PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Hearing-impaired sound perception: What can we learn from a biophysical model of the human auditory periphery?

Alejandro OSSES VECCHI[(1)], Frauke ERNST[(2)], Sarah VERHULST[(1)]

[(1)]Department of Information Technology, Ghent University, Belgium, alejandro.osses@ugent.be, s.verhulst@ugent.be

[(2)]Department of Otorhinolaryngology, Head and Neck Surgery, University of Giessen, Germany, frauke.ernst@hno.med.uni-giessen.de,

ABSTRACT

Auditory modelling provides a powerful quantitative framework to study performance in psychoacoustic listening tasks such as signal or modulation detection, and speech perception. Biophysically-inspired auditory models provide the straightforward possibility to account for different aspects of sensorineural hearing loss (e.g., hair cell loss or synaptopathy) and are well suited to study the origin of individual performance differences on behavioural tasks. Here, we simulated how detection thresholds to low- and high-frequency supra-threshold stimuli changed for various profiles of sensorineural hearing loss: (1) normal or high-frequency sloping audiograms ($>$ 1 kHz, 30 dB HL at 8 kHz), and (2) normal or gradual loss of auditory-nerve synapses. The tasks included the detection of amplitude-modulated tones at carrier frequencies of 500 Hz (f_{mod}=5 Hz) and 4 kHz (f_{mod}=100 Hz) presented with or without background noise. The simulations were compared with experimental data collected from listeners with normal or high-frequency sloping audiograms. The simulations accounted for up to 8.6 dB of the experimental variability, where the simulated variability was derived from the maximum shift between simulated thresholds in each condition.

Keywords: Auditory modelling, hearing impairment, psychoacoustics

1 INTRODUCTION

Historically, simulations of signal processing along the auditory pathway have been used as a tool to study specific aspects of the auditory physiology or to explain experimental results obtained in psychoacoustic experiments [e.g., 2, 5]. In both cases, artificial sounds with controlled properties can be simulated to determine the expected outcomes of a given experimental task that will be conducted with human listeners, either using an objective or behavioural experimental approach.

The set of processing stages accounted for by such auditory models mimic with greater or lesser accuracy constitutive parts of the human ear including: Outer and middle ear, which are often simulated as a combined bandpass filter; Inner ear, simulated as a set of bandpass filters covering the audible range (cochlear filtering), followed by some type of transformation that resembles the transduction from mechanical waves into neural responses, and; Central processing stage that approximates the way in which higher-level neural representations are further processed and compared to other representations, as it is done in a psychoacoustic detection task.

Following the model classification proposed by Saremi et al. [11], existing auditory models can be classified into perceptually- or mechanically-inspired models. We refer here to this latter type as biophysical models.

One benefit of using a biophysical model is that aspects of hearing impairment, such as cochlear gain loss (that leads to elevated hearing thresholds) or synaptopathy (hypothesised to decrease supra-threshold hearing performance), can be directly accounted for by setting the mechanical model parameters that affect the basilar membrane motion (sensitivity and tuning of cochlear filtering) or by disconnecting auditory nerve fibres systematically, to account for cochlear gain loss and synaptopathy, respectively.

In this study, we adopt the biophysical model described by Verhulst et al. [13, 14] to study individual differences in psychoacoustic performance[1]. This model was earlier adopted to simulate a tone-in-noise and amplitude-modulation (AM) detection thresholds for high carrier frequencies (f_c =4 kHz) [16]. We extend this work here to explore how much of the performance variability in a range of detection conditions collected from the same listeners can be explained by (i) adopting different profiles of cochlear gain loss and synaptopathy, and by (ii) stimulus changes across conditions.

[1]Examples of published perceptual models can be found in [2, 6, 17] and of biophysically-inspired models in [7, 18].

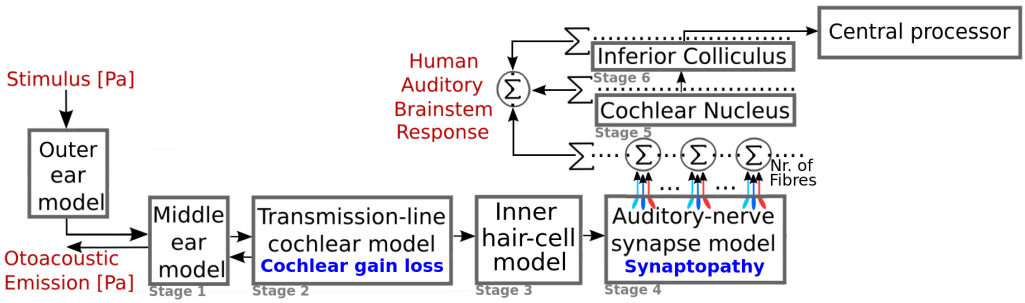

Figure 1. Block diagram of the auditory model. Two stages were added or modified from the original description [13, 14]: "Outer ear model" and "central processor".

2 MATERIALS AND METHODS

2.1 MODEL DESCRIPTION

2.1.1 Peripheral processing stages

The adopted model offers a quantitative description of the different functional components of the hearing system and provides simulated neural representations along the ascending (human) auditory pathway up to the inferior colliculus (auditory brainstem) [13, 14]. The block diagram is depicted in Fig. 1 and illustrates that sounds are fed into either the outer-ear model [10] (as implemented in the AMT toolbox for MATLAB [12]), to account for the transfer function from circumaural headphones to the tympanic membrane, or directly passed through a middle-ear filter, if the sounds are presented over in-ear earphones (Stage 1). The signal is then fed into a transmission-line cochlear filter bank (Stage 2) where basilar membrane vibrations of 401 equidistant locations along the cochlea are simulated according to the human cochlear place-frequency map. Each partition is a bandpass filter matching human otoacoustic emission tuning. A non-linear compression is applied to the cochlear vibrations to simulate the level dependence of cochlear filter tuning. The filter outputs are then converted into inner hair cell receptor potentials (Stage 3) which are used to generate spike-rate patterns at the level of the auditory nerve (Stage 4). These patterns are related to three types of fibres that have high- (HSR, 70 spikes/s), medium- (MSR, 10 spikes/s), and low-spontaneous rates (LSR, 1 spikes/s). These patterns are eventually added up (13 HSR, 3 MSR, and 3 LSR fibres, i.e., 13-3-3) to form the input to a functional model of the spherical bushy cell in the cochlear nucleus (Stage 5) and inferior colliculus (Stage 6). These stages capture the modulation and onset sensitivity of bushy cells in the auditory brainstem [8]. The information across the simulated characteristic frequencies available at the outputs of Stages 4, 5, and 6 offer a correlate of human brainstem responses recorded from scalp electrodes. In this paper we only use simulated neural responses at the output of Stage 6 to assess the detection information that human listeners have available when conducting a psychoacoustic experiment. This process is described further in Section 2.3.2.

2.1.2 Hearing loss

As highlighted in Fig. 1 the auditory model can account for two types of inner-ear dysfunction: cochlear gain loss and synaptopathy. Cochlear gain loss is caused by damaged outer hair cells and leads to elevated hearing thresholds and reduced frequency selectivity. In the model (Stage 2), individual audiograms can be simulated by changing the gain parameters at each frequency partition [15]. In this study we adopted two cochlear gain loss profiles that reflect: (i) a normal audiogram, and (ii) a high-frequency sloping audiogram above 1 kHz, with a hearing threshold of 30 dB HL at 8 kHz. In the model, synaptopathy can be introduced by disconnecting auditory nerve fibres (Stage 4). In this study we adopted three synaptopathy profiles: 13-3-3, 13-0-0, and 6-0-0.

2.2 EXPERIMENTAL PROCEDURES

The experiments were conducted in a double-walled sound-proof booth. The stimuli had a sampling frequency of 44.1 kHz and were presented monaurally either in quiet or using a background noise via Sennheiser HD200 circumaural headphones. The participants' responses were collected using the AFC toolbox for MATLAB [3]. The experimental procedures complied with the ethical procedures of the University of Oldenburg.

2.2.1 Protocol

The stimuli were presented using three-alternative forced-choice (3-AFC) trials where the adjustable variable, the modulation depth m, was changed using an adaptive tracking rule (two-down, one-up). The starting m value was set to -12 dB and the step size was adapted every two reversals (10, 5, 3, 1 dB). Modulation-detection thresholds were calculated as the mean of the last 6 reversals with a 1-dB step size. Each experimental condition was measured three times for each participant, and their median was used for further analysis.

Table 1. List of conditions and information about the background noises used in the AM detection tasks.

Noise type	f_c, f_{mod} [Hz]	Bandwidth BW (f_{lower}, f_{upper}) [Hz]	$BW_{1\text{-ERB}}$ [Hz]	SpecL [dB]	BL_{ERB} [dB]	L_{tot} [dB]	Stimulus duration [s]
Q	500, 5	-	79	-	-	-	1.0
$N_{2/3\text{-Oct.}}$	500, 5	233 (397, 630)	79	41.0	60.0	64.7	1.0
$N_{8\text{-Hz}}$	500, 5	8 (496, 504)	79	50.0	59.0	59.0	1.0
Q	4000, 100	-	456	-	-	-	0.5
$N_{2/3\text{-Oct.}}$	4000, 100	1865 (3175, 5040)	456	25.0	51.6	57.7	0.5
$N_{160\text{-Hz}}$	4000, 100	160 (3920, 4080)	456	45.0	67.0	67.0	0.5

2.2.2 Stimuli

The reference sound (presented twice) was a pure-tone centred at either $f_c = 500$ Hz (duration of 1 s) or $f_c = 4$ kHz (duration of 0.5 s) and included 25-ms and 12.5-ms raised cosine ramps, respectively. The target sound (presented once) was an AM tone of the same f_c, modulation depth m, modulated at rates of 5 and 100 Hz, respectively. All sounds were set to have an RMS level of 70 dB. AM detection thresholds were collected in three conditions: (i) in quiet (Q), (ii) in a background noise which had a bandwidth (BW) of 80% the sideband-to-sideband width of the modulation tone (condition labelled as $N_{8\text{-Hz}}$ for 500-Hz tones and $N_{160\text{-Hz}}$ for 4-kHz tones), and (iii) in a 2/3 octave-band wide noise around f_c ($N_{2/3\text{-Oct.}}$). The noises had a spectrum level (SpecL) equal to 50 and 45 dB, respectively (Table 1). The $N_{2/3\text{-Oct.}}$ noise level was set to resemble relative SpecLs found in speech, resulting in SpecLs of 41 dB at 500 Hz and 25 dB at 4 kHz. It is important to note that this choice of parameters does not allow us to draw conclusions about the masking power of the background noises across conditions. Therefore our focus is on evaluating the power of the auditory model to predict the individual performance on the different conditions.

2.2.3 Participants

Thirty-four participants took part in this experiment. Based on their age and audiometric hearing thresholds, each participant was classified into one of three groups: young normal hearing (yNH), older normal hearing (oNH), or older hearing impaired (oHI). The yNH group had 14 participants (6 females, 8 males) who were between 18 and 29 years old (avg=24.6, std=3.1). The oNH group had 10 participants (9 females, 1 male) who were between 56 and 77 y. (avg=66.1, std=7.7). In both groups, the participants had hearing thresholds better than 20 dB HL for frequencies between 125 Hz and 4 kHz. The oHI group had 10 participants (4 females, 6 males) who were between 61 and 78 y. (avg=70.6, std=6.3), and had hearing thresholds with a high-frequency slope up to 45 dB HL at 4 kHz.

2.3 SIMULATIONS

2.3.1 Protocol

The AM detection task was implemented as previously described, with the exception that a constant stimulus instead of an adaptive procedure was adopted. The stimuli were generated within the AFC toolbox [3], resampled to a 100-kHz rate, and then fed into the auditory model using (i) two cochlear gain profiles that approximate a normal audiogram ('Flat00') and a high-frequency sloping audiogram (for $f > 1$ kHz, 30 dB HL at 8 kHz, 'Slope30'), and (ii) three synaptopathy profiles: 13-3-3, 13-0-0, and 6-0-0. The simulations were run based on 20 stored trials for each condition (Table 1) at four modulation depths m ($-28, -20, -12, -4$ dB). The same trials were used in combination with each simulated hearing profile (Flat00+13-3-3, 13-0-0, 6-0-0; Slope30+13-3-3, 13-0-0, 6-0-0).

2.3.2 Central processor: RMS difference detector

The simulated neural representation of the intervals in the 3-AFC trial were designated as r_1, r_2, and r_3, with interval 1 containing the target sound (AM tone) and intervals 2 and 3 containing the reference sounds. The neural representations r_1, r_2, and r_3 can be compared by computing RMS differences as follows:

$$\Delta r_{12} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(r_1[n]-r_2[n])^2}, \quad \Delta r_{13} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(r_1[n]-r_3[n])^2}, \quad \Delta r_{23} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(r_2[n]-r_3[n])^2} \tag{1}$$

where n indicates that r_1, r_2, and r_3 are digital signals with a total of N samples. The RMS differences were assessed from the corresponding neural representations at each characteristic frequency. Based on our definitions, the larger the difference Δr, the larger the probability that the target sound is correctly detected by a (human) listener. With this rationale, the simulated detection cue (det_{cue}) can be formulated as:

$$det_{cue} = \min\left(20\cdot\log_{10}\Delta r_{12}, 20\cdot\log_{10}\Delta r_{13}\right) - 20\cdot\log_{10}\Delta r_{23} \tag{2}$$

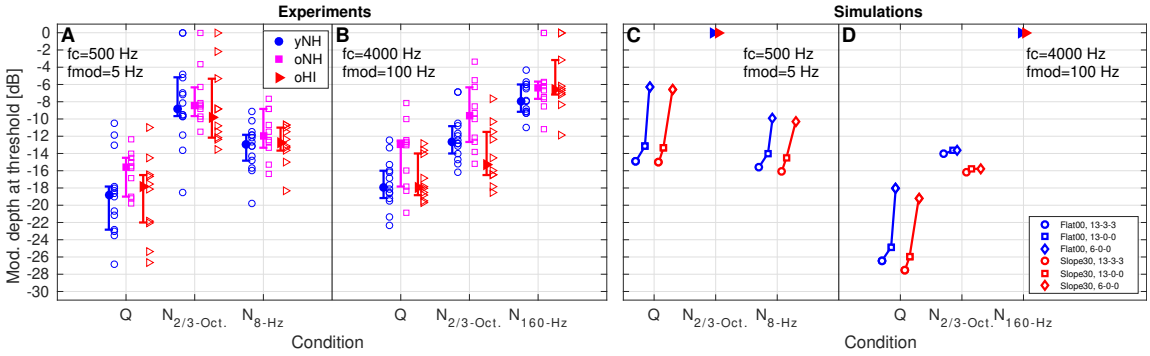

Figure 2. **A,B**: Experimental AM detection thresholds for f_cs at 500 Hz and 4 kHz, respectively. The results are shown as medians (filled markers) with their corresponding interquartile ranges (IQR). The data for the yNH, oNH, and oHI groups are indicated in blue, magenta, and red, respectively. Open markers indicate individual data. **C,D** Simulated AM thresholds for the same tasks, derived as the median detection cues across 20 repetitions for each condition. The simulation results are shown here to facilitate their comparison with the experimental thresholds and are explained more comprehensively in the text.

To determine whether a given det_{cue} value is actually detectable, a threshold $\text{det}_{\text{thres}}$ has to be determined such that a correct (simulated) detection is obtained when $\text{det}_{\text{cue}} > \text{det}_{\text{thres}}$. For this study we chose to use a different $\text{det}_{\text{thres}}$ value for each experiment and we made assumptions to motivate this calibration procedure: (i) the simulated representations should have a (neural) floor level, below which participants are not able to integrate any information, and (ii) $\text{det}_{\text{thres}}$ can be estimated from the experimental data, more specifically from the det_{cue} value derived from the model at a better-than-average experimental performance. This is in line with the classical view that the information available in simulations can be optimally integrated and should therefore compare to the results of best performance participants [see, e.g., 5, Chapter 6]. We used the yNH threshold at the lower limit of the interquartile range (IQR) in the easiest noise condition (-14.8 dB in $\text{N}_{8\text{-Hz}}$, Fig. 2**A**; -14 dB in $\text{N}_{2/3\text{-Oct.}}$, Fig. 2**B**). The floor level used in the model was set to -98 dB re. 1μV and corresponded to the lowest difference value ($\min(\Delta r_{12}, \Delta r_{13})$) obtained across all conditions at $m = -28$ dB, which was found for the Q condition at $f_c = 500$ Hz (not shown).

3 RESULTS

3.1 EXPERIMENTS

The experimental results for the six conditions of the AM detection task are shown in Fig. 2 for carrier frequencies f_c of 500 Hz (panel **A**) and 4 kHz (panel **B**), where higher detection thresholds indicate poorer performance. The filled markers indicate median thresholds and the error bars indicate their corresponding interquartile ranges (IQRs). The open markers indicate individual thresholds averaged (median) across 3 repetitions.

At $f_c = 500$ Hz, $f_{\text{mod}} = 5$ Hz, the Q performance was better than that of the $\text{N}_{8\text{-Hz}}$, and $\text{N}_{2/3\text{-Oct.}}$ conditions. Although the IQRs had similar values in all three groups, the spread of the data (Fig. 2**A**) was larger for the yNH group, smaller for oNH and in between for the oHI group. Furthermore the Q performance was overall 3 dB worse for the oNH group compared to both yNH and oHI, while the overall performance in the other conditions ($\text{N}_{2/3\text{-Oct.}}$, $\text{N}_{8\text{-Hz}}$) was similar across groups.

At $f_c = 4$ kHz, $f_{\text{mod}} = 100$ Hz, the experimental conditions sorted from better to worse performance were: Q, $\text{N}_{2/3\text{-Oct.}}$, and $\text{N}_{160\text{-Hz}}$. The IQRs were consistent within each participant group, with an approximate length of 3 dB for yNH participants, 6 dB for oNH participants (but 2 dB at $\text{N}_{160\text{-Hz}}$), and about 4 dB in $\text{N}_{2/3\text{-Oct.}}$. The individual data points (open markers in Fig. 2**B**) had a similar spread across conditions in the yNH and oHI groups, while it was larger for the oNH group, especially in the Q and $\text{N}_{2/3\text{-Oct.}}$ conditions. In the Q condition, yNH and oHI had a similar overall performance, with a 6-dB poorer performance for oNH participants. The best performance in the $\text{N}_{2/3\text{-Oct.}}$ condition was obtained for the oHI group (threshold at $m = -15.25$ dB) followed by the yNH and oNH groups with thresholds 3 and 6 dB poorer, respectively. A similar overall performance was obtained for all groups in the $\text{N}_{160\text{-Hz}}$ condition.

In the most difficult conditions ($\text{N}_{2/3\text{-Oct.}}$ at $f_c = 500$ Hz, $\text{N}_{160\text{-Hz}}$ at $f_c = 4$ kHz) there were a number of participants who were not able to perform the AM task. Their results are shown as data points at $m = 0$ dB (Fig.2**A**, 2 data points from each participant group; Fig.2**B**, 2 data points from oNH and 2 from oHI).

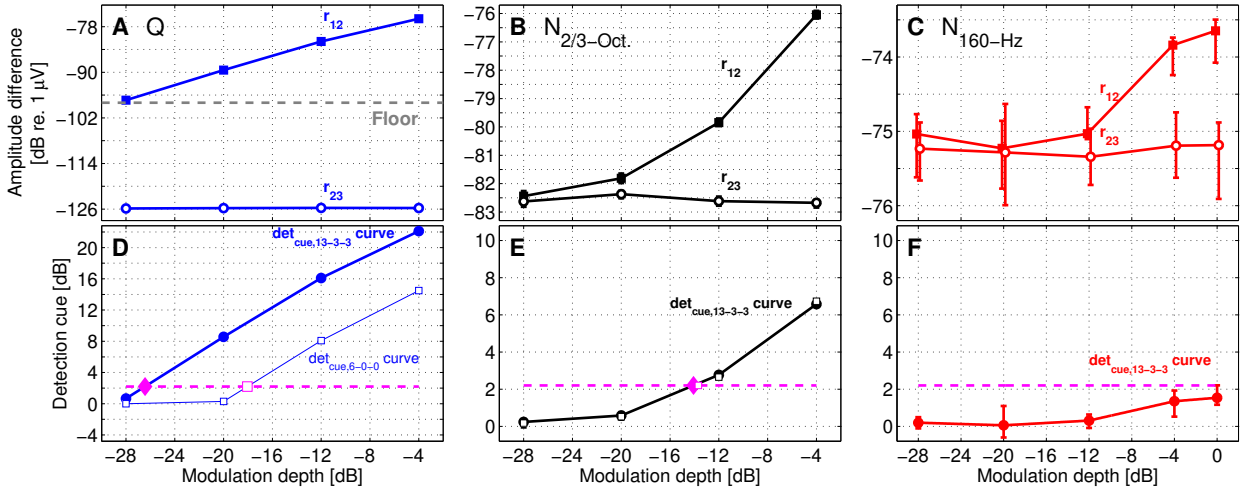

Figure 3. Simulation amplitude differences (top) and detection cues (bottom) for AM detection at $f_c = 4$ kHz for the normal hearing model (Flat00, 13-3-3) for the conditions in quiet (**A**,**D**), $N_{2/3\text{-Oct.}}$ (**B**,**E**), and $N_{160\text{-Hz}}$ (**C**,**F**). Panels **A–C**: Differences Δr_{12} (filled markers) and Δr_{23} (open markers). Panels **D–F**: Detection curves derived using Eq. 1. The simulated thresholds (magenta diamonds) were -26.4, -14.1, and > 0 dB (no estimation), respectively. The $\text{det}_{\text{thres}}$ was set to 2.2 dB (magenta dashed line) and $\text{det}_{\text{floor}}$ was set to -98 dB re. 1μV (grey line in **A**). Each of the curves was derived from 20 difference values and the error bars indicate the IQRs of those values. For the curves in panels **D–F**, the average IQRs across modulation depths are 0 (no variability), 0.4 dB, and 1.1 dB, respectively (clearly visible in panel **F** only). In panels **D–E**, the square markers indicate the detection curves for one synaptopathy profile (Flat00,6-0-0). The 6-0-0 thresholds (open magenta markers) were -18.1, -13.7 (shifts of 8.3 dB and 0.4 dB with respect to Flat00, 13-3-3). Note that the ordinate axes in panels **D–F** are not to scale and that simulations at $m = 0$ dB were added to the $N_{160\text{-Hz}}$ conditions.

3.2 SIMULATIONS

Simulated difference values Δr_{12} and Δr_{23} (Eq. 1) are shown in Fig. 3[2] using a normal hearing model (Flat00, 13-3-3) for the AM task at $f_c = 4$ kHz. The procedure used to derive the simulated thresholds is depicted in Fig. 3, where the simulated detection curves (panels **D–F**) were assessed as the distance between Δr_{23} (or floor level) and Δr_{12} (or Δr_{13}), and det_{cue} was obtained as the amount of detection cue that produces a simulated threshold at $m = -14$ dB (see Section 2.3.2) yielding a det_{cue} value of 2.2 dB. A similar calibration procedure was adopted to derive simulated thresholds in the AM task at $f_c = 500$ Hz and $f_{\text{mod}} = 5$ Hz, obtaining a det_{cue} value of 0.54 dB. The resulting simulated thresholds are shown in Fig. 2**C–D** and Table 2.

In the high-frequency task, the simulated thresholds worsened from Q, to $N_{2/3\text{-Oct.}}$ and $N_{160\text{-Hz}}$, which coincides with the order of difficulty observed experimentally. For the Flat00 13-3-3 profile and fixed det_{cue}, the assessed values were -26.4 dB, -14.1 dB, and > 0 dB (no estimation), respectively. The simulations with cochlear gain loss (Slope30) improved these thresholds by 1.1 and 2.1 dB in the Q and $N_{2/3\text{-Oct.}}$ conditions (see Table 2), respectively. The different synaptopathy profiles increased the Q thresholds considerably by up to 8.3 dB (square markers along the $\text{det}_{\text{cue,6-0-0}}$ curve, Fig. 2**D**) and only marginally in the $N_{2/3\text{-Oct.}}$ condition (square markers in Fig. 2**E**). Although not shown here, a similar shift in the detection curves in the $N_{160\text{-Hz}}$ condition was found (but the cues never exceeded $\text{det}_{\text{thres}} = 2.2$ dB).

In the low-frequency task, the simulated thresholds ranked from better to worse were found for the $N_{8\text{-Hz}}$, Q, and $N_{2/3\text{-Oct.}}$ conditions, respectively. Although this is not exactly the same difficulty order found experimentally (where Q had lower thresholds), the $N_{8\text{-Hz}}$ and Q thresholds differ only by about 1 dB: For normal hearing model (Flat00, 13-3-3) the assessed values were -15.6 dB, -14.9 dB, and > 0 dB (no estimation), respectively. Simulations with the Slope30 profile showed an (almost) marginal improvement of up to 0.5 dB in the $N_{8\text{-Hz}}$ condition (see Table 2). Different synaptopathy profiles increased the simulated thresholds in the Q and $N_{8\text{-Hz}}$ conditions considerably by 8.6 dB and 5.7 dB, respectively (see the shift from square to diamond markers in Fig. 2**C** and also row "Flat00, 13-3-3–6-0-0" in Table 2). A marginal 0.5-dB shift of the detection curves was found in the $N_{2/3\text{-Oct.}}$ condition (not shown here).

[2]The difference r_{13} was omitted for clarity but the obtained values were in a similar range than Δr_{12}.

Table 2. Simulated thresholds $\text{thres}_{\text{sim}}$ using the RMS difference detector for six hearing profiles. The thresholds (also shown in Fig. 2**C-D**) were assessed using the procedures visualised in Fig. 3. Simulated threshold shifts when introducing either cochlear gain loss or synaptopathy are shown in the bottom rows of the table.

	f_c=500 Hz, f_{mod}=5 Hz Threshold $\text{thres}_{\text{sim}}$ [dB]			f_c=4 kHz, f_{mod}=100 Hz Threshold $\text{thres}_{\text{sim}}$ [dB]		
Hearing profile	Q	$N_{2/3\text{-Oct.}}$	$N_{8\text{-Hz}}$	Q	$N_{2/3\text{-Oct.}}$	$N_{160\text{-Hz}}$
Flat00, 13-3-3	-14.9	>0	-15.6	-26.4	-14.1	>0
Flat00, 13-0-0	-13.2	>0	-14.0	-24.9	-13.7	>0
Flat00, 6-0-0	-6.3	>0	-9.9	-18.0	-13.7	>0
Slope30, 13-3-3	-15.0	>0	-16.1	-27.5	-16.2	>0
Slope30, 13-0-0	-13.4	>0	-14.5	-26.0	-15.8	>0
Slope30, 6-0-0	-6.6	>0	-10.3	-19.2	-15.8	>0
	$\text{thres}_{\text{shift}}$ [dB]			$\text{thres}_{\text{shift}}$ [dB]		
Compared profiles	Q	$N_{2/3\text{-Oct.}}$	$N_{8\text{-Hz}}$	Q	$N_{2/3\text{-Oct.}}$	$N_{160\text{-Hz}}$
Slope30-Flat00, 13-3-3	-0.1	–	-0.4	-1.1	-2.1	–
Slope30-Flat00, 13-0-0	-0.2	–	-0.5	-1.1	-2.1	–
Slope30-Flat00, 6-0-0	-0.3	–	-0.4	-1.2	-2.1	–
13-0-0−13-3-3, Flat00	1.7	–	1.6	1.5	0.4	–
6-0-0−13-3-3, Flat00	8.6	–	5.7	8.4	0.4	–
13-0-0−13-3-3, Slope30	1.4	–	1.6	1.5	0.4	–
6-0-0−13-3-3, Slope30	8.4	–	5.8	8.3	0.4	–

Summarising, cochlear gain loss caused smaller threshold shifts than synaptopathy, which introduced simulated threshold shifts between 5.7 and 8.6 dB in three of the six simulated conditions, two of them in the low-frequency task. In all three cases, the shift was caused by an increase in the neural noise floor.

4 DISCUSSION

The supra-threshold AM detection task was conducted for a high- ($f_c = 4$ kHz) and low-frequency carrier ($f_c =$ 500 Hz) where different stimulus characteristics–temporal envelope and temporal fine structure–are believed to influence the (measured or simulated) detection cues in auditory neural representations. Our analysis of both experimental and simulated results was focused on understanding how the participants' results in the high- and low-frequency regions were affected by cochlear gain loss or synaptopathy and, therefore, our analysis could be linked to how temporal fine structure and temporal envelope cues are affected by sensorineural hearing loss.

According to our group classification, oHI participants were expected to suffer from cochlear gain loss while yNH and oNH were not. There is also a low risk that yNH participants suffered from synaptopathy, while oNH and oHI participants might have been at risk for synaptopathy given their advanced age [9].

4.1 EXPERIMENTS

When comparing the results of the yNH and oHI groups, we observed that the AM detection task for a high-frequency carrier ($f_c = 4$ kHz, $f_{mod} = 100$ Hz) is more affected by cochlear gain loss than the low-frequency task ($f_c = 500$ Hz, $f_{mod} = 5$ Hz), given that the best performing oHI participants in Fig. 2**B** had lower thresholds than the yNH participants. This performance improvement is also reflected in one (median) threshold at $m = -15.25$ dB in the $N_{2/3\text{-Oct.}}$ condition, which was 3 and 6 dB below the performance of yNH and oNH participants, respectively. The elevated detection thresholds for oNH participants seem to stem from a reduced high-frequency coding despite their normal hearing audiograms.

In the low-frequency AM detection task, oNH participants had an elevated detection threshold in quiet only (by 3 dB, Fig. 2**A**), while for all other conditions, the three groups had a similar overall performance. Therefore, it seems that cochlear gain loss alone does not play a prominent role for low-frequency AM detection.

4.2 SIMULATIONS

Threshold accuracy: The simulated thresholds (Fig. 2**C-D**) did not fully agree with the obtained experimental thresholds (2**A-B**) with a larger range of values (overall $\text{thres}_{\text{sim}}$ were between -16.4 dB to >0 dB, $\text{thres}_{\text{exp}}$ values were between -18.8 dB and -6.4 dB). At $f_c = 500$ Hz, $f_{mod} = 5$ Hz, the $\text{thres}_{\text{sim}}$ values did not decrease in the Q condition with respect to $N_{8\text{-Hz}}$ as was expected from the experimental data (expected decrease by 4-6 dB) and increased by 0.7 dB instead. The non-estimated $\text{thres}_{\text{sim}}$ in $N_{2/3\text{-Oct.}}$ ($m > 0$ dB) denotes that less detection cues can be extracted using the auditory model in this condition and, therefore, the AM detection is harder than in the $N_{8\text{-Hz}}$ and Q conditions. This tendency was also observed in the experimental data, where 6 participants could not finish the task. At $f_c = 4$ kHz, $f_{mod} = 100$ Hz, the $\text{thres}_{\text{sim}}$ in Q were about 12 dB better

than for $N_{2/3\text{-Oct.}}$ while the experimental thresholds denoted an overall shift of 6 dB. No estimation was possible in the $N_{160\text{-Hz}}$ condition, indicating that in this frequency range this is the most difficult condition. This trend was also observed in the experimental data. The deviations between $thres_{sim}$ and $thres_{exp}$ may be caused by a reduced sensitivity of the model to inherent modulations of the $N_{2/3\text{-Oct.}}$ and $N_{160\text{-Hz}}$ noises for 500-Hz and 4-kHz tones, respectively, when superimposed with the corresponding modulation frequencies of 5 and 100 Hz.

Threshold variability: As summarised in Table 2, cochlear gain loss (Slope30 profile) introduces a small threshold improvement of up to 2.1 dB in the $N_{2/3\text{-Oct.}}$ condition for 4-kHz AM tones. In this condition, the experimental data showed a threshold shift of 3 dB (between yNH and oHI). The simulated threshold shift became (almost) marginal in the 500-Hz AM task where an average improvement of 0.4 dB or less was observed. Simulations for different synaptopathy profiles did show a threshold shift between 5.7 and 8.6 dB for 500-Hz AM-tones (from 13-3-3 to 6-0-0 in the Q and $N_{2/3\text{-Oct.}}$ conditions), while for 4-kHz tones, a significant effect was only found in the Q condition (maximum threshold shift of 8.4 dB). These observations support that cochlear gain loss introduces a smaller detection threshold shift (2 dB in our simulations) than synaptopathy (up to 8.6 dB). The increment of experimental thresholds in Q for 500-Hz (3-dB shift) and 4-kHz tones (6-dB shift) for oNH participants with respect to yNH, could be explained by the simulated threshold shift ($thres_{shift}$) due to synaptopathy, given that in both conditions $thres_{shift}$ was about 8 dB (more than the average experimental shift). Noisy conditions ($N_{8\text{-Hz}}$, $N_{160\text{-Hz}}$, and both $N_{2/3\text{-Oct.}}$) contain representations $r_{1,2,3}$ that are stochastic, introducing variability in the assessed detection cues due to the statistical properties of the background noises. For the AM detection of 4-kHz tones in $N_{160\text{-Hz}}$ (Fig. 3**F**), this external variability would affect the detection sensitivity by $\pm$1.5 dB ("vertical shift" of the detection curve). For the $N_{2/3\text{-Oct.}}$ condition (Fig.3**E**), the external variability (IQR of 0.4 dB) can shift the estimated threshold by 1 dB (i.e., $\pm$0.5 dB). Comparable threshold shifts were assessed for the AM task at 500 Hz (not shown here).

4.2.1 Limitations of the model

Based on the obtained results (Fig. 2**C-D**), it is possible to infer that the auditory model is not sensitive enough in the most difficult AM conditions and that the model is more sensitive to the 4-kHz task at f_{mod}=100 Hz than to the 500-Hz task at f_{mod}= 5 Hz. This is also evidenced by the det_{thres} values of 0.54 dB and 2.2 dB (1.46 dB difference) obtained during the model calibration, where a more balanced sensitivity to low- and high-frequencies should have lead to values that are closer to each other (difference close to zero). This sensitivity variation might be related to the evaluated modulation frequency range. Stages 5 and 6 of the model act as a modulation filter with a passband centred at f_{mod}= 100 Hz, while other models in the literature claim that more modulation filters are needed primarily for frequencies below f_{mod}< 250 Hz [1, 2, 17]. If the auditory model is made more sensitive in a given modulation frequency region, the simulated thresholds in that range may spread more than reported here. In our low-frequency experiment this could lead to a higher effect of synaptopathy than the (up to) 8.6 and 5.8 dB in the Q and $N_{8\text{-Hz}}$ conditions.

4.2.2 Limitations of the central processor

In this paper we adopted a central processor that compares within-trial intervals to derive simulated detection cues. Our detector corresponds to an extension of the cue extraction procedure used in a previous study [16], and is compatible with AM detection tasks with target and reference intervals affected by the stimulus external variability. Despite the moderate success of our simulations, where we could derive simulation thresholds for only four of the six test conditions, a careful interpretation of simulation and experimental data can provide interesting insights about the underlying auditory decision processes. Other approaches to derive detection cues include the use of a memory component (template) within the model, which has been found relevant to account for a number of AM detection tasks [see, e.g., 4].

Our simulation results where synaptopathy played a relevant role rely on the assumption that a neural floor level exists. This assumption is similar to the use of an internal noise to limit the model performance [e.g., 2]. It could be tested whether a more elaborate neural noise model would improve the simulations across conditions.

5 CONCLUSION

In this paper the auditory model described by Verhulst et al. [13] was used to simulate AM detection performance at two carrier and modulation frequencies f_c = 500 Hz modulated at f_{mod}=5 Hz and f_c = 4 kHz modulated at f_{mod}=100 Hz. AM detection thresholds were derived using a within-trial comparison of simulated neural representations (model output of Stage 6, Fig. 1). The simulation results shown in Fig. 2**C–D** reveal that synaptopathy importantly reduces AM performance in 3 of the 6 conditions tested with threshold shifts between 5.7 and 8.6 dB for those conditions. In turn, cochlear gain loss (high-frequency sloping audiogram,

hearing threshold of 30 dB HL at 8 kHz) introduces a slight increment in performance (less than 2.1 dB), in line with the findings published in [16]. We found, however, that our auditory model seems to have a reduced sensitivity in some of the conditions we tested, especially in the low-frequency task and this reduced sensitivity may be related to the adopted modulation rate f_{mod}=5 Hz. If the model is made more sensitive, then the actual variability attributed to cochlear gain loss or synaptopathy may be larger than reported here.

ACKNOWLEDGEMENTS

We would like to thank Sabrina Pieper for collecting the experimental data. This study received support from the European Research Council (ERC) (RobSpear, grant agreement Nr. 678120: AO, SV) and was further supported by the DFG Cluster of Excellence EXC 1077/1 “Hearing4all” (FE).

REFERENCES

[1] L. Carney, T. Li, and J. McDonough. Speech coding in the brain: Representation of vowel formants by midbrain neurons tuned to sound fluctuations. *eNeuro*, 2(4):1–12, 2015.

[2] T. Dau, B. Kollmeier, and A. Kohlrausch. Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J. Acoust. Soc. Am.*, 102(5):2892–2905, 1997.

[3] S. Ewert. AFC - A modular framework for running psychoacoustic experiments and computational perception models. In *Proc. of the International Conference on Acoustics AIA-DAGA*, pages 1326–1329, 2013.

[4] S. Ewert and T. Dau. External and internal limitations in amplitude-modulation processing. *J. Acoust. Soc. Am.*, 116(1):478–490, 2004.

[5] D. Green and J. Swets. *Signal detection theory and psychophysics*. John Wiley & Sons Inc., 1966.

[6] M. Jepsen, S. Ewert, and T. Dau. A computational model of human auditory signal processing and perception. *J. Acoust. Soc. Am.*, 124(1):422–438, 2008.

[7] R. Meddis and M. Hewitt. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *J. Acoust. Soc. Am.*, 89(6):2866–2882, 1991.

[8] P. Nelson and L. Carney. A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J. Acoust. Soc. Am.*, 116(4):2173–2186, 2004.

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

[10] D. Pralong and S. Carlile. The role of individualized headphone calibration for the generation of high fidelity virtual auditory space. *J. Acoust. Soc. Am.*, 100(6):3785–3793, 1996.

[11] A. Saremi, R. Beutelmann, M. Dietz, G. Ashida, J. Kretzberg, and S. Verhulst. A comparative study of seven human cochlear filter models. *J. Acoust. Soc. Am.*, 140(3):1618–1634, 2016.

[12] P. Søndergaard and P. Majdak. The Auditory Modeling Toolbox. In Jens Blauert, editor, *The technology of binaural listening*, chapter 2, pages 33–56. Springer Berlin Heidelberg, 2013.

[13] S. Verhulst, A. Altoè, and V. Vasilkov. Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss. *Hear. Res.*, 360:55–75, 2018.

[14] S. Verhulst, H. Bharadwaj, G. Mehraei, C. Shera, and B. Shinn-Cunningham. Functional modeling of the human auditory brainstem response to broadband stimulation. *J. Acoust. Soc. Am.*, 138(3):1637–1659, 2015.

[15] S. Verhulst, A. Jagadeesh, M. Mauermann, and F. Ernst. Individual differences in auditory brainstem response wave characteristics: Relations to different aspects of peripheral hearing loss. *Trends in Hearing*, 20:1–20, 2016.

[16] S. Verhulst, F. Ernst, M. Garrett, and V. Vasilkov. Supra-threshold psychoacoustics and envelope-following response relations: normal-hearing , synaptopathy and cochlear gain loss. In *Acta Acust. united Ac.* Pages 800–803, 2018.

[17] N. Wallaert, B. Moore, S. Ewert, and C. Lorenzi. Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation. *J. Acoust. Soc. Am.*, 141(2):971–980, 2017.

[18] M. Zilany, I. Bruce, P. Nelson, and L. Carney. A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics. *J. Acoust. Soc. Am.*, 126(5):2390–2412, 2009.