From derived-band envelope-following responses to individualized models of near- and supra-threshold hearing deficits

SARINEH KESHISHZADEH* AND SARAH VERHULST

Hearing Technology @WAVES, Department of Information Technology, Ghent University, Belgium

Auditory models which include frequency-dependent profiles of near and supra-threshold hearing deficits can aid the design of individualized hearing-aid algorithms. However, determining individual auditory-nerve (AN) fiber loss parameters is controversial as diagnostic metrics are presently based on auditory brainstem responses (ABRs) or envelope following responses (EFRs). These measures do not necessarily yield a frequency-specific quantification and might be affected by both outer-hair-cell and AN damage. We developed a derived-band EFR (DBEFR) metric to offer a frequency-specific assessment and complemented these with click-evoked otoacoustic emissions and audiometry. Cochlear-gain-loss profiles were derived from the latter measurements and inserted into individualized models, in which different synaptopathy profiles were introduced and DBEFRs simulated. Using a clustering technique, the best match between experimental and simulated synaptopathy profiles was determined and validated using the ABR data collected from the same listener. Results showed promise in offering a method to determine individualized sensorineural hearing-loss profile given a limited number of objective metrics.

INTRODUCTION

The number of auditory nerve (AN) fibers which synapse onto inner-hair-cells is an important factor for auditory processing of supra-threshold sound. A reduction of synapses as a consequence of noise exposure or ageing, i.e. cochlear synaptopathy (CS), degrades temporal encoding fidelity of supra-threshold sound, while leaving the audiometric thresholds unaffected (Kujawa and Liberman, 2009; Bharadwaj *et al.*, 2014). Evidence from animal studies have shown that the amplitude of auditory evoked potentials (AEPs), such as auditory brainstem responses (ABR) and envelope following responses (EFR) are sensitive markers of histologically verified CS (Kujawa and Liberman, 2009; Furman *et al.*, 2013; Shaheen *et al.*, 2015). CS compromises the number of AN fibers which can fire synchronously to a stimulus and hence reduces the supra-threshold ABR wave-I amplitude (Kujawa and Liberman, 2009) and its growth-slope as a function of increasing stimulus level (Furman *et al.*, 2013; Möhrle *et al.*, 2016). Moreover, noise-induced CS degrades the phase-locking strength to the envelope of a modulated tone (Shaheen *et al.*, 2015; Bharadwaj *et al.*,

*Corresponding author: sarineh.keshishzadeh@ugent.be

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 7: Auditory Learning in Biological and Artificial Systems, August 2019, Nyborg, Denmark. Edited by A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau. The Danavox Jubilee Foundation, 2019. ISSN: 2596-5522.

2014, 2015; Parthasarathy and Kujawa, 2018). However, direct assessment of AN synapses requires invasive procedures, therefore quantification of CS in humans is challenging, as scalp-recorded AEPs represent summed activity of large populations of neurons which can be influenced by multiple sources, such as outer-hair-cell (OHC) loss and CS (Gorga *et al.*, 1985; Verhulst *et al.*, 2016). Furthermore, the spread of basilar membrane (BM) excitation caused by the stimulation paradigm causes off-frequency channels to contribute and complicate frequency-specific AEP-based CS diagnostics (Bharadwaj *et al.*, 2014; Encina-Llamas *et al.*, 2019). To address these issues and make a precise quantification of CS in humans possible, this study adopts a combined experimental and modelling approach. A comparison between simulated and experimental frequency-specific EFRs are used to find the best matching AN-damage profile among different simulated CS profiles, which can explain the experimental observations. Cochlear gain model parameters (OHC damage) were set based on experimental audiometric thresholds and click-evoked otoacoustic emissions (CEOAE), while individual CF-specific CS profiles were the outcome parameter, given the EFR recordings.

METHOD

A. Experimental approach

Participants

Two groups were recruited for the experiment, (i) a normal-hearing (NH: 24.21 $\pm$ 4.10 years, N = 16) group and (ii) a group of listeners with normal audiometric thresholds ($<$ 20 dB HL at 4 kHz and $<$ 25 dB HL at frequencies above 4 Hz), but self-reported hearing difficulties in noisy environments (NHSR: 33.78 $\pm$ 8.57 years, N = 9). The latter participant group was recruited using flyers asking *"Do you experience problems when communicating in noisy environments?"*. The underlying assumption was that the NHSR group could suffer from supra-threshold hearing deficits and yield EFR metrics representative of those recorded in synaptopathy animals. Audiometric thresholds at half-octave frequencies between 250 and 8000 Hz were assessed and the best audiometric ear at 4 kHz was chosen for the experiment. Measurements were performed in an acoustically and electrically shielded booth, while subjects were watching a silent movie with subtitles. Participants were informed about the experiment details and an informed consent was received according to the ethical commission at Ghent University.

Derived-band envelope following responses (DBEFRs)

EFRs were recorded to 120-Hz amplitude-modulated white-noise carriers with a bandwidth of [2-22] or [4-22] kHz and a modulation depth of 100%. Stimuli were presented monaurally with an equal spectral level of 70 dB SPL, yielding a lower loudness percept for the narrower band stimulus. Alternate polarity 1.25 s-length epochs were presented 370 times (185 of each polarity) using the same experimental setup as described in (Keshishzadeh *et al.*, 2019a). 1s-long epochs were extracted

from Cz-channel EFRs, starting from 0.25 s after the trigger onset. Employing the same steps adopted in Keshishzadeh *et al.* (2019b), peak to noise-floor spectral absolute values at the fundamental frequency ($f_0 = 120$Hz) and the two following harmonics were extracted (EFR_{PtN}). Afterwards, the DBEFR magnitude was defined by subtracting the EFR_{PtN} to different bandwidths using:

$$DBEFR_{[2-4]} = \begin{cases} (EFR_{PtN})_{[2-22]} - (EFR_{PtN})_{[4-22]}, & (EFR_{PtN})_{[2-22]} > (EFR_{PtN})_{[4-22]} \\ 0, & \text{else} \end{cases} \quad \text{(Eq. 1)}$$

Auditory brainstem responses (ABRs)

Alternate polarity 80 μs-length clicks were presented with a 11.38 Hz rate using the setup described in Keshishzadeh *et al.* (2019a). ABRs were recorded to 90 and 100 dB peSPL clicks and 3000 repetitions with a 10% jitter around the inter-click interval (ICI). The waveforms from nine central channels, i.e. F1, Fz, F2, FC1, FCz, FC2, C1, Cz and C2 were bandpass-filtered using an 800-th order FIR-filter (zero-phase filtering with *filtfilt* function of MATLAB) in [100-1500]-Hz bandwidth and were epoched between -5 to 20-ms relative to the onset. A baseline drift correction was applied to each epoch by subtracting corresponding means. Afterwards, every positive epoch was averaged with the following negative one and the subtraction between the peak and trough of each averaged pair was calculated as the criterion for rejection of noisy trials. Ninety paired-averages (i.e., 180 trials) with the highest peak-to-trough values were assumed as artifact-contaminated pairs and removed. The peak-to-trough subtracted values of the remaining trials did not exceed 25 μV. This approach was adopted to avoid the possible unequal averaging of epochs of each polarity after artifact rejection and was the same for all listeners. Finally, ABRs were calculated by averaging the remaining paired-averages across the nine channels. Wave-V amplitudes were identified manually from the wave-V peak to the next trough in the waveform.

Click-evoked otoacoustic emissions (CEOAEs)

CEOAEs were recorded in response to 83.33 μs duration clicks and presented at 70 d -peSPL with a rate of 25 Hz with ICI of 39.92 ms and 2000 repetitions. Clicks were generated in MATLAB and sent via a Fireface UCX external sound card (RME) and headphone driver to an ER10X Extended-Bandwidth Etymotic Research Probe System. The ER10X recorded ear-canal pressure and responses were digitized via the external sound card (RME). First, time-domain raw recordings were converted to pressure by using the microphone sensitivity (50 $\frac{mV}{Pa}$) and amplifier gain (40dB), and were then bandpass-filtered between 250 and 6000 Hz using a 32^{nd} order zero-phase FIR-filter. Ten percent of the trials with amplitudes larger than twice the standard deviation of the mean, were rejected. The noise-floor was calculated by subtracting the odd and even trials, assuming that the residual noise was not correlated to the stimulus. Afterwards, a linear windowing method was adopted to extract the OAE

using a tenth-order recursive exponential window (Kalluri and Shera, 2001; Shera and Zweig, 1993).

B. Individualized auditory periphery model

A computational model of the auditory periphery (Verhulst *et al.*, 2018) was employed to simulate the experimental data and derive individualized frequency-specific CS profiles to simulate individual sensorineural hearing-loss (SNHL) profiles. Two main steps were necessary to simulate individualized SNHL profiles: (i) Introduce frequency-specific OHC damage profiles using the CEOAE and audiometric data and (ii) simulate EFRs for a variety of CS profiles. In the first step, we first simulated individual cochlear BM impedance discontinuities (Shera and Guinan Jr, 2007), which were derived from the 70 dB peSPL CEOAE recordings. The randomized discontinuities across CF in the model, i.e. the CF-dependent roughness parameters, were individualized by matching the spectral peaks and troughs in the [250-6000] Hz frequency range of the recorded CEOAEs. Spectral bins without peaks or troughs were set to zero and the generated CF-dependent vector was normalized between -1 and +1. Next, individual audiometric thresholds were used to adjust the cochlear-gain-loss parameters by determining the pole values of the BM admittance function across CF (Verhulst *et al.*, 2016). This operation translates a dB-HL loss into a corresponding cochlear filter with lower gain and wider bandwidth. In the second step, the individualized model for OHC-damage-related parameters was used to simulate EFR/ABRs for a range of loss profiles.

To model the spread of total AN fibers per CF, the CF-dependence of counted synapses for rhesus monkey reported in Valero *et al.* (2017) were adopted and mapped to the human cochlea using the Greenwood function (Greenwood, 1990). This resulted in a non-uniform distribution of AN fibers across the CF channels, where a specific CF encompasses $N_{HSR} = 68\%, N_{MSR} = 16\%$ and $N_{LSR} = 16\%$ of the total population of AN fibers in that channel. The population of AN fibers were reduced in all CF channels to simulate different degrees of CS (Table 1). Lastly, models with different CS profiles were used to simulate EFRs. DBEFRs were extracted from the steady-state part of the response, using the same method as for the experimental data. The simulations were run for five different individualized CS profiles introduced in Table 1. To determine the CS profile which best matched the simulated and recorded $DBEFR_{[2-4]}$, a Fuzzy C-Means (FCM)-based clustering technique was adopted to

AN Type	**Total number of AN fibers per CF for no CS *(N)***	**Simulated CS Profile**				
		N	**A**	**B**	**C**	**D**
HSR	13	100%	100%	54%	31%	8%
MSR	3	100%	–	–	–	–
LSR	3	100%	–	–	–	–

Table 1: Simulated CS profiles

explain the uncertainty with which a recorded DBEFR pair belonged to a certain CS profile (using a Fuzzy membership function). Different from original FCM clustering, where initial clusters centers are unknown and updated in each iteration, the cluster centers were defined as the simulated $\text{DBEFR}_{[2-4]}$ for each CS profile. Therefore, a single iteration was run to calculate the partition matrix. Algorithm 1 presents a pseudocode to cluster the experimental $\text{DBEFR}_{[2-4]}$ using model simulations.

RESULTS AND VALIDATION

Experimental EFRs (Fig. 1a) showed individual variabilities and overall lower group-means for the NHSR group ($\text{EFR}_{[2-22]}$: $t(19) = 3.36$, $p \approx 0.003$ and $\text{EFR}_{[4-22]}$: $t(19) = 2.76$, $p \approx 0.012$). Extracting the $\text{DBEFRs}_{[2-4]}$ (Fig. 1b) reduced the group-mean differences ($\text{DBEFR}_{[2-4]}$: $t(19) =$ -0.90, $p \approx 0.338$). Individual DBEFRs were employed to find the best match between simulated/recorded personalized CS profiles. Table 2 shows ranked clustering results based on the individual Fuzzy membership degrees (u_k). The first column refers to the best-matched CS profile and the last column corresponds to the lowest ranked CS profile. To validate the predicted DBEFR-based CS-profiles, an independent measure was used which was also recorded (the ABR), but not adopted in the model fit procedure. Although evidence from animal studies point to decreased ABR wave-I growth-slopes as a consequence of CS, robust wave-I peak-picking is controversial in humans. Given that the ABR growth-slope is a relative metric, we assumed that any hearing deficit reflecting on ABR wave-I, would travel through the auditory pathway to inferior colliculus (IC) and reflect on ABR wave-V, as well. The ABR wave-V has similar generator sources in the vicinity of the IC as the EFR and can be recorded with higher signal-to-noise ratios. ABR wave-V amplitudes to 90 and 100 dB peSPL clicks were simulated for the same five CS profiles in Table 1, and the corresponding wave-V amplitudes growth-slope was calculated. Note that subjects without experimentally growing ABR slopes (N = 7) were dropped, since validation was impossible for these individuals. Model-predicted slopes for each CS profile were compared to the experimental ABR slopes and best matching CS profiles were determined based

Algorithm 1 FCM-based Clustering

Fix the number of clusters c to 5 and fuzzifier m to 1.2

$x = x_1, x_2, ..., x_n$, experimental $\text{DBEFR}_{[2-4]}$ for subject n

$v_n = v_{n1}, v_{n2}, ..., v_{nc}$, simulated $\text{DBEFR}_{[2-4]}$ as cluster centers for subject n

for $k = 1$ to n **do**
  for $j = 1$ to c **do**
    $d[v_{kj}, x_k] = abs(x_k - v_{kj})$
    Calculate membership matrix $u_{kj} = \frac{1}{\sum_{i=1}^{c} (\frac{d[v_{kj}, x_k]}{d[v_{ki}, x_k]})^{\frac{2}{m-1}}}$
  end for
end for

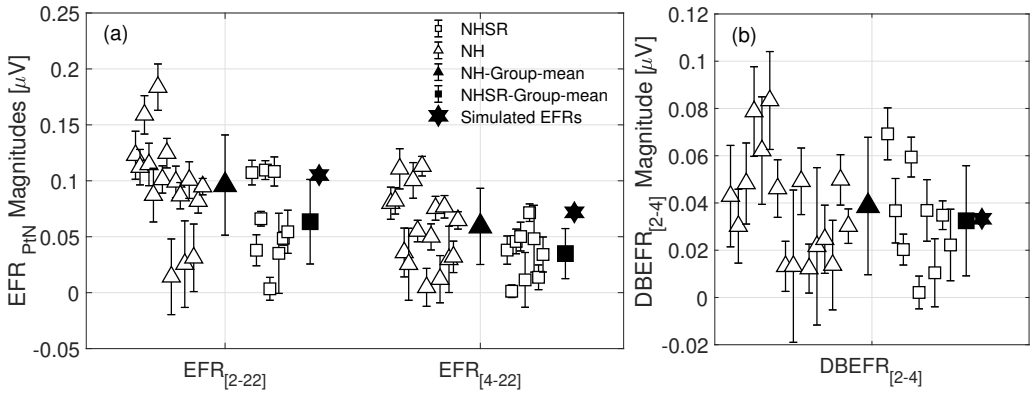


Fig. 1: Experimental and simulated (a) EFRs and (b) $\text{DBEFR}_{[2-4]}$. Error bars indicate the standard deviation of the mean-values.

on the lowest absolute difference between experimental and simulated growth-slopes (indicated with * in Table 2).

DISCUSSION

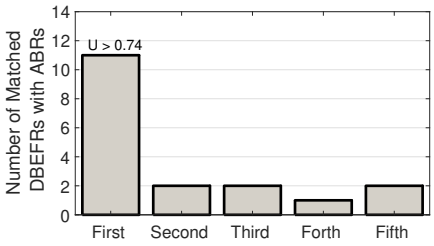
The employed relative derived-band metric was designed to yield a frequency-specific EFR marker of supra-threshold temporal envelope coding which would suppress individual variability (Fig.1a) stemming from subject-specific factors, such as head-size and gender. In a previous study, we used the same DBEFR method (Keshishzadeh *et al.*, 2019a), which revealed statistically significant difference between young normal-hearing (yNH) and old normal-hearing (oNH) DBEFRs, and assigned the degraded DBEFRs in the oNH group to age-induced CS. By extension, one could

NH	**Ranked Predicted CS Profile**					**NHSR**	**Ranked Predicted CS Profile**				
	First	Second	Third	forth	Fifth		First	Second	Third	Forth	Fifth
1	N^*	A	B	C	D	**1**	N^*	A	B	C	D
2	N^*	A	B	C	D	**2**	N	A^*	B	C	D
3	N^*	A	B	C	D	**3**	A	B	C	N^*	D
4	N^*	A	B	C	D	**4**	N^*	A	B	C	D
5	N^*	A	B	C	D	**5**	N^*	A	B	C	D
6	B	C	A	D	N^*	**6**	B	C	D	A	N^*
7	N^*	A	B	C	D	**7**	N^*	A	B	C	D
8	A^*	N	B	C	D	**8**	A	B	N^*	C	D
9	D	C^*	B	A	N						
10	N	A	B^*	C	D						

* *ABR-based predicted CS profiles*

Table 2: DBEFR-based predictions of individual CS profiles validated using experimental ABR growth-slopes derived from the same listeners. *(N: Normal)*

assume that the group mean differences between NH and NHSR group could be explained by noise-induced CS. However differently, we did not observe significant DBEFR group-mean differences and obtained *Normal* CS profiles for most participants of both groups together. Taking into account our previous findings, and the insensitivity of the DBEFR metric to head-size and DPOAE threshold differences (Keshishzadeh *et al.*, 2019a, see Fig.7), we believe that the present results reflect that not all listeners with self-reported hearing difficulties suffer from CS. Additionaly, the unquantifiable separation criterion between the two tested groups, i.e. self-reported hearing difficulties in noisy environments (Coughlin, 1990) and insufficient number of participants can explain the absense of significant differences between the groups. To evaluate the quality of our CS profile predictions, Fig. 2 depicts the performance of our method in correctly predicting either the ABR or DBEFR experimental results using the individual CS profile extracted from the DBEFR clustering method. A coincidence of 61.11% validated profiles with the first-ranked $\text{DBEFR}_{[2-4]}$ based predictions shows promising results. In the future our method can be improved by including other AEP-derived metrics and OHC-deficit related measures in the clustering to yield more robust predictions of individual CS degrees.



Rank	Coincidence (%)
1	61.11
2	11.11
3	11.11
4	5.56
5	11.11

Fig. 2: Evaluation of the method based on the coincidence between first-ranked $\text{DBEFR}_{[2-4]}$ based predictions and ABR wave-V growth-slopes

ACKNOWLEDGEMENT

Work supported by European Research Council grant ERC-StG-678120 (RobSpear).

REFERENCES

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., and Shinn-Cunningham, B. G. (**2015**), "Individual differences reveal correlates of hidden hearing deficits," J. Neurosci., **35**(5), 2161–2172.

Bharadwaj, H. M., Verhulst, S., Shaheen, L., Liberman, M. C., and Shinn-Cunningham, B. G. (**2014**), "Cochlear neuropathy and the coding of supra-threshold sound," Front. Syst. Neurosci., **8**, 26.

Encina-Llamas, G., Harte, J. M., Dau, T., Shinn-Cunningham, B., and Epp, B. (**2019**), "Investigating the Effect of Cochlear Synaptopathy on Envelope Following Responses Using a Model of the Auditory Nerve," J. Assoc. Res. Oto., 1–20.

Furman, A. C., Kujawa, S. G., and Liberman, M. C. (**2013**), "Noise-induced cochlear

neuropathy is selective for fibers with low spontaneous rates," J. Neurophysiol., **110**(3), 577–586.

Gorga, M. P., Worthington, D. W., Reiland, J. K., Beauchaine, K. A., and Goldgar, D. E. (**1985**), "Some comparisons between auditory brain stem response thresholds, latencies, and the pure-tone audiogram." Ear Hearing, **6**(2), 105–112.

Greenwood, D. D. (**1990**), "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Am., **87**(6), 2592–2605.

Kalluri, R. and Shera, C. A. (**2001**), "Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation," J. Acoust. Soc. Am., **109**(2), 622–637.

Keshishzadeh, S., Garrett, M., and Verhulst, S. (**2019**a), "The Derived-Band Envelope Following Response and its Sensitivity to Sensorineural Hearing Deficits," bioRxiv, 820704.

Keshishzadeh, S., Vasilkov, V., and Verhulst, S. (**2019**b), "Tonotopic Sensitivity to Supra-Threshold Hearing Deficits of the Envelope Following Response Evoked by Broadband Stimuli," 23rd International Congress on Acoustics, 6548–6553.

Kujawa, S. G. and Liberman, M. C. (**2009**), "Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss," J. Neurosci., **29**(45), 14077–14085.

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

Parthasarathy, A. and Kujawa, S. G. (**2018**), "Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing," J. Neurosci., **38**(32), 7108–7119.

Shaheen, L. A., Valero, M. D., and Liberman, M. C. (**2015**), "Towards a diagnosis of cochlear neuropathy with envelope following responses," J. Assoc. Res. Oto., **16**(6), 727–745.

Shera, C. A. and Guinan Jr, J. J. (**2007**), "Cochlear traveling-wave amplification, suppression, and beamforming probed using noninvasive calibration of intracochlear distortion sources," J. Acoust. Soc. Am., **121**(2), 1003–1016.

Shera, C. A. and Zweig, G. (**1993**), "Noninvasive measurement of the cochlear traveling-wave ratio," J. Acoust. Soc. Am., **93**(6), 3333–3352.

Valero, M., Burton, J., Hauser, S., Hackett, T., Ramachandran, R., and Liberman, M. (**2017**), "Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta)," Hearing Res., **353**, 213–223.

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

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