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

Encina-Llamas, Gerard; Harte, James; Dau, Torsten; Shinn-Cunningham, Barbara; Epp, Bastian

Published in:
J A R O

Publication date:
2018

Document Version
Early version, also known as pre-print

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Abstract: The healthy auditory system enables communication in challenging situations with high levels of background noise. Despite normal sensitivity to pure-tones, many listeners complain about having difficulties in such situations. Recent animal studies demonstrated that noise over-exposure that produces temporary threshold shifts can cause the loss of auditory nerve (AN) fiber synapses (i.e., cochlear synaptopathy), assumed to be selective for medium- and low-spontaneous rate (SR) fibers. In the present study, envelope following response (EFR) level-growth functions were recorded in normal-hearing (NH) threshold and mildly hearing-impaired (HI) listeners at frequencies above 2 kHz. EFRs were elicited by sinusoidally amplitude modulated (SAM) tones with a carrier frequency of 2 kHz that was modulated at 93 Hz with modulation depths of 85% (strong) or 25% (shallow). Whereas the EFR level-growth functions for strongly modulated tones were similar for all listeners, EFR level-growth functions for shallowly modulated tones were reduced at medium stimulation levels in some of the NH threshold listeners and saturated in HI the listeners for the whole level range. A phenomenological model of the AN was used to investigate the effects of off-frequency contributions (i.e., displaced from the characteristic place of the stimulus) and the differential loss of different AN fiber types on EFR level-growth functions. The model simulations suggest that: (1) EFRs are dominated by the activity of high-SR fibers at all stimulus intensities, and (2) EFRs at medium-to-high stimulus levels are dominated by the off-frequency contribution. Postulated synaptopathy led to...
simulations generally consistent with the recorded data; however, there had to be a substantial reduction in the number of all types of AN fibers to account for the results.
To,

Prof. Michael Heinz,
Associate editor,
Journal of the Association for Research in Otolaryngology

Subject: Manuscript Submission

Dear Editor,

Please find attached our manuscript titled “Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve” which we would like to submit for publication in the Journal of the Association for Research in Otolaryngology.

The manuscript describes our investigation on cochlear synaptopathy in human listeners through the use of envelope following responses (EFR). EFR level-growth functions for strongly and shallowly sinusoidally amplitude modulated tones were recorded in normal-hearing (NH) and hearing-impaired (HI) listeners. A model of the auditory nerve on which synaptopathy can be simulated was used to analyze and investigate the different trends obtained in the individual empirical results. We have also suggested prospective referees for the scientific review of the manuscript, namely Dr. M. Charles Liberman, Dr. Roland Schaette and Dr. Ian C. Bruce.

The manuscript consists of original work and has not been published nor submitted to another journal. All the authors have read and agreed on the manuscript.

Yours sincerely,

Gerard Encina-Llamas, James M. Harte, Torsten Dau, Barbara Shinn-Cunningham and Bastian Epp
Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve

Gerard Encina-Llamas · James M. Harte · Torsten Dau · Barbara Shinn-Cunningham · Bastian Epp

Received: date / Accepted: date

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

**Keywords** Cochlear synaptopathy · "hidden" hearing loss · envelope following responses · auditory steady-state responses · auditory nerve modeling
1 Introduction

It is well known that noise over exposure can impair the auditory system by producing a sensorineural hearing loss that results in a permanent elevation of the pure-tone thresholds. This has led to the interpretation that sound stimulation only producing a temporary threshold shift (TTS), but not a permanent threshold shift (PTS), does not permanently damage the auditory system. However, recent animal studies have shown that noise over exposure producing TTS can in fact lead to the loss of auditory-nerve (AN) fiber synapses without damaging the sensitive hair cells in the cochlea (Kujawa and Liberman, 2009). As this neuronal degeneration does not result in a PTS, it has been termed "hidden" hearing loss (Schaette and McAlpine, 2011). Kujawa and Liberman (2009) demonstrated in mice that "hidden" hearing loss, or more accurately "cochlear synaptopathy" (for a review, see Liberman and Kujawa, 2017), resulting from carefully controlled noise exposure, did not alter hearing thresholds. It was further shown that, in the same mice, the level-growth function of distortion-product otoacoustic emissions (DPOAE) remained unaffected, indicating that the outer hair cells (OHC) were not damaged as a result of the noise exposure. However, the amplitude of the auditory brainstem response (ABR) wave-I was reduced at supra-threshold sound pressure levels. Wave-I reflects the action potentials of the auditory nerve, and should therefore be sensitive to a loss of AN fiber synapses. A selective loss of medium- and low-spontaneous rate (SR) fibers could account for the reduction of supra-threshold ABR wave-I magnitudes, while still preserving normal thresholds (Furman et al, 2013).
However, even in the case of a substantial loss of inner hair cells (IHC) and AN fibers, behavioral pure-tone thresholds can remain unchanged, suggesting that a major loss of high-SR fibers would not be reflected as a PTS either (Lobarinas et al, 2013).

Noise-induced synaptopathy has to date been observed in several non-human mammalian species, such as mice (Kujawa and Liberman, 2009; Furman et al, 2013), guinea pigs (Lin et al, 2011; Liu et al, 2012), rats (Lobarinas et al, 2016) and rhesus macaques (Valero et al, 2017). Cochlear synaptopathy has also been reported as a natural phenomena in the normally aging (non-exposed) mouse ear (Sergeyenko et al, 2013). Noise exposure seems to accelerate such a natural degeneration of the AN (Fernandez et al, 2015). In humans, there is some evidence of such age-related synaptopathy (Makary et al, 2011; Viana et al, 2015). In addition, Viana et al (2015) suggested that, as in mice (Sergeyenko et al, 2013), the loss of peripheral axons in normal-aging humans is significantly greater than the loss of spiral ganglion cells (SGC). The argument is that SGC survive for months after the loss of their peripheral axons (Kujawa and Liberman, 2015).

Despite normal sensitivity to pure tones, some listeners complain about having difficulties in challenging acoustical situations (Saunders and Haggard, 1989; Kumar et al, 2007; Hind et al, 2011; Tremblay et al, 2015), suggesting that NH threshold human listeners might also suffer from cochlear synaptopathy. However, direct evidence of noise-induced synaptopathy in humans has not yet been reported, and the potential perceptual consequences are unknown.
(Plack et al, 2014), despite attempts in large studies to identify them (e.g., Prendergast et al, 2016a; Lopez-Poveda et al, 2017; Le Prell et al, 2017; Grose, 2017).

Animal studies suggest that synaptopathy is reflected in electroencephalographic (EEG) evoked response measurements, such as ABR wave-I (Kujawa and Liberman, 2009; Furman et al, 2013) or envelope following responses (EFRs) (Shaheen et al, 2015). On the one hand, some researchers have attempted to relate changes in evoked responses to self-reported estimates of noise exposure in humans (e.g., Prendergast et al, 2016a,b; Lopez-Poveda et al, 2017; Le Prell et al, 2017; Grose, 2017). To date, no correlation has been found on this first group of studies, with the remarkable exception of Liberman et al (2016). Of course, noise exposure scores derived from self-reported questionnaires of lifetime noise exposure rely on the subjective recall of noisy events and are based on numerous assumptions limiting their reliability (Coughlin, 1990).

On the other hand, other studies have found correlations between evoked responses and behavioral measures of temporal processing at supra-threshold levels in individual normal-hearing (NH) threshold listeners (Bharadwaj et al, 2015; Mehraei et al, 2016). Based on animal studies, those researchers hypothesized that poor performance was linked to the loss of medium- and low-SR fibers, but proof of this idea is not definitive. The inconclusive outcome of the human studies reflects the challenge of directly assessing the status of AN synapses in humans where one cannot do invasive studies. This contrasts with studies in other animals. Evoked responses measured using surface (scap...
electrodes represent the far-field sum of the activity of large populations of neurons, which might not be sensitive to specific local neuronal damage. This could explain why human studies provide sometimes inconclusive evidence.

In the present study, EFRs were measured as a function of stimulus level (EFR level-growth functions) for both strongly and shallowly modulated pure tones, and for both listeners with NH thresholds and in hearing-impaired (HI) listeners. It was hypothesized (via heuristic reasoning) that a preferential loss of medium- and low-SR fibers would reduce the EFR magnitudes at high supra-threshold stimulus levels, but not at lower levels. Thus, if a listener has lost medium- and low-SR fibers, the slope in their EFR level-growth functions should differ from that obtained from the listeners who have not suffered such loss. It was further hypothesized that such a reduction or slope change should be more pronounced in the EFR responses elicited by the shallowly modulated tones than the strongly modulated tones, because medium- and low-SR fibers may be especially important in coding of modulation for high-intensity (Bharadwaj et al, 2014, 2015). For HI listeners, EFR level-growth functions were recorded at both modulation depths, but the stimulus was presented only at frequencies where their audiogram was within the normal range to increase the likelihood of synaptopathy. It has been proposed that cochlear synaptopathy might be a precursor of posterior hair-cell damage (Sergeyenko et al, 2013; Kujawa and Liberman, 2015; Liberman and Kujawa, 2017). Thus, listeners who already show a threshold elevation (and therefore hair-cell dys-
function) at higher audiometric frequencies may be especially likely to suffer from synaptopathy at lower frequencies, where they have normal thresholds.

As the history of noise exposure in both the NH threshold listeners and the HI listeners in this study is unknown, and previous studies have failed to find a link between estimates of lifetime noise exposure and cochlear synaptopathy in humans (e.g.; Prendergast et al, 2016a,b), the present study focused on individual differences in EFR level-growth functions and their potential relation to cochlear synaptopathy. In order to assist with interpreting any potential effects of synaptopathy on the obtained EFRs, a state-of-the-art computational model of the AN was used to study how a differential loss of different AN fiber types influenced EFR level-growth functions.

2 Methods

2.1 Listeners

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

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

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

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

2.3 EFR recordings and analysis

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

The stimulus was a single sinusoidally amplitude modulated (SAM) tone,
with \( SAM(t) = A \cdot \sin(2\pi f_c t) \cdot \left( \frac{1 + m \sin(2\pi f_m t)}{2} \right) \), where \( A, f_c, f_m, m \in [0,1] \)
and \( t \) represent the amplitude, the carrier frequency, the modulation frequency,
the modulation index, and time, respectively. The SAM tone had a carrier
frequency \( f_c \) of 2005 Hz and a modulation frequency \( f_m \) of 93 Hz. Two
modulation depths \( m \) were used: "strong" \( m = 85\% \) and "shallow" \( m =
25\% \). The stimuli were calibrated using a B&K 4157 ear simulator to the
desired root mean squared (RMS) level. The stimuli were digitally generated
as 1-s long epochs and continuously presented to the listener in a loop, where
a trigger signal marked the beginning of a new epoch for later averaging. The
total stimulus duration depended on the stimulus intensity, and was chosen
to achieve a statistically significant EFR signal-to-noise ratio (SNR). Table 1
shows the stimulus duration used for each input level in the EFR recordings.

The recorded EEG data were analyzed as described in Encina-Llamas et al
(2017a). In short, the recorded EEG epochs were 1) band-pass filtered between 60 and 400 Hz, 2) rejected if the absolute value of their voltage amplitude exceeded ±80 µV, 3) averaged with a weighted average to increase the EFR SNR as in John et al. (2001), and 4) concatenated in epochs of 16-s duration to achieve a higher frequency resolution in the EFR spectrum. The recorded EFR magnitude was detected from an estimate of the background EEG noise in the range of 3 Hz below and above the modulation frequency (96 bins). The EFR was considered significantly above the noise floor if $p \leq 0.01$ was achieved in an F-test statistical measure (Dobie and Wilson, 1996; Picton et al., 2003).

Table 1: Duration of EFR stimuli for each used input level.

<table>
<thead>
<tr>
<th>Input level [dB SPL]</th>
<th>34</th>
<th>40</th>
<th>46</th>
<th>54</th>
<th>60</th>
<th>66</th>
<th>71</th>
<th>77</th>
<th>81</th>
<th>87</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration [min]</td>
<td>10.0</td>
<td>8.5</td>
<td>8.5</td>
<td>7.0</td>
<td>7.0</td>
<td>6.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
</tr>
</tbody>
</table>

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

A humanized phenomenological AN model, implemented in MATLAB, was used to simulate the activity of the AN (Zilany et al., 2009, 2014). Individual model fibers were tuned to one of 200 characteristic frequencies (CF) ranging from 0.2 to 20 kHz; these corresponding to equally spaced positions along the basilar membrane (BM), according to the cochlear frequency map for humans (Greenwood, 1990). A non-uniform distribution of AN fibers per CF (or IHC) was implemented to mimic the distribution reported in Spoendlin and Schrott (1989), with a total of about 32000 AN fibers in a healthy auditory system. About 190 AN fibers synapses were independently computed at each CF. Synaptopathy was simulated by using a smaller number of AN fiber synapses at each CF in the model. Frequency-specific synaptic loss was simulated by fixing a given percentage of loss of fibers at particular CFs, which were fit using a shape-preserving piecewise cubic Hermite interpolating polynomial evaluated over the complete range of modeled CFs. Hair-cell impairment was approximated by fitting the listener’s audiogram using the fitaudiogram MATLAB function written by Zilany et al. (2009). As the distribution of the different AN fiber types at each CF is unknown in humans, the distribution reported from cats was used: 61% of high-SR fibers, 23% of medium-SR fibers and 16% of low-SR fibers (Liberman, 1978). Model simulations were performed using the same stimuli as in the human EFR recordings, but with a duration of 1.2-s. Stimulus levels ranged from 10 to 100 dB SPL, in steps of 5 dB.
The model allows IHC and outer hair-cell (OHC) functionality to be controlled independently, and sets a deterministic IHC voltage and synaptic output of each AN fiber type separately. The same IHC voltage at each CF was used to drive the stochastic synapse- and spike-generator models (see Zi- lany et al, 2009), which was executed independently once for each of the AN fibers for each IHC. The resulting synaptic outputs for each AN fiber type were summed to obtain the population response of this fiber type at each CF, which can be compared directly with the peri-stimulus time histogram (PSTH) used to summarize the experimental data. In order to analyze the steady-state encoding of the modulation, the 1-s long steady-state response was analyzed, excluding the transient onsets and offsets. A Fast Fourier Transform (FFT) was performed on the resulting synaptic output. The magnitude of the FFT at the modulation frequency bin was used to quantify the simulated EFR.

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

Previous studies have attempted to simulate steady-state responses, such as EFRs (Rønne, 2013) or frequency following responses (FFRs) (Dau, 2003), by convolving the simulated response of an AN model with a unitary response that reflected the contributions of different neural population along the auditory brainstem to the far-field evoked potential (e.g., Melcher and Kiang, 1996). In the present study, only AN activity was considered, both because synaptopathy occurs at this level and because this analysis is simpler than alternatives that consider higher-level contributions. It was assumed that the coding of the envelope at the level of the AN would be similar to the recorded EFRs. However, it has been suggested that EFRs to 80-100 Hz modulations are primarily generated at the level of the brainstem (Herdman et al, 2002).

However, Parthasarathy et al (2016) showed good consistency between EFR recordings in rats and simulated EFRs using the cat version of the AN model (Zilany et al, 2009, 2014).

3 Results

The data reported in this study is publicly available here (Encina-Llamas et al, 2017b).
Fig. 1: Simulated synaptic output of the AN model obtained using a strongly modulated SAM tone at 80 dB SPL with $f_c = 2$ kHz and $f_m = 93$ Hz. A) Simulated AN synaptic output at CFs from 0.2 to 20 kHz for three cycles of the $f_m$ at the steady-state part of the response. The green rectangles denote the populations used to compute the on-frequency (centered on 2 kHz) and the off-frequency bands (centered on 3 and 7 kHz). B-D) Synaptic outputs for the three different $\frac{1}{3}$-octave bands. E) Synaptic output after summing across CFs.

3.1 EFR level-growth functions in human listeners

Figure 2 shows the complete set of EFR level-growth functions for the NH threshold (A) and the HI (B) listeners. The recorded EFR magnitudes are shown as circles, represented in dB relative to 1µV (blue for $m = 85\%$, red for...
$m = 25\%$). Filled symbols represent statistically significant EFRs magnitudes (positive F-test), open symbols represent responses that are not significantly above the noise floor (negative F-test). The estimated EEG background noises for each modulation depth are depicted as thin lines with consistent color labeling.

For both listener groups, EFR magnitudes obtained with the strongly modulated stimuli (blue) were larger than those obtained with the shallowly modulated tones (red). However, different trends were observed in the EFR level-growth functions across listeners, particularly for the shallowly modulated tones.

In the case of the NH threshold listeners (Fig. 2, A), the results have been organized gradually from patterns showing monotonic and parallel EFR level-growth functions (i.e., listeners NH01 or NH02) to patterns showing non-monotonic level-growth functions (i.e., listeners from NH07 to NH09). In particular, for listener NH09, the EFR magnitudes for the strongly modulated tones grew monotonically with a linear slope throughout the whole level range. This subject was considered as a potentially synaptopathic listener within the NH threshold group. In contrast, the responses to the shallowly modulated tones initially grew with a single slope up to 55 dB SPL, showed a decrease of the EFR magnitudes from 55 to 70 dB SPL, and then a recovery above 70 dB SPL, with comparable EFR magnitudes between 80 to 90 dB SPL, similar to results for the strongly modulated tones.
Fig. 2: EFR level-growth function recorded in A) NH threshold and B) HI listeners using strongly (blue) and shallowly (red) modulated tones. EFR magnitudes (dB relative to 1µV) are represented as filled circles in case of a statistically significant response (positive F-test), and as open circles in cases where responses are not statistically significant (negative F-test). Estimates of the EEG background noise floors for each modulation depth are shown as thin lines using consistent color labeling.
For the HI listeners (Fig. 2, B), the EFR level-growth functions for the strongly modulated tones grew monotonically with a slope similar to that of the NH threshold listeners. The EFR level-growth function for the shallowly modulated tones showed, however, a strongly compressive or even saturating growth. Figure 3 shows boxplots of the fitted slopes for both modulation depths in the NH threshold and the HI listeners. A post-hoc statistical analysis using a mixed linear model revealed that the estimated slopes in the NH threshold listeners were not statistically different for the two modulation depths ($t_{148} = 0.723, p = 0.4704$). However, the EFR level-growth function for the shallowly modulated tones in the HI listeners tones were significantly lower than the slopes for the strongly modulated tones ($t_{48} = 3.646, p = 0.00042$).

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

Figure 4 shows the simulated EFR level-growth functions for A) a NH listener, B) a HI listener accounting for the threshold elevation with dysfunctional OHCs only, and C) a HI listener accounting for the impairment with a combination of one third of IHCs and two thirds of OHCs, as suggested in the literature (Spongr et al, 1997; Lopez-Poveda and Johannesen, 2012). The representation is similar to Figure 2, but with the simulated EFR magnitudes expressed in arbitrary units (a.u.) in dB. The circles represent the simulated EFRs for a NH listener (reprinted in all panels for comparison), while the squares represent the simulated EFR for a HI listeners. The blue and red sym-
Fig. 3: Estimated slopes of the EFR level-growth functions recorded in the NH threshold and HI listeners using strongly (blue) and shallowly (red) modulated tones. Tukey boxplots, where the bottom and top of the box are the first and third quartiles respectively, and the band inside the box is the median. Whiskers show 1.5 of the interquartile range (IQR) of the lower and upper quartile (*** corresponds to a $p$-value $\leq 0.001$, ns corresponds to a $p$-value $> 0.05$).

Bols represent EFR magnitudes obtained with the strongly and the shallowly modulated tones, respectively. The audiogram of listener HI04, who was selected as a representative HI listeners, falling in the middle of the range of the tested HI participants, was taken into account by adjusting the parameters of the IHC and OHC processes in the model.

The simulated EFR level-growth functions for the NH listener (panel A) showed a parallel and monotonic growth over the input level range used in the EFR recordings (35-90 dB SPL, un-shaded area). The EFR magnitudes were larger for the strongly modulated tones than for the shallowly modulated tones. In general, the model simulations were able to capture the trend observed in
Fig. 4: Simulated AN EFR level-growth functions in NH and HI listeners. Squares indicate the simulated EFRs with different combinations of hair-cell dysfunction, whereas the circles show the EFR level-growth function in a simulated NH listener as reference. A) Simulation for a NH listener. B) Simulation for a HI listener accounting for the threshold elevation with a pure OHC dysfunction. C) Simulation for a HI listener accounting for the threshold elevation with a combination of one third of IHC and two thirds of OHC dysfunction.
the recorded EFR level-growth functions in some NH threshold listeners (i.e. NH01 and NH02, see Fig. 2).

The simulated EFR level-growth functions for the HI listener with a purely OHC dysfunction (panel B), did not differ noticeably from the simulated results for the NH listener. If at least one third of the threshold elevation was assigned to IHC dysfunction and the remaining two thirds to OHC dysfunction (panel C), the only difference from the NH simulation was a small reduction at input level between 60 and 85 dB SPL, amounting to about 3 dB both for the strongly and shallowly modulated tones. A similar result was found when assuming a hearing loss due to IHCs dysfunction only (not shown). This is at odds with the recorded EFR level-growth functions in HI listeners where both a loss of sensitivity and a saturation were found in the response to shallowly modulated tones (see Fig. 2).

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

Figure 5 shows the simulated EFR for A) a NH threshold listener after assuming a complete loss of medium- and low-SR fibers at all CFs (an AN with only high-SR fibers); for B) a NH threshold listener including an empirically chosen synaptic loss that best approximated the results obtained from the NH threshold listener NH09; and for C) a HI listener including an empirically chosen loss of synapses that best approximated the results obtained from the HI
The simulated EFR level-growth functions with a loss of 100% of medium- and low-SR fibers (panel A, squares) were nearly the same as the EFR level-growth functions in the reference simulation (circles), with a small decrement that was less than 1.5 dB for both modulation depths. The simulation that best approximated the non-monotonic growth found for some NH threshold listeners (i.e., NH09) required a frequency-specific loss of all types of AN fibers (panel B) and, more specifically, a loss of up to 85% in the range from 3000 to 4000 Hz. To be able to simulate EFR level-growth functions that are similar to those of the listener HI04 (panel C), a substantial loss of AN fiber synapses was required in addition to the hair-cell dysfunction. A total loss of all three types of AN fiber synapses above 6 kHz was considered as well as significant losses of synapses (of > 60%) at frequencies above 600 Hz. While the simulated EFRs in the NH threshold and the HI cases (panels B and C respectively) fitted the trend in the recorded data (see Fig. 2), the impact was the same for the strongly and shallowly modulated tones. In contrast to the model simulations, the recorded EFRs showed a stronger reduction for the shallowly modulated tones compared to the strongly modulated tones.
Fig. 5: Simulated AN EFR level-growth functions in NH threshold and HI listeners with additional postulated synaptopathy. Squares indicate the simulated EFRs with synaptopathy, whereas the circles show the NH listener simulated response as reference (same as Fig. 4, A)). A) Simulation with a loss of 100% of medium- and low-SR fibers. B) Simulation with a frequency-specific loss of all types of fiber synapses to better approximate the response obtained in NH09 in Figure 2, A. C) Simulation with a frequency-specific loss of all types of fiber synapses to better approximate the response obtained in HI04 in Figure 2, B.
Table 2: Percentage of all three types of additional AN fiber loss at different CFs implemented in the model in the NH threshold and the HI listeners simulations.

<table>
<thead>
<tr>
<th>Frequency [kHz]</th>
<th>0.1</th>
<th>0.3</th>
<th>0.6</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>10</th>
<th>13</th>
<th>16</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss of AN fibers [%] in NH simulations</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>65</td>
<td>85</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Loss of AN fibers [%] in HI simulations</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>70</td>
<td>70</td>
<td>80</td>
<td>95</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

4 Discussion

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

It was hypothesized, based on an heuristic argument, that synaptopathy produces differences in the EFR level-growth functions within a homogeneous group of young NH threshold listeners. It was further hypothesized that synaptopathy leads to non-monotonic or saturating EFR level-growth functions for the shallowly modulated SAM tones. This hypothesis was based on the assumption that high-intensity sounds are encoded by the activity of medium- and low-SR fibers because high-SR are saturated at those high stimulation levels (Liberman, 1978; Yates, 1990; Bharadwaj et al, 2014, 2015). Indeed, the individual results of the NH threshold listeners demonstrated different EFR level-growth functions for shallowly modulated tones (Fig. 2, A); and more similar functions for strongly modulated tones. For instance, the EFR level-growth
function for the shallowly modulated tones grew monotonically with a single slope for listener NH01, whereas in NH09 the EFRs grew non-monotonically with reduced magnitudes at medium stimulus levels but recovered at higher levels. The EFR magnitude reduction at 65-70 dB SPL was about 10 dB. This difference might reflect a true physiological difference between those two subjects, especially because this difference is probably larger than the intrinsic EFR variability. Encina-Llamas et al (2017a) reported EFR test-retest variability results at 70 dB SPL from SAM tones at $f_c = 2000$ Hz, $f_m = 93$ Hz and $m = 85\%$ of no larger than 5 dB, much smaller than the difference seen across NH threshold listeners in Figure 2. It is assumed though that the difference in modulation depth in both studies would not affect the test-retest variability, although there is no available data to our knowledge showing test-retest variability in EFRs from SAM tones of different modulation depths. Whether the different patterns observed in the EFR level-growth functions in listener NH01 versus NH09 are due to a loss of AN fiber synapses remains unknown.

Individual differences in the EFR level-growth functions for shallowly modulated tones were also observed for the HI listeners (Fig. 2, B). Whereas the EFR functions for strongly modulated tones grew monotonically with a single slope (similarly to the strongly modulated EFRs in the NH threshold listeners), the EFR magnitudes for shallowly modulated tones did not vary much across stimulus level, leading to a saturated growth function. The change in slope was shown to be statistically significant (linear mixed-effects model, Fig. 3). The question as to whether the different patterns observed from the EFR level-
growth functions in both NH threshold and HI listeners could be due to the loss of AN fiber synapses was investigated using a computational model.

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

Under the assumption that the EFR is sufficiently described by the summation of the instantaneous firing rate of AN fibers across frequency and fiber type, our simulations should shed light on the contributions of CF bands and fiber type to the overall response. Figure 6 shows the simulated EFR level-growth functions for a NH listener, separately for each fiber type (rows) and CF bands (columns). Column A (panels A1-A4) shows the synaptic output summed across all CF bands. Column B (panels B1-B4) shows the output of the band centered at 2 kHz (on-frequency band). Column C (panels C1-C4) shows the output of the band centered at 3 kHz, and column D (panels D1-D4) the output of the band centered at 7 kHz.

The analysis of the simulated AN activity in different CF bands (columns) showed that the EFRs at medium-to-high stimulus levels were not driven solely by activity in the on-frequency band (column B), but also reflected strong contributions of responses of AN neural populations located more basally (i.e., at higher CF bands, columns C and D). Within each frequency band, the response showed a bell-shaped curve, horizontally shifted along the stimulus level axis for more distal CF bands, consistent with the synchrony-level functions recorded from single neurons in the AN of the cat (Joris, 1992). Hence, at
Fig. 6: Simulated AN EFR level-growth functions in a NH listener, separately for different CF bands and for each AN fiber type. Solid circles represent statistically significant EFR magnitudes in dB and open circles represent non-significant responses. Blue markers show responses for strongly modulated stimuli and red markers for shallowly modulated stimuli. The thin lines represent the noise floor in the simulations. Columns show the EFR level-growth functions centered at different CF bands, while rows show the simulated results for different AN SR fiber types.
Investigating the effect of synaptopathy on EFRs using an AN model

higher stimulation levels, the off-frequency contributions dominated the total
EFR magnitudes in the model framework, leading to the overall monotonic
growth observed when summing across CFs.

Similarly, when analyzing the contributions of the different types of AN
fibers (rows), the simulated EFRs were dominated by high-SR fibers across
the whole stimulus level range. For simplicity, let us focus on the on-frequency
band (column B). Even though the strength of the responses of the medium-
SR (B2) and low-SR (B3) fibers level-growth functions did not decrease as
much with level as did the high-SR (B1) fibers, their overall firing rates were
much lower than those of the high-SR fibers (notice the different vertical axis
scales). In other words, the medium- and low-SR fibers contribute very little
to the summed response (B4). The underlying reason for this is mainly due to
the uneven distribution of fiber types assumed in the model, with the high-
SR fibers dominating in number (Liberman, 1978). In conclusion, the model
simulations suggest that for low stimulus intensities, the envelope was encoded
by the high-SR fibers tuned to on-frequency CFs. As stimulus level increases,
the modulation coding shifted to being dominated by the high-SR fibers at
off-frequency bands (see also Fig. 1).

The above model observations explain why the simulated EFR level-growth
functions do not change significantly when a pure OHC dysfunction is imposed
(Fig. 4, B). A loss of gain (OHC dysfunction) might only affect the response
when presenting a stimulus with spectral content at CFs with OHC dam-
age (on-frequency processing). The AN fiber tuning curves associated with
CFs where there is substantial OHC dysfunction will each show a reduction of sensitivity at the tip of the tuning curve (broader frequency selectivity). However, the tails (off-frequency responses) will be largely unaffected (Liberman and Dodds, 1984). Thus, the off-frequency excitation of high CF AN fibers through stimuli at a lower CF with normal sensitivity will be the same, regardless of OHC loss. In contrast, when at least one third of the hearing threshold elevation was assigned to dysfunction of the IHCs (Fig. 4, C), a small reduction in the simulated EFR level-growth function was found at stimulus levels corresponding to CFs bands at which the off-frequency contributions were maximal. Nevertheless, no combination of hair-cell impairment led to simulated EFR level-growth functions like those recorded from HI listeners.

The model framework was used to explore how synaptic loss might be used to explain the non-monotonic patterns observed in some NH threshold listeners (Fig. 2, A) and the strong saturation in the HI listeners (Fig. 2, B). We do not intend to claim that a given listener suffers from synaptopathy. Our purpose was to investigate the potential effects of synaptopathy on the EFR level-growth functions by using computational models. The simulated EFR level-growth functions were insensitive to a selective loss of medium- and low-SR fibers (Fig. 5, A); for instance, EFR magnitudes decreased by less than 1.5 dB after including a complete loss of the medium- and low-SR fibers at all CFs. However, it was possible to account for the non-monotonic growth found in some NH threshold listeners (i.e., NH09) by reducing the number of all types of AN fibers (including high-SR fibers loss), mainly centered in...
the CF band at which the response peaks at medium levels (Fig. 6, column D). To obtain simulated EFRs that better approximated the results from the HI listeners, a large loss of all three types of AN fiber synapses had to be included across a broad CFs. This was in addition to the posited hair-cell dysfunction (see Table 2). These observations are at odds with previous interpretation, which assume that only medium- and low-SR fibers are affected by synaptopathy (e.g., Furman et al, 2013). However, the model simulations are consistent with previous empirical findings (Paul et al, 2016), where a certain degree of high-SR fiber loss had to be included in order for model simulations to account for the differences observed in the EFR magnitudes recorded in NH threshold listeners with and without tinnitus, which has also been related to synaptopathy (Schaette and McAlpine, 2011).

Our model simulations were not able to capture all the details of the EFR level-growth functions observed in the recorded data. The implementation of either hair-cell dysfunction and synaptopathy within the model framework affected the EFR level-growth functions for both strongly and shallowly modulated tones similarly (see Figs. 4 and 5). In contrast, the recorded EFR level-growth functions for strongly modulated tones were similar across all listeners; large individual differences were observed only on the EFRs for shallowly modulated tones (see Fig. 2). Consistent with the AN model simulations, Shaheen et al (2015) reported a significant reduction in the EFR level-growth functions of synaptopathic mice when using strongly modulated SAM tones with modulation frequencies between 800 to 1000 Hz. The EFRs showed group delays...
consistent with generators between the AN and the cochlear nucleus. In the present study, a modulation frequency of 93 Hz was used to elicit the EFRs in humans, which are assumed to be dominated by brainstem sources (Kuwada et al, 2002; Herdman et al, 2002). Brainstem processing, such as central gain mechanisms (Möhrle et al, 2016; Chambers et al, 2016), may differentially affect results from strongly and shallowly modulated stimuli, which may explain the inconsistencies between human data and animal data and the model simulations.

Nevertheless, the model simulation suggested two main conclusions: (1) the EFR level-growth functions at medium-to-high stimulation levels are strongly dominated by the contributions from off-frequency neuronal activity, and (2) there must be a significant degree of loss of high-SR fibers for synaptopathy to be reflected in the EFR. First, the interpretation of the role of the medium- and low-SR fibers on encoding temporal fluctuations at high stimulus levels, based on the rate-level curves of the different AN fibers types (Liberman, 1978; Yates, 1990), has led to different hypotheses about how best to study synaptopathy in humans (including the present study and others like Bharadwaj et al (2014, 2015); Mehraei et al (2016)). Rate-level functions are derived from direct recordings in single AN neurons, and therefore provide information regarding AN neuronal activity at on-frequency stimulation. However, electrophysiological evoked responses are driven by large neural populations that encode and interpret sounds. As synaptopathy affects supra-threshold processing, studies of the effects of synaptopathy generally use high sound stimulation levels that
produce a broad excitation of the AN. The contribution of AN neurons tuned
to off-frequency CFs are not well understood, and thus should be carefully con-
sidered in the design of future hypotheses. Second, animal results suggesting
that synaptopathy is selective to medium- and low-SR fibers (Furman et al,
2013) have been interpreted as if noise exposure did not damage high-SR fibers
at all. Incorporating only medium- and low-SR fibers in the model produces
no significant change in the simulated EFRs. However, EFRs were shown to
be reduced in synaptopathic mice relative to the unexposed animals (Shaheen
et al, 2015). One explanation could be that the effect of synaptopathy on the
EFRs was not well captured in the model simulations. Another explanation is
that the lack of vulnerability of the high-SR fibers to synaptopathy reported
in mice may not be directly transferable to other species, such as humans.
An alternative explanation is that high- versus medium- and low-SR fibers
are more evenly distributed in mice (Taberner and Liberman, 2005) than in
cats (Liberman, 1978). In turn, such a distribution will result in a stronger
impact of the loss of medium- and low-SR fiber in the EFR obtained in mice.

In order to study the potential consequences of synaptopathy in humans, audi-
itory evoked potential studies must be in parallel in humans and in non-human
species where synaptopathy has been characterized, together with the use of
species-specific computational models.
5 Conclusions

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

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

Overall, the simulations suggest that EFRs are dominated by high-SR fibers, and that off-frequency neurons increasingly contribute to the EFR as stimulus level increases. The finding that when SAM tones are presented at high stimulus levels, the envelope is better encoded at off-frequency CF responses (rather than on-frequency) is a key concept to consider when using EFRs to investigate supra-threshold coding.
Acknowledgements The authors want to thank Le Wang and Graham Voysey from Boston University (BU) for their valuable comments and discussions regarding the model implementation and analysis. I would like to express our sincere gratitude to Laurel Carney for her wise advice regarding the analysis of the model outcome and the fascinating discussions during her stay in Denmark about the role of the different types of AN fibers, the interpretation of the rate-level functions and the processing of modulations in the AN, the CN and the IC. This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CheSS) at the Technical University of Denmark (DTU). The authors want to acknowledge the support provided by the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience.

References


Investigating the effect of synaptopathy on EFRs using an AN model


Lobarinas E, Salvi R, Ding D (2013) Insensitivity of the audiogram to carbo-

platin induced inner hair cell loss in chinchillas. Hear Res 302:113–20

Lobarinas E, Spankovich C, Le Prell CG (2016) Evidence of hidden hearing

loss following noise exposures that produce robust TTS and ABR wave-I

amplitude reductions. Hear Res

Lopez-Poveda EA, Johannesen PT (2012) Behavioral estimates of the contri-
bution of inner and outer hair cell dysfunction to individualized audiometric


Lopez-Poveda EA, Johannesen PT, Buzo BC, Renne FM, Pontoppidan NH,

Harte JM (2017) On the value of brief sound audiometry as a diagnostic tool


Makary CA, Shin J, Kujawa SG, Liberman MC, Merchant SN (2011) Age-

related primary cochlear neuronal degeneration in human temporal bones.


Mehraei G, Hickox AE, Bharadwaj HM, Goldberg H, Verhulst S, Liberman

MC, Shinn-Cunningham BG (2016) Auditory Brainstem Response Latency


3764

Melcher JR, Kiang NYS (1996) Generators of the brainstem auditory evoked

potential in cat III: Identified cell populations. Hear Res 93(1-2):52–71

Möhrle D, Ni K, Varakina K, Bing D, Lee SC, Zimmermann U, Knipper M,

Rüttiger L (2016) Loss of auditory sensitivity from inner hair cell synapto-

thy can be centrally compensated in the young but not old brain. Neurobiol Aging 44:173–184


Paul BT, Bruce IC, Roberts LE (2016) Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus. Hear Res


Investigating the effect of synaptopathy on EFRs using an AN model


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

Gerard Encina-Llamas · James M. Harte · Torsten Dau · Barbara Shinn-Cunningham · Bastian Epp

Received: date / Accepted: date

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

**Keywords** Cochlear synaptopathy · "hidden" hearing loss · envelope following responses · auditory steady-state responses · auditory nerve modeling
1 Introduction

It is well known that noise over exposure can impair the auditory system by producing a sensorineural hearing loss that results in a permanent elevation of the pure-tone thresholds. This has led to the interpretation that sound stimulation only producing a temporary threshold shift (TTS), but not a permanent threshold shift (PTS), does not permanently damage the auditory system. However, recent animal studies have shown that noise over exposure producing TTS can in fact lead to the loss of auditory-nerve (AN) fiber synapses without damaging the sensitive hair cells in the cochlea (Kujawa and Liberman, 2009). As this neuronal degeneration does not result in a PTS, it has been termed "hidden" hearing loss (Schaette and McAlpine, 2011). Kujawa and Liberman (2009) demonstrated in mice that "hidden" hearing loss, or more accurately "cochlear synaptopathy" (for a review, see Liberman and Kujawa, 2017), resulting from carefully controlled noise exposure, did not alter hearing thresholds. It was further shown that, in the same mice, the level-growth function of distortion-product otoacoustic emissions (DPOAE) remained unaffected, indicating that the outer hair cells (OHC) were not damaged as a result of the noise exposure. However, the amplitude of the auditory brainstem response (ABR) wave-I was reduced at supra-threshold sound pressure levels. Wave-I reflects the action potentials of the auditory nerve, and should therefore be sensitive to a loss of AN fiber synapses. A selective loss of medium- and low-spontaneous rate (SR) fibers could account for the reduction of supra-threshold ABR wave-I magnitudes, while still preserving normal thresholds (Furman et al, 2013).
However, even in the case of a substantial loss of inner hair cells (IHC) and AN fibers, behavioral pure-tone thresholds can remain unchanged, suggesting that a major loss of high-SR fibers would not be reflected as a PTS either (Lobarinas et al, 2013).

Noise-induced synaptopathy has to date been observed in several non-human mammalian species, such as mice (Kujawa and Liberman, 2009; Furman et al, 2013), guinea pigs (Lin et al, 2011; Liu et al, 2012), rats (Lobarinas et al, 2016) and rhesus macaques (Valero et al, 2017). Cochlear synaptopathy has also been reported as a natural phenomena in the normally aging (non-exposed) mouse ear (Sergeyenko et al, 2013). Noise exposure seems to accelerate such a natural degeneration of the AN (Fernandez et al, 2015). In humans, there is some evidence of such age-related synaptopathy (Makary et al, 2011; Viana et al, 2015). In addition, Viana et al (2015) suggested that, as in mice (Sergeyenko et al, 2013), the loss of peripheral axons in normal-aging humans is significantly greater than the loss of spiral ganglion cells (SGC). The argument is that SGC survive for months after the loss of their peripheral axons (Kujawa and Liberman, 2015).

Despite normal sensitivity to pure tones, some listeners complain about having difficulties in challenging acoustical situations (Saunders and Haggard, 1989; Kumar et al, 2007; Hind et al, 2011; Tremblay et al, 2015), suggesting that NH threshold human listeners might also suffer from cochlear synaptopathy. However, direct evidence of noise-induced synaptopathy in humans has not yet been reported, and the potential perceptual consequences are unknown.
Investigating the effect of synaptopathy on EFRs using an AN model

(Plack et al., 2014), despite attempts in large studies to identify them (e.g., Prendergast et al., 2016a; Lopez-Poveda et al., 2017; Le Prell et al., 2017; Grose, 2017).

Animal studies suggest that synaptopathy is reflected in electroencephalographic (EEG) evoked response measurements, such as ABR wave-I (Kujawa and Liberman, 2009; Furman et al., 2013) or envelope following responses (EFRs) (Shaheen et al., 2015). On the one hand, some researchers have attempted to relate changes in evoked responses to self-reported estimates of noise exposure in humans (e.g., Prendergast et al., 2016a,b; Lopez-Poveda et al., 2017; Le Prell et al., 2017; Grose, 2017). To date, no correlation has been found on this first group of studies, with the remarkable exception of Liberman et al (2016). Of course, noise exposure scores derived from self-reported questionnaires of lifetime noise exposure rely on the subjective recall of noisy events and are based on numerous assumptions limiting their reliability (Coughlin, 1990).

On the other hand, other studies have found correlations between evoked responses and behavioral measures of temporal processing at supra-threshold levels in individual normal-hearing (NH) threshold listeners (Bharadwaj et al., 2015; Mehraei et al., 2016). Based on animal studies, those researchers hypothesized that poor performance was linked to the loss of medium- and low-SR fibers, but proof of this idea is not definitive. The inconclusive outcome of the human studies reflects the challenge of directly assessing the status of AN synapses in humans where one cannot do invasive studies. This contrasts with studies in other animals. Evoked responses measured using surface (scalp)
electrodes represent the far-field sum of the activity of large populations of neurons, which might not be sensitive to specific local neuronal damage. This could explain why human studies provide sometimes inconclusive evidence.

In the present study, EFRs were measured as a function of stimulus level (EFR level-growth functions) for both strongly and shallowly modulated pure tones, and for both listeners with NH thresholds and in hearing-impaired (HI) listeners. It was hypothesized (via heuristic reasoning) that a preferential loss of medium- and low-SR fibers would reduce the EFR magnitudes at high supra-threshold stimulus levels, but not at lower levels. Thus, if a listener has lost medium- and low-SR fibers, the slope in their EFR level-growth functions should differ from that obtained from the listeners who have not suffered such loss. It was further hypothesized that such a reduction or slope change should be more pronounced in the EFR responses elicited by the shallowly modulated tones than the strongly modulated tones, because medium- and low-SR fibers may be especially important in coding of modulation for high-intensity (Bharadwaj et al, 2014, 2015). For HI listeners, EFR level-growth functions were recorded at both modulation depths, but the stimulus was presented only at frequencies where their audiogram was within the normal range to increase the likelihood of synaptopathy. It has been proposed that cochlear synaptopathy might be a precursor of posterior hair-cell damage (Sergeyenko et al, 2013; Kujawa and Liberman, 2015; Liberman and Kujawa, 2017). Thus, listeners who already show a threshold elevation (and therefore hair-cell dys-
function) at higher audiometric frequencies may be especially likely to suffer from synaptopathy at lower frequencies, where they have normal thresholds.

As the history of noise exposure in both the NH threshold listeners and the HI listeners in this study is unknown, and previous studies have failed to find a link between estimates of lifetime noise exposure and cochlear synaptopathy in humans (e.g.; Prendergast et al, 2016a,b), the present study focused on individual differences in EFR level-growth functions and their potential relation to cochlear synaptopathy. In order to assist with interpreting any potential effects of synaptopathy on the obtained EFRs, a state-of-the-art computational model of the AN was used to study how a differential loss of different AN fiber types influenced EFR level-growth functions.

2 Methods

2.1 Listeners

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

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

EFRs were recorded using a Biosemi ActiveTwo system (sampling rate $f_{s|EFR} = 4096$ Hz, 24 bits). Sixty-four active pin-type electrodes were used following the 10-20 system (American Clinical Neurophysiology Society, 2006). The results shown in this study represent the Cz-P10 potential in response to right-ear stimulation, and the Cz-P9 potential in response to left-ear stimulation (similar to vertex to ipsi- and to contra-mastoid montage respectively). Ground was a "Common Mode voltage" driven by two electrodes (DRL and CMS) placed at the center of the parieto-occipital coronal line (on either side of electrode POz). Conductive electrode gel was applied and the offset voltage was stabilized at $< 20$ mV for each electrode. The recorded EEG signals were downsampled by a factor of 2, using a hardware implemented 5th order
sinc-response low-pass filter with a -3 dB point at approximately 410 Hz. The EEG data were stored to hard disk for offline analysis.

2.3 EFR recordings and analysis

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

The stimulus was a single sinusoidally amplitude modulated (SAM) tone, with \( \text{SAM}(t) = A \cdot \sin(2\pi f_c t) \cdot \left( \frac{1 + m \sin(2\pi f_m t)}{2} \right) \), where \( A, f_c, f_m, m \in [0, 1] \) and \( t \) represent the amplitude, the carrier frequency, the modulation frequency, the modulation index, and time, respectively. The SAM tone had a carrier frequency \( f_c \) of 2005 Hz and a modulation frequency \( f_m \) of 93 Hz. Two modulation depths \( (m) \) were used: "strong" \( (m = 85\%) \) and "shallow" \( (m = 25\%) \). The stimuli were calibrated using a B&K 4157 ear simulator to the desired root mean squared (RMS) level. The stimuli were digitally generated as 1-s long epochs and continuously presented to the listener in a loop, where a trigger signal marked the beginning of a new epoch for later averaging. The total stimulus duration depended on the stimulus intensity, and was chosen to achieve a statistically significant EFR signal-to-noise ratio (SNR). Table 1 shows the stimulus duration used for each input level in the EFR recordings.

The recorded EEG data were analyzed as described in Encina-Llamas et al.
In short, the recorded EEG epochs were 1) band-pass filtered between 60 and 400 Hz, 2) rejected if the absolute value of their voltage amplitude exceeded ±80 µV, 3) averaged with a weighted average to increase the EFR SNR as in John et al (2001), and 4) concatenated in epochs of 16-s duration to achieve a higher frequency resolution in the EFR spectrum. The recorded EFR magnitude was detected from an estimate of the background EEG noise in the range of 3 Hz below and above the modulation frequency (96 bins). The EFR was considered significantly above the noise floor if \( p \leq 0.01 \) was achieved in an F-test statistical measure (Dobie and Wilson, 1996; Picton et al, 2003).

Table 1: Duration of EFR stimuli for each used input level.

<table>
<thead>
<tr>
<th>Input level [dB SPL]</th>
<th>34</th>
<th>40</th>
<th>46</th>
<th>54</th>
<th>60</th>
<th>66</th>
<th>71</th>
<th>77</th>
<th>81</th>
<th>87</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration [min]</td>
<td>10.0</td>
<td>8.5</td>
<td>8.5</td>
<td>7.0</td>
<td>7.0</td>
<td>6.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
</tr>
</tbody>
</table>

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

A humanized phenomenological AN model, implemented in MATLAB, was used to simulate the activity of the AN (Zilany et al, 2009, 2014). Individual model fibers were tuned to one of 200 characteristic frequencies (CF) ranging from 0.2 to 20 kHz; these corresponding to equally spaced positions along the basilar membrane (BM), according to the cochlear frequency map for humans (Greenwood, 1990). A non-uniform distribution of AN fibers per CF (or IHC) was implemented to mimic the distribution reported in Spoendlin and Schrott (1989), with a total of about 32000 AN fibers in a healthy auditory system. About 190 AN fibers synapses were independently computed at each CF. Synaptopathy was simulated by using a smaller number of AN fiber synapses at each CF in the model. Frequency-specific synaptic loss was simulated by fixing a given percentage of loss of fibers at particular CFs, which were fit using a shape-preserving piecewise cubic Hermite interpolating polynomial evaluated over the complete range of modeled CFs. Hair-cell impairment was approximated by fitting the listener’s audiogram using the fitaudiogram MATLAB function written by Zilany et al (2009). As the distribution of the different AN fiber types at each CF is unknown in humans, the distribution reported from cats was used: 61% of high-SR fibers, 23% of medium-SR fibers and 16% of low-SR fibers (Liberman, 1978). Model simulations were performed using the same stimuli as in the human EFR recordings, but with a duration of 1.2-s. Stimulus levels ranged from 10 to 100 dB SPL, in steps of 5 dB.
The model allows IHC and outer hair-cell (OHC) functionality to be controlled independently, and sets a deterministic IHC voltage and synaptic output of each AN fiber type separately. The same IHC voltage at each CF was used to drive the stochastic synapse- and spike-generator models (see Zilany et al, 2009), which was executed independently once for each of the AN fibers for each IHC. The resulting synaptic outputs for each AN fiber type were summed to obtain the population response of this fiber type at each CF, which can be compared directly with the peri-stimulus time histogram (PSTH) used to summarize the experimental data. In order to analyze the steady-state encoding of the modulation, the 1-s long steady-state response was analyzed, excluding the transient onsets and offsets. A Fast Fourier Transform (FFT) was performed on the resulting synaptic output. The magnitude of the FFT at the modulation frequency bin was used to quantify the simulated EFR.

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

Previous studies have attempted to simulate steady-state responses, such
as EFRs (Rønne, 2013) or frequency following responses (FFRs) (Dau, 2003),
by convolving the simulated response of an AN model with a unitary response
that reflected the contributions of different neural population along the au-
ditory brainstem to the far-field evoked potential (e.g., Melcher and Kiang,
1996). In the present study, only AN activity was considered, both because
synaptopathy occurs at this level and because this analysis is simpler than
alternatives that consider higher-level contributions. It was assumed that the
coding of the envelope at the level of the AN would be similar to the recorded
EFRs. However, it has been suggested that EFRs to 80-100 Hz modulations
are primarily generated at the level of the brainstem (Herdman et al, 2002).
However, Parthasarathy et al (2016) showed good consistency between EFR
recordings in rats and simulated EFRs using the cat version of the AN model

3 Results

The data reported in this study is publicly available here (Encina-Llamas et al,
2017b).
Fig. 1: Simulated synaptic output of the AN model obtained using a strongly
modulated SAM tone at 80 dB SPL with $f_c = 2$ kHz and $f_m = 93$ Hz. A)
Simulated AN synaptic output at CFs from 0.2 to 20 kHz for three cycles of
the $f_m$ at the steady-state part of the response. The green rectangles denote
the populations used to compute the on-frequency (centered on 2 kHz) and
the off-frequency bands (centered on 3 and 7 kHz). B-D) Synaptic outputs for
the three different 1/3-octave bands. E) Synaptic output after summing across
CFs.

3.1 EFR level-growth functions in human listeners

Figure 2 shows the complete set of EFR level-growth functions for the NH
threshold (A) and the HI (B) listeners. The recorded EFR magnitudes are
shown as circles, represented in dB relative to $1\mu V$ (blue for $m = 85\%$, red for
m = 25%). Filled symbols represent statistically significant EFRs magnitudes (positive F-test), open symbols represent responses that are not significantly above the noise floor (negative F-test). The estimated EEG background noises for each modulation depth are depicted as thin lines with consistent color labeling.

For both listener groups, EFR magnitudes obtained with the strongly modulated stimuli (blue) were larger than those obtained with the shallowly modulated tones (red). However, different trends were observed in the EFR level-growth functions across listeners, particularly for the shallowly modulated tones.

In the case of the NH threshold listeners (Fig. 2, A), the results have been organized gradually from patterns showing monotonic and parallel EFR level-growth functions (i.e., listeners NH01 or NH02) to patterns showing non-monotonic level-growth functions (i.e., listeners from NH07 to NH09). In particular, for listener NH09, the EFR magnitudes for the strongly modulated tones grew monotonically with a linear slope throughout the whole level range. This subject was considered as a potentially synaptopathic listener within the NH threshold group. In contrast, the responses to the shallowly modulated tones initially grew with a single slope up to 55 dB SPL, showed a decrease of the EFR magnitudes from 55 to 70 dB SPL, and then a recovery above 70 dB SPL, with comparable EFR magnitudes between 80 to 90 dB SPL, similar to results for the strongly modulated tones.
Fig. 2: EFR level-growth function recorded in A) NH threshold and B) HI listeners using strongly (blue) and shallowly (red) modulated tones. EFR magnitudes (dB relative to 1µV) are represented as filled circles in case of a statistically significant response (positive F-test), and as open circles in cases where responses are not statistically significant (negative F-test). Estimates of the EEG background noise floors for each modulation depth are shown as thin lines using consistent color labeling.
For the HI listeners (Fig. 2, B), the EFR level-growth functions for the strongly modulated tones grew monotonically with a slope similar to that of the NH threshold listeners. The EFR level-growth function for the shallowly modulated tones showed, however, a strongly compressive or even saturating growth. Figure 3 shows boxplots of the fitted slopes for both modulation depths in the NH threshold and the HI listeners. A post-hoc statistical analysis using a mixed linear model revealed that the estimated slopes in the NH threshold listeners were not statistically different for the two modulation depths ($t_{148} = 0.723, p = 0.4704$). However, the EFR level-growth function for the shallowly modulated tones in the HI listeners tones were significantly lower than the slopes for the strongly modulated tones ($t_{48} = 3.646, p = 0.00042$).

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

Figure 4 shows the simulated EFR level-growth functions for A) a NH listener, B) a HI listener accounting for the threshold elevation with dysfunctional OHCs only, and C) a HI listener accounting for the impairment with a combination of one third of IHCs and two thirds of OHCs, as suggested in the literature (Spongr et al, 1997; Lopez-Poveda and Johannesen, 2012). The representation is similar to Figure 2, but with the simulated EFR magnitudes expressed in arbitrary units (a.u.) in dB. The circles represent the simulated EFRs for a NH listener (reprinted in all panels for comparison), while the squares represent the simulated EFR for a HI listeners. The blue and red sym-
Fig. 3: Estimated slopes of the EFR level-growth functions recorded in the NH threshold and HI listeners using strongly (blue) and shallowly (red) modulated tones. Tukey boxplots, where the bottom and top of the box are the first and third quartiles respectively, and the band inside the box is the median. Whiskers show 1.5 of the interquartile range (IQR) of the lower and upper quartile (*** corresponds to a \( p\)-value \( \leq 0.001 \), ns corresponds to a \( p\)-value > 0.05).

The simulated EFR level-growth functions for the NH listener (panel A) showed a parallel and monotonic growth over the input level range used in the EFR recordings (35-90 dB SPL, un-shaded area). The EFR magnitudes were larger for the strongly modulated tones than for the shallowly modulated tones. In general, the model simulations were able to capture the trend observed in
Fig. 4: Simulated AN EFR level-growth functions in NH and HI listeners. Squares indicate the simulated EFRs with different combinations of hair-cell dysfunction, whereas the circles show the EFR level-growth function in a simulated NH listener as reference. A) Simulation for a NH listener. B) Simulation for a HI listener accounting for the threshold elevation with a pure OHC dysfunction. C) Simulation for a HI listener accounting for the threshold elevation with a combination of one third of IHC and two thirds of OHC dysfunction.
the recorded EFR level-growth functions in some NH threshold listeners (i.e. NH01 and NH02, see Fig. 2).

The simulated EFR level-growth functions for the HI listener with a purely OHC dysfunction (panel B), did not differ noticeably from the simulated results for the NH listener. If at least one third of the threshold elevation was assigned to IHC dysfunction and the remaining two thirds to OHC dysfunction (panel C), the only difference from the NH simulation was a small reduction at input level between 60 and 85 dB SPL, amounting to about 3 dB both for the strongly and shallowly modulated tones. A similar result was found when assuming a hearing loss due to IHCs dysfunction only (not shown). This is at odds with the recorded EFR level-growth functions in HI listeners where both a loss of sensitivity and a saturation were found in the response to shallowly modulated tones (see Fig. 2).

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

Figure 5 shows the simulated EFR for A) a NH threshold listener after assuming a complete loss of medium- and low-SR fibers at all CFs (an AN with only high-SR fibers); for B) a NH threshold listener including an empirically chosen synaptic loss that best approximated the results obtained from the NH threshold listener NH09; and for C) a HI listener including an empirically chosen loss of synapses that best approximated the results obtained from the HI
Investigating the effect of synaptopathy on EFRs using an AN model

The representation is the same as in Figure 4, but the squares represent the simulated EFRs including a postulated synaptic loss.

The simulated EFR level-growth functions with a loss of 100% of medium- and low-SR fibers (panel A, squares) were nearly the same as the EFR level-growth functions in the reference simulation (circles), with a small decrement that was less than 1.5 dB for both modulation depths. The simulation that best approximated the non-monotonic growth found for some NH threshold listeners (i.e., NH09) required a frequency-specific loss of all types of AN fibers (panel B) and, more specifically, a loss of up to 85% in the range from 3000 to 4000 Hz. To be able to simulate EFR level-growth functions that are similar to those of the listener HI04 (panel C), a substantial loss of AN fiber synapses was required in addition to the hair-cell dysfunction. A total loss of all three types of AN fiber synapses above 6 kHz was considered as well as significant losses of synapses (of > 60%) at frequencies above 600 Hz. While the simulated EFRs in the NH threshold and the HI cases (panels B and C respectively) fitted the trend in the recorded data (see Fig. 2), the impact was the same for the strongly and shallowly modulated tones. In contrast to the model simulations, the recorded EFRs showed a stronger reduction for the shallowly modulated tones compared to the strongly modulated tones.
Fig. 5: Simulated AN EFR level-growth functions in NH threshold and HI listeners with additional postulated synaptopathy. Squares indicate the simulated EFRs with synaptopathy, whereas the circles show the NH listener simulated response as reference (same as Fig. 4, A)). A) Simulation with a loss of 100% of medium- and low-SR fibers. B) Simulation with a frequency-specific loss of all types of fiber synapses to better approximate the response obtained in NH09 in Figure 2, A. C) Simulation with a frequency-specific loss of all types of fiber synapses to better approximate the response obtained in HI04 in Figure 2, B.
Table 2: Percentage of all three types of additional AN fiber loss at different CFs implemented in the model in the NH threshold and the HI listeners simulations.

<table>
<thead>
<tr>
<th>Frequency [kHz]</th>
<th>0.1</th>
<th>0.3</th>
<th>0.6</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>10</th>
<th>13</th>
<th>16</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss of AN fibers [%] in NH simulations</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>65</td>
<td>85</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Loss of AN fibers [%] in HI simulations</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>70</td>
<td>70</td>
<td>80</td>
<td>95</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

4 Discussion

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

It was hypothesized, based on an heuristic argument, that synaptopathy produces differences in the EFR level-growth functions within a homogeneous group of young NH threshold listeners. It was further hypothesized that synaptopathy leads to non-monotonic or saturating EFR level-growth functions for the shallowly modulated SAM tones. This hypothesis was based on the assumption that high-intensity sounds are encoded by the activity of medium- and low-SR fibers because high-SR are saturated at those high stimulation levels (Liberman, 1978; Yates, 1990; Bharadwaj et al, 2014, 2015). Indeed, the individual results of the NH threshold listeners demonstrated different EFR level-growth functions for shallowly modulated tones (Fig. 2, A); and more similar functions for strongly modulated tones. For instance, the EFR level-growth...
function for the shallowly modulated tones grew monotonically with a single
slope for listener NH01, whereas in NH09 the EFRs grew non-monotonically
with reduced magnitudes at medium stimulus levels but recovered at higher
levels. The EFR magnitude reduction at 65-70 dB SPL was about 10 dB. This
difference might reflect a true physiological difference between those two sub-
jects, especially because this difference is probably larger than the intrinsic
EFR variability. Encina-Llamas et al (2017a) reported EFR test-retest vari-
ability results at 70 dB SPL from SAM tones at $f_c = 2000$ Hz, $f_m = 93$ Hz and
$m = 85\%$ of no larger than 5 dB, much smaller than the difference seen across
NH threshold listeners in Figure 2. It is assumed though that the difference
in modulation depth in both studies would not affect the test-retest variabil-
ity, although there is no available data to our knowledge showing test-retest
variability in EFRs from SAM tones of different modulation depths. Whether
the different patterns observed in the EFR level-growth functions in listener
NH01 versus NH09 are due to a loss of AN fiber synapses remains unknown.

Individual differences in the EFR level-growth functions for shallowly mod-
ulated tones were also observed for the HI listeners (Fig. 2, B). Whereas the
EFR functions for strongly modulated tones grew monotonically with a single
slope (similarly to the strongly modulated EFRs in the NH threshold listeners),
the EFR magnitudes for shallowly modulated tones did not vary much across
stimulus level, leading to a saturated growth function. The change in slope was
shown to be statistically significant (linear mixed-effects model, Fig. 3). The
question as to whether the different patterns observed from the EFR level-
growth functions in both NH threshold and HI listeners could be due to the loss of AN fiber synapses was investigated using a computational model.

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

Under the assumption that the EFR is sufficiently described by the summation of the instantaneous firing rate of AN fibers across frequency and fiber type, our simulations should shed light on the contributions of CF bands and fiber type to the overall response. Figure 6 shows the simulated EFR level-growth functions for a NH listener, separately for each fiber type (rows) and CF bands (columns). Column A (panels A1-A4) shows the synaptic output summed across all CF bands. Column B (panels B1-B4) shows the output of the band centered at 2 kHz (on-frequency band). Column C (panels C1-C4) shows the output of the band centered at 3 kHz, and column D (panels D1-D4) the output of the band centered at 7 kHz.

The analysis of the simulated AN activity in different CF bands (columns) showed that the EFRs at medium-to-high stimulus levels were not driven solely by activity in the on-frequency band (column B), but also reflected strong contributions of responses of AN neural populations located more basally (i.e., at higher CF bands, columns C and D). Within each frequency band, the response showed a bell-shaped curve, horizontally shifted along the stimulus level axis for more distal CF bands, consistent with the synchrony-level functions recorded from single neurons in the AN of the cat (Joris, 1992). Hence, at
Fig. 6: Simulated AN EFR level-growth functions in a NH listener, separately for different CF bands and for each AN fiber type. Solid circles represent statistically significant EFR magnitudes in dB and open circles represent non-significant responses. Blue markers show responses for strongly modulated stimuli and red markers for shallowly modulated stimuli. The thin lines represent the noise floor in the simulations. Columns show the EFR level-growth functions centered at different CF bands, while rows show the simulated results for different AN SR fiber types.
higher stimulation levels, the off-frequency contributions dominated the total EFR magnitudes in the model framework, leading to the overall monotonic growth observed when summing across CFs.

Similarly, when analyzing the contributions of the different types of AN fibers (rows), the simulated EFRs were dominated by high-SR fibers across the whole stimulus level range. For simplicity, let us focus on the on-frequency band (column B). Even though the strength of the responses of the medium-SR (B2) and low-SR (B3) fibers level-growth functions did not decrease as much with level as did the high-SR (B1) fibers, their overall firing rates were much lower than those of the high-SR fibers (notice the different vertical axis scales). In other words, the medium- and low-SR fibers contribute very little to the summed response (B4). The underlying reason for this is mainly due to the uneven distribution of fiber types assumed in the model, with the high-SR fibers dominating in number (Liberman, 1978). In conclusion, the model simulations suggest that for low stimulus intensities, the envelope was encoded by the high-SR fibers tuned to on-frequency CFs. As stimulus level increases, the modulation coding shifted to being dominated by the high-SR fibers at off-frequency bands (see also Fig. 1).

The above model observations explain why the simulated EFR level-growth functions do not change significantly when a pure OHC dysfunction is imposed (Fig. 4, B). A loss of gain (OHC dysfunction) might only affect the response when presenting a stimulus with spectral content at CFs with OHC damage (on-frequency processing). The AN fiber tuning curves associated with
CFs where there is substantial OHC dysfunction will each show a reduction of sensitivity at the tip of the tuning curve (broader frequency selectivity).

However, the tails (off-frequency responses) will be largely unaffected (Liberman and Dodds, 1984). Thus, the off-frequency excitation of high CF AN fibers through stimuli at a lower CF with normal sensitivity will be the same, regardless of OHC loss. In contrast, when at least one third of the hearing threshold elevation was assigned to dysfunction of the IHCs (Fig. 4, C), a small reduction in the simulated EFR level-growth function was found at stimulus levels corresponding to CFs bands at which the off-frequency contributions were maximal. Nevertheless, no combination of hair-cell impairment led to simulated EFR level-growth functions like those recorded from HI listeners.

The model framework was used to explore how synaptic loss might be used to explain the non-monotonic patterns observed in some NH threshold listeners (Fig. 2, A) and the strong saturation in the HI listeners (Fig. 2, B). We do not intend to claim that a given listener suffers from synaptopathy. Our purpose was to investigate the potential effects of synaptopathy on the EFR level-growth functions by using computational models. The simulated EFR level-growth functions were insensitive to a selective loss of medium- and low-SR fibers (Fig. 5, A); for instance, EFR magnitudes decreased by less than 1.5 dB after including a complete loss of the medium- and low-SR fibers at all CFs. However, it was possible to account for the non-monotonic growth found in some NH threshold listeners (i.e., NH09) by reducing the number of all types of AN fibers (including high-SR fibers loss), mainly centered in
the CF band at which the response peaks at medium levels (Fig. 6, column D). To obtain simulated EFRs that better approximated the results from the HI listeners, a large loss of all three types of AN fiber synapses had to be included across a broad CFs. This was in addition to the posited hair-cell dysfunction (see Table 2). These observations are at odds with previous interpretation, which assume that only medium- and low-SR fibers are affected by synaptopathy (e.g., Furman et al, 2013). However, the model simulations are consistent with previous empirical findings (Paul et al, 2016), where a certain degree of high-SR fiber loss had to be included in order for model simulations to account for the differences observed in the EFR magnitudes recorded in NH threshold listeners with and without tinnitus, which has also been related to synaptopathy (Schaette and McAlpine, 2011).

Our model simulations were not able to capture all the details of the EFR level-growth functions observed in the recorded data. The implementation of either hair-cell dysfunction and synaptopathy within the model framework affected the EFR level-growth functions for both strongly and shallowly modulated tones similarly (see Figs. 4 and 5). In contrast, the recorded EFR level-growth functions for strongly modulated tones were similar across all listeners; large individual differences were observed only on the EFRs for shallowly modulated tones (see Fig. 2). Consistent with the AN model simulations, Shaheen et al (2015) reported a significant reduction in the EFR level-growth functions of synaptopathic mice when using strongly modulated SAM tones with modulation frequencies between 800 to 1000 Hz. The EFRs showed group delays
consistent with generators between the AN and the cochlear nucleus. In the present study, a modulation frequency of 93 Hz was used to elicit the EFRs in humans, which are assumed to be dominated by brainstem sources (Kuwada et al., 2002; Herdman et al., 2002). Brainstem processing, such as central gain mechanisms (Möhrle et al., 2016; Chambers et al., 2016), may differentially affect results from strongly and shallowly modulated stimuli, which may explain the inconsistencies between human data and animal data and the model simulations.

Nevertheless, the model simulation suggested two main conclusions: (1) the EFR level-growth functions at medium-to-high stimulation levels are strongly dominated by the contributions from off-frequency neuronal activity, and (2) there must be a significant degree of loss of high-SR fibers for synaptopathy to be reflected in the EFR. First, the interpretation of the role of the medium- and low-SR fibers on encoding temporal fluctuations at high stimulus levels, based on the rate-level curves of the different AN fibers types (Liberman, 1978; Yates, 1990), has led to different hypotheses about how best to study synaptopathy in humans (including the present study and others like Bharadwaj et al. (2014, 2015); Mehraei et al. (2016)). Rate-level functions are derived from direct recordings in single AN neurons, and therefore provide information regarding AN neuronal activity at on-frequency stimulation. However, electrophysiological evoked responses are driven by large neural populations that encode and interpret sounds. As synaptopathy affects supra-threshold processing, studies of the effects of synaptopathy generally use high sound stimulation levels that
produce a broad excitation of the AN. The contribution of AN neurons tuned
to off-frequency CFs are not well understood, and thus should be carefully con-
sidered in the design of future hypotheses. Second, animal results suggesting
that synaptopathy is selective to medium- and low-SR fibers (Furman et al,
2013) have been interpreted as if noise exposure did not damage high-SR fibers
at all. Incorporating only medium- and low-SR fibers in the model produces
no significant change in the simulated EFRs. However, EFRs were shown to
be reduced in synaptopathic mice relative to the unexposed animals (Shaheen
et al, 2015). One explanation could be that the effect of synaptopathy on the
EFRs was not well captured in the model simulations. Another explanation is
that the lack of vulnerability of the high-SR fibers to synaptopathy reported
in mice may not be directly transferable to other species, such as humans.
An alternative explanation is that high- versus medium- and low-SR fibers
are more evenly distributed in mice (Taberner and Liberman, 2005) than in
cats (Liberman, 1978). In turn, such a distribution will result in a stronger
impact of the loss of medium- and low-SR fiber in the EFR obtained in mice.
In order to study the potential consequences of synaptopathy in humans, audi-
tory evoked potential studies must be in parallel in humans and in non-human
species where synaptopathy has been characterized, together with the use of
species-specific computational models.
5 Conclusions

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

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

Overall, the simulations suggest that EFRs are dominated by high-SR fibers, and that off-frequency neurons increasingly contribute to the EFR as stimulus level increases. The finding that when SAM tones are presented at high stimulus levels, the envelope is better encoded at off-frequency CF responses (rather than on-frequency) is a key concept to consider when using EFRs to investigate supra-threshold coding.
Acknowledgements The authors want to thank Le Wang and Graham Voysey from Boston University (BU) for their valuable comments and discussions regarding the model implementation and analysis. I would like to express our sincere gratitude to Laurel Carney for her wise advice regarding the analysis of the model outcome and the fascinating discussions during her stay in Denmark about the role of the different types of AN fibers, the interpretation of the rate-level functions and the processing of modulations in the AN, the CN and the IC. This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) at the Technical University of Denmark (DTU). The authors want to acknowledge the support provided by the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience.

References


Investigating the effect of synaptopathy on EFRs using an AN model


Investigating the effect of synaptopathy on EFRs using an AN model.


Lobarinas E, Spankovich C, Le Prell CG (2016) Evidence of hidden hearing loss following noise exposures that produce robust TTS and ABR wave-I amplitude reductions. Hear Res


thy can be centrally compensated in the young but not old brain. Neurobiol Aging 44:173–184


Paul BT, Bruce IC, Roberts LE (2016) Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus. Hear Res


Schaette R, McAlpine D (2011) Tinnitus with a normal audiogram: physiological
evidence for hidden hearing loss and computational model. J Neurosci
31(38):13,452–7

synaptopathy: an early-onset contributor to auditory functional decline. J
Neurosci 33(34):13,686–94

Shaheen LA, Valero MD, Liberman MC (2015) Towards a Diagnosis of
Cochlear Neuropathy with Envelope Following Responses. J Assoc Res Oto-
laryngol


of hair cell loss in CBA and C57BL/6 mice throughout their life spans. J
Acoust Soc Am 101(6):3546–53

Taberner AM, Liberman MC (2005) Response properties of single auditory

Tremblay KL, Pinto A, Fischer ME, Klein BEK, Klein R, Levy S, Tweed TS,
Cruickshanks KJ (2015) Self-reported hearing difficulties among adults with
normal audiograms: the beaver dam offspring study. Ear Hear 36(6):e290–9,


AN synaptic output

A) AN synaptic output

B) OFF-frequency (7 kHz)

C) OFF-frequency (3 kHz)

D) ON-frequency (2 kHz)

E) Σ across CFs

Stimulus level: 5 dB SPL

Click here to download Supplemental Material (Not to be Published)
AN synaptic output

Characteristic frequency [kHz]

Stimulus level: 10 dB SPL

Figures to compile the animation in LaTeX.

Click here to download Supplemental Material (Not to be Published)
AN synaptic output

Characteristics:

A) Stimulus level: 15 dB SPL

- Time [s]
- Synaptic response [spikes]

- OFF-frequency (7 kHz)
- OFF-frequency (3 kHz)
- ON-frequency (2 kHz)
- Σ across CFs

E) Σ across CFs

Figures to compile the animation in LaTeX

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 20 dB SPL

**AN synaptic output**

- **Characteristic frequency [kHz]**
  - A)
    - 0.408
    - 0.416
    - 0.424
    - 0.432
    - 0.44

- **Time [s]**
  - 0.2
  - 0.4
  - 0.7
  - 1.2
  - 2.0
  - 3.15
  - 5.0
  - 8.0
  - 12.5
  - 20.0

**OFF-frequency (7 kHz)**

- B)
  - 0
  - 10
  - 20
  - 30
  - 40

**OFF-frequency (3 kHz)**

- C)
  - 0
  - 10
  - 20
  - 30
  - 40

**ON-frequency (2 kHz)**

- D)
  - 0
  - 10
  - 20
  - 30
  - 40

**Σ across CFs**

- E)
  - 0
  - 10
  - 20
  - 30
  - 40

Stimulus level: 20 dB SPL

~Figures to compile the animation in latex~

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 25 dB SPL

Figures to compile the animation in latex.

A) AN synaptic output

B) OFF-frequency (7 kHz)

C) OFF-frequency (3 kHz)

D) ON-frequency (2 kHz)

E) Σ across CFs

Stimulus level: 25 dB SPL
Stimulus level: 30 dB SPL

**A)** AN synaptic output

![Graph showing AN synaptic output with characteristic frequencies marked at 2.0, 3.15, 5.0, and 8.0 kHz.]

**B)** OFF-frequency (7 kHz)

![Graph showing synaptic response at 7 kHz.]

**C)** OFF-frequency (3 kHz)

![Graph showing synaptic response at 3 kHz.]

**D)** ON-frequency (2 kHz)

![Graph showing synaptic response at 2 kHz.]

**E)** $\Sigma$ across CFs

![Graph showing the sum of synaptic responses across characteristic frequencies.]

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 35 dB SPL

Figures to compile the animation in latex

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 40 dB SPL

AN synaptic output

OFF-frequency (7 kHz)

OFF-frequency (3 kHz)

ON-frequency (2 kHz)

Σ across CFs

Figures to compile the animation in LaTeX: 08

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 45 dB SPL

A) AN synaptic output

B) OFF-frequency (7 kHz)
C) OFF-frequency (3 kHz)
D) ON-frequency (2 kHz)
E) $\Sigma$ across CFs

Stimulus level: 45 dB SPL
AN synaptic output

A) Stimulus level: 50 dB SPL

B) OFF-frequency (7 kHz)

C) OFF-frequency (3 kHz)

D) ON-frequency (2 kHz)

E) Σ across CFs

Synaptic response [spikes]

Time [s]

Characteristics frequency [kHz]
Stimulus level: 55 dB SPL

AN synaptic output

Characteristic frequency [kHz]

OFF-frequency (7 kHz)

OFF-frequency (3 kHz)

ON-frequency (2 kHz)

Σ across CFs

Figures to compile the animation in latex.

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 60 dB SPL

Figures to compile the animation in latex. Click here to download Supplemental Material (Not to be Published)
AN synaptic output

Characteristic frequency [kHz]

Time [s]

Stimulus level: 65 dB SPL

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 70 dB SPL

A) AN synaptic output

B) OFF-frequency (7 kHz)

C) OFF-frequency (3 kHz)

D) ON-frequency (2 kHz)

E) Σ across CFs

Characteristics of frequency [kHz]

0.2 0.4 0.6 0.8 1 1.2 1.4 1.6 1.8 2

Time [s]

0.408 0.416 0.424 0.432 0.44
Stimulus level: 75 dB SPL

A) AN synaptic output

B) OFF-frequency (7 kHz)

C) OFF-frequency (3 kHz)

D) ON-frequency (2 kHz)

E) Σ across CFs

Synaptic response [spikes]

Characteristics frequency [kHz]

Time [s]

Stimulus level: 75 dB SPL

Click here to download Supplemental Material (Not to be Published)
Stimulus level: 80 dB SPL

A) AN synaptic output

B) OFF-frequency (7 kHz)

C) OFF-frequency (3 kHz)

D) ON-frequency (2 kHz)

E) Σ across CFs

Stimulus level: 80 dB SPL

Figures to compile the animation in latex  
Click here to download Supplemental Material (Not to be Published)
AN synaptic output

Stimulus level: 90 dB SPL

Figures to compile the animation in latex

Click here to download Supplemental Material (Not to be Published)
Supplemental Material (to be Published)
1_fig__method_an__gram_anim__mov.mpeg
Click here to access/download

Supplemental Material (to be Published)
Encina-Llamas_efr_human_model_2017__supl_matrl_anim.pdf
Click here to access/download

Supplemental Material (to be Published)
Encina-
LlamasEFRhuman_model_2017__supl_matrl_anim.tex
Subject: Conflict of Interest disclosure statement

Authors declare no conflicts of interests.

Gerard Encina-Llamas, James M. Harte, Torsten Dau, Barbara Shinn-Cunningham and Bastian Epp