

Assessment, modeling, and compensation of inner and outer hair cell damage

STEFFEN KORTLANG* AND STEPHAN D. EWERT

Medizinische Physik, Universität Oldenburg and Cluster of Excellence 'Hearing4all', Oldenburg, Germany

Reduced temporal fine structure (TFS) sensitivity is proposed to accompany cochlear hearing loss even if audibility and loudness perception are compensated for by hearing aids, or can be present in elderly listeners with unremarkable audiometric thresholds. In both cases, inner hair cell (IHC) damage or neuronal degeneration of subsequent stages can be assumed to play a role. To investigate psychoacoustic measures for assessment of IHC loss, random frequency modulation (FM) detection thresholds in quiet and in background noise were collected for six young normal-hearing (NH) listeners, six older NH listeners, and eleven HI listeners. Two possible detection mechanisms based on phase-locking and amplitude modulation (AM) were assessed in a probabilistic, 'spiking' auditory model [Meddis, *J Acoust Soc Am* 119, 406 (2006)]. IHC and outer hair cell (OHC) damage were incorporated and adapted to predict the psychoacoustic data. The resulting hearing-impaired (HI) model was then used to simulate the auditory nerve (AN) response in aided conditions with an improved model-based dynamic compression algorithm [based on Ewert and Grimm, *ISAAR*, 393 (2011)]. Comparison to simulated normal-hearing AN responses revealed partial compensation of OHC damage while IHC damage resulted in supra-threshold 'internal noise' which might contribute to the limited benefit from compensation strategies in hearing aids.

INTRODUCTION

Even if audibility and loudness perception are restored by dynamic compression strategies in hearing aids, supra-threshold processing deficits may persist. Recently, it has been shown that sound exposure can lead to a permanent impairment of auditory-nerve (AN) fibers with low spontaneous rate (LSR) in the absence of elevated audiometric thresholds (Kujawa and Libermann, 2009). Such a degeneration of AN fibers or losses of synaptic elements in the inner hair cells (IHC) might reduce the redundancy of neural coding (Henry and Heinz, 2012), acting as a source of 'internal noise' in the signal representation. Particularly, the usability of temporal fine structure (TFS) information in the signal might be reduced as consequence of IHC damage. TFS sensitivity was shown to decline with hearing loss and age (e.g., Hopkins and Moore, 2011).

As a measure of TFS sensitivity, low-rate frequency modulation (FM) detection thresholds are proposed here and assessed in three different subject groups. FM

*Corresponding author: steffen.kortlang@uni-oldenburg.de

detection requires accurate coding of temporal and spectral cues. However, FM detection tasks are assumed to involve less central or complex stages such as higher-level language processes which are involved in, e.g., speech perception tasks. FM detection therefore appears suited as a measure of early (peripheral) damage in auditory perception. Here it is hypothesized that auditory deafferentiation results in an undersampled representation of the signal at the level of the AN (see also Lopez-Poveda and Barrios, 2013), whereas a loss of outer hair cells (OHC) primarily results in a loss of compression and broader auditory-filter bandwidths. The effects of these two independent impairments are simulated in an FM detection model. By reducing the amount of AN fibers in the model, the neural coding fidelity is diminished ('internal noise'), while filter broadening as a consequence of OHC damage increases the effect of 'external noise' in conditions with noise maskers. The goal of this study is to better understand and to predict perceptual consequences of IHC and OHC damage with regard to FM detection that may also contribute to poor speech-in-noise performance.

RANDOM FREQUENCY MODULATION DETECTION THRESHOLDS

Method

Random frequency modulation detection thresholds (RFMDTs) were examined in six normal-hearing young listeners (NH-Y), six normal-hearing older listeners (NH-O), and eleven hearing-impaired listeners (HI) with sloping sensorineural hearing loss. In an adaptive three-interval, three-alternative, forced-choice (3-AFC) procedure, the FM interval had to be detected against a pure-tone reference at 500 Hz, 2 kHz, and 6 kHz. The random frequency modulation (RFM) tones were generated by imposing a bandpass noise (1-4 Hz) as instantaneous frequency deviation to the pure tone's frequency (f_c). The frequency modulation depth was expressed as root-mean-square (RMS) deviation of the instantaneous frequency from f_c . To reduce amplitude modulation (AM) based detection cues, an additional 1-4-Hz bandpass-noise AM was applied with an RMS modulation depth of -12 dB. After a training run, four threshold runs were performed. To ensure a comfortable level and comparable loudness among all subjects, NH listeners were measured at 65 dB HL, while signals were presented to the HI group at the level of medium loudness (L_{25}) obtained from categorical loudness scaling (CLS; Brand and Hohmann, 2002). To test the impact of external noise on FM detection, Gaussian white noise with 5-ERB bandwidth centred around f_c was added at signal-to-noise ratios (SNRs) of +3 dB and -3 dB in two further conditions. The signals were 500 ms in duration including 25-ms Hann ramps. Further details are provided in Ewert *et al.* (2013).

Results and discussion

Average RFMDTs for the different conditions, test frequencies, and listener groups are shown in Fig. 1. Empirical results are represented by the bars of different shades of grey indicating +/- one standard deviation. Detection thresholds increase with additional external noise and significantly differ between listener groups. In line with literature, RFMDTs at medium loudness were elevated for the NH-O (grey) and HI (dark grey) group (e.g., Strelcyk *et al.*, 2009). However, it is apparent that the differences between

the listener groups are less systematic at 6 kHz, where a larger spread per group is observed. This aspect is particularly interesting given that, in this frequency region, audiometric thresholds, loudness growth (characterized by CLS), and filter bandwidths (as estimated in an additional notched noise measurement) differed most between the NH-Y and HI group. Thus, OHC-related processing deficits may play a smaller role and RFMDTs appear to be a meaningful choice to examine IHC-related impairment. In particular, in the absence of elevated audiometric thresholds, altered loudness growth, or increased filter bandwidth for the NH-O listeners, results indicate independent supra-threshold processing disorders reducing the temporal coding fidelity. A three-way repeated-measures analysis of variance (ANOVA) with factors subject group, centre frequency, and condition confirmed that all factors were significant ($p < 0.001$).

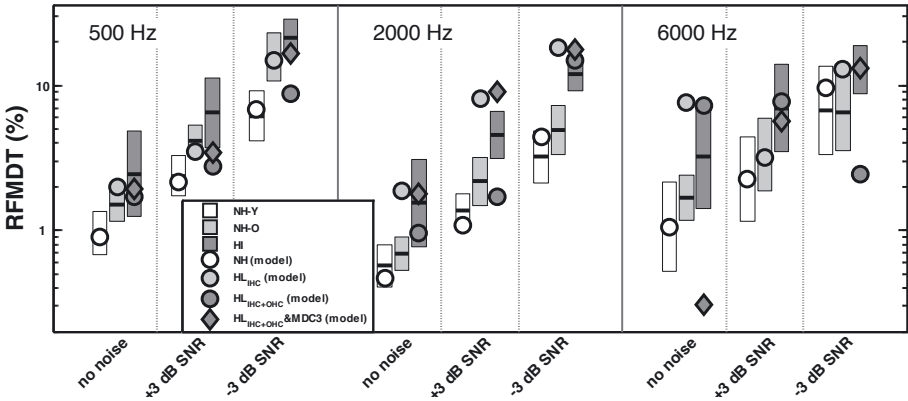


Fig. 1: RFMDTs as a function of test frequency and noise condition. Within each listener group, the thick horizontal line indicates the geometric mean. Model results are indicated by the different symbols (see legend).

MODELING

The RFMDTs presented above can be assumed to rely on two independent detection cues: At small modulation and carrier frequencies, distinctness from a pure sinusoid is thought to rely on gentle fluctuations in the timing of neural spikes (Strelcyk *et al.*, 2009), i.e., TFS cues. At higher rates, FM detection is thought to be based primarily on a place mechanism due to FM-induced AM (‘FM-to-AM conversion’). Both cues were considered here at the level of the AN, including timing (spike phaselocking, TFS) and level (spike density, AM) information.

Model structure

The auditory model (MAP) by Meddis (2006; MAP1_14g release) was used to simulate AN responses. In the model, stimuli are filtered with a linear bandpass to model the

outer and middle ear and then subjected to the dual-resonance-non-linear (DRNL) filter (Meddis *et al.*, 2001) accounting for non-linear basilar membrane (BM) processing. The DRNL output is then converted to membrane potentials in the IHC stage, to generate either an AN spiking probability or probabilistic AN spike responses including refractory effects and adaptation. To keep interpretation as simple as possible, feedback paths available in the model (acoustic and medial olivocochlear reflex) were deactivated here. As an example of the MAP model output, Fig. 2 shows the mean AN spiking probability of a high-spontaneous-rate (HSR) fiber as a function of time and characteristic frequency (CF) for the three intervals of the 3-AFC measurement (2 kHz). Here, the target was in the middle interval containing an FM of 5%-RMS modulation depth. It is apparent that the FM introduces additional AM to the AN activity. For illustration purposes, the pattern at the 2-kHz channel is marked with the grey plane.

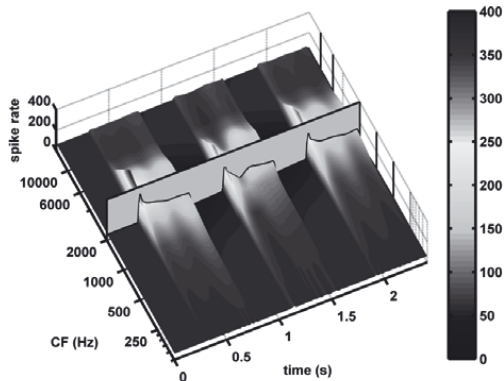


Fig. 2: Mean spike rate (spike probability output of MAP) for a HSR fiber in response to three intervals of a 3-AFC measurement.

In the further modeling, the ‘spiking’ mode was used to generate spike trains at BM characteristic frequencies of 125 Hz up to 16 kHz in octave steps, including the intermediate frequencies (187.5 Hz, 375 Hz, 750 Hz, ...). As illustrated in Fig. 3, the FM target and unmodulated reference signal were passed through the model. Spike trains of 100 AN fibers were simulated for each auditory channel. As the more sensitive HSR fibers outnumber the LSR fibers by about 4 to 1 (e.g., Schnupp, 2011), they were separated into 80 HSR and 20 LSR fibers. The spike patterns were cut to match the steady-state part of the signal (450 ms excluding 25 ms on- and offset ramps). For the two above-mentioned detection mechanisms, two independent paths were modeled: i) To account for the AM cue (upper pathway in Fig. 3), post-stimulus time histograms (PSTH) were formed by summing the output of the 100 fibers. These PSTHs form a similar pattern to the spiking probability used in Fig. 2, but underly stochastic fluctuations due to the spiking process. A 4th-order zero-phase bandpass filter (1-8 Hz) was applied to the PSTHs to extract low-rate AM information. Finally,

for each BM channel, the AM cue was calculated as the variance of the bandpass-filtered PSTH (PSTH-VAR) and transformed to log domain. The FM tone usually shows higher variance than the pure tone. ii) To calculate the TFS cue (lower pathway in Fig. 3), first-order inter-spike-interval (ISI) histograms were calculated to examine the distribution of the observed times between spikes merged over all AN fibers. Phase-locking to f_c is represented by the local maxima in the distribution separated by one period ($1/f_c$) of the signal. The strength of phase-locking was quantified by the vector strength of the ISI histogram (ISI-VS) to f_c . Here, the FM tone usually shows lower values of ISI-VS caused by smearing of the maxima.

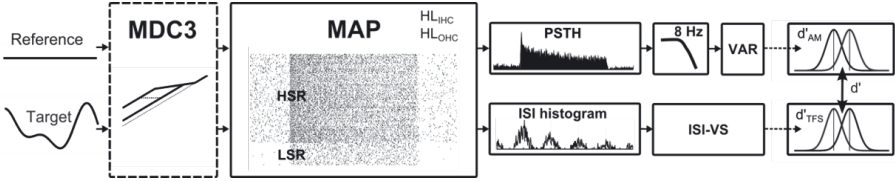


Fig. 3: Block diagram of the model to predict RFMDTs including AM and TFS information extraction in the upper and lower pathway, respectively.

The calculations were repeated N times to estimate mean and standard deviation for PSTH-VAR and ISI-VS, from which two detectability measures d'_{AM} and d'_{TFS} (Cohen's d) were calculated as RMS over the individual d 's in each BM channel. The final detectability measure d' was calculated by combination of d'_{AM} and d'_{TFS} using a weighting factor α :

$$d' = \sqrt{\alpha \cdot d'_{AM}{}^2 + (1 - \alpha) \cdot d'_{TFS}{}^2} \quad (\text{Eq. 1})$$

Method

Different model versions were used to predict the RFMDT data. The NH model (representing group NH-Y) used the standard settings proposed by Meddis (2006) despite the above-mentioned changes with 100 AN fibers for each BM channel. Similar to Lopez-Poveda and Barrios (2013), IHC loss (HL_{IHC}) was modeled as a reduction of AN fibers from 100 to 10 (reducing the spike rate by a factor of 10, equivalent to 20 dB linear attenuation). This reduction of AN fibers results in decreased d' values (equivalent to internal noise reducing the acuity of spike pattern analysis). The stimulus level for the two models was 65 dB HL as in the experiment. The HI group was modeled as a combination of IHC and OHC loss ($HL_{IHC+OHC}$). Here, a reduction of gain (gain loss, GL) in the non-linear pathway of the DRNL stage (see, e.g., Jepsen and Dau, 2011) was additionally introduced. GL was estimated for all subjects of the HI group from audiometric thresholds (HL) and the lower slope of the loudness function (m_{low}) derived from the CLS measurement, as suggested in Ewert and Grimm (2011). Finally, GL was averaged across all HI subjects resulting in values ranging from, e.g., 18 dB at 500 Hz, 30 dB at 2 kHz, and 32 dB at 6 kHz. For the HI group simulations, the mean stimulus

level of all HI listeners in the FM detection measurement was used (77, 78, and 79 dB HL for 500 Hz, 2 kHz, and 6 kHz, respectively).

RFMDTs were simulated by estimating d' values for $N = 100$ repetitions for all experimental conditions and 20 different RMS FM depths, separated equally on log space between 0.2 and 36%. Third-order polynomial functions were fitted to the d' functions in a least-squares sense. For all model versions, a single d' threshold per frequency was selected, so that the NH model accounted best for the data of the NH-Y group. A weighting factor of $\alpha = 0.27$ was chosen.

Comparison of predicted thresholds and data

Predicted RFMDTs are shown as the black symbols in Fig. 1. It is apparent that the NH model (white circles) can reproduce the main trends of the NH-Y data. At high frequencies, the HL_{IHC} model (light grey circles) tends to overestimate the performance of the HI_{IHC} group, showing the need for frequency-dependant IHC loss estimates. Additional OHC damage (dark grey circles) did mainly influence the results at the higher two frequencies and in the presence of external noise. For the noise conditions, broadening of auditory filters as a consequence of OHC damage (gain loss) in the model should reduce the salience of the TFS cue as more (external) noise energy falls into the filter. However, this TFS cue is dominant at low frequencies where the applied gain loss was low and the overall effect of OHC damage was thus low in the model simulations. In some cases, OHC damage led to lower RFMDTs in the model simulations, contradicting data. This might be related to changes in the model's AN pattern for the higher signal levels applied the HI group simulations.

MODEL-BASED DYNAMIC COMPRESSION

The effect of dynamic compression on RFMDTs was assessed using the physiologically-motivated model-based dynamic compression algorithm (MDC3, Fig. 4) which is based on the algorithm of Ewert and Grimm (2011) and Ewert *et al.* (2013). The input signal was analysed in a 4th-order Gammatone filterbank (30 bands, one ERB wide). In each filter channel, the instantaneous level (L_{inst}), the instantaneous frequency (F_{inst}), and a smoothed (50-ms 1st-order lowpass) broad-band, 'long-term' level (L_{lt}), estimated over five adjacent frequency bands, were calculated. A model for BM compression for NH and HI subjects was computed in real-time, based on a combination of L_{lt} and L_{inst} . Off-frequency component suppression was realized using the F_{inst} estimate. The difference between the modeled NH and individual HI BM-I/O function was applied as gain per frequency band. The output signal of the algorithm was generated by a 2nd-order Gammatone resynthesis filterbank including delay compensation between the channels (Hohmann, 2002). In comparison to Ewert *et al.* (2013), the main new stage was the 'HI broadband compensation' (see Fig. 4) prior to resynthesis. It estimates the intensities in the normal and impaired system, taking into account filter widening. A level correction is applied based on the effect of widened filters, resulting in a slightly reduced gain for broadband input signals.

The compressor was fitted for the $HL_{IHC+OHC}$ model using the same GL estimates as described above. Reference and target signals with an input level of 65 dB HL were processed. Model results for simulated aided RFMDTs are shown in Fig. 1 (dark grey diamonds). In some of the external-noise conditions dynamic compression led to higher RFMDTs, most likely due to reduction of AM cues. Overall the results are not clear-cut, however, it is obvious that dynamic compression is not suited to significantly improve thresholds of the $HL_{IHC+OHC}$ model compared to the unaided condition (dark grey circles).

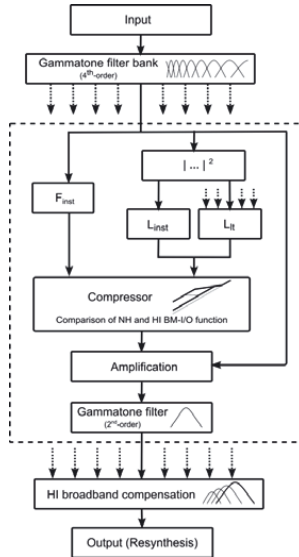


Fig. 4: Block diagram of the model-based dynamic compression algorithm MDC3. Details are described in the text.

GENERAL DISCUSSION

The proposed model for the simulation of RFMDTs is able to reproduce NH data within good accuracy. Neither the model of HL_{OHC} , nor dynamic compression (as can be found in most hearing aids) showed consistent effects in the model simulations. Reduced TFS sensitivity was empirically found in both HI and elderly NH listeners. This could be accounted for by modeled IHC damage which resulted in supra-threshold ‘internal noise’. The effect of increased external noise on the AN representation as a consequence of broadened filters in case of OHC damage could not be mimicked by the model. A possible confound comes from ‘unrealistic’ changes in the model’s AN representation depending on the signal level. It is obvious that the salience of both cues available in the model, TFS and AM, could not be improved by dynamic compression, which might explain limited benefit of hearing aids when audibility is not the sole problem. Taken together, a first promising step towards a

framework for (aided) performance prediction based on stochastic AN responses with individually adjustable IHC and OHC loss was suggested. While the model is able to mimic supra-threshold processing deficits in HI listeners in the TFS and AM domain, further work is required to, e.g., determine frequency-dependent IHC-loss estimates and to assess the effect of signal level on the model's AN representation.

ACKNOWLEDGMENTS

This work was supported by the German Federal Ministry of Education and Research BMBF, research unit 'Modellbasierte Hörgeräte' (13EZ1127D), and the German research funding organization DFG (SFB TRR 31).

REFERENCES

- Brand, T., and Hohmann, V. (2002). "An adaptive procedure for categorical loudness scaling," *J. Acoust. Soc. Am.*, **112**, 1597-1604.
- Ewert, S.D., and Grimm, G. (2011). "Model-based hearing aid gain prescription rule" in Proceedings of ISAAR 2011: *Speech Perception and Auditory Disorders*. Nyborg, Denmark, pp. 393-400.
- Ewert, S.D., Kortlang, S., and Hohmann, V. (2013). "A Model-based hearing aid: Psychoacoustics, models and algorithms," *ASA, ICA 2013 Montréal*, **19**, 1.
- Henry, K., and Heinz, M.G. (2012). "Diminished temporal coding with sensorineural hearing loss emerges in background noise," *Nat. Neurosci.*, **15**, 1362-1364.
- Hohmann, V. (2002). "Frequency analysis and synthesis using a Gammatone filterbank," *Acta Acustica*, **88**, 433-442.
- Hopkins, K., and Moore, B.C.J. (2011). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech perception in noise," *J. Acoust. Soc. Am.*, **130**, 334-349.
- Jepsen, M., and Dau, T. (2011). "Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.*, **129**, 262-281.
- Kujawa, S.G., and Liberman, M.C. (2009). "Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss," *J. Neurosci.*, **29**, 14077-14085.
- Kortlang, S., Mauermann, M., Kollmeier, B., and Ewert, S.D. (2012). "Characterization of IHC loss and its relevance to hearing aid gain prescription rules," Poster at IHCON, Lake Tahoe, USA.
- Lopez-Poveda, E.A., and Barrios, P. (2013). "Perception of stochastically undersampled sound waveforms: A model of auditory deafferentiation," *Front. Neurosci.*, **7**, 124.
- Meddis, R. (2001). "A computational algorithm for computing auditory frequency selectivity," *J. Acoust. Soc. Am.*, **109**, 2852-2861.
- Meddis, R. (2006). "Auditory-nerve first-spike latency and auditory absolute threshold: a computer model," *J. Acoust. Soc. Am.*, **119**, 406-417.
- Schnupp, J., Nelken, I., and King, A. (2011). *Auditory Neuroscience – Making Sense of Sound*, MIT Press.
- Strelyck, O., and Dau, T. (2009). "Relations between frequency selectivity, temporal fine-structure processing and speech reception in impaired hearing," *J. Acoust. Soc. Am.*, **125**, 3328-3345.