2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Individualized Cochlear Models Based on Distortion Product Otoacoustic Emissions

Sarineh Keshishzadeh and Sarah Verhulst[1]

Abstract— **Auditory models have been adopted for years to simulate characteristics of the human auditory processing for normal and hearing-impaired listeners. However, individual differences due to varying degrees of frequency-dependent hearing damage hinders the simulation of auditory processing on an individualized basis. Here, with a view on precise auditory profiling, recorded distortion product otoacoustic emission (DPOAE) metrics are used to determine individual parameters of cochlear non-linearity to yield individualized human cochlear models, which can be used as pre-processors for hearing-aid and machine-hearing applications. We test whether individualized cochlear models based on DPOAE measurements can simulate the measured DPOAEs and audiograms of normal-hearing and hearing-impaired listeners. Results showed that cochlear models individualized based on DPOAE-grams measured at low stimulus levels or DPOAE thresholds, yield the smallest simulation errors.**

Clinical Relevance—**The outcomes of this study can improve individualized model predictions of auditory function for mixed hearing pathologies, e.g. cochlear synaptopathy in presence of outer-hair-cell loss, and can enhance future individualized hearing-aid algorithms.**

I. INTRODUCTION

Individualized models of hearing-impaired auditory processing have been widely adopted in the design of hearing-aid algorithms. The hitherto developed hearing-aid fitting procedures are largely based on audiometric thresholds or psychoacoustic metrics (e.g. speech intelligibility and loudness perception) [1]–[4], and do not account for the cochlear synaptopathy (CS) aspect of the sensorineural hearing-loss (SNHL). However, determining individual CS parameters is controversial, as diagnostic metrics of CS are presently based on auditory evoked potentials (AEPs). These AEP measures do not necessarily yield a frequency-specific quantification of CS and might be affected by both outer-hair-cell (OHC) and auditory nerve (AN) damages, i.e. sensory and neural hearing-loss, respectively [5]–[8]. Hence, incorporating such individualized cochlear models within the biophysically-inspired auditory model framework that accounts for both OHC-loss and AN-damage aspects of SNHL, will enable us to simulate how either of the aspects affects AEP markers, and conversely will enable to use recorded AEPs to derive frequency-specific personalized SNHL profiles.

Individual cochlear-gain-loss (CGL) parameters can be derived from audiograms measured within the standard audiometric frequencies or at extended high-frequency regions, as well as metrics derived from distortion product otoacoustic emission (DPOAE) recordings, e.g. DPOAE thresholds (DPTHs) and DPOAE-grams. While the audiogram reflects a behavioural response, DPOAE-based features yield objective metrics that are byproduct of cochlear amplification and are informative of OHC-damage [9], [10]. For the first time, in [11], the applicability of audiograms and DPTHs as candidate metrics for developing individualized SNHL profiles were assessed and DPTH-based individualized cochlear models yielded a better accuracy in predicting personalized CS profiles of a cohort consisted of normal-hearing and hearing-impaired young and older listeners. However, an accurate estimation of DPTH at a given frequency, requires DPOAE amplitudes measured over a range of primary levels, which is time-consuming. In addition, achieving an appropriate curve-fit that leads to a robust estimation of the DPTHs is challenging (see Fig. 1). To address these issues, in this study we adopted the computational model of the auditory periphery developed in [18] and evaluated the potential of the DPOAE-gram, as an alternative metric for DPTH to develop individualized cochlear models. A DPOAE-gram provides a frequency-dependent objective measure of OHC-damage within a more reasonable acquisition time than DPTHs. Therefore, we aim to determine which of the individualized cochlear models, based on either DPTHs or DPOAE-grams of different stimulation levels, (1) yields the least DPOAE simulation error, and (2) simulates the behavioural audiograms of the study participants with an acceptable error with regard to the measurement resolution, i.e. 5 dB. Lastly, we explore whether the primary level, at which DPOAE-grams are measured, plays a role in developing accurate individualized cochlear models.

II. METHOD

The adopted database in this study consists of 12 young normal-hearing (yNH: 25.08±1.93 years), 12 older normal-hearing (oNH: 64.58 ±1.88 years) and 11 older hearing-impaired (oHI: 65.27±1.95 years) listeners. Audiograms of the participants were measured at 12 standard audiometric frequencies, between 0.125 and 10 kHz, and the ear with the lower threshold at 4 kHz was chosen for DPOAE measurements. The DPOAE recording paradigms and analysis were performed using a custom-made *MATLAB* program described in [16]. The Ethical Review Board of Oldenburg University approved all experimental procedures involving human subjects in this study. Further

[1]Sarineh Keshishzadeh and Sarah Verhulst are with Faculty of Architecture and Engineering, Information Technology Department, Hearing Technology Group, Ghent University, 9000 Ghent, Belgium. sarineh.keshishzadeh@ugent.be, s.verhulst@ugent.be

details regarding the experiment design can be found in [11].

A. DPOAE threshold extraction

To record DPOAEs, two pure-tones with f_1 and f_2 primary frequencies (f_2/f_1=1.2) were simultaneously presented to the ear through ER-2 earphones coupled to the ER-10B+ microphone system (Etymotic Research). During the stimulation, the primary f_2 frequency was swept continuously [12] with a rate of 2s/octave over a 1/3 octave-range around the determined center frequencies, i.e. 0.8, 1, 2 and 4 kHz [13]. The L_2 primary level varied between 30 and 60 (for yNH and oNH listeners) or 72 (for oHI listeners) dB-SPL, with a 6-dB step, and L_1 primary level was determined based on the *scissors* paradigm, $L_1 = 0.4L_2 + 39$ [14]. For a given primary frequency and level, the distortion product amplitude (L_{DP}) was calculated at $2f_1 - f_2$ frequency. To extract DPTH at a given f_2, a cubic function was fit to L_{DP} values measured at different L_2 levels, and the L_2 level at which the curve reaches to L_{DP} of -25 dB SPL, was estimated as the DPTH [13], [15]. Fig. 1 shows an exemplar DPOAE input-output (I/O) function measured at $f_2 = 0.8$ kHz and illustrates the implemented method for extracting the DPOAE threshold.

B. DPOAE-gram extraction

A DPOAE-gram is characterized by L_{DP}s measured across the f_2 values at fixed levels. For each subject DPOAE-grams were extracted at six primary L_2 levels.

C. Individualization of the cochlear model parameters

The individualization of the cochlear transmission-line (TL) model [18] was implemented by predicting characteristic frequency (CF) dependent CGL parameters, based on the recorded DPOAEs. To simulate CGL caused by OHC-damage, poles of the basilar membrane (BM) admittance function [17], [18] were estimated using recorded DPTHs and DPOAE-grams of six primary levels, independently. Each pole-function comprises of 1001 CF-channels and corresponding pole values (α^* in [18]) range between 0.036 and 0.32, which determine the gain and width of every simulated cochlear filter at low stimulation levels. Pole-function with constant across-CF pole-values of 0.036 or 0.32, accounts for an intact ($flat_{min}$) or completely damaged cochlea ($flat_{max}$: 35 dB-HL), respectively. At each CF, α^* follows a stimulus level-dependent trajectory that accounts for BM compression and wider cochlear filters with level increment [17]. Therefore, considering a constant stimulation level, by increasing α^* without changing the level-dependent pole-trajectory function, the cochlear model output sensitivity will reduce [13], [18]. Accordingly, to set CF-dependent pole-values, a neural network (NN) was trained by 26 random pole-functions and corresponding DPTHs or DPOAE-gram simulations (see Fig. 2a) [11]. To derive DPTH and DPOAE-gram from model simulations (sDPTH and sDP-gram), we

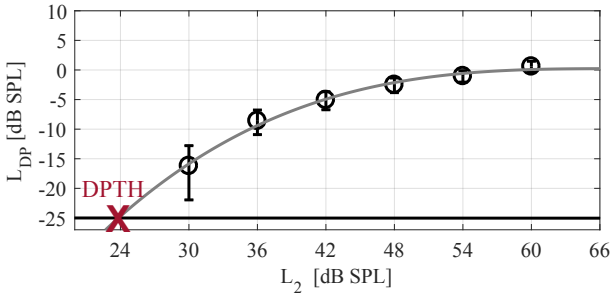

Fig. 1. Measured distortion product otoacoustic emission (DPOAE) input-output function of a young normal-hearing (yNH) subject at $f_2 = 0.8kHz$. The cubic curve-fit is shown in grey and the intersection with $L_{DP} = -25$ dB SPL is specified with a red cross, which indicates the DPOAE-threshold (DPTH) at 0.8 kHz.

followed a similar approach as explained in Sections II-A and II-B, respectively. However, we chose an L_{DP} of -10 dB SPL to estimate the DPTHs from model simulations, since applying the L_{DP}=-25 dB SPL to the simulated DPOAE I/O functions, yielded inconclusive sDPTHs, in particular for pole-values associated with greater CGLs.

The simulated sDP-grams at six primary levels and sDPTHs of the 26 random pole-functions were used to train seven different NNs with identical architectures that invert the model (Fig. 2a). In this way, trained NNs can be adopted to predict the individualized pole-functions of the experimental cohort (Fig. 2b), given their measured DPOAE-grams or DPTHs (mDPTHs or mDP-grams).

The sDP-grams of a $flat_{min}$ profile overestimated the mDP-grams of the yNH listeners. In addition, unlike the mDPTH etimation, we chose an L_{DP} of -10 dB-SPL (instead of -25 dB SPL) to estimate sDPTHs. Therefore, first, mDP-grams and mDPTHs were mapped to the range of associated sDP-grams and sDPTHs of the $flat_{min}$ and $flat_{max}$ profiles ($\text{mDP-gram}_{\text{map}}$ and $\text{mDPTH}_{\text{map}}$). Then, respective Z-scores

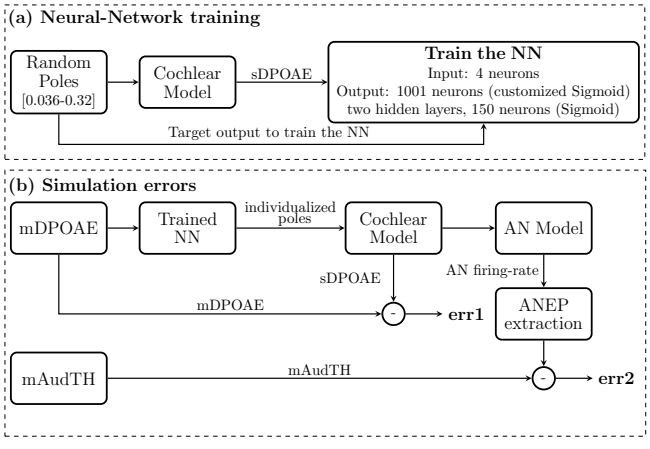

Fig. 2. Block-diagram of the implemented method. (a) The applied input and output data to train the neural-network. (b) The steps of cochlear transmission-line (TL) model individualization and quantification of the distortion product otoacoustic emission and audiogram simulation errors (*err1* and *err2*).

were fed into the trained NNs and predicted individualized pole-functions were used to simulate the individualized DPOAE metrics and evaluate the simulation errors (Fig. 2b).

D. DPTH and DPOAE-gram simulation errors

To quantify the DPTH and DPOAE-gram simulation errors, first mDP-gram$_{map}$/mDPTH$_{map}$ and sDP-gram/sDPTH were referenced to the corresponding DPOAE metrics simulated for a normal-hearing profile (sDP_{NH}) as follows:

$$mDP_{ref} = (-1)^n (sDP_{NH} - mDP_{map}) \quad (1)$$

$$sDP_{ref} = (-1)^n (sDP_{NH} - sDP) \quad (2)$$

In (1) mDP_{map} refers to measured DPOAEs which were mapped to the range of $flat_{min}$ and $flat_{max}$ profiles and in (2), *sDP* stands for simulated DPOAE metrics. If *DP* stands for DPTHs, n equals to one, and otherwise $n = 0$. In this way, the type-1 error was defined as follows.

$$err1(f_2) = |mDP_{ref}(f_2) - sDP_{ref}(f_2)| \quad (3)$$

In this regard, Fig. 3a compares the *err1* mean of DPTH simulations across the frequency and experiment groups. Type-1 errors of DPOAE-gram simulations for different stimulation levels are shown in the first row of Fig. 4 and each panel represents the respective averaged *err*1 of each group (a: yNH, b: oNH and c: oHI).

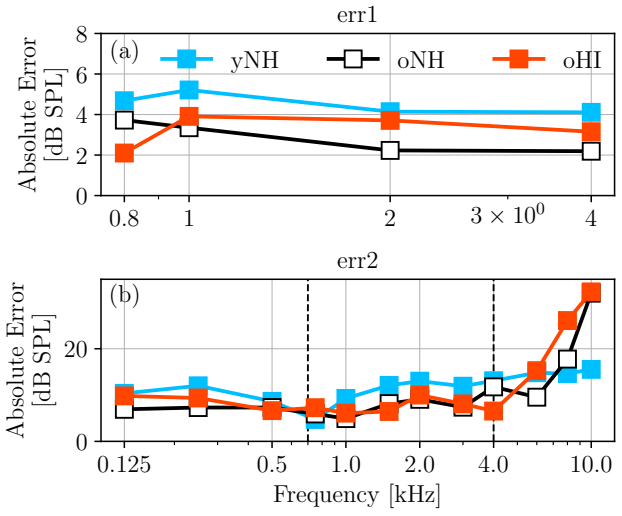

Fig. 3. Averaged (a) type-1 and (b) type-2 errors of the distortion product otoacoustic emission threshold (DPTH) based cochlear model individualization across the young normal-hearing (yNH), older NH (oNH) and older hearing-impaired (oHI) listeners. Two vertical dashed lines in panel (b), specify the frequency region of the audiogram which matches with the primary frequencies of the recorded DPTHs.

E. Audiogram simulation error

In section II-D, the *err*1 was used to evaluate how well trained NNs with sDPTH or sDP-gram simulate the mDPTH or mDP-gram. It is also worthwhile to assess whether NNs trained using DPOAE metrics, can be used to simulate the measured audiograms. This would offer an additional and independent evaluation technique to assess the efficiency of the proposed DPOAE-based cochlear model individualization method. For this purpose, the audiometric thresholds of experiment participants were simulated by determining the stimulation intensity that was required to minimize the energy difference between the AN excitation-pattern (ANEP) of a considered participant and the ANEP of a NH-profile [11]. To quantify the audiogram simulation error, the type-2 error (*err*2) was defined as the absolute difference between simulated and measured audiometric thresholds ($sAudTH_{DP}$ and $mAudTH_{DP}$) of the study participants.

$$err2(f_{aud}) = |sAudTH_{DP}(f_{aud}) - mAudTH_{DP}(f_{aud})| \quad (4)$$

In (4), f_{aud} refers to the frequency points at which audiometric thresholds were measured. The group-mean of type-2 errors are shown in Fig. 3b and the second row of Fig. 4.

III. Results

Comparing the type-1 errors of the DPTHs with those of DPOAE-grams (Fig. 3a vs. Fig. 4a to 4c), shows that cochlear models individualized by DPOAE-grams measured with low primary levels and DPTHs performed equally well. In each method, the *err1* values were averaged over the frequency (err1$_{avg}$). Pooling the err1$_{avg}$ of all groups together, a one-way ANOVA test confirmed the significant effect of the stimulation level on the DPOAE-gram simulation errors ($F_{(5,204)}$=9.95, p<0.000). Moreover, paired-sample t-test with Bonferroni correction showed that the only significant difference between the err1$_{avg}$ values of the DPTH and DPOAE-grams was existent for L_2=60 dB SPL ($t_{(34)}$=-3.71, p<0.000).

Considering the audiogram simulation errors, elevated high-frequency *err2* values were observed, especially for oNH and oHI groups in both DPTH and DPOAE-gram based methods. This high-frequency error increment relates to the model limitation in simulating CGLs higher than 35 dB HL, since the maximal amount of the cochlear mechanical filter gain is 35 dB [18]. Consistent with the type-1 error, we observed lower type-2 errors when DPOAE-grams were recorded to lower primary levels. Given that DPOAEs were measured at four frequency points (between 0.75 and 4 kHz), only audiometric thresholds between 0.75 and 4 kHz frequency region were considered for further analysis of audiogram simulation errors. This frequency region is specified by vertical dashed lines in Fig. 3b and the second row of Fig. 4. The *err2* values within the mentioned frequency region were averaged (err2$_{avg}$) and all subjects pooled together, a one-way ANOVA test showed a significant

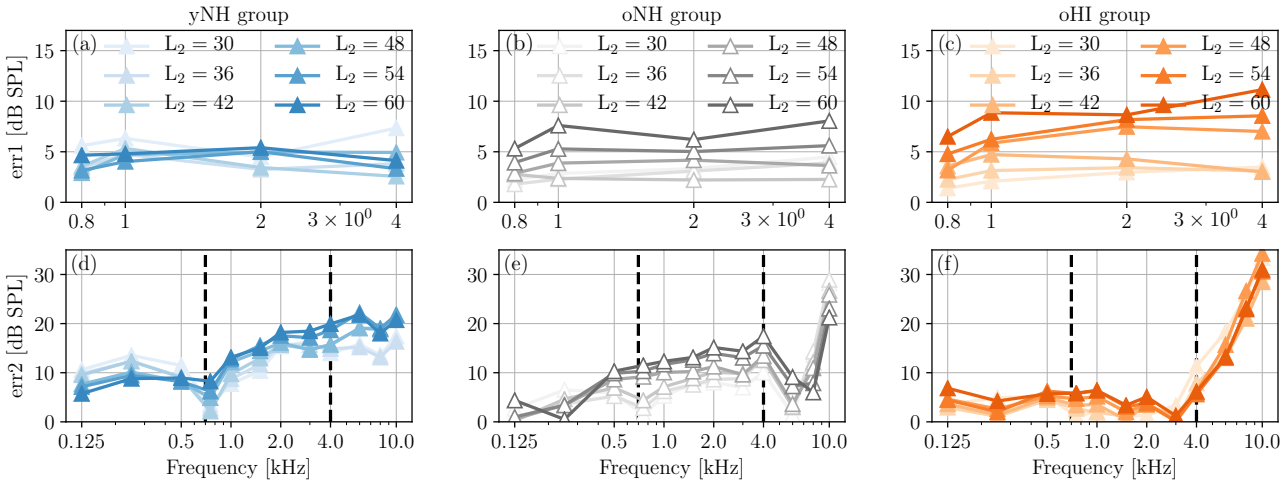

Fig. 4. Averaged type-1 and type-2 errors of (a,d) young normal-hearing (yNH), (b,e) older NH (oNH) and (c,f) older hearing-impaired (oHI) listeners for distortion product otoacoustic emission (DPOAE) gram based individualized cochlear models. In each panel, simulation errors based on DPOAE-grams measured at L_2 levels between 30 (light colors) and 60 (dark colors) dB SPL are presented. Two vertical dashed lines in panel d, e and f, specify the frequency region of the audiogram which matches with the primary frequencies of the recorded DPOAE-grams.

effect of the stimulation level on err2_{avg} of the DPOAE-gram ($F_{(5,204)}$=3.14, p=0.009). A paired-sample t-test with Bonferroni correction showed that the err2_{avg} of DPTH was significantly different from that of the DPOAE-gram measured at all levels, except at L_2=30 dB SPL ($t_{(34)}$=-1.43, p=0.16). This finding is consistent with that of the type-1 error, and indicates that individualized cochlear models based on either DPOAE-grams measured at low stimulus levels or DPTHs not only yielded the lowest *err1* value, but also they simulated individual audiograms with a smaller *err2* value than DPOAE-grams of higher stimulus levels.

Lastly, to make an overall comparison between type-1 and type-2 errors of each method, err1_{avg} and err2_{avg} values were averaged across the study groups ($\text{err1}_{\text{avg,groups}}$ and $\text{err2}_{\text{avg,groups}}$). According to Fig. 5, the $\text{err2}_{\text{avg,groups}}$ values are always ≈5 dB higher than the $\text{err1}_{\text{avg,groups}}$, and that an accurate simulation of DPOAE measurements, results in a more precise simulation of an individual audiogram.

IV. Discussion and Conclusions

Our analysis showed that DPOAE-grams measured at low primary levels, as well as the DPTHs, yield the lowest type-1 and type-2 errors. To understand the reason, the following two aspects were considered: (i) At low stimulation levels, DPOAE amplitudes are maximally impacted by OHC amplification and corresponding losses. Therefore, near-threshold stimulation levels may provide more reliable DPOAE measurements [14]. This may partly explain the relatively high *err1* and *err2* values for some yNH subjects even at low stimulation levels, since their hearing thresholds were below 20 dB HL at all frequencies. At the same time, the smaller simulation errors of the HI listeners can be explained by their higher audiometric thresholds within the frequency range of [0.75-4] kHz, which are closer to the low stimulation levels (Fig. 3b and 4f). (ii) The range of active pole values in the adopted cochlear TL model was constrained based on the stimulation levels below 35 dB SPL, since low stimulation intensities lead to sharper cochlear filters [18]. Consequently, using DPOAE-grams of higher stimulation levels to set the CF-dependent α^* of the cochlear model, might have resulted in the higher simulation errors.

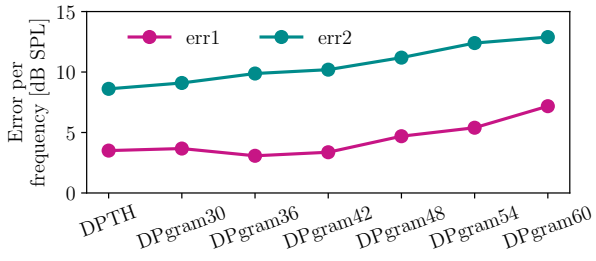

Fig. 5. Comparison between the type-1 (purple) and type-2 (green) errors of different cochlear individualization methods, which are averaged across the frequency and study groups.

In all implemented methods, type-2 errors were greater than the type-1 (Fig. 5). The first reason for the observed difference can be related to the fact that cochlear models, individualized by an objective measure of OHC-loss (DPOAE-based methods), were used to simulate audiogram, which is a behaviourally measured metric. Secondly, AN excitation patterns in response to pure-tones were used to simulate individualized audiometric thresholds, which are affected by the contribution of a large number of off-CF channels. Hence, compared to DPOAE-based metrics (i.e. more localized response), yielded less accurate simulations of the audiometric thresholds.

Taken together, our results suggest that DPOAE-grams recorded at stimulation levels below 36 dB SPL perform as well as DPTH-based methods for developing individualized cochlear models with an approximate error of 5-

dB per frequency. Additionally, the developed method can simulate individual audiograms at frequencies involved in DPOAE measurements with an average error of less than 10 dB/frequency. Considering the audiogram measurement resolution, i.e. 5 dB, the generated prediction error is acceptable and in this way, an additional independent test supports the efficiency of the proposed method. Despite these promising results, we suggest that conducting DPOAE measurements at more primary frequencies in the future may lower the simulation errors and improve the functionality of personalized cochlear model simulations.

ACKNOWLEDGMENT

This work was supported by European Research Council under the Horizon 2020 Research and Innovation Program, grant agreement no. 678120 RobSpear. Authors would also like to thank Markus Garrett and Manfred Mauermann for their contribution to the data collection.

REFERENCES

[1] H. Dillon, "Nal-nl1: A new procedure for fitting non-linear hearing aids," *The Hearing Journal*, vol. 52, no. 4, pp. 10–12, 1999.

[2] D. Byrne, H. Dillon, T. Ching, R. Katsch, and G. Keidser, "Nal-nl1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures.," *Journal of the American academy of audiology*, vol. 12, no. 1, 2001.

[3] L. E. Cornelisse, R. C. Seewald, and D. G. Jamieson, "The input/output formula: A theoretical approach to the fitting of personal amplification devices," *The Journal of the Acoustical Society of America*, vol. 97, no. 3, pp. 1854–1864, 1995.

[4] J. E. Marriage, D. A. Vickers, T. Baer, B. R. Glasberg, and B. C. Moore, "Comparison of different hearing aid prescriptions for children," *Ear and hearing*, vol. 39, no. 1, pp. 20–31, 2018.

[5] S. G. Kujawa and M. C. Liberman, "Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss," *Journal of Neuroscience*, vol. 29, no. 45, pp. 14077–14085, 2009.

[6] A. Parthasarathy and S. G. Kujawa, "Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing," *Journal of Neuroscience*, vol. 38, no. 32, pp. 7108–7119, 2018.

[13] 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*, vol. 20, 2016.

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

[8] V. Vasilkov, M. Garrett, M. Mauermann, and S. Verhulst, "Enhancing the sensitivity of the envelope-following response for cochlear synaptopathy screening in humans: The role of stimulus envelope," *Hearing Research*, vol. 400, p. 108132, 2021.

[9] S. A. Gaskill and A. M. Brown, "The behavior of the acoustic distortion product, 2 f 1- f 2, from the human ear and its relation to auditory sensitivity," *The Journal of the Acoustical Society of America*, vol. 88, no. 2, pp. 821–839, 1990.

[10] F. Harris, "Distortion-product otoacoustic emissions in humans with high frequency sensorineural hearing loss," *Journal of Speech, Language, and Hearing Research*, vol. 33, no. 3, pp. 594–600, 1990.

[11] S. Keshishzadeh, M. Garrett, and S. Verhulst, "Towards personalized auditory models: Predicting individual sensorineural hearing-loss profiles from recorded human auditory physiology," *Trends in Hearing*, vol. 25, 2021.

[12] G. R. Long, C. L. Talmadge, and J. Lee, "Measuring distortion product otoacoustic emissions using continuously sweeping primaries," *The Journal of the Acoustical Society of America*, vol. 124, no. 3, pp. 1613–1626, 2008.

[14] P. Kummer, T. Janssen, and W. Arnold, "The level and growth behavior of the 2 f1- f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," *The Journal of the Acoustical Society of America*, vol. 103, no. 6, pp. 3431–3444, 1998.

[15] P. Boege and T. Janssen, "Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission i/o-functions in normal and cochlear hearing loss ears," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1810–1818, 2002.

[16] M. Mauermann, "Improving the usability of the distortion product otoacoustic emisssions (dpoae)-sweep method: An alternative artifact rejection and noise-floor estimation," in *Proceedings of Meetings on Acoustics ICA2013*, vol. 19, ASA, 2013.

[17] S. Verhulst, T. Dau, and C. A. Shera, "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission," *The Journal of the Acoustical Society of America*, vol. 132, no. 6, pp. 3842–3848, 2012.

[18] S. Verhulst, A. Altoè, and V. Vasilkov, "Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss," *Hearing research*, vol. 360, pp. 55–75, 2018.