growth behavior of the 2 f1-f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss. *The Journal of the Acoustical Society of America*, *103*, 3431–3444. doi:10.1121/1.423054.

Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., & D'Angelo, W. R. (2002). Sources of the scalp-recorded amplitude-modulation following response. *Journal of the American Academy of Audiology*, *13*, 188–204.

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W.-D., & Pok, S.-M. (2011). Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *130*, 1515–1529.

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

Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016). Toward a differential diagnosis of hidden hearing loss in humans. *PloS one*, *11*.

Lin, H. W., Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2011). Primary neural degeneration in the guinea pig cochlea after reversible noise-induced threshold shift. *Journal of the Association for Research in Otolaryngology*, *12*, 605–616.

Lobarinas, E., Salvi, R., & Ding, D. (2013). Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hearing research*, *302*, 113–120.

Long, G. R., Talmadge, C. L., & Lee, J. (2008). Measuring distortion product otoacoustic emissions using continuously sweeping primaries. *The Journal of the Acoustical Society of America*, *124*, 1613–1626. doi:10.1121/1.2949505.

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *Journal of the Association for Research in Otolaryngology*, *12*, 711–717.

Mauermann, M. (2013). Improving the usability of the distortion product otoacoustic emisssions (DPOAE)-sweep method: An alternative artifact

rejection and noise-floor estimation. *The Journal of the Acoustical Society of America*, *133*, 3376. doi:10.1121/1.4805803.

Mehraei, G., Hickox, A. E., Bharadwaj, H. M., Goldberg, H., Verhulst, S., Liberman, M. C., & Shinn-Cunningham, B. G. (2016). Auditory brainstem response latency in noise as a marker of cochlear synaptopathy. *Journal of Neuroscience*, *36*, 3755–3764.

Möhrle, D., Ni, K., Varakina, K., Bing, D., Lee, S. C., Zimmermann, U., Knipper, M., & Rüttiger, L. (2016a). Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain. *Neurobiology of aging*, *44*, 173–184.

Möhrle, D., Ni, K., Varakina, K., Bing, D., Lee, S. C., Zimmermann, U., Knipper, M., & Rüttiger, L. (2016b). Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain. *Neurobiology of aging*, *44*, 173–184.

Nelson, P. C., & Carney, L. H. (2004). A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America*, *116*, 2173–2186.

Nimon, K., Lewis, M., Kane, R., & Haynes, R. M. (2008). An r package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods*, *40*, 457–466.

Oxenham, A. J. (2016). Predicting the perceptual consequences of hidden hearing loss. *Trends in hearing*, *20*, 2331216516686768.

van de Par, S., & Kohlrausch, A. (1997). A new approach to comparing binaural masking level differences at low and high frequencies. *The Journal of the Acoustical Society of America*, *101*, 1671–1680.

Parthasarathy, A., Bartlett, E. L., & Kujawa, S. G. (2018). Age-related changes in neural coding of envelope cues: peripheral declines and central compensation. *Neuroscience*, .

Parthasarathy, A., & Kujawa, S. G. (2018). Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing. *Journal of Neuroscience*, (pp. 3240–17).

Picton, T. W. (2010). *Human auditory evoked potentials.*. San Diego, CA: Plural Publishing.

Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual consequences of "hidden" hearing loss. *Trends in hearing*, *18*, 2331216514550621.

Plack, C. J., Léger, A., Prendergast, G., Kluk, K., Guest, H., & Munro, K. J. (2016). Toward a diagnostic test for hidden hearing loss. *Trends in hearing*, *20*, 2331216516657466.

Prendergast, G., Guest, H., Munro, K. J., Kluk, K., Léger, A., Hall, D. A., Heinz, M. G., & Plack, C. J. (2017a). Effects of noise exposure on young

adults with normal audiograms i: Electrophysiology. *Hearing research*, *344*, 68–81.

Prendergast, G., Millman, R. E., Guest, H., Munro, K. J., Kluk, K., Dewey, R. S., Hall, D. A., Heinz, M. G., & Plack, C. J. (2017b). Effects of noise exposure on young adults with normal audiograms ii: Behavioral measures. *Hearing research*, *356*, 74–86.

Prendergast, G., Tu, W., Guest, H., Millman, R. E., Kluk, K., Couth, S., Munro, K. J., & Plack, C. J. (2018). Supra-threshold auditory brainstem response amplitudes in humans: Test-retest reliability, electrode montage and noise exposure. *Hearing Research*, *364*, 38–47.

Purcell, D. W., John, S. M., Schneider, B. A., & Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *The Journal of the Acoustical Society of America*, *116*, 3581–3593.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. URL: `https://www.R-project.org/`.

Schaette, R., & McAlpine, D. (2011). Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. *Journal of Neuroscience*, *31*, 13452–13457.

Schoof, T., & Rosen, S. (2016). The role of age-related declines in subcor-

tical auditory processing in speech perception in noise. *Journal of the Association for Research in Otolaryngology*, *17*, 441–460.

Schuknecht, H. F., & Woellner, R. C. (1955). An experimental and clinical study of deafness from lesions of the cochlear nerve. *The Journal of Laryngology & Otology*, *69*, 75–97.

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *Journal of Neuroscience*, *33*, 13686–13694.

Shaheen, L. A., Valero, M. D., & Liberman, M. C. (2015). Towards a diagnosis of cochlear neuropathy with envelope following responses. *Journal of the Association for Research in Otolaryngology*, *16*, 727–745.

Sheets, L. (2017). Excessive activation of ionotropic glutamate receptors induces apoptotic hair-cell death independent of afferent and efferent innervation. *Scientific reports*, *7*, 1–14.

Stürzebecher, E., Cebulla, M., & Neumann, K. (2003). Click-evoked abr at high stimulus repetition rates for neonatal hearing screening: Potenciales auditivos evocados (abr) por clicks a tasas altas de estimulación para la identificación de problemas auditivos en neonatos. *International journal of audiology*, *42*, 59–70.

Valero, M., Burton, J., Hauser, S., Hackett, T., Ramachandran, R., & Liber-

man, M. (2017). Noise-induced cochlear synaptopathy in rhesus monkeys (macaca mulatta). *Hearing research*, *353*, 213–223.

Van Canneyt, J., Hofmann, M., Wouters, J., & Francart, T. (2019). The effect of stimulus envelope shape on the auditory steady-state response. *Hearing research*, *380*, 22–34.

Vasilkov, V., & Verhulst, S. (2019). Towards a differential diagnosis of cochlear synaptopathy and outer-hair-cell deficits in mixed sensorineural hearing loss pathologies. *medRxiv*, (p. 19008680).

Vecchi, A. O., & Verhulst, S. (2019). Calibration and reference simulations for the auditory periphery model of verhulst et al. 2018 version 1.2. *arXiv preprint arXiv:1912.10026*, .

Verhulst, S., Altoè, A., & Vasilkov, V. (2018a). Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss. *Hearing research*, *360*, 55–75.

Verhulst, S., Bharadwaj, H. M., Mehraei, G., Shera, C. A., & Shinn-Cunningham, B. G. (2015). Functional modeling of the human auditory brainstem response to broadband stimulation. *The Journal of the Acoustical Society of America*, *138*, 1637–1659.

Verhulst, S., Dau, T., & Shera, C. A. (2012). Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission. *The Journal of the Acoustical Society of America*, *132*, 3842–3848.

Verhulst, S., Ernst, F., Garrett, M., & Vasilkov, V. (2018b). Suprathreshold psychoacoustics and envelope-following response relations: Normal-hearing, synaptopathy and cochlear gain loss. *Acta Acustica united with Acustica*, *104*, 800–803.

Verhulst, S., Jagadeesh, A., Mauermann, M., & Ernst, F. (2016). Individual differences in auditory brainstem response wave characteristics: relations to different aspects of peripheral hearing loss. *Trends in hearing*, *20*, 2331216516672186.

Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A., Santos, F., Merchant, S. N., Liberman, L. D., & Liberman, M. C. (2015). Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue. *Hearing research*, *327*, 78–88.

Wu, P., Liberman, L., Bennett, K., De Gruttola, V., O'Malley, J., & Liberman, M. (2018). Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. *Neuroscience*, .

Zhu, L., Bharadwaj, H., Xia, J., & Shinn-Cunningham, B. (2013). A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. *The Journal of the Acoustical Society of America*, *134*, 384–395. doi:10.1121/1.4807498.

## Figure captions

**Figure 1. a-d** Two cycles of the amplitude-modulated stimuli with different envelope shapes but the same modulation rate of 120 Hz. All stimuli were presented with the same RMS SPL (black) and stimuli with rectangular envelopes were additionally presented in an equal ptp amplitude to the reference SAM tone (cyan). **e-h** Simulated ANF responses at the 4-kHz CF evoked by the corresponding stimuli (equal-RMS). Solid traces depict responses summed across 19 AN fibers per IHC (i.e. intact ANF profile: 3 Low, 3 Medium, 13 High SR fibers) and dotted black lines represent summed responses from three fibers per IHC (i.e. severe synaptopathy $HI_{CS}$: 0 Low, 0 Medium, 3 High SR fibers). $HI_{OHC:10@4K}$ and $HI_{OHC:35@4K}$ traces represent responses for simulated sloping audiometric hearing loss with 10 dB or 35 dB threshold elevation at 4 kHz. NH shows the responses without simulated hearing deficits. **i-l** Time-domain representation of simulated and recorded EFRs in response to stimuli with different envelope shapes (same RMS). Open traces depict simulated EFRs for normal-hearing (NH), extreme synaptopathy ($HI_{CS:0L,0M,3H}$) and audiometric ($HI_{OHC:35@4K}$) profiles. Solid traces depict recorded EFRs averaged across young (yNH) and old (oNH) normal-hearing, and old hearing-impaired (oHI) participants groups.

**Figure 2. a** Pure-tone hearing thresholds measured at frequencies between 0.125 and 8 kHz. Dashed traces depict mean values across yNH, oNH and oHI groups. Solid lines represent simulated cochlear gain loss profiles

with the corresponding dB HL sloping hearing loss (NH, $HI_{OHC:10@4K}$ and $HI_{OHC:35@4K}$). **b** Pure-tone hearing thresholds and **c** distortion-product otoacoustic emission thresholds at 4 kHz.

**Figure 3.** Schematic of the adopted computational model of the auditory periphery which simulates subcortical sources of human AEPs in response to acoustic stimuli (Verhulst et al., 2018a; Vecchi & Verhulst, 2019).

**Figure 4.** Illustration of how the $EFR_{PtN}$ was computed from the raw EEG recordings. **a** Magnitude spectrum (gray) of the AEP recorded in response to $RAM25_{rms}$ stimulation and averaged within one bootstrap run. Red dash markers depict the estimated noise floor (NF) and blue vertical arrows indicate peak ($F_n$) to $NF_n$ magnitudes at the modulation frequency and its harmonics. **b** 100 ms-scaled time-domain representation of the recorded AEP (gray) and reconstructed time-domain EFR waveform (blue) based on noise corrected energy at the fundamental and available harmonic components ($F_n$-$NF_n$) with the corresponding phase angle values. The EFR magnitude was defined as half the peak-to-peak amplitude of the reconstructed signal in the time domain (blue arrow) for each bootstrap run.

**Figure 4.** Simulated (open symbols) and recorded (filled symbols) individual AEPs for different SNHL profiles and stimulus types. **a** $EFR_{PtN}$ magnitudes evoked by the sustained amplitude-modulated stimuli with dif-

ferent envelope shapes. **b** Comparison between EFRs and transient ABR waveform features to 10-Hz click trains of 70 and 100 dB-peSPL amplitudes.

**Figure 5.** Linear regression plots for EFR magnitudes evoked by AM stimuli with different envelope shapes and 4-kHz DPOAE thresholds. Top and left error bars indicate groups means and standard deviations of normal (yNH&oNH) and elevated (oHI) audiometric thresholds at 4 kHz groups (downward triangles) or young (yNH) and older (oNH&oHI) groups (upward triangles).

**Figure 6.** Simulated EFR magnitudes for 10 Hz (**a**) and 120 Hz (**b**) modulation rate RAM stimuli which had the same peak-to-peak amplitude as the reference 70-dB-SPL SAM tone. EFR magnitudes are shown for the NH model as well as for models with sloping cochlear gain loss ($HI_{OHC:10@4K}$ and $HI_{OHC:35@4K}$).

**Figure 7.** Simulated and recorded EFR magnitudes for reference SAM and RAM stimuli with different stimulus modulation frequencies. The duty-cycle was 25% for the RAM stimuli and the stimulus level was 70 dB SPL in all cases.

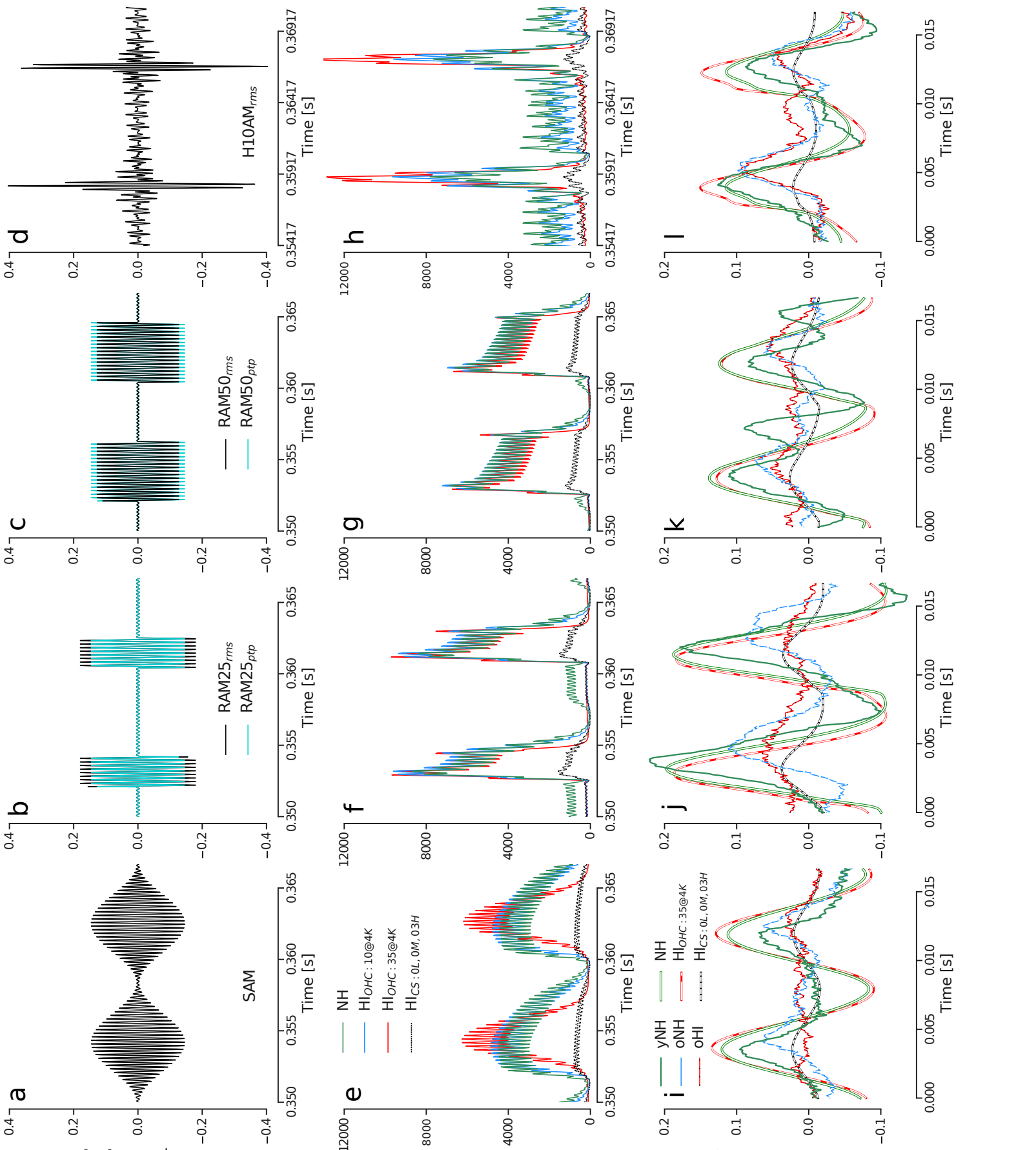


Figure 1:

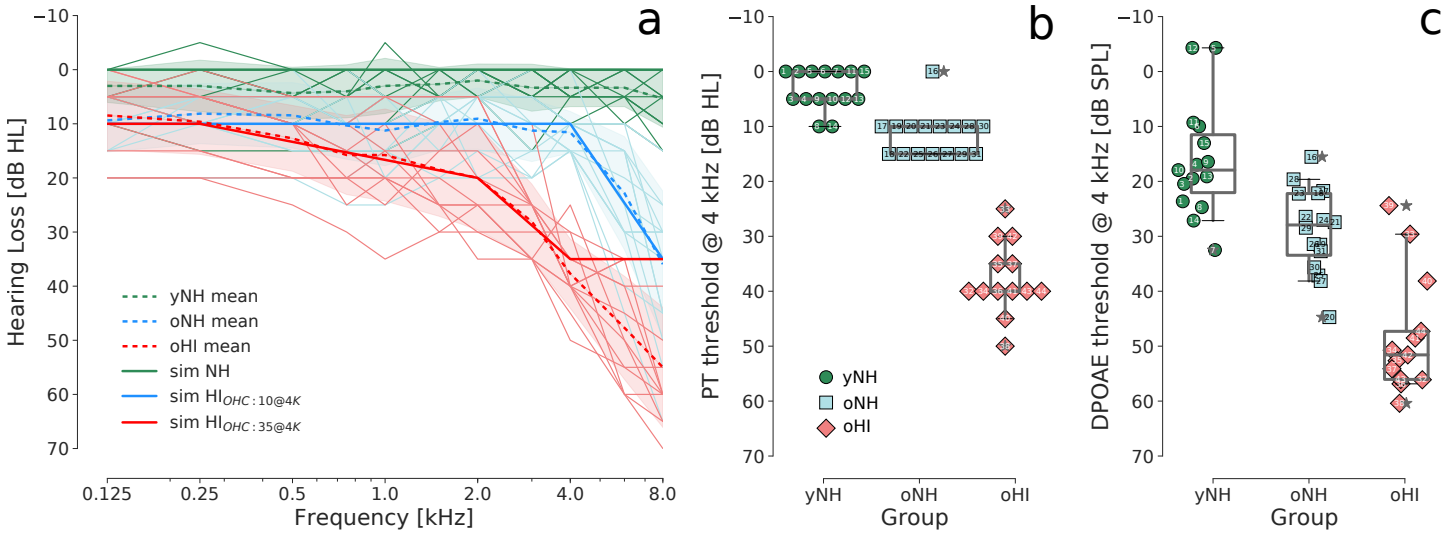


Figure 2:

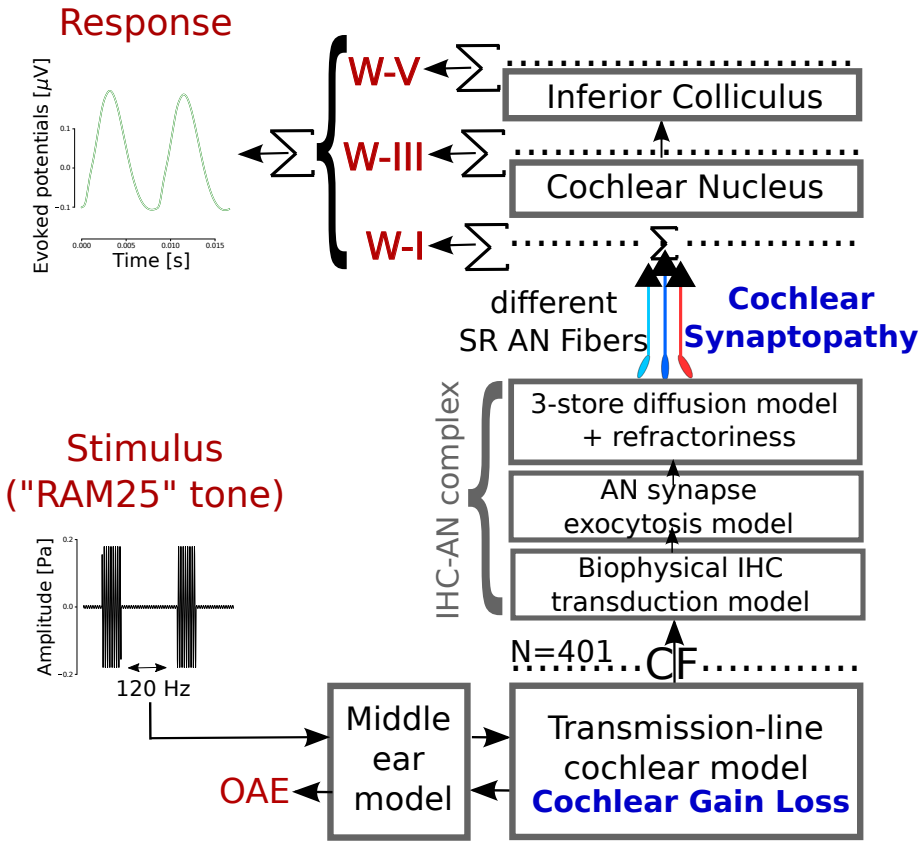


Figure 3:

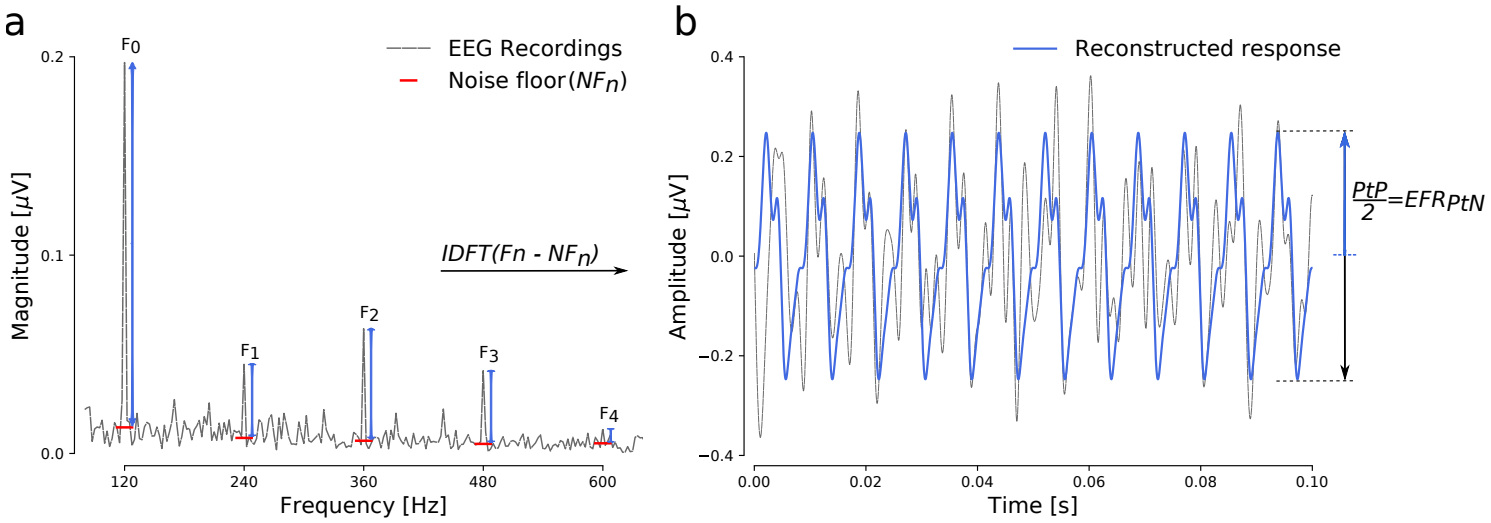


Figure 4:

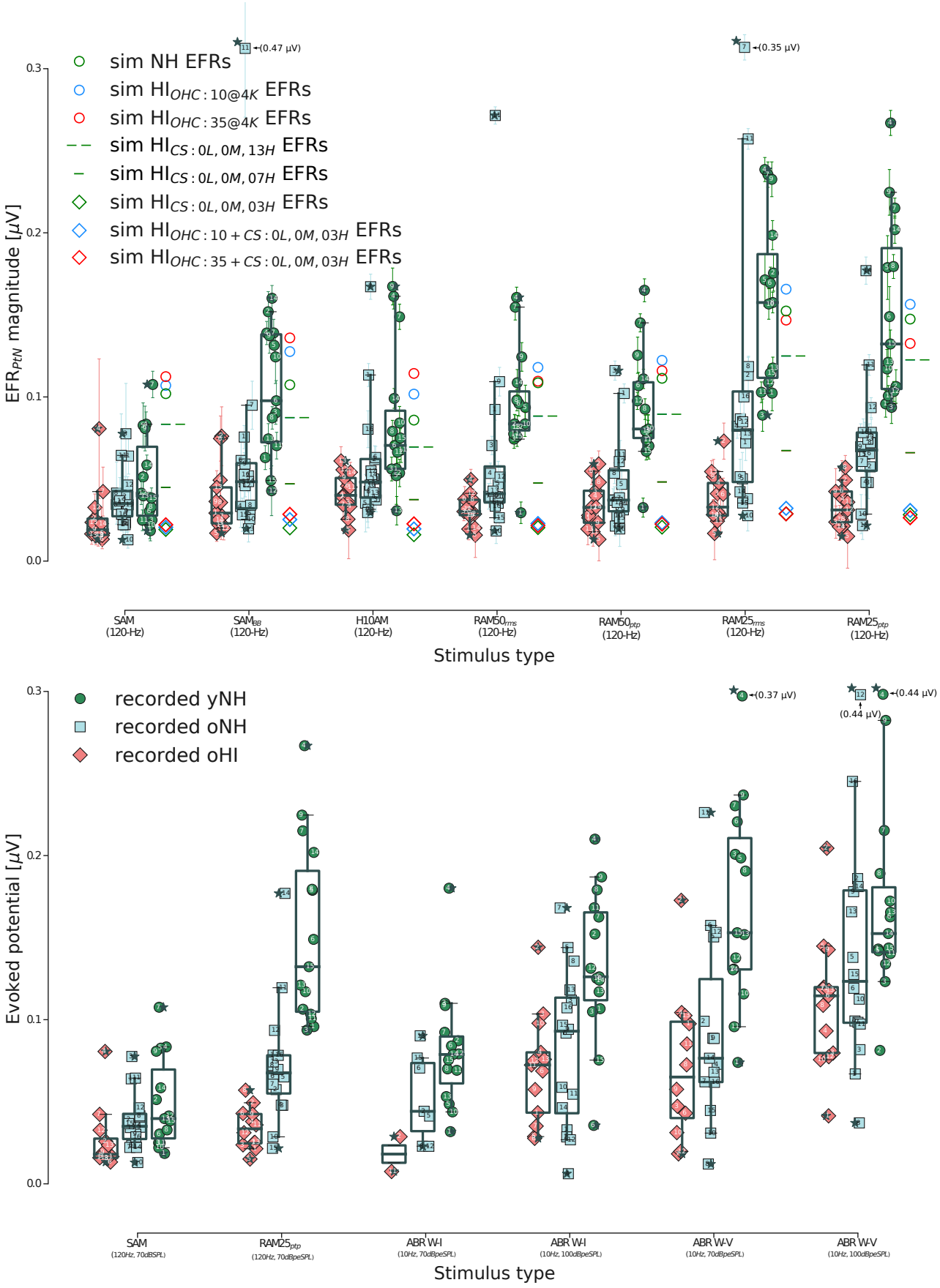


Figure 5:

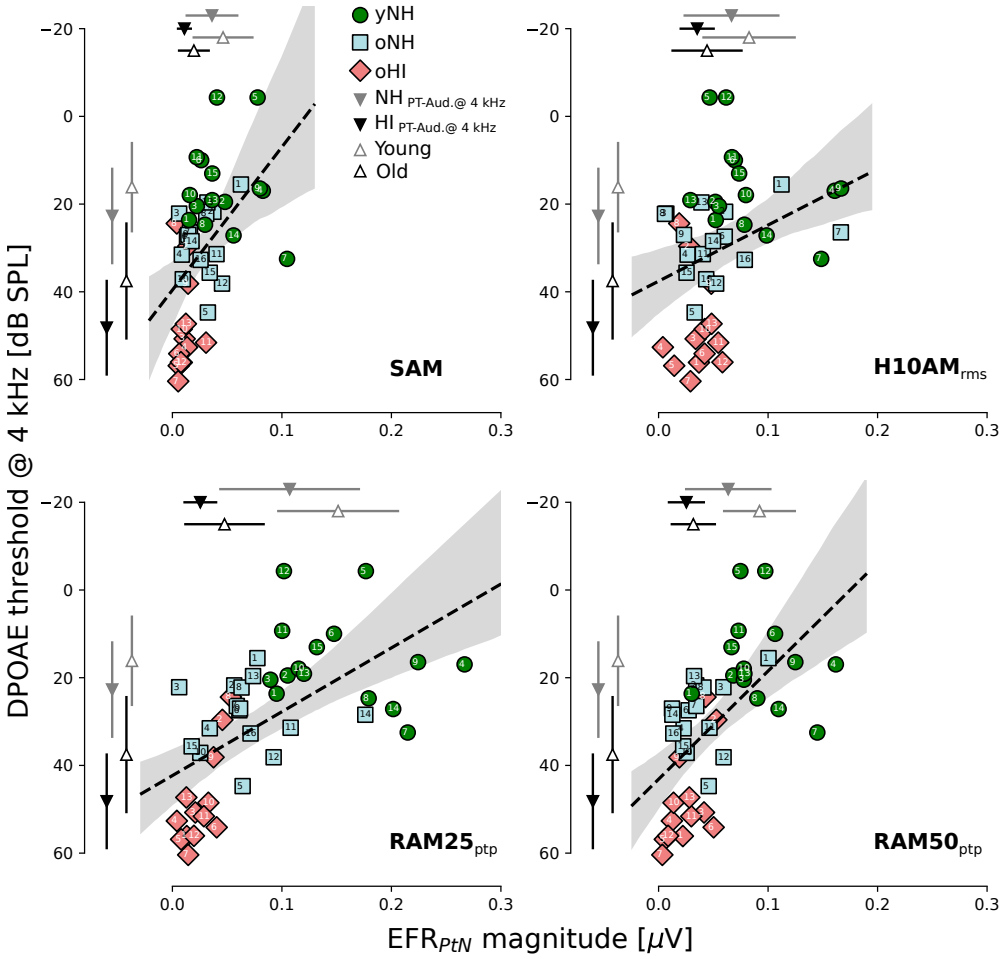


Figure 6:

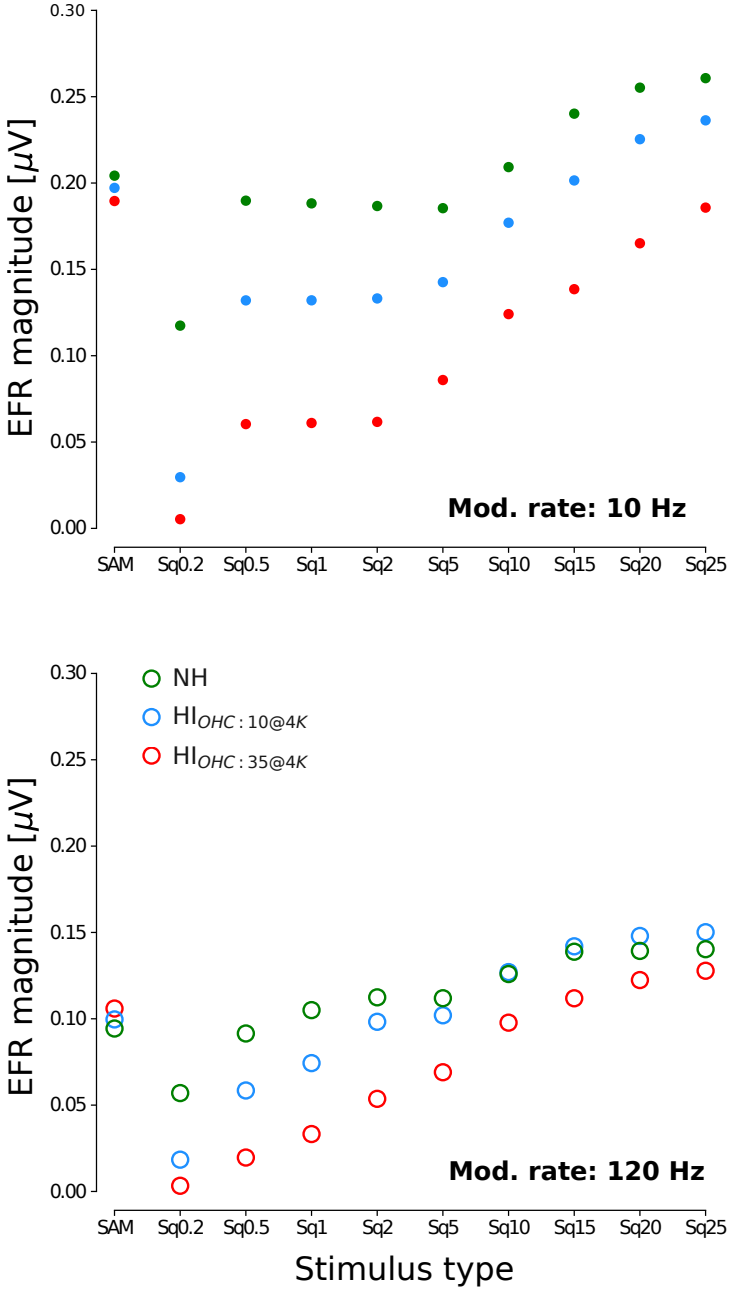


Figure 7:

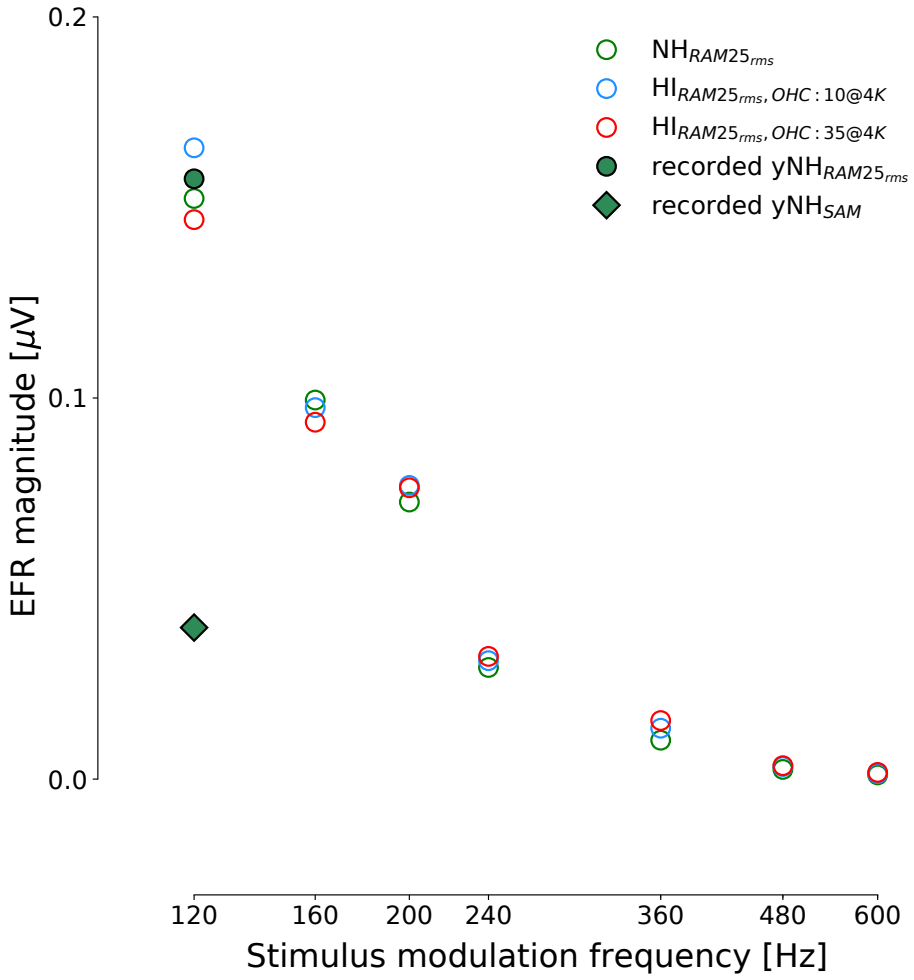


Figure 8:

| Subj# | G | A | 0.125 | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 3 | **4** | 6 | 8 | PTA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Young participants with normal hearing (yNH) | | | | | | | | | | | | | | |
| 1 | m | 27 | 5 | 5 | 10 | 5 | -5 | 5 | 0 | 5 | **0** | 10 | 5 | 4.1 |
| 2 | f | 27 | 5 | 10 | 10 | 10 | 10 | 0 | 5 | 0 | **0** | 0 | 0 | 4.5 |
| 3 | m | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **5** | 0 | 5 | 0.9 |
| 4 | f | 22 | 0 | 0 | 5 | 5 | 0 | 0 | 5 | 5 | **5** | 5 | 10 | 3.6 |
| 5 | f | 27 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** | 5 | 5 | 1.4 |
| 6 | f | 20 | 0 | -5 | 0 | 0 | 0 | 5 | 0 | 5 | **0** | 0 | 0 | 0.5 |
| 7 | f | 24 | 5 | 5 | 5 | 0 | 5 | 0 | 0 | 0 | **0** | 0 | 0 | 1.8 |
| 8 | f | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | **10** | 5 | 10 | 2.7 |
| 9 | f | 28 | 0 | 0 | 0 | 0 | 0 | 5 | 5 | 0 | **5** | 0 | 0 | 1.4 |
| 10 | f | 23 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | **5** | 5 | 10 | 2.3 |
| 11 | m | 22 | 0 | 0 | 0 | 0 | 5 | 5 | 0 | 0 | **0** | 0 | 0 | 0.9 |
| 12 | m | 24 | 5 | 5 | 0 | 5 | 0 | 0 | 0 | 0 | **5** | 5 | 0 | 2.3 |
| 13 | m | 25 | 5 | 5 | 5 | 10 | 10 | 5 | 5 | 15 | **5** | 5 | 15 | 7.7 |
| 14 | m | 23 | 10 | 15 | 15 | 15 | 15 | 10 | 10 | 10 | **10** | 10 | 15 | 12.3 |
| 15 | m | 26 | 5 | 5 | 10 | 10 | 5 | 5 | 0 | 5 | **0** | 0 | 5 | 4.5 |
| Older participants with normal hearing(oNH) | | | | | | | | | | | | | | |
| 1 | f | 63 | 15 | 10 | 5 | 10 | 5 | 15 | 5 | 0 | **0** | 25 | 30 | 10.9 |
| 2 | f | 61 | 10 | 10 | 10 | 5 | 5 | 5 | 5 | 10 | **10** | 15 | 45 | 11.8 |
| 3 | m | 63 | 5 | 5 | 10 | 15 | 15 | 5 | 10 | 10 | **15** | 45 | 65 | 18.2 |
| 4 | m | 62 | 5 | 5 | 5 | 10 | 5 | 5 | 5 | 5 | **10** | 10 | 15 | 7.3 |
| 5 | f | 65 | 10 | 10 | 15 | 15 | 15 | 10 | 10 | 20 | **10** | 15 | 30 | 14.5 |
| 6 | m | 64 | 0 | 0 | 0 | 0 | 5 | 5 | 5 | 5 | **10** | 10 | 35 | 6.8 |
| 7 | f | 62 | 10 | 10 | 15 | 15 | 20 | 10 | 10 | 10 | **15** | 30 | 30 | 15.9 |
| 8 | m | 65 | 5 | 5 | 5 | 5 | 5 | 5 | 0 | 10 | **10** | 30 | 35 | 10.5 |
| 9 | m | 65 | 10 | 5 | 5 | 5 | 5 | 10 | 5 | 15 | **10** | 25 | 50 | 13.2 |
| 10 | f | 66 | 5 | 10 | 10 | 10 | 15 | 20 | 15 | 20 | **15** | 30 | 55 | 18.6 |
| 11 | m | 66 | 20 | 20 | 20 | 20 | 15 | 15 | 15 | 15 | **15** | 40 | 45 | 21.8 |
| 12 | f | 65 | 20 | 20 | 20 | 25 | 25 | 20 | 20 | 20 | **15** | 30 | 30 | 22.3 |
| 13 | f | 62 | 0 | 0 | 0 | 10 | 10 | 10 | 5 | 5 | **10** | 10 | 15 | 6.8 |
| 14 | m | 67 | 15 | 5 | 5 | 10 | 10 | 5 | 20 | 15 | **15** | 15 | 45 | 14.5 |
| 15 | m | 67 | 10 | 5 | 5 | 5 | 5 | 5 | 5 | 10 | **10** | 15 | 25 | 9.1 |
| 16 | f | 65 | 10 | 10 | 5 | 5 | 20 | 10 | 10 | 10 | **15** | 20 | 25 | 12.7 |
| Older participants with elevated hearing thresholds (oHI) | | | | | | | | | | | | | | |
| 1 | f | 64 | 0 | 5 | 10 | 10 | 10 | 10 | 20 | 25 | **40** | 35 | 35 | 18.2 |
| 2 | f | 67 | 20 | 20 | 25 | 30 | 35 | 30 | 30 | 35 | **25** | 50 | 60 | 32.7 |
| 3 | f | 64 | 0 | 5 | 10 | 15 | 10 | 25 | 25 | 35 | **40** | 50 | 65 | 25.5 |
| 4 | m | 66 | 5 | 5 | 10 | 10 | 15 | 15 | 15 | 25 | **35** | 40 | 60 | 21.4 |
| 5 | m | 64 | 5 | 5 | 10 | 15 | 20 | 25 | 25 | 25 | **40** | 45 | 50 | 24.1 |
| 6 | f | 66 | 10 | 10 | 10 | 10 | 5 | 5 | 5 | 25 | **35** | 40 | 40 | 17.7 |
| 7 | f | 66 | 10 | 10 | 10 | 6925 | 20 | 25 | 35 | 35 | **50** | 55 | 60 | 30.5 |
| 8 | f | 61 | 10 | 15 | 20 | 20 | 15 | 25 | 30 | 30 | **30** | 35 | 35 | 24.1 |
| 9 | m | 65 | 20 | 20 | 20 | 20 | 20 | 20 | 25 | 35 | **45** | 60 | 65 | 31.8 |
| 11 | m | 64 | 5 | 10 | 10 | 20 | 25 | 15 | 10 | 25 | **40** | 60 | 60 | 25.5 |
| 12 | m | 66 | 5 | 0 | 5 | 5 | 5 | 5 | 5 | 25 | **30** | 60 | 70 | 19.5 |
| 13 | f | 68 | 15 | 15 | 20 | 20 | 20 | 30 | 25 | 35 | **40** | 35 | 60 | 28.6 |
| 14 | f | 67 | 5 | 5 | 5 | 5 | 5 | 10 | 10 | 10 | **40** | 55 | 55 | 18.6 |

Table 3: Median, mean, standard deviation (SD), 1st and 3rd quartile (Q), 5th and 95th percentile (P) of the EFR magnitudes (in $\mu$V) evoked by different stimuli.

| Stimulus | Group | Median | Mean | SD | $Q_1$ | $Q_3$ | $P_5$ | $P_{95}$ |
|---|---|---|---|---|---|---|---|---|
| SAM | yNH | 0.036 | 0.046 | 0.027 | 0.024 | 0.067 | 0.016 | 0.089 |
| | oNH | 0.028 | 0.027 | 0.015 | 0.014 | 0.035 | 0.008 | 0.05 |
| | oHI | 0.009 | 0.011 | 0.007 | 0.006 | 0.013 | 0.005 | 0.021 |
| $SAM_{BB}$ | yNH | 0.095 | 0.102 | 0.039 | 0.07 | 0.137 | 0.042 | 0.153 |
| | oNH | 0.044 | 0.067 | 0.106 | 0.023 | 0.056 | 0.011 | 0.185 |
| | oHI | 0.018 | 0.028 | 0.021 | 0.014 | 0.034 | 0.008 | 0.072 |
| H10AM | yNH | 0.07 | 0.083 | 0.041 | 0.054 | 0.089 | 0.041 | 0.163 |
| | oNH | 0.042 | 0.052 | 0.039 | 0.026 | 0.061 | 0.007 | 0.126 |
| | oHI | 0.037 | 0.035 | 0.015 | 0.028 | 0.048 | 0.01 | 0.056 |
| $RAM50_{rms}$ | yNH | 0.083 | 0.094 | 0.032 | 0.079 | 0.103 | 0.06 | 0.156 |
| | oNH | 0.033 | 0.053 | 0.061 | 0.024 | 0.047 | 0.015 | 0.148 |
| | oHI | 0.022 | 0.02 | 0.012 | 0.009 | 0.03 | 0.006 | 0.037 |
| $RAM50_{ptp}$ | yNH | 0.078 | 0.092 | 0.032 | 0.074 | 0.108 | 0.056 | 0.15 |
| | oNH | 0.033 | 0.037 | 0.022 | 0.023 | 0.046 | 0.013 | 0.07 |
| | oHI | 0.022 | 0.025 | 0.016 | 0.012 | 0.041 | 0.005 | 0.051 |
| $RAM25_{rms}$ | yNH | 0.157 | 0.157 | 0.05 | 0.111 | 0.187 | 0.097 | 0.236 |
| | oNH | 0.076 | 0.094 | 0.085 | 0.043 | 0.1 | 0.02 | 0.277 |
| | oHI | 0.026 | 0.032 | 0.016 | 0.021 | 0.039 | 0.012 | 0.057 |
| $RAM25_{ptp}$ | yNH | 0.132 | 0.151 | 0.054 | 0.104 | 0.19 | 0.093 | 0.237 |
| | oNH | 0.062 | 0.066 | 0.038 | 0.051 | 0.075 | 0.015 | 0.125 |
| | oHI | 0.021 | 0.025 | 0.015 | 0.013 | 0.038 | 0.006 | 0.049 |

Table 4: Median, mean, standard deviation (SD), 1st and 3rd quartile (Q), 5th and 95th percentile (P) of the ABR amplitudes (in $\mu$V) evoked by different stimuli.

| Stimulus | Group | Median | Mean | SD | $Q_1$ | $Q_3$ | $P_5$ | $P_{95}$ |
|---|---|---|---|---|---|---|---|---|
| ABR W-I$^{70\dagger}$ | yNH | 0.079 | 0.081 | 0.034 | 0.061 | 0.09 | 0.04 | 0.131 |
| | oNH | 0.044 | 0.053 | 0.025 | 0.032 | 0.074 | 0.023 | 0.086 |
| | oHI | 0.018 | 0.018 | 0.011 | 0.013 | 0.023 | 0.009 | 0.028 |
| ABR W-I$^{100\dagger}$ | yNH | 0.126 | 0.134 | 0.043 | 0.112 | 0.165 | 0.063 | 0.194 |
| | oNH | 0.093 | 0.083 | 0.046 | 0.043 | 0.113 | 0.022 | 0.15 |
| | oHI | 0.072 | 0.07 | 0.032 | 0.043 | 0.08 | 0.028 | 0.12 |
| ABR W-V$^{70\dagger}$ | yNH | 0.153 | 0.176 | 0.07 | 0.131 | 0.211 | 0.089 | 0.276 |
| | oNH | 0.077 | 0.093 | 0.055 | 0.062 | 0.125 | 0.025 | 0.178 |
| | oHI | 0.065 | 0.071 | 0.043 | 0.04 | 0.099 | 0.019 | 0.135 |
| ABR W-V$^{100\dagger}$ | yNH | 0.152 | 0.179 | 0.083 | 0.141 | 0.181 | 0.111 | 0.331 |
| | oNH | 0.123 | 0.149 | 0.091 | 0.098 | 0.179 | 0.06 | 0.294 |
| | oHI | 0.115 | 0.111 | 0.039 | 0.08 | 0.12 | 0.062 | 0.169 |
| ABR W-V$^{100\dagger}_{\text{(P5-P3)}}$ | yNH | 0.267 | 0.271 | 0.067 | 0.233 | 0.314 | 0.167 | 0.371 |
| | oNH | 0.213 | 0.201 | 0.062 | 0.153 | 0.233 | 0.119 | 0.313 |
| | oHI | 0.165 | 0.169 | 0.061 | 0.118 | 0.207 | 0.091 | 0.265 |

$^{\dagger}$ dB peSPL