Deciphering compromised speech-in-noise intelligibility in older listeners: the influence of cochlear synaptopathy

Markus Garrett[a,*], Viacheslav Vasilkov[b,*], Manfred Mauermann[a], John L. Wilson[c,e], Leslie Gonzales[e], Kenneth S. Henry[c,d,e], Sarah Verhulst[b,**]

[a]*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*
[b]*Hearing Technology @ WAVES, Department of Information Technology, Ghent University, Technologiepark 216, 9052 Zwijnaarde, Belgium*
[c]*Department of Otolaryngology, University of Rochester, Rochester, NY USA*
[d]*Department of Biomedical Engineering, University of Rochester, Rochester, NY USA*
[e]*Department of Neuroscience, University of Rochester, Rochester, NY USA*

Abstract

Speech intelligibility declines with age and sensorineural hearing damage (SNHL) but to date it is unclear whether a recently-discovered form of SNHL, cochlear synaptopathy (CS), plays a crucial role in this hearing problem. CS refers to damaged auditory-nerve synapses that innervate the inner hair cells and there is currently no go-to diagnostic test for CS. Furthermore, age-related hearing damage can comprise various aspects (e.g., hair cell damage, CS) that each can play a role in impaired sound perception. To address this disconnect between cochlear damage and speech intelligibility deficits, this study investigates to which degree CS contributes to impaired, low-cognitive-effort, speech intelligibility in older listeners. To quantify CS, we selected an envelope-following response (EFR) marker and first verified its sensitivity to CS in a Budgerigar model. We then adopted the marker in our human experiments, where we restricted the frequency content of the speech-material to ensure that both the EFR and the behavioral task relied on auditory processing in similar cochlear frequency regions. Following this approach, we identified the relative contribution of hearing sensitivity and CS to speech

*These authors contributed equally to the work
**Corresponding author
Email address: s.verhulst@ugent.be (Sarah Verhulst)

intelligibility in two age-matched (65-year-old) groups with clinically normal (n=16, 8 females) or impaired audiograms (n=13, 8 females). Compared to a young normal-hearing control group (n = 13, 7 females), the older groups demonstrated lower EFR responses and impaired speech reception thresholds, irrespective of their hearing sensitivity. We conclude that age-related CS reduces supra-threshold temporal envelope coding with subsequent speech coding deficits in noise that cannot be explained based on hearing sensitivity alone.

Abbreviated title: Synaptopathy and speech intelligibility

Number of pages: 73

Number of figures: 7

Number of tables: 3

Number of words: Abstract: 248, Significance Statement: 120, Introduction: 1025, Discussion: 2150

Significance Statement

Temporal bone histology reveals that cochlear synaptopathy (CS), characterized by damage to inner hair cell auditory nerve fiber synapses, precedes sensory cell damage and hearing sensitivity decline. Despite this, clinical practice primarily evaluates hearing status based on audiometric thresholds, potentially overlooking a prevalent aspect of sensorineural hearing damage due to aging, noise exposure, or ototoxic drugs—all of which can lead to CS. To address this gap, we employ a novel and sensitive EEG-based marker of CS to investigate its relationship with speech intelligibility. This study addresses a crucial unresolved issue in hearing science: whether CS significantly contributes to degraded speech intelligibility as individuals age. Our study-outcomes are pivotal for identifying the appropriate target for treatments aimed at improving impaired speech perception.

Abbreviations

AN(F) - auditory nerve (fiber);

BB - broadband;

CF - characteristic frequency;

CN - cochlear nucleus;

DPOAE - distortion product otoacoustic emission;

EEG - electroencephalography;

EFR - envelope following response;

H/M/LSR fibers - high/medium/low spontaneous rate fibers;

HI - hearing impaired;

HP - high-pass filtered;

IC - inferior colliculus;

IHC - inner hair cell;

I/O function - input/output function;

LP - low-pass filtered;

NH - normal hearing;

OHC - outer hair cell;

peSPL - peak-equivalent sound pressure level;

PT - pure-tone;

SAM - sinusoidally amplitude-modulated;

SiN - speech in noise;

SiQ - speech in quiet;

SNHL - sensorineural hearing loss;

SRT - speech reception threshold;

SPL - sound pressure level;

RAM - rectangularly amplitude-modulated;

TENV - temporal envelope;

TFS - temporal fine structure;

Introduction

The cochlea is deeply embedded within the temporal bone, rendering a direct histological assessment of sensorineural hearing loss (SNHL) impossible in living humans. While outer-hair-cell (OHC) deficits are routinely diagnosed using indirect metrics such as the pure-tone-audiogram or distortion-product-otoacoustic-emission (DPOAE) threshold, damage to auditory-nerve-fiber (ANF) synapses between the inner hair cells (IHCs) and spiral ganglion cells, i.e., *cochlear synaptopathy, CS* (Kujawa & Liberman, 2009), can so far only be quantified via post-mortem temporal bone histology (Makary et al., 2011; Viana et al., 2015; Wu et al., 2018). As observed through reduced envelope-following responses (EFRs) in animals with histologically verified synaptopathy (Parthasarathy et al., 2014; Shaheen et al., 2015; Parthasarathy & Kujawa, 2018), synaptopathy compromises temporal envelope (TENV) coding of supra-threshold stimuli. The EFR is an indirect and non-invasive auditory-evoked potential (AEP) type that reflects the phase-locked neural response of a population of peripheral and brainstem neurons to a stimulus envelope (Kraus et al., 2017). Its frequency-domain response exhibits peaks coinciding with the stimulus modulation frequency and its harmonics. Given that EFRs can be recorded reliably in both animals and humans, its spectral magnitude has been proposed as a non-invasive

marker of synaptopathy which can be used across species (e.g., Shaheen et al., 2015). Synaptopathy precedes hearing sensitivity reductions in the progression of noise- or age-induced SNHL (Sergeyenko et al., 2013; Fernandez et al., 2015; Parthasarathy & Kujawa, 2018). Due to its association with compromised supra-threshold temporal envelope coding (Bharadwaj et al., 2014), CS is thought to contribute to reduced speech intelligibility in individuals with otherwise normal audiograms (Mepani et al., 2021). This connection might provide insights into why hearing sensitivity alone falls short as a predictor for individual speech-intelligibility scores (Festen & Plomp, 1983; Papakonstantinou et al., 2011).

In order to develop effective treatments for SNHL, it is crucial to understand whether CS compromises speech intelligibility. This endeavor has prompted numerous studies investigating this presumed relationship in humans. However, these efforts have yielded a body of mixed and inconclusive results: while the ABR wave-I amplitude, a marker of CS in animal studies (Kujawa & Liberman, 2009; Möhrle et al., 2016), has been shown to predict speech intelligibility in some studies (Bramhall et al., 2015; Liberman et al., 2016), it has not shown consistent results in others (Johannesen et al., 2019; Prendergast et al., 2017). Similarly, the same inconsistency is observed for EFR markers of CS (Shaheen et al., 2015; Parthasarathy et al., 2019): while some studies demonstrate that certain EFR markers of CS predict speech intelligibility in individuals with normal hearing (Mepani et al., 2021), others

do not (Guest et al., 2018; Grose et al., 2017).

Several factors could be responsible for these inconsistent outcomes across studies: (i) individual differences in absolute ABR and EFR strength can, to a certain degree, reflect non-hearing-related factors in humans (Trune et al., 1988; Mitchell et al., 1989; Plack et al., 2016) and require a relative metric design (e.g. Bharadwaj et al., 2015; Mehraei et al., 2016; Hickox et al., 2017; Le Prell, 2019). (ii) As the ABR and EFR can reflect a mixture of synaptopathy and OHC deficits, an unambiguous interpretation of the ABR or EFR in terms of synaptopathy is difficult when the histopathology is unknown (Verhulst et al., 2016b; Garrett & Verhulst, 2019; Van Der Biest et al., 2023). (iii) The recording paradigms established in animal studies might not be sensitive enough for application in humans and hence overlook synaptopathy (Hickox et al., 2017; Bramhall et al., 2019). Lastly, (iv) whereas speech is a broadband signal, the EFR mostly reflects auditory TENV coding mechanisms associated with ANF fibers above the phase-locking limit (Joris & Yin, 1992; Verschooten et al., 2015; Henry et al., 2016). Aside from TENV cues, speech recognition is also reliant on the frequency content below the phase-locking limit, which is encoded as temporal fine-structure (TFS) information by the ANFs. This TFS information serves as a robust perceptual cue (e.g. Lorenzi et al., 2006; Hopkins et al., 2008; Hopkins & Moore, 2010; Henry et al., 2016; Mai & Howell, 2023; Borjigin & Bharadwaj, 2023). However, conventional EFR markers primarily assess TENV coding mechanisms and their impairments, rather than specifically targeting temporal TFS coding. Therefore,

the absence of a correlation between the EFR (TENV-driven) and speech recognition (involving both TFS and TENV coding) does not definitively indicate that CS does not influence speech intelligibility. It could also occur if distinct mechanisms drive each metric independently.

This study aims to clarify factors that may have complicated the interpretation of the relationship between speech perception and AEP markers of CS. Firstly, our focus is on EFRs rather than ABRs. This choice is based on prior model simulations and data collection, which suggested that potential confounding OHC deficits would have a more pronounced effect on the ABR amplitude compared to the amplitude of pure-tone-evoked EFRs.(Verhulst et al., 2016a; Garrett & Verhulst, 2019; Vasilkov et al., 2021). Secondly, auditory model simulations have demonstrated that optimizing the EFR stimulus envelope to be rectangular, rather than sinusoidal, can enhance the sensitivity of the EFR marker to CS (Vasilkov et al., 2021). Building on these findings, our study initially assesses the sensitivity of an optimized EFR marker of CS in a Budgerigar model of Kainic-Acid induced ANF damage. Subsequently, we validate its predicted superiority in detecting individual CS differences over conventional EFR markers of CS.

We proceed to examine the correlation between this EFR marker and speech intelligibility using both low and high-pass filtered speech material in human subjects. Our hypothesis posits that the EFR marker provides

more accurate predictions of speech recognition thresholds when both measures depend on TENV mechanisms and their associated impairments. To delineate the individual contributions of CS and hearing sensitivity to speech intelligibility, we investigated the connections between objective markers of hearing and speech intelligibility in three groups: young individuals with normal hearing, and two age-matched groups of older participants: one with normal audiograms and the other exhibiting impaired audiograms. It is our presumption that the older participants may be afflicted by age-related CS (Wu et al., 2018), potentially compounded by additional pathologies affecting OHCs.

Materials and Methods

Study Participants

Three participant groups were recruited using age and audiometric pure-tone thresholds as the classification criteria. Young: $20 \leq$ age ≤ 30, Older: $60 \leq$ age ≤ 70, normal-hearing: ≤ 25 dB HL for every audiometric frequency up to 4 kHz, hearing-impaired: > 25 dB HL. The grouping criterion did not account for individual variations in the degree of synaptopathy, an unquantified variable at the start of the study. Using these criteria, fifteen young normal-hearing (yNH: 24.5 ± 2.3 y/o, 8 females), 16 older normal-hearing (oNH: 64.3 ± 1.9 y/o, 8 females) and 13 older hearing-impaired (oHI: 65.2 ± 1.8 y/o, 8 females) participants were recruited. There were no significant age

differences between the oNH and oHI participant groups ($t_{(27)}$ = -1.41, p = 0.17). Otoscopy was performed prior to data collection to ensure that participants had no obstructions or other visible outer or middle ear pathologies. The experiments were approved by the ethics committee of the University of Oldenburg. Participants gave written informed consent and were paid for their participation.

Speech Reception Thresholds (SRTs)

Speech intelligibility was assessed by applying a standard German Matrix sentence test which determines the speech reception threshold (SRT) in quiet or in a fixed-level noise background (Matrix-test or OLSA; Wagener et al., 1999; Brand & Kollmeier, 2002). The OLSA test consists of 5-word sentences (male speaker: name - verb - numeral - adjective - object) using a corpus of 50 possible words. The speech-shaped background noise was generated from the sentences and matched the long-term spectrum of the speech material. The SRT for 50% correctly identified words (i.e., SRT_{50}) was determined using a 1-up/1-down adaptive procedure with varying step size based on word scoring (20 sentences per condition).

The SRT_{50} was determined in quiet (SiQ) and noise (SiN) for unfiltered/broadband (BB), low-pass filtered (LP) and high-pass filtered (HP) audio. The speech and noise signals in the LP and HP conditions were

generated by applying a 1024th order FIR filter with respective cut-off frequencies of 1.5 and 1.65 kHz to the OLSA test material (i.e., BB condition). The idea behind using filtered speech stimuli is that the listener can be made to rely on specific auditory processing cues associated with low or high-frequency hearing. In this context, the perception of the LP condition is predominantly based on available temporal fine structure (TFS) cues (Lorenzi et al., 2006), whereas the HP condition mostly relies on temporal envelope (TENV) cues. It is well known that auditory-nerve phase-locking to TFS declines with increasing frequency (i.e., 1.4 kHz in humans based on ITD sensitivity; Joris & Verschooten, 2013), and that for frequencies beyond this limit, the auditory system can only rely on temporal envelope processing. Since the EFR marker we utilize to investigate the impact CS on speech intelligibility predominantly captures high-frequency TENV processing, it is logical to incorporate a speech condition that similarly depends on cochlear TENV processing, such as the SRT_{HP}.

Both SiQ and SiN conditions were included in our study because we anticipate that speech processing in the presence of a fixed-level background noise undermines speech perception more significantly than the SiQ condition. This is due to the principles of deafferentation and coding redundancy associated with cochlear synaptopathy (CS), as highlighted by Lopez-Poveda & Barrios (2013). Additionally, the SiQ condition may also be affected by CS-compromised TENV processing, as indicated by studies such as Shaheen et al. (2015); Parthasarathy & Kujawa (2018). We considered both condi-

tions because it is possible that CS does not affect SiQ and SiN processing to the same extent. The initial speech level in the OLSA test was 70 dB SPL and the noise level was kept fixed at 70 dB SPL while the speech level varied adaptively to yield the SRT in the SiN conditions. In the SiQ conditions, the speech level was varied and the dB SPL at which the 50% correct threshold was reached, was reported. The six conditions (3 SiQ, 3 SiN) were presented in pseudo-random order. Participants completed three training runs in which a SiN_{BB} with a fixed SNR of 5 dB was followed by a regular SiN_{BB} condition with SRT tracking for training purposes. All possible words were displayed on the screen during those runs to familiarize participants with the stimulus material. The third training run was a SiN_{HP} condition with SRT tracking but without visual aid. During the experiment, answers were logged by the experimenter and no visual feedback was provided. Measurements were conducted in a double-walled sound-insulated booth using Sennheiser HDA200 headphones in combination with the Earbox Highpower ear 3.0 sound card (Auritec) and the Oldenburg Measurement Platform (HörTech gGmbH). The setup was calibrated using an artificial ear type 4153, microphone type 4134, preamplifier type 2669 and sound level meter type 2610 (Brüel & Kjær).

Behavioral and Physiological Markers of SNHL

We adopted two measures that are specific to quantifying OHC or CS integrity using non-invasive physiological techniques, namely the distortion-

product otoacoustic emission threshold (TH_{DP}) and the envelope-following response (EFR) amplitude. We concentrated on these measures within the 4-kHz cochlear region, aligning with the carrier frequency of the EFR stimulus. Additionally, we included a standard-clinical audiogram to assess the behavioral hearing sensitivity for pure-tone frequencies up to 8 kHz. The audiogram and TH_{DP} values were used as selection criterion to assign listeners to specific groups. Young normal-hearing (yNH) control group: audiometric pure-tone average (PTA) for frequencies up to 4 kHz $\leq$ 10 dB HL and $TH_{DP,\ 4\ kHz} \leq 25$ dB SPL. Older normal-hearing (oNH) or hearing-impaired (oHI) groups had a PTA upto 4 kHz $\leq$ or $>$ 20 dB HL, respectively. We applied a more stringent criterion for selecting participants for the control group compared to the original selection of young normal-hearing study participants. Our aim was to minimize the risk of OHC damage as much as possible for the conclusions drawn in our study. As a result, two participants from the young normal-hearing group were excluded from the control group based on their $TH_{DP,\ 4\ kHz}$. Their datapoints were, however, included in the multiple regression analysis conducted across all individuals, and were identified with distinct markers in the corresponding regression figures. Finally, for certain sections of the post-hoc analysis, the entire population was divided into groups with either normal (NH DP) or impaired (HI DP) otoacoustic emission thresholds, reflecting $TH_{DP,\ 4\ kHz} \leq$ or $>$ 25 dB SPL, respectively.

Pure-tone audiometric thresholds were collected for frequencies be-

tween 0.125-8 kHz using Sennheiser HDA200 headphones and a clinical audiometer AT900 (Auritec). Figure 1 shows audiograms of the tested ears (the right ear was used in 10 yNH, 11 oNH and 10 oHI subjects). The test ear was chosen based on the better of the left and right audiograms. At 4 kHz, yNH participants had a mean (± std) threshold of 3.3 ± 3.6. The older groups had thresholds of 11.6 ± 4.0 dB HL (oNH) and 37.7 ± 6.7 dB HL (oHI) respectively.

Distortion-product otoacoustic emissions were recorded at 4 kHz to assess OHC integrity more directly. Stimuli were presented over ER-2 speakers (Etymotic Research) using foam ear tips and recordings were made using the ER10B+ OAE microphone system (Etymotic Research) and custom-made MATLAB scripts (Mauermann, 2013). Two pure tones (f_1, f_2) were simultaneously presented at a fixed f_2/f_1 ratio of 1.2 using a primary frequency sweep method (Long et al., 2008). Frequencies were exponentially swept up (2 s/octave) over a $1/3^{rd}$ octave range around the geometric mean of 4 kHz. L_1 levels followed the Scissors paradigm (L_1=0.4 L_2+39; Kummer et al., 1998) given a primary L_2 level of 30-60 dB SPL in steps of 6 dB (in oHI listeners, L_2s of 66 and 72 dB SPL were additionally collected). The distortion component (L_{DC}) was extracted using a sharp 2-Hz-wide least-squares-fit filter and the center frequency of the measured frequency range was used to construct L_{DC} growth functions. Individual L_{DC} data points and their standard deviations were used in a bootstrapping procedure to

fit an adapted cubic function through the L_{DC} datapoints as described in Verhulst et al. (2016a) to yield the DPOAE threshold (TH_{DP}) and its standard deviation. L_{DC} growth functions typically increase monotonically or saturate with increasing L_2 (Mauermann & Kollmeier, 2004; Abdala et al., 2021). We thus constrained our bootstrapping procedure to only include random L_{DC} draws (i.e. from within the confidence interval of each mean L_{DC} measurement point) to impose monotonous growth in each L_{DC} (L_2,b) bootstrap run. We used an automated algorithm which eliminated adjacent data points at either end of the growth function (never intermediate points) that compromised monotonicity. TH_{DP} was determined in each bootstrap run as the L_2 at which the extrapolated fitting curve reached a level of -25 dB SPL (Neely et al., 2009). This bootstrapping procedure yielded the median threshold (TH_{DP}) and its standard deviation. TH_{DP} of two yNH participants with reasonable growth functions but very low thresholds (-30.9, -5.4 dB SPL) were set to 1.5-times the interquartile range (IQR) of the yNH TH_{DP} threshold distribution (-4.3 dB) to avoid high-leverage data-points. Audiogram and DPOAE thresholds at 4 kHz showed similar trends across groups and a significant correlation ($\rho = 0.84$; $p < 0.0001$, $n = 44$). This latter observation corroborates earlier results of a large cohort study (Boege & Janssen, 2002), and supports the notion that both metrics assess hearing sensitivity.

Stimulation. EFRs were recorded to two amplitude-modulated pure-tone stimuli with the same carrier frequency (f = 4 kHz) and rate ($\mathrm{f_m}$ = 120 Hz, starting phase (φ) of $\frac{3\pi}{2}$). The only difference between the stimuli was their envelope shape: the SAM stimulus had a sinusoidal modulator (Eq.1), as commonly used in animal studies of CS (e.g., Shaheen et al., 2015; Parthasarathy & Kujawa, 2018) while the second modulator was a rectangular-wave with a duty-cycle (τ) of 25% (RAM; Vasilkov et al., 2021). A modulation depth (md) of 0.95 (95%, -0.45 dB re. 100%) was applied. MATLAB's 'square' function was adopted to create the RAM stimulus, and both stimuli are formulated as follows:

$$\mathrm{SAM}: \; x(t) = [1 + \mathrm{md} \cdot \sin(2\pi \mathrm{f_m} t + \varphi)] \sin(2\pi \mathrm{f} t) \tag{1}$$

$$\begin{aligned} &\mathrm{RAM}: \; y(t) = [2 + 2\mathrm{md} \cdot m(t)] \sin(2\pi \mathrm{f} t)\,, \text{ with} \\ &m(t) = 2 \sum_{n=0}^{d \cdot \mathrm{f_m} - 1} \left[u\left(t \cdot \mathrm{f_m} - n + \frac{\varphi}{2\pi}\right) - u\left(t \cdot \mathrm{f_m} - n + \frac{\varphi}{2\pi} - \tau\right) \right] - 1, \text{ and} \\ &u[p] = \begin{cases} 0,\, p < 0, \\ 1,\, p \geq 0, \end{cases} \quad \text{(unit step function)} \end{aligned} \tag{2}$$

Figure 2 depicts two cycles of each stimulus type. Stimuli were windowed

using a 2.5% tapered-cosine window, had a duration (d) of 0.4 seconds and were repeated 1000 times each (500 per polarity). The inter stimulus interval consisted of a uniformly distributed random silence jitter (100 ms ± 10 ms). The SAM tone was presented at 70 dB SPL and the RAM stimulus at the same peak-to-peak amplitude, which corresponded to 68 dB SPL. A previous study compared a broad range of possible ABR and EFR markers of CS (Vasilkov et al., 2021), and we selected the most promising RAM stimulus for use in the present study.

EFR recordings and analysis. Recordings took place in a double-walled electrically shielded measurement booth (IAC acoustics) and participants sat in a reclining chair while watching a silent movie. EEG signals were recorded using a 64-channel cap with equidistant electrode spacing (Easycap) and active Biosemi Ag/AgCl electrodes were connected to a Biosemi amplifier. A sampling rate (f_s) of 16384 Hz and 24-bit analog-to-digital conversion were used to store the raw data traces. A common-mode-sense (CMS) active electrode was placed on the fronto-central midline and a driven-right-leg (DRL) passive electrode was placed on the tip of the nose. Reference electrodes were placed on each earlobe. Electrode offsets (DC values of the common mode signal) were kept below 25 mV. Stimuli were generated in MATLAB (R2015b) at a sampling rate of 48 kHz and calibrated using the same equipment as for the speech recognition test (ear simulator type 4157 (Brüel & Kjær) for insert earphones). A Fireface UCX sound card (RME) and

TDT-HB7 headphone driver (Tucker-Davis) were used to drive the ER-2 insert earphones (Etymotic Research) using the open-source portaudio playrec ASIO codec (Humphrey, 2008). Stimuli were presented monaural to the test ear.

Raw EEG recordings were extracted in Python (version 2.7.10 | Anaconda 2.3.0 (64-bit), (`www.python.org`) and MNE-Python (version 0.9.0; Gramfort et al., 2013, 2014) and all channels were re-referenced to the offline-averaged earlobe electrodes. Data were epoched in 400 ms windows starting from the stimulus onset and baseline corrected by the average amplitude per epoch. We only present results from the vertex channel (Cz) in this study which is known to yield good signal strength for subcortical auditory sources (Picton, 2010).Pre-processing was performed in MATLAB (R2014b). The EFR estimates for each stimulus condition and participant were computed based on the energy at the modulation frequency and its first four harmonics (h_0-$\mathrm{h}_4 = k \times \mathrm{f}_m$, k=[1..5]) to account for all envelope related energy in the EEG signal (Vasilkov et al., 2021). Equation 3 was used in a bootstrap routine to obtain the mean EFR amplitude and corresponding standard deviation (Zhu et al., 2013):

$$
\begin{aligned}
\mathrm{EFR} = \frac{1}{2}\,\text{peak-to-peak of}\left(\frac{1}{B}\sum_{b=1}^{B}\left(W_{\mathrm{b}}\right)\right),\ \text{with} \\
W_{\mathrm{b}} = \left(\frac{1}{N}\sum_{n=0}^{N-1}\left(\mathrm{F}_{n,b} - \mathrm{NF}_{n,b}\right)e^{i\theta_{n,b}}\right); \\
\text{if } n \neq \frac{k\mathrm{f}_m}{\mathrm{f}_s}N,\ \text{then } \mathrm{F}_{n,b}, \mathrm{NF}_{n,b} = 0, \text{for } k = [1..5]
\end{aligned}
\tag{3}
$$

where N corresponds to the length of the magnitude spectrum. The process is visualized in Fig. 3. First, a mean spectral estimate of the EEG recording for each frequency component (n) was computed by averaging the complex discrete Fourier transform values of 1000 randomly drawn epochs using 500 epochs of each polarity (with replacement) in each bootstrap run (b; $\mathrm{F_{n,b}}$). Epochs were windowed using a 2% tapered-cosine window before the frequency-domain transformation. The electrophysiological noise floor ($\mathrm{NF_{n,b}}$) at frequencies $\mathrm{h_0}$-$\mathrm{h_4}$ was computed as the average magnitude of the ten frequency-bins surrounding the respective frequency (five bins each side). The noise floor estimates were then subtracted from the signal components $\mathrm{h_0}$-$\mathrm{h_4}$, to yield peak-to-noise floor ($\mathrm{PtN_{n,b}}$) magnitude estimates (see Fig. 3, left panels). All frequency components apart from the harmonic frequencies ($\mathrm{h_0}$-$\mathrm{h_4}$) were removed from the spectrum before it was transformed back to

the time domain using the inverse discrete Fourier transform and the original phase information ($\theta_{\mathrm{n,b}}$) of the harmonic frequencies. This procedure was repeated 200 times to yield $\mathrm{W_b}\,\mathrm{b} = 200$ reconstructed time-domain estimates of the EFR waveform. The $\mathrm{W_b}$ waveforms were then averaged and the EFR was defined as half the peak-to-peak amplitude of the averaged reconstructed time-domain waveform (see Fig. 3, left panels - reconstructed EFR). The metric defined in Eq.3 corresponds to the EFR peak-to-noise-floor amplitude and is further referred to as the EFR amplitude (in μV) or our EFR marker.

Budgerigar EFR recordings and Kainic-acid induced synaptopathy

EFRs were also recorded in eighteen young adult budgerigars (Melopsittacus undulatus), approximately two years of age, before and after induction of CS using kainic acid (KA). The budgerigar is a small parrot species with human-like behavioral sensitivity to many simple and complex sounds in operant conditioning studies (Dent et al., 2000; Dooling et al., 2000), and was selected based on its use in ongoing behavioral studies of CS (Wong et al., 2019). All procedures were performed at the University of Rochester and approved by the University Committee on Animal Resources. Three animals were tested before and after KA administration and monitored over time, whereas the others in the cohort either belonged to a control or KA group. KA is a glutamate analog that damages cochlear afferent synapses between hair cells and auditory-nerve fibers through excitotoxicity in mammals and birds (Bledsoe et al., 1981; Pujol et al., 1985; Zheng et al., 1997; Sun et al.,

2001). In budgerigars, bilateral cochlear infusion of KA has been shown to permanently reduce ABR wave-I amplitude by up to 70% without impacting behavioral audiometric thresholds or DPOAEs generated by sensory hair cells (Henry & Abrams, 2018; Wong et al., 2019). We used the methods described in Henry & Abrams (2018); Wong et al. (2019) to induce synaptopathy in budgerigars. Briefly, animals were anesthetized with ketamine (5-6 mg/kg) and dexmedetomidine (0.1 mg/kg; subcutaneous bolus injection) and placed in a custom stereotaxic device. Ketamine/dexmedetomidine were administered subcutaneously for anesthetic induction, but anesthesia was maintained throughout the surgery (approximately 1-2 hours) by a continuous anesthetic infusion pump (Razel Scientific; Fairfax, VT, USA). The middle-ear space was accessed surgically using a posterior approach to expose the basal prominence of the cochlear duct, where a 0.15-mm diameter cochleostomy was made using gentle rotating pressure on a small manual drill. Thereafter, 2.5 μL of 2-mM KA (Abcam ab 144490; Cambridge, UK) in hanks balanced salt solution (Sigma-Aldrich H8264; St. Louis, MO, USA) was infused into cochleostomy over 90 seconds using a microinjection needle. Compound action potentials of the auditory nerve were recorded before and after infusion in response to clicks. Excitotoxic synaptic injury was confirmed by observing >90% reduction of compound action potentials within 10-20 minutes following KA exposure. The left and right ears were treated with KA during different surgical procedures four weeks apart to minimize operating time and to avoid excessive anesthetic exposure. DPOAEs were

recorded using a swept-tone paradigm (see Wong et al., 2019) before and after surgeries to confirm no adverse impact of the procedures on sensory hair cells. Prior to KA exposure, wave-I amplitude of the ABR in response to 90-dB p.e. SPL clicks was 24.56 ± 2.31 μV in animal K20 and 23.69 ± 2.60 μV in animal K25 (means ± SD). Reduction of wave I, based on ABRs recorded four or more weeks post KA (during the steady-state period), was 68.5% in animal K20 and 64.9% in animal K25. EFRs were measured at two to three time points before kainic acid exposure at four time points after (4, 6, 8, and 12 weeks) the second infusion. Repeated measurements assessed within-subject variability of responses, since synaptic injury is relatively stable one month following KA exposure (Sun et al., 2000; Henry & Abrams, 2018; Wong et al., 2019). Anesthesia was performed as described above using ketamine and dexmedetomidine (for recording sessions, anesthesia was only administered subcutaneously), and body temperature was maintained in the normal range for this species of 39-41 degrees C. Stimuli were generated in MATLAB (The MathWorks, Natick, MA, USA) at a sampling frequency of 50 kHz and converted to analog using a data acquisition card (PCIe-6251; National Instruments, Austin, TX, USA), which also digitized response activity using the same internal clock. Stimuli were presented free-field through a loudspeaker (MC60; Polk Audio, Baltimore, MD, USA) positioned 20 cm from the animals head in the dorsal direction (the rostral surface of the head faced downward in the stereotaxic apparatus; thus the loudspeaker and animal were located in the same horizontal plane). Level

was controlled digitally (by scaling stimulus waveforms in MATLAB) and by up to 60 dB of analog attenuation applied by a programmable attenuator (PA5; Tucker Davis Technologies, Alachua, FL, USA). Calibration was based on the output of a $\frac{1}{4}$" precision microphone (model 4938; Brüel and Kjær, Marlborough, MA USA) in response to pure tones. Electrophysiological activity was recorded differentially between a stainless steel electrode implanted at the vertex (M0.6 threaded machine screw; advanced through the skull to the level of the dura) and a platinum needle electrode (model F-E2; Natus Manufacturing, Gort, Co. Galway, Ireland) inserted at the base of the skull near the nape of the neck. A second needle electrode in the animal's back served as ground. Activity was amplified by a factor of 50,000 and filtered from 30-10,000 Hz (P511; Grass Instruments, West Warwick, RI USA) prior to sampling (50 kHz) by the data acquisition card.

Stimuli were SAM and RAM tones presented at 75 dB SPL with 10-ms cosine-squared onset and offset ramps, 300-ms duration, and a 130-ms silent interval between successive stimuli. Carrier frequency and modulation frequency were 2830 Hz and 100 Hz, respectively, and the polarity of the carrier signal was alternated between stimulus repetitions. Depth of modulation was 100%, and the duty-cycle for RAM modulation was fixed at 25%. Responses to 300 repetitions of the same stimulus (including both polarities) were averaged to produce each EFR waveform and amplitude, following the procedure described in Eq.3.

Statistical and post-hoc analysis

To disentangle the effects of CS and OHC damage on the EFR markers and speech reception thresholds, we considered both the dataset as a whole in a multiple regression analysis (n=44), as well as performed group statistics using subgroups. These subgroups included the $yNH_{control}$ group (n=13) and the recruited age-matched oNH (n=16) or oHI groups (n=13) when investigating main effects of age and OHC damage respectively. It is noted once more that two recruited yNH participants did not meet the strict $yNH_{control}$ post-hoc criterion for normal DP_{TH}s, causing them to be omitted from the $yNH_{control}$ group in the statistical group analysis but leaving them included in the regression analysis on the whole dataset. In our grouping analysis, we also considered the pooled older group (n=29) when investigating main effects of age-related CS in comparison to the $yNH_{control}$ group, and post-hoc separated the older group into subgroups with normal or impaired TH_{DP}s when parsing out effects of OHC damage.

The 'SciPy' python package for scientific computing (Oliphant, 2007; Millman & Aivazis, 2011) was used for two-sample inference statistics. All correlations reported refer to the Pearson correlation coefficient (r) if both variables were normally-distributed, otherwise the Spearman's rank correlation coefficient (ρ) was reported. The R programming environment (R Core Team, 2019) and 'lsmeans' package (Russell, 2016) was used for analyses of variance (ANOVA) and regression analyses. Reported p-values for multiple

comparisons were Bonferroni adjusted to control for the family-wise error rate.

Partial spearman correlation analyses were conducted using the 'PResiduals' package (Dupont et al., 2018) and partial pearson correlations were performed using the 'ppcor' package (Kim, 2015). Additionally, we performed commonality analysis using the 'yhat' package (Nimon et al., 2008). Commonality analysis combines linear regressions on the dependent variable and allows for the decomposition of the explained variance (R^2) of the linear predictors into subcomponents explained by the unique and the common/shared variance of predictors and all their possible combinations (Newton & Spurrell, 1967). This technique also works in the presence of multicollinearity (Ray-Mukherjee et al., 2014).

Results

Envelope-following-response sensitivity to cochlear synaptopathy

Panels C and D in Fig. 3 show the effect of KA on budgerigar SAM and RAM EFR magnitude spectra and reconstructed waveforms, respectively. Energy at all stimulus repetition frequencies and harmonics was reduced after KA administration (left panels), leading to an overall reduction in the reconstructed EFR waveform and its corresponding amplitude (right panels). EFR amplitude reductions occurred consistently across the three longitudinally-

monitored budgerigars (Fig.4A), and was attributed to a histology-verified reduction in auditory-nerve peripheral axons and cell bodies (Fig.4B). The histology confirms that KA introduces cochlear synaptopathy in budgerigars, and the additionally performed DPOAE analysis in Wang et al. (2023); Wilson et al. (2021) furthermore demonstrates the selectivity of KA to AN synapses and cells without damaging the OHCs. As Fig.4A depicts, EFR amplitude reductions occur instantly after KA administration, and recover slightly to an overall reduced amplitude over the following weeks.

The post-KA EFR amplitudes shown in Fig.4C correspond to average EFR amplitudes over the different post-KA measurement time points for each animal (i.e., a minimum of 4 different recordings per animal) and are compared to control EFR amplitudes. Connected lines refer to data points stemming from the same budgerigar, and SAM and RAM EFRs were recorded during the same session. KA reduced the SAM and RAM magnitudes significantly (d.f. = 19, t_{SAM}=2.8, p=0.006 and t_{RAM}=6.72, p<1e-5), and the effect size was greater for the RAM than SAM stimulus (Cohens d_{SAM}= 1.3 d_{RAM}=3.1). The latter observation relates to the almost five-times larger pre-KA RAM amplitudes. These recordings confirm the positive effect the stimulus envelope has on the EFR signal-to-noise ratio. Single-unit AN recordings show more synchronized AN responses to faster-rising stimulus envelopes (Dreyer & Delgutte, 2006), and the AN and EFR model simulations performed in Vasilkov et al. (2021) show that this effect also impacts the neural generators of the EFRs. We conclude that the RAM EFR is a sensitive and

selective non-invasive marker of cochlear synaptopathy in budgerigar and that the RAM EFR has an improved sensitivity over the SAM EFR in identifying individual differences in CS. This renders the RAM EFR a suitable candidate for use in human studies for whom the intrinsic EEG signal-to-noise ratio is inherently smaller than in research animals.

Human EFR recordings: OHC and age-related deficits

Figure 4D depicts human SAM and RAM EFRs for yNH control (white symbols) and older subjects (gray symbols) with or without hearing sensitivity loss based on the TH_{DP} above 25 dB SPL criterion. The human EFR amplitudes are in agreement with both model predictions (Vasilkov et al., 2021) and budgerigar findings (panel C) in showing overall 3.7 times larger RAM than SAM amplitudes. Older subject showed SAM and RAM EFR amplitude reductions in the order of 6 and 47%, respectively. Only the RAM EFR reduction was significant (d.f. = 40, t=3.92, p=0.0004) in older listeners, and can be interpreted as caused by age-related cochlear synaptopathy. However, humans differ from the KA budgerigar model of CS in that it cannot be excluded that our human cohort had other forms of SNHL that interacted with the RAM EFR marker of CS. Prior model simulations of the EFR generators showed that for the RAM stimulus, a potential confounding effect of OHC damage is eliminated at the level of individual AN responses and limited to 5-10% of the total EFR response. Simulated SAM EFRs were more sensitive to coexisting OHC damage, as the SAM stimulus

affected individual AN responses considerably (Fig.1 in Vasilkov et al., 2021).

To investigate the potential influence of OHC-damage on an age-related CS interpretation of the observed EFR reductions, we divided our human subjects into groups with either normal (DP<25) or impaired (DP>25) hearing sensitivity at 4 kHz, i.e., the carrier frequency of the EFR stimulus. Figure 4D illustrates that among the listeners with normal TH_{DP}s below 25 dB SPL, the RAM-EFRs of the younger subjects (white symbols) were significantly larger than the older listeners (d.f. = 17, t=3.9, p=0.0005). Even though there were only two young subjects in the DP>25 group, their EFR amplitudes were significantly larger than those of the older subjects (gray symbols) in the same group (d.f. = 23, t=9, p=0.0021). At the same time, there were no significant EFR amplitude differences between the older subjects with normal or impaired TH_{DP}s (d.f = 27, t=1.1, p=0.29), implying the mean TH_{DP} difference of 20.9 dB between the groups had no effect on the EFR amplitude in the older group. Taken together, these findings support an age-related CS interpretation of the RAM-EFR reductions observed in the older human subjects. Going further, we investigate the degree to which the RAM-EFR marker can predict speech intelligibility declines in older listeners.

Speech recognition thresholds

Individual and group speech reception thresholds (SRTs) are depicted in Fig.5 for quiet (SiQ; panel A) and noise (SiN, panel B) backgrounds. For each filtered condition, SRTs of the $yNH_{control}$ group are compared to the older group, which was further subdivided into oNH and oHI study groups. The pooled data is also shown as a post-hoc separation into subjects with normal or impaired 4-kHz TH_{DP}s.

A two-way (3x3) mixed-design ANOVA analysis investigated the role of participant group ($yNH_{control}$, oNH, oHI) and filter-condition (LP,HP,BB) on the SRT (Table 1). Apart from significant main effects of group (SiQ: $F_{(2,39)} = 32$, $p < 0.0001$; SiN: $F_{(2,39)} = 32.6$, $p < 0.0001$) and filter-condition (SiQ: $F_{(2,78)} = 475.84$, $p < 0.0001$; SiN: $F_{(2,78)} = 714.8$, $p < 0.0001$), the interaction terms were also significant (SiQ: $F_{(4,82)} = 30.34$, $p < 0.0001$; SiN: $F_{(4,82)} = 18.34$, $p < 0.0001$), indicating that group SRTs were differently affected by the filtering. The analysis was repeated for a different grouping using the $yNH_{control}$ group and $TH_{DP} = 25$ dB criterion to divide the older listeners into a normal or hearing-impaired subgroups, but the outcomes remained the same.

To investigate whether the quiet or noise background affected the group results, the SRTs were standardised using z-scores and included in a 3x6 mixed-design ANOVA. A subsequent pairwise post-hoc comparison (Bonfer-

roni corrected, 153 tests), showed that adding background noise particularly affected the performance of the oNH group in the LP condition (t=-9.8, $p<0.0001$), in line with how age-related synaptopathy can be expected to affect speech processing in the presence of normal thresholds. The other conditions and groups did not show significant z-score differences when adding the noise, pointing to a noise-robustness in the $yNH_{control}$ group and a possible dual degraded SRT performance in the oHI group. Support for this latter statement stems from the observation that the $SRT_{SiQ\text{-}HP}$ was significantly different between the oNH and oHI groups (t=7.07, $p<0.0001$) and the oNH group performed as poorly as the oHI group in the $SRT_{SiN\text{-}HP}$ condition (t=3.5, p=0.16), while the oNH performed significantly poorer than the $yNH_{control}$ group in both conditions ($p<0.0001$).

A cursory examination Fig.5 further reveals that the SRT_{HP} was overall worse than SRT_{LP}. This effect was not dependent on the use of background noise, which indicates that German speech intelligibility relies more on the speech frequency information below 1.5 kHz. In all tested conditions, the older group had worse SRTs than the $yNH_{control}$ group, and when considering the SRTs of NH_{DP} and HI_{DP} groups, it is clear that the younger participants (white markers) had the best SRTs in each subgroup. This indicates that age had a stronger effect on the SRT than did the TH_{DP}, and this bears resemblance to the earlier observed trends for the RAM-EFR amplitudes. Next, we investigate whether there is a relationship between individual speech

intelligibility and the RAM-EFR marker of CS.

Table 1:

SiQ ($yNH_{control}$,oNH,oHI)	Df	Sum Sq	Mean Sq	F	p
Filter(BB,LP,HP)	2	10641	5320	475.48	< 2e-16 ***
Group:Filter	4	1358	340	30.34	3.15e-15 ***
Residuals	78	873	11		
SiN ($yNH_{control}$,oNH,oHI)					
Filter(BB,LP,HP)	2	1540.5	770.3	714.76	< 2e-16 ***
Group:Filter	4	79.1	19.8	18.34	1.17e-10 ***
Residuals	78	84.1	1.1		

Results of a two-way ANOVA analysis conducted in R using group (yNH, oNH, oHI) and filter condition (BB, LP, HP) as predictor variables and the SRT-in-quiet (SiQ) or SRT-in-noise (SiN) as the outcome variable.

Speech reception thresholds: individual differences

Figures 6 and 7 depict the relationship between RAM (top) or SAM (bottom) EFR amplitudes and SRT_{SiQ} or SRT_{SiN}, respectively. Overall, the SRT related most strongly to the RAM, not SAM, EFR amplitude. After correcting for multiple comparisons (n=12, p<0.0042), all SRT_{SiQ} conditions correlated significantly to the RAM EFR, while for the SRT_{SiN} conditions, only the BB and HP condition remained significant. This illustrates that as the EFR marker became more sensitive to detect individual CS differences (RAM vs. SAM), the marker became better at predicting individual speech recognition differences. Secondly, the EFR markers particularly targeted

TENV coding mechanisms given its 4-kHz carrier frequency. Consequently, it better predicted the SRT_{HP} than SRT_{LP}. To support a hypothesis wherein RAM EFR and SRTs are simultaneously affected by an underlying CS cause, it is necessary to rule out other factors that may have affected their relationship. Such confounding factors relate to aspects which only impact speech recognition (e.g., aging) but not the RAM EFR, or to a dominance of other inter-correlated SNHL pathologies (e.g., OHC damage) which could drive both metrics in the relationship.

The role of OHC damage and age in predicting the SRT_{HP}

To factor out the mediating effect of OHC damage on the strongest observed relationship between the RAM EFR and $SRT_{SiN\text{-}HP}$, we computed partial correlation coefficients to control for TH_{DP}. The relationship between RAM EFR and SRT_{SiN} decreased from ρ = -0.73 to -0.45 but remained significant ($p < 0.006$, $n = 44$). For the SRT_{SiQ}, the significant relationship disappeared (ρ = -0.26, $p = 0.11$), reflecting a dominant influence of hearing sensitivity on processing speech in quiet. When correcting for TH_{DP} in the subgroup with normal hearing thresholds, the correlation remained strong (ρ_{SiN} = -0.82 , $p < 0.0001$ and ρ_{SiQ} = -0.60 , $p < 0.017$, $n = 19$), suggesting that for people with normal hearing sensitivity, the RAM EFR marker is a good predictor of individual speech recognition. The correlation disappeared completely when considering only older listeners (n=29). The individual variability in SRTs was rather large compared to that of the RAM EFRs in

this subgroup, which could be interpreted as follows: for those with similarly (large) degrees of CS, OHC damage is the dominant factor affecting both metrics in the relationship.

Aging could have affected the SRT and RAM EFR differently, hence we considered its independent contribution. Whereas a prominent age effect was observed on the RAM EFR when pooling the data (ρ = -0.66, p<0.0001, n=44), the age effect disappeared when considering the subgroup of older participants. At the same time, there was only a weak age effect within the $\text{yNH}_{\text{control}}$ subgroup (ρ = -0.57, p<0.041, n=13). These results indicate there is a general age-related trend for CS exists, whereas two individuals within the same age decade can still have different degrees of CS.

A role for CS in predicting the HP SRT

Table 2 shows the results from a multiple regression and commonality analyses which considered the entire cohort (n=44) and three independent variables RAM EFR (β_1), TH_{DP} (β_2) and age (β_3) using the following equation for the dependent variable $\text{SRT}_{\text{SiN-HP}}$:

$$\text{SRT}_{\text{SiN-HP}} = \beta_0 + \beta_1 * \text{RAM EFR}_{\text{PtN}} + \beta_2 * \text{TH}_{\text{DP}} + \beta_3 * \text{Age} + \epsilon. \quad (4)$$

When considering the variables separately, all had a significant impact on $\text{SRT}_{\text{SiN-HP}}$ with R^2 in range of 0.59 (RAM EFR), 0.49 TH_{DP} and 0.38 (Age),

respectively. Because Age is not independent from either RAM EFR nor TH_{DP}, its inclusion into any considered multiple regression analysis always dominated the model. The only regression model that considered peripheral hearing factors without general age effects, indicates that the RAM EFR was the stronger predictor of $SRT_{SiN\text{-}HP}$ ($p<0.001$ with unique contribution of 30.37%) over the TH_{DP}. Adding more weight to the statistical independence of these two independent variables in the regression model, we earlier showed that the RAM EFR was not significantly different between older groups with normal or impaired audiograms. Secondly, auditory model simulations of the human EFR generators performed in Vasilkov et al. (2021); Van Der Biest et al. (2023) show that simulated OHC damage only has a 5-10 % impact on the 4-kHz RAM EFR amplitude whereas CS heavily impacts the response (upto 81%). Taken together, the RAM-EFR marker has an independent and strong contribution in explaining the individual variability in $SRT_{SiN\text{-}HP}$.

Speech processing in quiet is not to the same degree impacted by CS than is speech processing in background noise. Even though both HP-filtered conditions rely mostly on TENV processing, which in several animal studies has been shown to be compromised due to synaptopathy (Parthasarathy & Kujawa, 2018; Shaheen et al., 2015), the relationship between the RAM-EFR marker of CS and $SRT_{SiQ\text{-}HP}$ is weaker compared to that with $SRT_{SiN\text{-}HP}$. As the multiple regression analysis for $SRT_{SiQ\text{-}HP}$ in Table 3 shows, TH_{DP} becomes the dominant predictor of performance when both $SRT_{SiQ\text{-}HP}$ and RAM-EFR are considered. This confirms a strong relation between SRT_{SiQ}

and audibility or OHC integrity (Festen & Plomp, 1983; Papakonstantinou et al., 2011), whereas $SRT_{SiN\text{-}HP}$ performance may be much more driven by supra-threshold TENV deficits associated with (CS-induced) temporal coding deficits (Lopez-Poveda & Barrios, 2013).

Table 2:

$SRT_{SiN\text{-}HP}$	R^2	Adj. R^2	F (df_1,df_2)	p	parameters	unique	common
RAM EFR	0.49	0.47	39.63 (1,42)	<0.0001	β_1=-32.31***		
					ϵ=5.01***		
TH_{DP}	0.38	0.36	25.84 (1,42)	<0.0001	β_2=0.12***		
					ϵ=-1.24		
Age	0.59	0.58	59.87 (1,42)	<0.0001	β_3=0.12***		
					ϵ=-3.86***		
RAM EFR	0.62	0.60	33.57 (2,41)	<0.0001	β_1=-12.82 .	5.34%	72.85%
Age					β_3=0.09***	21.81%	
					ϵ=-1.11*		
TH_{DP}	0.61	0.58	31.9 (2,41)	<0.0001	β_2=0.04	3.46%	59.12%
Age					β_3=0.1***	37.42%	
					ϵ=-3.94***		
RAM EFR	0.55	0.53	24.76 (2,41)	<0.0001	β_1=-23.65***	30.37%	58.38%
TH_{DP}					β_2=0.06*	11.26%	
					ϵ=2.50*		
RAM EFR	0.63	0.6	22.99 (3,40)	<0.0001	β_1=-11.18	3.82%	49.04%
TH_{DP}					β_2=0.03	1.90%	β_1 & β_2: 1.42%
Age					β_3=0.08**	13.57%	β_2 & β_3: 7.82%
					ϵ=-1.53		β_1 & β_3: 22.43%

Regression models for SRT_{HP} in noise. Significance codes: *** : p≤0.001, ** : p≤0.01, * : p≤0.05, . : p≤ 0.1

Table 3:

$SRT_{SiQ\text{-}HP}$	R^2	Adj. R^2	F (df_1,df_2)	p	parameters	unique	common
RAM-EFR	0.38	0.36	25.33 (1,42)	<0.0001	β_1=-102.68***		
					ϵ=54.52***		
TH_{DP}	0.52	0.51	46.21 (1,42)	<0.0001	β_2=0.50***		
					ϵ=30.87***		
Age	0.56	0.54	52.41 (1,42)	<0.0001	β_3=0.43***		
					ϵ=24.28***		
RAM-EFR	0.56	0.54	26.26 (2,41)	<0.0001	β_1=-20.34	1.14%	64.85%
Age					β_3=0.38***	33.01%	
					ϵ=28.63***		
TH_{DP}	0.65	0.63	37.94 (2,41)	<0.0001	β_2=0.28**	14.49%	66.21%
Age					β_3=0.27***	19.30%	
					ϵ=23.63***		
RAM-EFR	0.57	0.55	27.57 (2,41)	<0.0001	β_1=-46.71*	8.67%	56.92%
TH_{DP}					β_2=0.38***	34.42%	
					ϵ=38.24***		
RAM-EFR	0.65	0.62	24.7 (3,40)	<0.0001	β_1=-4.40	0.04%	49.32%
TH_{DP}					β_2=0.28**	13.54%	β_1 & β_2: 0.94%
Age					β_3=0.26**	11.69%	β_2 & β_3: 16.86%
					ϵ=24.59***		β_1 & β_3: 7.61%

Regression models for SRT_{HP}. Significance codes: *** : p≤0.001, ** : p≤0.01, * : p≤0.05, . : p≤ 0.1

Discussion

Understanding how different aspects of sensorineural hearing damage contribute to speech recognition deficits is still a major unsolved problem in hearing science, especially given the challenges associated with diagnosing synaptopathy non-invasively in humans (Plack et al., 2016; Hickox et al., 2017; Kobel et al., 2017; Bramhall et al., 2019). However, we have to consider synaptopathy as an important facet of SNHL based on animal and human temporal-bone studies on the progression of synaptopathy with age (Sergeyenko et al., 2013; Parthasarathy & Kujawa, 2018; Wu et al., 2018). This importance is further supported by the observation that synaptopathy sets in before OHC damage after noise-exposure (e.g. Kujawa & Liberman, 2009; Fernandez et al., 2015). Combined, these studies project its prevalence to largely exceed the WHO-estimated 5.3% world population, who suffer from a disabling hearing loss as diagnosed audiometrically (Stevens et al., 2013; WHO, 2019).

A role for synaptopathy in speech intelligibility deficits

Using a multiple regression and commonality analysis, we described to which degree two aspects of SNHL (synaptopathy and OHC-damage) were able to explain individual speech reception thresholds. We considered a DPOAE threshold variable as a well-known marker of OHC damage (e.g. Davis et al., 2005) and a RAM-EFR variable as a marker for the individual degree of synaptopathy. Model simulations and experimental findings

provided theoretical and circumstantial evidence that the two variables included in the regression analysis can be treated as independent markers of OHC damage and synaptopathy, respectively. Model simulations were crucial to the argument that the 4-kHz RAM-EFR marker is maximally sensitive to synaptopathy, even in the presence of co-existing OHC deficits (Vasilkov et al., 2021; Van Der Biest et al., 2023). We used these model simulations in combination with an experimental KA-induced synaptopathy approach to also establish the specificity of our RAM-EFR marker to CS.

Our human RAM-EFR results corroborate animal research findings which relate deficits in temporal coding at the earliest neural stages of the auditory pathway to progressive/noise-induced CS (Parthasarathy et al., 2014; Fernandez et al., 2015; Shaheen et al., 2015; Parthasarathy & Kujawa, 2018). Specifically, our results show an age-related decline in EFR amplitudes, even in the absence of OHC damage. By removing the low frequency content from the speech material to target TENV coding mechanisms in the behavioral task, we may have exacerbated the compromising effect of synaptopathy on speech recognition, as it is well known that TENV information is crucial for speech comprehension (Shannon et al., 1995; Ding & Simon, 2012).

Our speech recognition results fell in line with this prediction, as the performance-difference between the oNH and oHI group decreased from the SiQ to SiN condition, but only for the high-pass filtered speech material

(Fig. 5). In line with the prevailing hypothesis that CS would affect supra-threshold SiN coding more so than SiQ coding, we showed a strong predictive power of the RAM EFR on the high-pass, but not low-pass, filtered speech-in-noise recognition test. The perception of the low-pass filtered condition may have been influenced by competing perceptual cues such as temporal-fine-structure cues from cochlear CF regions < 1 kHz (Lorenzi et al., 2006) that were not captured by the RAM-EFR marker. Lastly, individual SiQ performance was more strongly predicted by the 4-kHz DPOAE threshold as a proxy of OHC integrity than by the RAM-EFR marker, suggesting a lesser impact of CS when the behavioral task is easier. Based on these results, we conclude that age-related synaptopathy plays an important role for speech intelligibility in noise.

The role of OHC damage in speech intelligibility declines

As permanent OHC damage is preceded by synaptopathy in the general aging process (Sergeyenko et al., 2013; Fernandez et al., 2015; Parthasarathy & Kujawa, 2018) or after noise exposure (Kujawa & Liberman, 2009; Furman et al., 2013), markers of OHC damage might also be a predictor of synaptopathy. Even though these two pathologies are unrelated in their pathogenesis and differ in their consequences on auditory processing, they share some of the same risk factors (e.g. age, noise exposure) and can both have an impact on EFR strength (Garrett & Verhulst, 2019). So even if the RAM-EFR marker primarily captures synaptopathy-related information (as the model

simulations and KA-manipulations predict), its strength might still relate to auditory threshold measures if people with impaired audiograms also suffer from synaptopathy.

At the same time, the SRT performance can also be affected by both synaptopathy and OHC damage (Hoben et al., 2017; Holmes & Griffiths, 2019) aspects, the latter known to result in compromised frequency selectivity (Glasberg & Moore, 1986). This aspect may have played a role in the absence of an EFR-SRT relationship within the oHI subgroup, wherein the SRT of all listeners was compromised by CS, but potentially more so for those with more severe degrees OHC damage. This effect was also seen in the group analysis which showed significantly worse $SRT_{SiN\text{-}HP}$ performance in the oHI than yNH group.

Context with prior human studies

This study considered and minimized several aspects that have troubled prior human studies on the causal relationship between speech-in-noise intelligibility deficits and EFR markers of synaptopathy. One reason for the mixed conclusions relates to the low sensitivity of the classically-used SAM-based EFR metrics. We show that, in the same individuals, the RAM EFR was overall larger, and thereby more strongly affected by KA-induced CS than the conventional SAM-EFR (Fig.4) This results in a greater sensitivity of the RAM compared to SAM-EFR to detect individual CS differences

and SRT differences (Fig.6 and 7). In addition, our prior model simulations have shown that the 4-kHz SAM EFR is more affected by OHC damage than the RAM EFR, reflecting a better specificity of the RAM EFR to CS (Vasilkov et al., 2021; Van Der Biest et al., 2023). A second factor that may have impacted prior studies relates to the mismatch in frequency content and associated auditory coding mechanisms between the compared metrics in the relationship. This mismatch may explain the absence of significant relationships between sound perception and physiological markers of reduced temporal coding in listeners with otherwise normal audiograms: Grose et al. (2017); Prendergast et al. (2017); Guest et al. (2018). To ensure that the perceptual measure was limited to the same TENV mechanisms that underlie the 4-kHz RAM-EFR generators, we limited the frequency content of the speech material to frequencies above 1.65 kHz, which roughly corresponds to the human phase-locking limit (Verschooten et al., 2015). Another aspect that hindered earlier studies relates to the possibility that detrimental effects of CS on auditory function might primarily be present in populations with co-existing OHC dysfunction (Bramhall et al., 2019). This complexity presumes that the CS marker needs to be independent from OHC damage markers (TH_{DP} or audiogram thresholds) to investigate the independent role of CS in sound perception.

Lastly, it remains difficult to dissociate general age effects from age-related CS in human studies. Earlier reports that included older participants

or those with extreme noise-exposure histories tended to find electrophysiological evidence for synaptopathy or reduced temporal coding fidelity in the considered test populations (e.g. Anderson et al., 2011, 2012; Konrad-Martin et al., 2012; Clinard & Tremblay, 2013; Schoof & Rosen, 2016; Bramhall et al., 2017; Valderrama et al., 2018; Bramhall et al., 2021). Within this context, our study confirms that age is inherently linked to the development of SNHL, but that the same age variable is a poor predictor of the CS degree in our age-restricted samples (older or $yNH_{control}$ subgroups). On a more general level, age is associated with the development of both OHC damage (Lin et al., 2011; ISO, 2017) and CS (Schmiedt et al., 1996; Makary et al., 2011; Konrad-Martin et al., 2012; Möhrle et al., 2016; Parthasarathy & Kujawa, 2018) and can hence be responsible for the observed stronger age vs RAM-EFR relationship when pooling the data across groups (ρ = -0.66, p $<$ 0.0001; n = 44). Because age was linked to *all* considered facets of SNHL and its biomarkers, its inclusion in statistical models exacerbates the interpretation of the unique underlying effects of CS and OHC damage.

Study limitations

Even though we demonstrate that the RAM-EFR marker captures synaptopathy-related variability in a budgerigar model, and used a human group and regression analysis to motivate that CS plays an important role in maintaining normal speech recognition in noise as we age, it is important to note the

limitations of our approach.

(i) Present human experimental limitations prevent us from establishing a direct relationship between CS and speech intelligibility in live humans and hence we need to rely on circumstantial evidence to support our conclusions. On the one hand, we relied on model simulations that studied the effect of various SNHL pathologies on the cochlear and neuronal generators of the EFR (Vasilkov et al., 2021; Van Der Biest et al., 2023), and on the other, we used an animal model to demonstrate the sensitivity of our EFR markers to synaptopathy. While the convergence of model and experimental results in several research studies (e.g. Verhulst et al., 2018; Encina-Llamas et al., 2019; Keshishzadeh et al., 2020; Vasilkov et al., 2021; Buran et al., 2022) lends credence to the validity of the model predictions for the purpose of this study, we need to keep in mind that state-of-the-art models are a good, but likely imperfect, representation of all cochlear and neural mechanisms involved in EFR generation. At the same time, we need to consider that species differences might pose a confound to the direct comparison between budgerigar and human EFRs. However, neural recordings from the budgerigar inferior colliculus suggest that they share similar processing strategies to the mammalian midbrain (Henry et al., 2017).

(ii) All oNH participants had normal audiometric thresholds up to 4 kHz, and DPOAE thresholds confirmed that the observed threshold differences stemmed from OHC damage. Nevertheless, the unfiltered and high-pass

filtered speech material included speech energy beyond 4-kHz, which implies that hearing sensitivity differences beyond 4 kHz may have influenced the SRT results, whereas this influence was smaller on the 4-kHz RAM EFR. Secondly, recent work on ANF coding (Henry et al., 2016; Encina-Llamas et al., 2019; Vasilkov & Verhulst, 2019) has shown that for supra-threshold stimulation, TENV information can be encoded through the tails of high-SR ANF tuning curves at basal cochlear regions. This implies that the generators of the 4-kHz EFR may also have spanned broader generator region than expected Encina-Llamas et al. (see also 2019). Accounting for these potential confounds (e.g. through high-frequency masking noise) in future work might further strengthen the relationship between RAM EFR and speech reception performance.

(iii) Significant correlations between the RAM EFR and $\mathrm{SRT}_{\mathrm{SiN\text{-}HP}}$ were only observed when pooling data from subgroups and thereby increasing the variance in outcome measures. Within the smaller groups ($\mathrm{yNH}_{\mathrm{control}}$, oNH, oHI), EFR-SRT relationships were only significant for the $\mathrm{yNH}_{\mathrm{control}}$ group wherein OHC function was normal. The large EFR amplitude range observed in this group corroborates findings from animal studies reporting that synaptopathy can be exacerbated by noise exposure levels which do not cause OHC damage (Fernandez et al., 2015). So within a $\mathrm{yNH}_{\mathrm{control}}$ group, individual RAM-EFR differences might be explained by previously accrued CS. The absence of significant EFR-SRT correlations within the older subgroups can be explained by the overall reduced EFR amplitudes

in older listeners (Parthasarathy & Kujawa, 2018; Dimitrijevic et al., 2016; Garrett & Verhulst, 2019), rendering it more difficult to observe coherent correlation trends between both metrics in those subgroups (Figs. 6 and 7). Future studies considering an age-gradient across a larger cohort of study participants may further shed light on the dynamics between age, CS and speech intelligibility.

(iv) Pooling data from groups which differ in more than one factor can introduce additional explanatory between-group-factors that were not controlled for. For example, as it is known that cognitive factors such as memory and attention decrease with age and are linked to speech-in-noise understanding (Humes et al., 2010; Humes, 2013; Yeend et al., 2017), it is likely that some of the unaccounted variance can be attributed to these factors, especially as our analyses included groups with considerable age differences. Nevertheless, the matrix test is designed to minimize memory effects by randomly generating word sequences and by minimizing the cognitive load through immediate recall of just five words at a time. As EFR recordings to stimuli with high modulation frequencies are believed to be free from top-down attention (e.g. Varghese et al., 2015) or memory effects, the strong correlations between the RAM-EFR and $SRT_{SiN\text{-}HP}$ scores suggest that cognitive contributions are not a unique dominating factor, but rather interact with peripheral encoding deficits (Johannesen et al., 2016). We argue that for the purpose of this study, pooling of data from different homogeneous groups is a valid, and even necessary, approach to be able to disentangle the relative contributions

of OHC damage and synaptopathy to speech intelligibility.

Conclusion

We conclude that age-related synaptopathy is an important hearing health issue, as we provide experimental and theoretical evidence that it strongly affects auditory TENV coding and speech intelligibility in noise. Sensitive hearing diagnostics is the first and foremost step towards understanding the role of synaptopathy for impaired sound perception. The adopted RAM-EFR marker can be considered a robust, widely applicable and selective marker for cochlear synaptopathy, and enabled us to draw these conclusions. With the established finding that synaptopathy is involved in speech perception declines in aging, future therapeutic interventions that aim to compensating for synaptopathy and its functional consequences can be developed.

Conflict of interest statement

Ghent University owns a patent (US Patent App. 17/791,985) related to the RAM-EFR methods adopted in this paper. Sarah Verhulst and Viacheslav Vasilkov are inventors.

Acknowledgments

This work was supported by the DFG Cluster of Excellence EXC 1077/1 "Hearing4all" (MG, MM, SV), the European Research Council (ERC) under the Horizon 2020 Research and Innovation Programme (grant agreement

No 678120 RobSpear; VV, SV), National Institutes of Health grant R01 DC017519 (KH) and a National Institutes of Health Predoctoral National Research Service Award Fellowship (TL1 TR002000) administered by the University of Rochester Clinical and Translational Science Institute (JW). The authors would like to thank the study participants as well as the Hörzentrum Oldenburg for helping with participant recruitment. Lastly, we thank Sarineh Keshishzadeh for help with the analysis scripts and data storage and labelling throughout the project.

Author Contributions

MG: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing: Original Draft, Visualization; VV: Methodology, Software, Investigation,; MM: Methodology, Software, Investigation; JW: Methodology, Software, Investigation; KH: Methodology, Software, Investigation; SV: Conceptualization, Methodology, Investigation, Resources, Writing: Original Draft, Writing: Review & Editing, Supervision, Project administration, Funding acquisition.

References

Abdala, C., Ortmann, A. J., & Guardia, Y. C. (2021). Weakened cochlear nonlinearity during human aging and perceptual correlates. *Ear and Hearing*, *42*, 832–845.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *The Journal of Neuroscience*, *32*, 14156–14164. doi:10.1523/JNEUROSCI.2176-12.2012.

Anderson, S., Parbery-Clark, A., Yi, H.-G., & Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear and Hearing*, *32*, 750–757. doi:10.1097/AUD.0b013e31822229d3.

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *The Journal of Neuroscience*, *35*, 2161–2172. doi:10.1523/JNEUROSCI.3915-14.2015.

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., & Shinn-Cunningham, B. G. (2014). Cochlear neuropathy and the coding of suprathreshold sound. *Frontiers in Systems Neuroscience*, *8*, 26. doi:10.3389/fnsys.2014.00026.

Bledsoe, S. C., Bobbin, R. P., & Chihal, D. M. (1981). Kainic acid: an evaluation of its action on cochlear potentials. *Hearing Research*, *4*, 109–120. doi:10.1016/0378-5955(81)90040-x.

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.

Borjigin, A., & Bharadwaj, H. M. (2023). Individual differences reveal the utility of temporal fine-structure processing for speech perception in noise. *bioRxiv*, .

Bramhall, N., Beach, E. F., Epp, B., Le Prell, C. G., Lopez-Poveda, E. A., Plack, C. J., Schaette, R., Verhulst, S., & Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing Research*, *377*, 88–103. doi:10.1016/j.heares.2019.02.016.

Bramhall, N., Ong, B., Ko, J., & Parker, M. (2015). Speech perception ability in noise is correlated with auditory brainstem response wave i amplitude. *Journal of the American Academy of Audiology*, *26*, 509–517.

Bramhall, N. F., Konrad-Martin, D., McMillan, G. P., & Griest, S. E. (2017). Auditory Brainstem Response Altered in Humans With Noise Exposure Despite Normal Outer Hair Cell Function. *Ear and Hearing*, *38*, e1–e12. doi:10.1097/AUD.0000000000000370.

Bramhall, N. F., McMillan, G. P., & Kampel, S. D. (2021). Envelope following response measurements in young veterans are consistent with noise-induced cochlear synaptopathy. *Hearing Research*, *408*, 108310.

Brand, T., & Kollmeier, B. (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *The Journal of the Acoustical Society of America*, *111*, 2801–2810. doi:10.1121/1.1479152.

Buran, B. N., McMillan, G. P., Keshishzadeh, S., Verhulst, S., & Bramhall, N. F. (2022). Predicting synapse counts in living humans by combining computational models with auditory physiology. *The Journal of the Acoustical Society of America*, *151*, 561–576.

Clinard, C. G., & Tremblay, K. L. (2013). Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *Journal of the American Academy of Audiology*, *24*, 590–599; quiz 643–644. doi:10.3766/jaaa.24.7.7.

Davis, B., Qiu, W., & Hamernik, R. P. (2005). Sensitivity of distortion product otoacoustic emissions in noise-exposed chinchillas. *Journal of the American Academy of Audiology*, *16*, 69–78. doi:10.3766/jaaa.16.2.2.

Dent, M. L., Dooling, R. J., & Pierce, A. S. (2000). Frequency discrimination in budgerigars (Melopsittacus undulatus): effects of tone duration and tonal context. *The Journal of the Acoustical Society of America*, *107*, 2657–2664. doi:10.1121/1.428651.

Dimitrijevic, A., Alsamri, J., John, M. S., Purcell, D., George, S., & Zeng, F.-G. (2016). Human Envelope Following Responses to Amplitude Modulation: Effects of Aging and Modulation Depth. *Ear and Hearing*, *37*, e322–335. doi:10.1097/AUD.0000000000000324.

Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National*

Academy of Sciences of the United States of America, *109*, 11854–11859. doi:10.1073/pnas.1205381109.

Dooling, R. J., Lohr, B., & Dent, M. L. (2000). Hearing in Birds and Reptiles. In R. J. Dooling, R. R. Fay, & A. N. Popper (Eds.), *Comparative Hearing: Birds and Reptiles* Springer Handbook of Auditory Research (pp. 308–359). New York, NY: Springer. URL: https://doi.org/10.1007/978-1-4612-1182-2_7. doi:10.1007/978-1-4612-1182-2_7.

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. doi:10.1152/jn.00326.2006.

Dupont, C., Horner, J., Li, C., Liu, Q., & Shepherd, B. (2018). PResiduals: Probability-Scale Residuals and Residual Correlations. URL: https://CRAN.R-project.org/package=PResiduals.

Encina-Llamas, G., Harte, J. M., Dau, T., Shinn-Cunningham, B., & Epp, B. (2019). Investigating the Effect of Cochlear Synaptopathy on Envelope Following Responses Using a Model of the Auditory Nerve. *Journal of the Association for Research in Otolaryngology: JARO*, *20*, 363–382. doi:10.1007/s10162-019-00721-7.

Fernandez, K. A., Jeffers, P. W. C., Lall, K., Liberman, M. C., & Kujawa, S. G. (2015). Aging after noise exposure: acceleration of cochlear synap-

topathy in "recovered" ears. *The Journal of Neuroscience*, *35*, 7509–7520. doi:10.1523/JNEUROSCI.5138-14.2015.

Festen, J. M., & Plomp, R. (1983). Relations between auditory functions in impaired hearing. *The Journal of the Acoustical Society of America*, *73*, 652–662. doi:10.1121/1.388957.

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. doi:10.1152/jn.00164.2013.

Garrett, M., & Verhulst, S. (2019). Applicability of subcortical EEG metrics of synaptopathy to older listeners with impaired audiograms. *Hearing Research*, *380*, 150–165. doi:10.1016/j.heares.2019.07.001.

Glasberg, B. R., & Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America*, *79*, 1020–1033. doi:10.1121/1.393374.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. doi:10.3389/fnins.2013.00267.

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. doi:10.1016/j.neuroimage.2013.10.027.

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

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. doi:10.1016/j.heares.2018.03.008.

Henry, K. S., & Abrams, K. S. (2018). Persistent Auditory Nerve Damage Following Kainic Acid Excitotoxicity in the Budgerigar (Melopsittacus undulatus). *Journal of the Association for Research in Otolaryngology: JARO*, *19*, 435–449. doi:10.1007/s10162-018-0671-y.

Henry, K. S., Abrams, K. S., Forst, J., Mender, M. J., Neilans, E. G., Idrobo, F., & Carney, L. H. (2017). Midbrain Synchrony to Envelope Structure Supports Behavioral Sensitivity to Single-Formant Vowel-Like Sounds in Noise. *Journal of the Association for Research in Otolaryngology: JARO*, *18*, 165–181. doi:10.1007/s10162-016-0594-4.

Henry, K. S., Kale, S., & Heinz, M. G. (2016). Distorted Tonotopic Coding of

Temporal Envelope and Fine Structure with Noise-Induced Hearing Loss. *The Journal of Neuroscience*, *36*, 2227–2237. doi:10.1523/JNEUROSCI.3944-15.2016.

Hickox, A. E., Larsen, E., Heinz, M. G., Shinobu, L., & Whitton, J. P. (2017). Translational issues in cochlear synaptopathy. *Hearing Research*, *349*, 164–171. doi:10.1016/j.heares.2016.12.010.

Hoben, R., Easow, G., Pevzner, S., & Parker, M. A. (2017). Outer Hair Cell and Auditory Nerve Function in Speech Recognition in Quiet and in Background Noise. *Frontiers in Neuroscience*, *11*. doi:10.3389/fnins.2017.00157.

Holmes, E., & Griffiths, T. D. (2019). 'Normal' hearing thresholds and fundamental auditory grouping processes predict difficulties with speech-in-noise perception. *Scientific Reports*, *9*, 16771. doi:10.1038/s41598-019-53353-5.

Hopkins, K., Moore, B., & Stone, M. (2008). Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *The Journal of the Acoustical Society of America*, *123*, 1140–53. doi:10.1121/1.2824018.

Hopkins, K., & Moore, B. C. J. (2010). The importance of temporal fine structure information in speech at different spectral regions for normal-

hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *127*, 1595–1608. doi:10.1121/1.3293003.

Humes, L. E. (2013). Understanding the speech-understanding problems of older adults. *American Journal of Audiology*, *22*, 303–305. doi:10.1044/1059-0889(2013/12-0066).

Humes, L. E., Kewley-Port, D., Fogerty, D., & Kinney, D. (2010). Measures of Hearing Threshold and Temporal Processing across the Adult Lifespan. *Hearing research*, *264*, 30–40. doi:10.1016/j.heares.2009.09.010.

Humphrey, R. (2008). Playrec, Multi-channel Matlab Audio. URL: http://www.playrec.co.uk.

ISO (2017). *Acoustics - Statistical distribution of hearing thresholds related to age and gender (7029)*. Switzerland: International Organization of Standardization, ISO. URL: https://www.iso.org/standard/42916.html.

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. doi:10.1016/j.heares.2019.01.017.

Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., & Lopez-Poveda, E. A. (2016). The Influence of Cochlear Mechanical Dysfunction, Temporal Processing Deficits, and Age on the Intelligibility of Audible

Speech in Noise for Hearing-Impaired Listeners. *Trends in Hearing*, *20*, 2331216516641055. doi:10.1177/2331216516641055.

Joris, P. X., & Verschooten, E. (2013). On the limit of neural phase locking to fine structure in humans. *Advances in Experimental Medicine and Biology*, *787*, 101–108. doi:10.1007/978-1-4614-1590-9_12.

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. doi:10.1121/1.402757.

Keshishzadeh, S., Garrett, M., Vasilkov, V., & Verhulst, S. (2020). The derived-band envelope following response and its sensitivity to sensorineural hearing deficits. *Hearing Research*, *392*, 107979. URL: http://www.sciencedirect.com/science/article/pii/S0378595519305064. doi:10.1016/j.heares.2020.107979.

Kim, S. (2015). ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *Communications for statistical applications and methods*, *22*, 665–674. doi:10.5351/CSAM.2015.22.6.665.

Kobel, M., Le Prell, C. G., Liu, J., Hawks, J. W., & Bao, J. (2017). Noise-induced cochlear synaptopathy: Past findings and future studies. *Hearing Research*, *349*, 148–154. doi:10.1016/j.heares.2016.12.008.

Konrad-Martin, D., Dille, M. F., McMillan, G., Griest, S., McDermott, D., Fausti, S. A., & Austin, D. F. (2012). Age-related changes in the auditory

brainstem response. *Journal of the American Academy of Audiology*, *23*, 18–75. doi:10.3766/jaaa.23.1.3.

Kraus, N., Anderson, S., & White-Schwoch, T. (2017). The Frequency-Following Response: A Window into Human Communication. In *The Frequency-Following Response: A Window into Human Communication* (pp. 1–15). doi:10.1007/978-3-319-47944-6_1.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. *The Journal of Neuroscience*, *29*, 14077–14085. doi:10.1523/JNEUROSCI.2845-09.2009.

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.

Le Prell, C. G. (2019). Effects of noise exposure on auditory brainstem response and speech-in-noise tasks: a review of the literature. *International Journal of Audiology*, *58*, S3–S32. doi:10.1080/14992027.2018.1534010.

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*, e0162726. doi:10.1371/journal.pone.0162726.

Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing loss prevalence and risk factors among older adults in the United States. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *66*, 582–590. doi:10.1093/gerona/glr002.

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.

Lopez-Poveda, E. A., & Barrios, P. (2013). Perception of stochastically undersampled sound waveforms: a model of auditory deafferentation. *Frontiers in Neuroscience*, *7*, 124. doi:10.3389/fnins.2013.00124.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 18866–18869. doi:10.1073/pnas.0607364103.

Mai, G., & Howell, P. (2023). The possible role of early-stage phase-locked neural activities in speech-in-noise perception in human adults across age and hearing loss. *Hearing Research*, *427*, 108647.

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: JARO*, *12*, 711–717. doi:10.1007/s10162-011-0283-2.

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.

Mauermann, M., & Kollmeier, B. (2004). Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source. *The Journal of the Acoustical Society of America*, *116*, 2199–2212. doi:10.1121/1.1791719.

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. *The Journal of Neuroscience*, *36*, 3755–3764. doi:10.1523/JNEUROSCI.4460-15.2016.

Mepani, A. M., Verhulst, S., Hancock, K. E., Garrett, M., Vasilkov, V., Bennett, K., de Gruttola, V., Liberman, M. C., & Maison, S. F. (2021). Envelope following responses predict speech-in-noise performance in normal-hearing listeners. *Journal of Neurophysiology*, *125*, 1213–1222.

Millman, K. J., & Aivazis, M. (2011). Python for Scientists and Engineers. *Computing in Science Engineering*, *13*, 9–12. doi:10.1109/MCSE.2011.36.

Mitchell, C., Phillips, D. S., & Trune, D. R. (1989). Variables affecting the auditory brainstem response: audiogram, age, gender and head size. *Hearing Research*, *40*, 75–85. doi:10.1016/0378-5955(89)90101-9.

Möhrle, D., Ni, K., Varakina, K., Bing, D., Lee, S. C., Zimmermann, U., Knipper, M., & Rüttiger, L. (2016). 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. doi:10.1016/j.neurobiolaging.2016.05.001.

Neely, S. T., Johnson, T. A., Kopun, J., Dierking, D. M., & Gorga, M. P. (2009). Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears. *The Journal of the Acoustical Society of America*, *126*, 728–738. doi:10.1121/1.3158859.

Newton, R. G., & Spurrell, D. J. (1967). A Development of Multiple Regression for the Analysis of Routine Data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *16*, 51–64. doi:10.2307/2985237.

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. doi:10.3758/BRM.40.2.457.

Oliphant, T. E. (2007). Python for Scientific Computing. *Computing in Science Engineering*, *9*, 10–20. doi:10.1109/MCSE.2007.58.

Papakonstantinou, A., Strelcyk, O., & Dau, T. (2011). Relations between perceptual measures of temporal processing, auditory-evoked brainstem responses and speech intelligibility in noise. *Hearing Research*, *280*, 30–37. doi:10.1016/j.heares.2011.02.005.

Parthasarathy, A., Bartlett, E. L., & Kujawa, S. G. (2019). Age-related Changes in Neural Coding of Envelope Cues: Peripheral Declines and Central Compensation. *Neuroscience*, *407*, 21–31. doi:10.1016/j.neuroscience.2018.12.007.

Parthasarathy, A., Datta, J., Torres, J. A. L., Hopkins, C., & Bartlett, E. L. (2014). Age-related changes in the relationship between auditory brainstem responses and envelope-following responses. *Journal of the Association for Research in Otolaryngology: JARO*, *15*, 649–661. doi:10.1007/s10162-014-0460-1.

Parthasarathy, A., & Kujawa, S. G. (2018). Synaptopathy in the Aging Cochlea: Characterizing Early-Neural Deficits in Auditory Temporal Envelope Processing. *The Journal of Neuroscience*, *38*, 7108–7119. doi:10.1523/JNEUROSCI.3240-17.2018.

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

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*. doi:10.1177/2331216516657466.

Prendergast, G., Millman, R. E., Guest, H., Munro, K. J., Kluk, K., Dewey, R. S., Hall, D. A., Heinz, M. G., & Plack, C. J. (2017). Effects of noise exposure on young adults with normal audiograms II: Behavioral measures. *Hearing Research*, *356*, 74–86. doi:10.1016/j.heares.2017.10.007.

Pujol, R., Lenoir, M., Robertson, D., Eybalin, M., & Johnstone, B. M. (1985). Kainic acid selectively alters auditory dendrites connected with cochlear inner hair cells. *Hearing Research*, *18*, 145–151. doi:10.1016/0378-5955(85)90006-1.

R Core Team (2019). R: A language and environment for statistical computing. URL: https://www.R-project.org/.

Ray-Mukherjee, J., Nimon, K., Mukherjee, S., Morris, D. W., Slotow, R., & Hamer, M. (2014). Using commonality analysis in multiple regressions: a tool to decompose regression effects in the face of multicollinearity. *Methods in Ecology and Evolution*, *5*, 320–328. doi:10.1111/2041-210X.12166.

Russell, V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*, 1–33. doi:doi:10.18637/jss.v069.i01.

Schmiedt, R. A., Mills, J. H., & Boettcher, F. A. (1996). Age-related loss of activity of auditory-nerve fibers. *Journal of Neurophysiology*, *76*, 2799–2803. doi:10.1152/jn.1996.76.4.2799.

Schoof, T., & Rosen, S. (2016). The Role of Age-Related Declines in Subcortical Auditory Processing in Speech Perception in Noise. *Journal of*

the Association for Research in Otolaryngology: JARO, *17*, 441–460. doi:10.1007/s10162-016-0564-x.

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *The Journal of Neuroscience*, *33*, 13686–13694. doi:10.1523/JNEUROSCI.1783-13.2013.

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: JARO*, *16*, 727–745. doi:10.1007/s10162-015-0539-3.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*, 303–304. doi:10.1126/science.270.5234.303.

Stevens, G., Flaxman, S., Brunskill, E., Mascarenhas, M., Mathers, C. D., Finucane, M., & Global Burden of Disease Hearing Loss Expert Group (2013). Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. *European Journal of Public Health*, *23*, 146–152. doi:10.1093/eurpub/ckr176.

Sun, H., Hashino, E., Ding, D. L., & Salvi, R. J. (2001). Reversible and irreversible damage to cochlear afferent neurons by kainic acid excitotoxi-

city. *The Journal of Comparative Neurology*, *430*, 172–181. doi:10.1002/1096-9861(20010205)430:2<172::aid-cne1023>3.0.co;2-w.

Sun, H., Salvi, R. J., Ding, D. L., Hashino, D. E., Shero, M., & Zheng, X. Y. (2000). Excitotoxic effect of kainic acid on chicken otoacoustic emissions and cochlear potentials. *The Journal of the Acoustical Society of America*, *107*, 2136–2142. doi:10.1121/1.428495.

Trune, D. R., Mitchell, C., & Phillips, D. S. (1988). The relative importance of head size, gender and age on the auditory brainstem response. *Hearing Research*, *32*, 165–174. doi:10.1016/0378-5955(88)90088-3.

Valderrama, J. T., Beach, E. F., Yeend, I., Sharma, M., Van Dun, B., & Dillon, H. (2018). Effects of lifetime noise exposure on the middle-age human auditory brainstem response, tinnitus and speech-in-noise intelligibility. *Hearing Research*, *365*, 36–48. doi:10.1016/j.heares.2018.06.003.

Van Der Biest, H., Keshishzadeh, S., Keppler, H., Dhooge, I., & Verhulst, S. (2023). Envelope following responses for hearing diagnosis: Robustness and methodological considerations. *The Journal of the Acoustical Society of America*, *153*, 191–208.

Varghese, L., Bharadwaj, H. M., & Shinn-Cunningham, B. G. (2015). Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses. *Brain Research*, *1626*, 146–164. doi:10.1016/j.brainres.2015.06.038.

Vasilkov, V., Garrett, M., Mauermann, M., & Verhulst, S. (2021). Enhancing the sensitivity of the envelope-following response for cochlear synaptopathy screening in humans: The role of stimulus envelope. *Hearing Research*, *400*, 108132.

Vasilkov, V., & Verhulst, S. (2019). Towards a differential diagnosis of cochlear synaptopathy and outer-hair-cell deficits in mixed sensorineural hearing loss pathologies. *medRxiv*, . doi:10.1101/19008680.

Verhulst, S., Ernst, F., Garrett, M., & Vasilkov, V. (2018). Suprathreshold psychoacoustics and envelope-following response relations: normal-hearing, synaptopathy and cochlear gain loss. *Acta Acustica united with Acustica*, *104*, 800–803. doi:10.3813/AAA.919227.

Verhulst, S., Jagadeesh, A., Mauermann, M., & Ernst, F. (2016a). Individual Differences in Auditory Brainstem Response Wave Characteristics: Relations to Different Aspects of Peripheral Hearing Loss. *Trends in Hearing*, *20*, 1–20. doi:10.1177/2331216516672186.

Verhulst, S., Piktel, P., Jagadeesh, A., & Mauermann, M. (2016b). On the Interplay Between Cochlear Gain Loss and Temporal Envelope Coding Deficits. *Advances in Experimental Medicine and Biology*, *894*, 467–475. doi:10.1007/978-3-319-25474-6_49.

Verschooten, E., Robles, L., & Joris, P. X. (2015). Assessment of the limits of neural phase-locking using mass potentials. *The Journal of Neuro-*

science: The Official Journal of the Society for Neuroscience, *35*, 2255–2268. doi:10.1523/JNEUROSCI.2979-14.2015.

Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A. C. P., 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. doi:10.1016/j.heares.2015.04.014.

Wagener, K. C., Kühnel, V., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache III: Evaluation des Oldenburger Satztests. *Zeitschrift für Audiologie / Audiological Acoustics*, *38*, 86–95.

Wang, Y., Abrams, K. S., Youngman, M., & Henry, K. S. (2023). Histological correlates of auditory nerve injury from kainic acid in the budgerigar (melopsittacus undulatus). *Journal of the Association for Research in Otolaryngology*, *24*, 473–485.

WHO (2019). Deafness and hearing loss. URL: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

Wilson, J. L., Abrams, K. S., & Henry, K. S. (2021). Effects of kainic acid-induced auditory nerve damage on envelope-following responses in the budgerigar (melopsittacus undulatus). *Journal of the Association for Research in Otolaryngology*, *22*, 33–49.

Wong, S. J., Abrams, K. S., Amburgey, K. N., Wang, Y., & Henry, K. S. (2019). Effects of selective auditory-nerve damage on the behavioral audiogram and temporal integration in the budgerigar. *Hearing Research*, *374*, 24–34. doi:10.1016/j.heares.2019.01.019.

Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., & Liberman, M. C. (2018). Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear. *Neuroscience*, *407*, 8–20. doi:10.1016/j.neuroscience.2018.07.053.

Yeend, I., Beach, E. F., Sharma, M., & Dillon, H. (2017). The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hearing Research*, *353*, 224–236. doi:10.1016/j.heares.2017.07.006.

Zheng, X. Y., Henderson, D., Hu, B. H., & McFadden, S. L. (1997). Recovery of structure and function of inner ear afferent synapses following kainic acid excitotoxicity. *Hearing Research*, *105*, 65–76. doi:10.1016/s0378-5955(96)00188-8.

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.

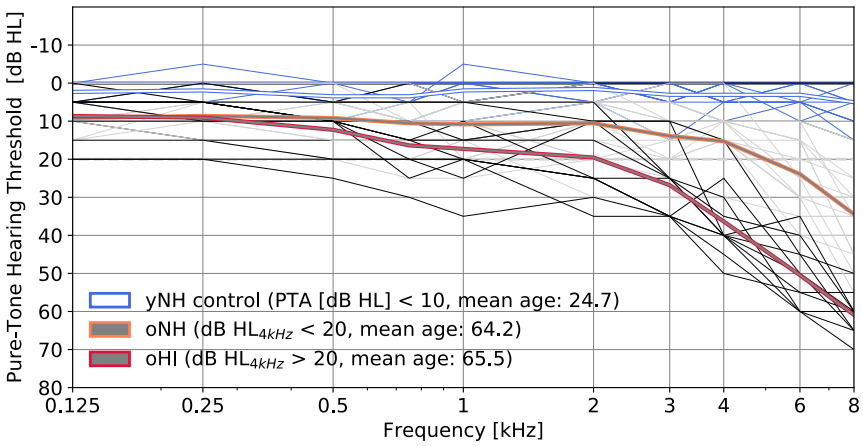

Figure 1: Pure-tone hearing thresholds (dB HL) at frequencies between 0.125-8 kHz. Groups were based on audiogram thresholds and age (yNH: young normal-hearing, oNH: old normal-hearing, oHI: old hearing impaired). Thick traces represent the group mean and thin traces represent individual audiogram profiles.

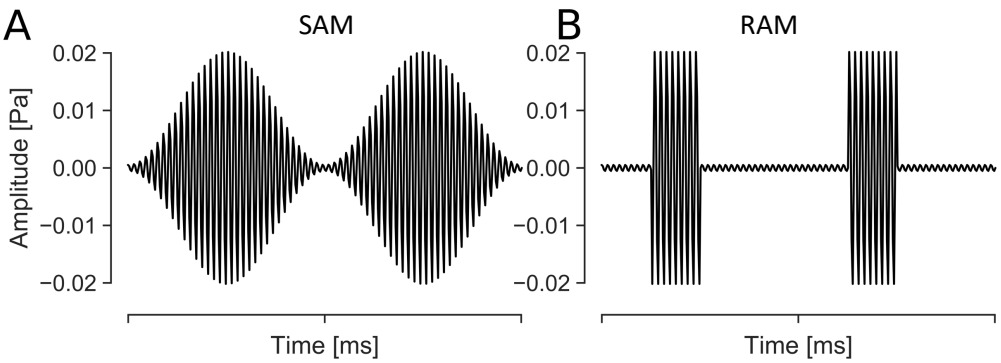

Figure 2: Two cycles of the SAM (A) and RAM (B) stimuli used for EFR recordings. The modulation depth was 95%, modulation frequency 120 Hz, and carrier frequency 4 kHz. The RAM stimulus had a duty cycle of 25%.

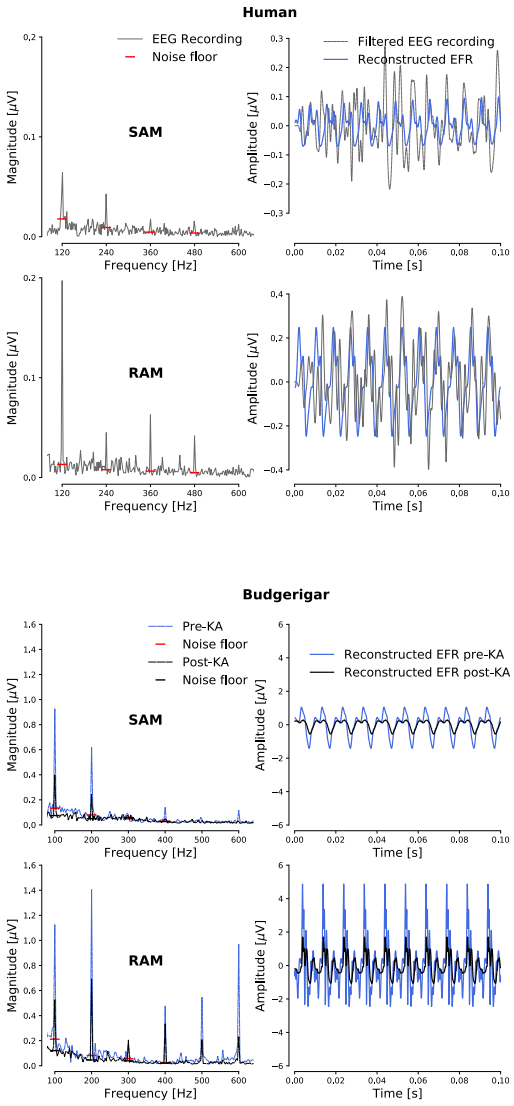

Figure 3: Comparison between human (top) and Budgerigar (bottom) EFR recordings and analysis for single subjects. For each species, the top row corresponds to the SAM stimulus and the bottom row to the RAM stimulus. Time-domain responses show the filtered EEG recording along with the reconstructed time-domain waveform that was based on five frequency components (h_0-h_4) and their respective phases, following Eq.3. The reported EFR amplitudes were extracted from the reconstructed time-domain EFR. The Budgerigar recordings are shown for the same animal before or after Kainic-Acid (KA) administration. Post-KA spectral peaks and reconstructed EFRs were smaller.

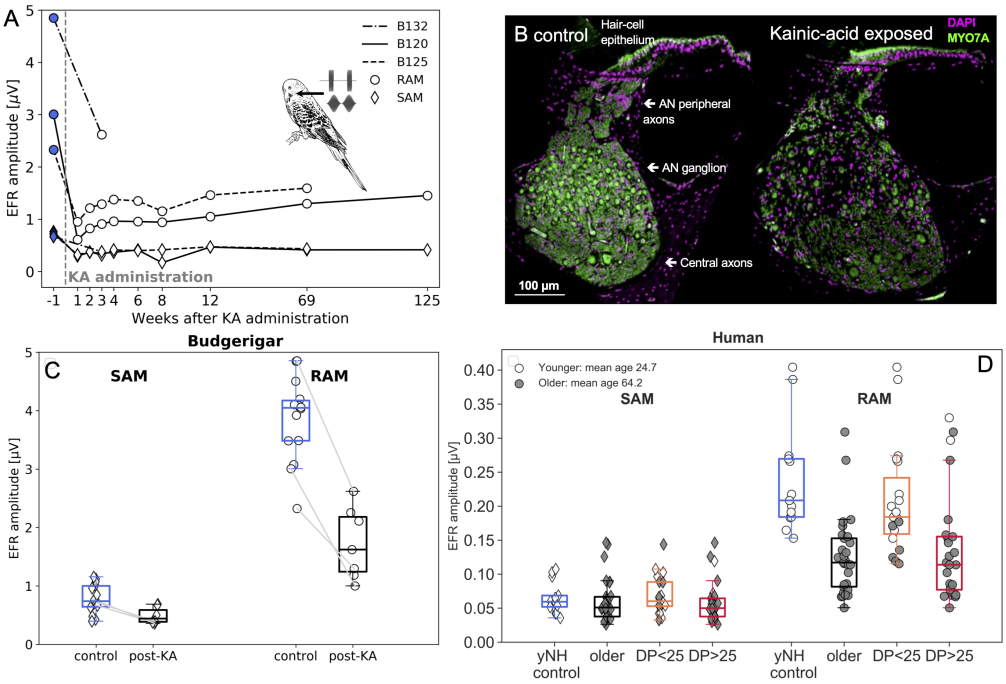

Figure 4: **A** EFR amplitudes from three budgerigars (B132, B120, B125) before or several weeks after the administration of kainic acid. **B** Representative cross sections of the budgerigar cochlea from a control ear (left) and from an ear exposed to 1-mM kainic-acid solution (12 weeks post exposure; right). Sections are stained for DAPI and Myosin 7A, and are from the location 50-60% of the distance from the apex to the base (≈2-kHz cochlear frequency; see Wang et al., 2023). Kainic-acid exposure causes marked reduction of auditory-nerve (AN) peripheral axons and cell bodies in the AN ganglion, without impacting the hair-cell epithelium. **C** Boxplots and individual data points of the budgerigar SAM and RAM EFR amplitudes before or after KA-administration. Connected lines correspond to data from the same animal. **D** Boxplots and individual data points of human SAM and RAM EFR amplitudes. Younger subjects are marked with white and older subjects with gray symbols. Data is show for different subgroups based on age or normal or impaired TH_{DP}s.

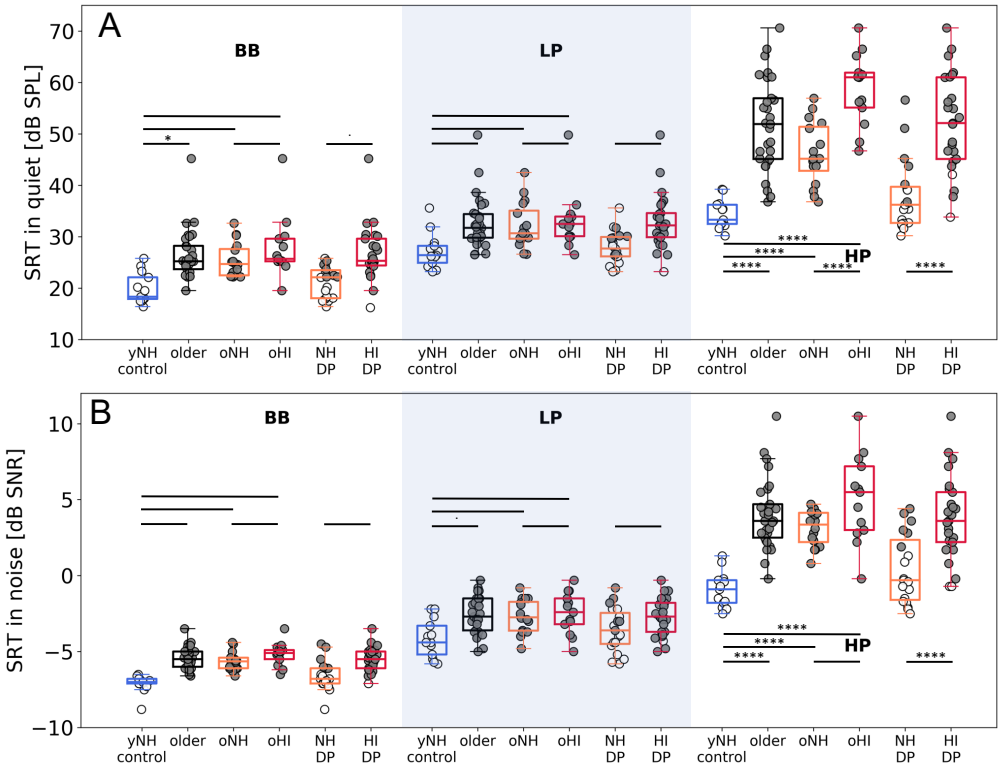

Figure 5: Speech reception thresholds for the OLSA matrix sentence test presented in quiet (A) and in speech-shaped noise (B) for three conditions: original (BB), low-pass filtered speech and noise material (f_c=1.5 kHz; LP), high-pass filtered speech and noise material (f_c=1.65 kHz; HP). SRTs are grouped by the selection groups ($yNH_{control}$, oNH, oHI), as well as pooled in an older group. Lastly, the entire cohort was subdivided in post-hoc groups with $TH_{DPOAE,4kHz}$ above (HI_{DP}) or below 25 dB SPL (HI_{DP}). Younger listeners (mean age: 24.7) are represented with white symbols and older listeners (mean age: 64.2) with gray symbols. Pairwise comparisons were computed for the quiet and noise conditions separate in an ANOVA with variables condition (BB, LP, HP) and group. The following group combinations were considered: ($yNH_{control}$, oNH, oHI), or ($yNH_{control}$, older), or (HI_{DP}, HI_{DP}). Bonferroni correction was performed for either 36 (three groups) or 15 (two groups) tests. Significance codes: **** : $p \leq p_{Bonferroni}$, *** : $p \leq 0.001$, ** : $p \leq 0.01$, * : $p \leq 0.05$, . : $p \leq 0.1$

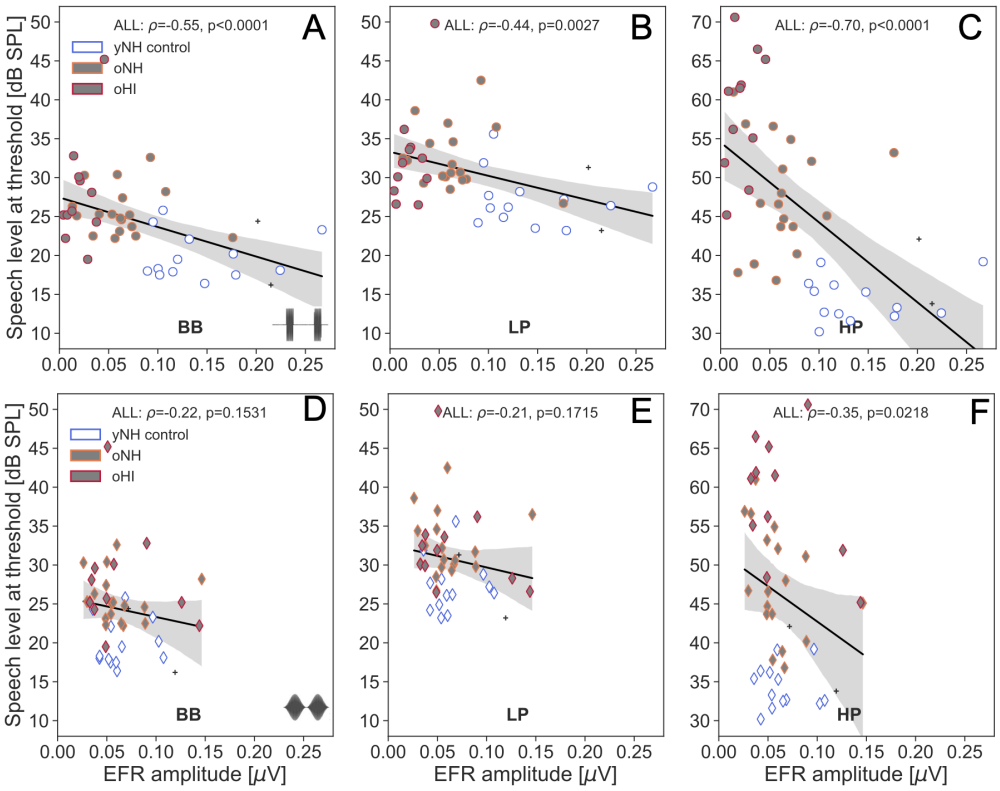

Figure 6: Regression plots between SRT_{SiQ} and the EFR amplitudes for RAM (top) and SAM stimuli (bottom). Analyses are performed for the broadband (**A,D**), low-pass (**B,E**) and high-pass filtered condition (**C,F**). Subjects belonging to the $yNH_{control}$, oNH, oHI groups are color-coded and the two yNH subjects who did not meet the TH_{DP} criterion to be included in the $yNH_{control}$ group are marked with blue dots.

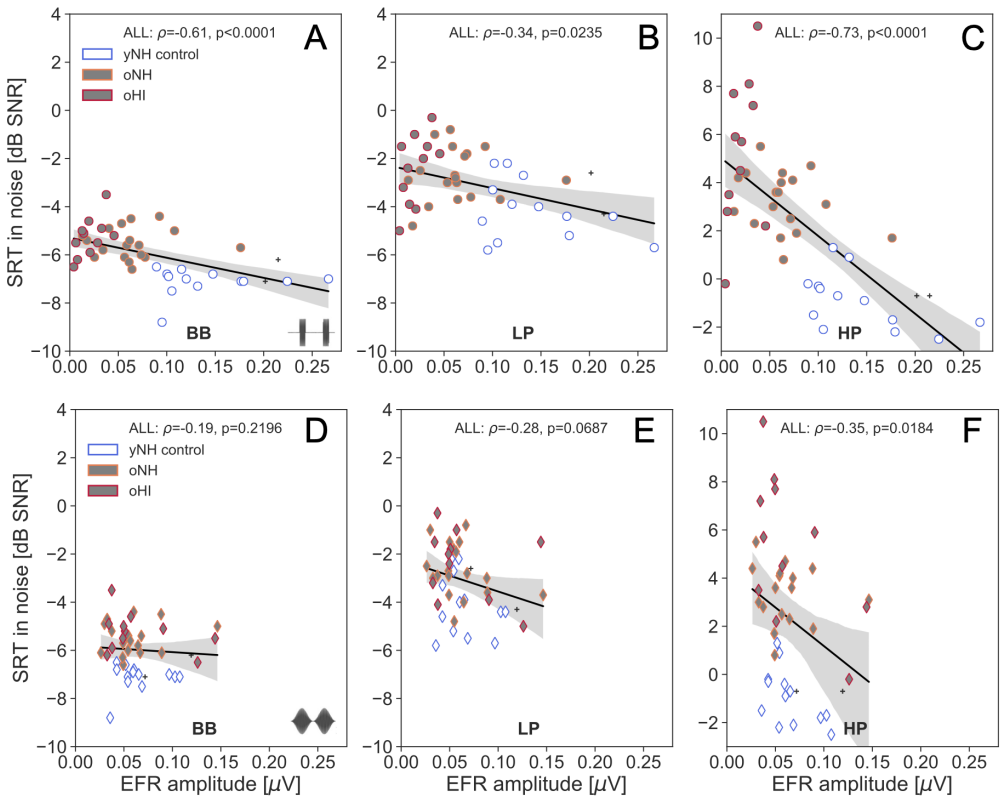

Figure 7: Regression plots between SRT_{SiN} and the EFR amplitudes for RAM (top) and SAM stimuli (bottom). Analyses are performed for the broadband (**A,D**), low-pass (**B,E**) and high-pass filtered condition (**C,F**). Subjects belonging to the $yNH_{control}$, oNH, oHI groups are color-coded and the two yNH subjects who did not meet the TH_{DP} criterion to be included in the $yNH_{control}$ group are marked with blue dots.