Analysis of a forward masking paradigm proposed to estimate cochlear compression using an auditory nerve model and signal detection theory

JENS THUREN LINDAHL, GERARD ENCINA-LLAMAS* AND BASTIAN EPP
Hearing Systems section, Department of Health Technology, Technical University of Denmark, DK-2800 Lyngby, Denmark

The healthy human auditory system has a large dynamic range. An "active mechanism", presumably due to the electromotility of the outer hair cells in the cochlea, leads to level-dependent amplification of basilar membrane (BM) vibration and a compressive BM input/output function. Different methods for estimating this compressive function based on behavioural forward masking have been suggested. These methods are based on the assumption that BM processing can be isolated from the response of the overall system and that the forward masking onto the probe is different for on- and off-frequency maskers. In the present study, a computational model of the auditory nerve (AN) in combination with methods from signal detection theory was used to test these assumptions. The simulated AN response was quantified in terms of rate and synchrony for different AN fibre types. Contribution of different tonotopic regions to overall sensitivity to the stimuli were analysed. The results show that on- and off-frequency maskers produce similar forward masking onto the probe. The simulation results suggest that the estimate of compression based on the behavioural experiment cannot be derived from sensitivity at the level of the AN but requires additional contributions, consistent with physiological studies.

INTRODUCTION

The healthy mammalian auditory system is remarkable in its ability to handle a large dynamic range of incoming sounds while having high sensitivity to low-intensity sounds. An underlying "active mechanism", presumably due to the electromotility of the outer hair cells (OHC) in the cochlea, leads to level-dependent amplification of basilar membrane (BM) vibrations and a compressive input/output (I/O) function over a narrow BM region (Ruggero *et al.*, 1997). Damage of OHCs leads to a hearing impairment (i.e., elevation of hearing threshold). Hence, compression is a proxy for the state of the OHCs and therefore for the health of the inner ear. No direct measurement of their state is currently possible in humans. Various methods have been developed to estimate BM compressive function based on behavioural forward masking paradigms (e.g., Nelson *et al.*, 2001). Important assumptions underlying these behavioural methods are, among others, that contributions of

*Corresponding author: encina@dtu.dk

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.

cochlear processing can be isolated from the behavioural response (the overall system), that sensitivity is dominated by the information at the tonotopic place corresponding to the probe frequency and that tonal maskers presented at on- and off-frequencies (near or away of the characteristic place of the tonal probe, respectively) produce different forward masking effect onto the probe (Jepsen and Dau, 2011). In the present study, a computational model of the auditory nerve (AN) in combination with methods from signal detection theory (SDT) was used to test the assumption that the "BM I/O function" derived using behavioural measures reflects the compressive growth of the BM at the characteristic place (on-frequency) corresponding to the probe frequency. The following assumptions were made: 1) Perceptual outcome measures are formed using the information present at the level of the AN in the form of spike rate and synchrony. 2) The information needed to resolve the task behaviourally can be encoded anywhere in the AN. 3) All three types of AN fibres (low, medium and high spontaneous rate, SR) can contribute to the perceptual outcome measure. 4) The information extracted in the behavioural outcome must be encoded at the level of the AN, as this is a mandatory stage.

METHODS

The experimental paradigm used in a behavioural study (Jepsen and Dau, 2011) was used to simulate the response of the AN. The AN model of Zilany *et al.* (2014) was used, implemented as described by Encina-Llamas *et al.* (2019).

Experimental paradigm

The stimulus was a tonal masker with a duration of 200 ms (5-ms rise/fall cosine ramps), a gap of varying length and a probe tone with a duration of 20 ms (10-ms rise/fall Hann window ramps with no steady portion) with a frequency of either 1000 or 4000 Hz. After computing the model threshold for each frequency, the probe level was fixed at 10 dB sensation level (SL). Thresholds in the AN model were calculated for each characteristic frequency (CF) comparing driven rate with SR using a non-parametric permutation test for equality of the means with a significance level of $\alpha = 1\%$. The masker was presented either at the probe frequency (on-frequency) or at a frequency 0.6 times that of the probe (off-frequency). The masker was presented at levels from 10 to 110 dB SPL in steps of 10 dB. The gap was set to 2, 5 and 10 to 100 ms in steps of 10 ms. First a block with the masker only was presented to the model followed by a block with the same masker plus the probe (see Fig. 1 top).

Two listeners from Jepsen and Dau (2011) were simulated: a normal-hearing (NH) listener represented by the averaged NH audiogram; and one mildly hearing-impaired (HI, listener HI01) with thresholds of 20 and 45 dB hearing level (HL) at 1 and 4 kHz, respectively. Hearing impairment, assuming a combination of 2/3 of OHC and 1/3 of inner hair cell (IHC) dysfunction, was implemented in the model by adjusting the parameters that control OHC gain (*cohc*) and IHC transduction (*cihc*).

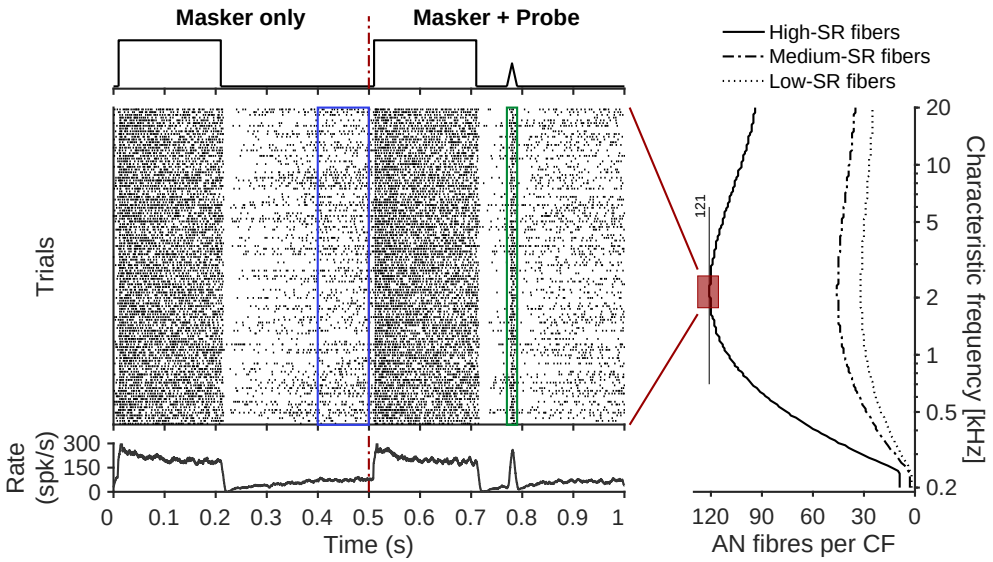

Fig. 1: Top: Stimulus paradigm. The masker was presented at the probe frequency (on-frequency) or at 0.6 times the probe frequency (off-frequency). The probe was presented following the second masker with a temporal gap of varying length. Middle: Trials of simulated raster plots for one CF. The number of simulated AN fibres per CF of each type differed across CFs (right panel). The rectangles indicate the analysis windows for the estimation of the spontaneous activity (blue) and the probe response (green). Bottom: Resulting post-stimulus time histogram (PSTH) derived from the response of all simulated nerve fibres at this CF; showing the onset response, adaptation, the reduction in rate following masker offset and the response to the probe.

Simulation and analysis using signal detection theory

The AN response was simulated for neurons with CFs from 200 Hz to 20 kHz, discretised into 200 CFs (cochlear segments) equally spaced along a logarithmic axis. The receptor potential of a single IHC was simulated for each CF. The number and type of nerve fibres at each CF followed the distribution used by Encina-Llamas *et al.* (2019), who considered a total number of 32000 AN fibres, based on physiological data from human temporal bones. At each CF and for each independently computed AN fibre, the resulting spike trains were stored and used to compute spike rate r_L and spike synchrony $\rho(f)$. These metrics were computed over two different time windows in the simulation: one centred on spontaneous activity (blue rectangle in Fig. 1 middle) and the other centred on the probe response (green rectangle in Fig. 1 middle). Spike rate (Eq. 1) was calculated as the number of spikes N_{spikes} normalised by the length of the analysis window L. Spike synchrony ρ (Eq. 2) was calculated as the vector sum of spike times projected onto a phasor rotating on the unit circle in the complex plane at the probe frequency f.

$$r_L = \frac{N_{\text{spikes}}}{L} \qquad \text{(Eq. 1)} \qquad \rho(f) = \left| \frac{1}{N} \sum_{n=1}^{N} e^{i2\pi f t_n} \right| \qquad \text{(Eq. 2)}$$

For each CF and fibre type, distributions of rate and synchrony were obtained for each window of analysis. The resulting distributions were used to calculate the sensitivity index d' using their mean and variance values, as shown in Eq. 3 (Jones, 2016).

$$d' = \frac{\mu_{\text{probe}} - \mu_{\text{SR}}}{\sqrt{\frac{1}{2}(\sigma^2_{\text{probe}} + \sigma^2_{\text{SR}})}} \qquad \text{(Eq. 3)} \quad d'_{\text{comb}} = \left[\sum_{m=1}^{2} \sum_{k=1}^{\text{CF}} \sum_{l=1}^{3} {}^{\text{SR}_l}_{\text{CF}_k} d'^{\,2}_{\text{m}} \right]^{\frac{1}{2}} \qquad \text{(Eq. 4)}$$

All AN fibres were treated as independent sources of information. Information in the form of d' was combined across fibre type, across CF and across metric (rate and synchrony). When combining across CF, three different integration bandwidths were considered: 1) a narrow range of CFs equivalent to 1/6-octave band, (d'_{narrow}); 2) a bandwidth centred at the probe CF to cover the region where direct measurements of BM motion suggest a compressive growth (Ruggero *et al.*, 1997) (lower bound 1/2-octave below CF of the probe and upper bound 1/3-octave above the CF equal to the probe frequency, $\left(d'_{\text{compr}}\right)$); and 3) a range of 2-octaves of simulated CFs centred at the probe frequency $\left(d'_{\text{wide}}\right)$. The final combined sensitivity d'_{comb} for each group of parameters was calculated as shown in Eq. 4, with $\text{SR} = 1, 2, 3$ denoting the AN fibre type (i.e., high-, medium- and low-SR fibres, respectively; assuming the distribution based on the physiology of the cat), CF_k denoting the CF index, and $\text{m} = 1, 2$ denoting the two derived metrics (rate and synchrony, respectively).

RESULTS

Figure 2 shows heatmaps of rate (panel A) and sensitivity (panel B) evaluated in the analysis window of the probe (green vertical rectangle in Fig. 1) as function of CF and masker level for the NH simulations, and the combined sensitivity d'_{comb} as a function of masker level (panel C) for a fixed gap length of 30 ms. For simplicity, only high-SR fibres were considered in this example.

At low masker levels, the rates were highest in a narrow CF region near on-frequency (within the band of 1/6-octave shown by the green dashed lines) and decreased with increasing masker level. The region of reduced rate after masker offset increased with increasing masker level. The off-frequency masker clearly showed an asymmetric spread of reduced rate towards higher CFs. Sensitivity also reduced with increasing masker level. The combined sensitivity strongly reduced over a range of masker levels of about 10-15 dB. Such drops in combined sensitivity occurred at low masker levels for the on-frequency masker and at much higher levels for the off-frequency masker. The combined sensitivity did not depend on the bandwidths of CF integration.

Panels A and D in Fig. 3 show sensitivity as a function of gap length for a fixed masker level of 60 dB SPL for the NH and HI listeners, respectively. Panels B and E show simulated temporal masking curves (TMC). Panels C and F show the corresponding

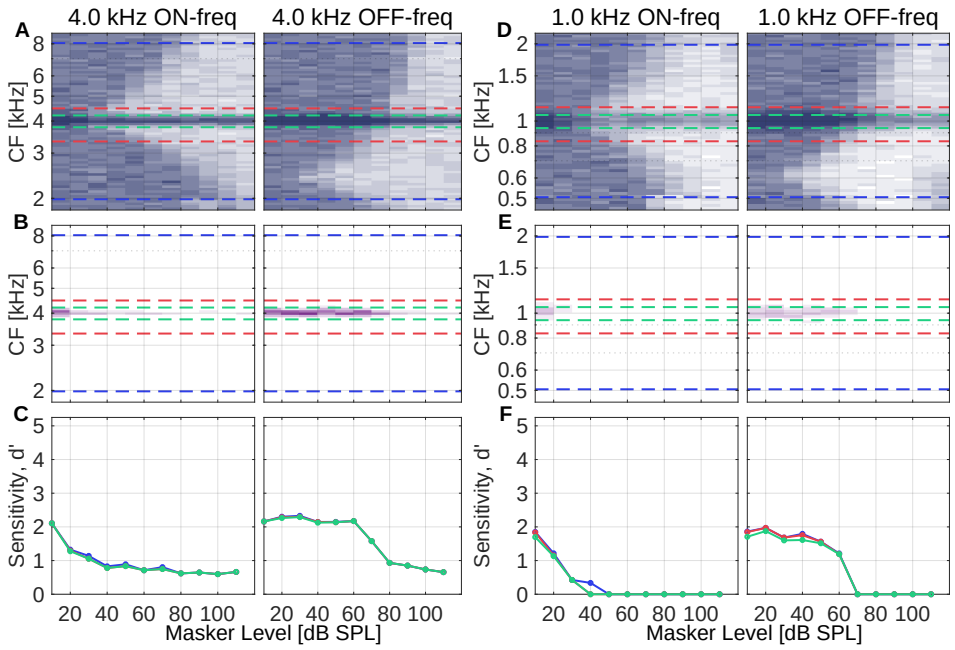

Fig. 2: Simulation results for the NH listener as a function of masker level for a fixed gap length of 30 ms. Panels A-C show results for a probe frequency of 4 kHz and panels D-F for a probe frequency of 1 kHz. Panels A and D and B and E show heatmaps of rate and sensitivity, respectively, as a function of CF and masker level. Panels C and F show combined sensitivity for different CF integration bands (green: d'_{narrow}, red: d'_{compr} and blue: d'_{wide}). Results show simulations considering only high-SR fibres.

derived "BM I/O" functions for the NH and the HI listener, respectively. The simulated "BM I/O" functions did not match the experimental data. In both cases, close to linear growth was predicted, in contrast to the behavioural data(black circles), because the TMCs for the on- and off-frequency maskers (panels B and E in Fig. 3) grew close to parallel. This finding was independent of the bandwidth of integration used to calculate the combined sensitivity. For the HI listener, the on- and off-frequency simulated TMCs (panel E) were shifted towards higher masker levels, also growing close to parallel, and more closely spaced. The broadening of the AN excitation for the HI listener (panel B) led to different results in the calculated combined sensitivity d'_{comb} for the different bandwidths of integration, resulting in higher values with increasing bandwidth of integration.

DISCUSSION

The simulation results show the reduction of AN response on SR activity across CFs due to the presence of the maskers (see panels A and D in Fig. 2). In agreement with direct BM measures (Ruggero *et al.*, 1997), the spread of such reduction is asymmetric towards higher CFs. For high masker levels, the response of high-SR fibres saturated

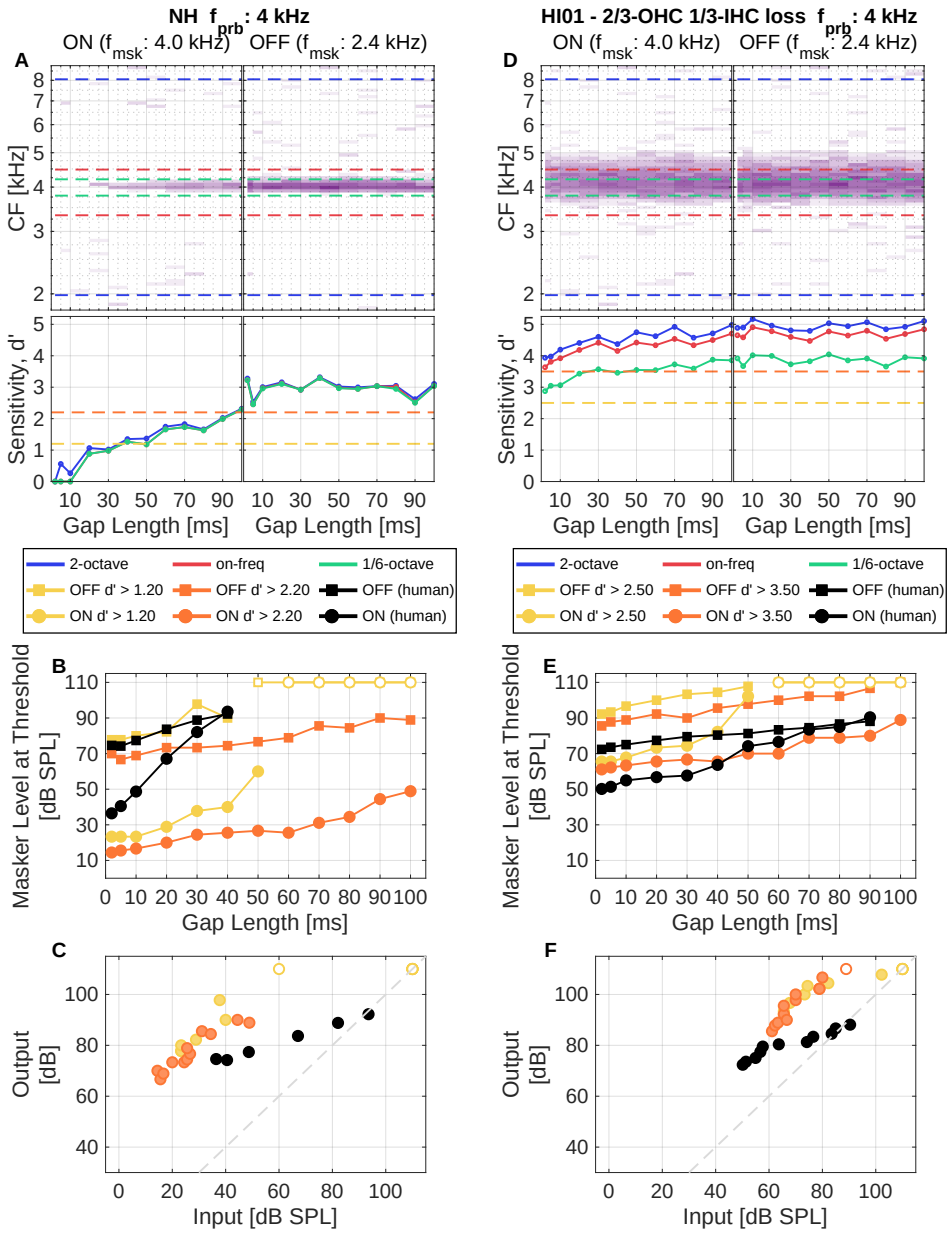

Fig. 3: Panels A and D show combined sensitivity as a function of gap length for NH and HI01, respectively, using a masker level of 60 dB SPL and a probe frequency of 4 kHz, with on- and off-frequency maskers. Dashed lines indicate integration bandwidths. Panels B and E show corresponding simulated TMC curves for two sensitivity criteria. Panels C and F show corresponding "BM I/O function" derived from the simulated TMC curves. In panels B, C, E and F, open symbols show ceiling responses using the maximum masker level of 110 dB SPL, and black solid lines and symbols show human results from Jepsen and Dau (2011), for comparison. The grey-dashed line indicates linear growth.

and hence the derived rate and sensitivity showed a plateau. The saturated rate of high-SR fibres limits the masking effect, because increasing the physical level of the masker does not lead to a change in the neuronal representation of the masker for this type of AN fibre. For the NH simulation, the results for the d'_{comb} did not show an effect of the bandwidth of integration (see panels C and F in Fig. 2), due to the narrow on-CF excitation produced by the low intensity probe. For the HI listener, the d'_{comb} grew with increasing integration bandwidth (see panel D in Fig. 3), consistent with the effect on the AN tuning curves (i.e., less sharply tuned tips) after reduced local (on-frequency) gain due to OHC damage (Liberman and Dodds, 1984). The intensity of the probe must be then increased to maintain excitation at 10 dB SL, leading to a larger populations of AN neurons encoding probe information and potentially contributing to detection. Even for the mild hearing impairment, the probe representation exceeded the definition of on-frequency bandwidth, d'_{compr}. In all cases, d'_{comb} was dominated by the high-SR fibres, as the probe was presented at 10 dB SL. The inclusion of synchrony had little effect on the results, probably due to the short duration of the probe.

The simulated TMCs did not match the behaviourally measured TMCs of Jepsen and Dau (2011). In contrast to the behavioural data, the simulated on- and off-frequency masking curves grew close to parallel, leading to almost linear growth in the derived "BM I/O functions". This implies that the effect of the tonal masker on the neuronal activity post-masker (and therefore on the probe) is essentially the same regardless of the frequency of the masker (on- vs off-frequency). Direct physiological recordings in AN neurons demonstrated that this is true particularly for the case of high-SR fibres (Yates *et al.*, 1990). Because the probe is kept at a very low stimulus level (10 dB SL), the encoding of the probe is dominated by high-SR fibres, which are not affected by the BM nonlinearity, and therefore cannot reflect the compressive growth. The AN model of Zilany *et al.* (2014) may not fully capture the effect of forward masking at the level of the AN, as this has been improved in a newer version of the model. Nevertheless, the model simulations seemed to generally agree with physiological evidence indicating that there is not sufficient forward masking at the level of the AN to account for the behavioural data, and that additional processing, which may include inhibitory mechanisms at the level of the inferior colliculus (IC), are needed to account for the behavioural responses (Nelson *et al.*, 2009). Hence, cochlear compression estimates measured using psychoacoustics might be strongly influenced by processing at these higher stages, compromising their power to infer only BM responses.

CONCLUSIONS

A model of the AN combined with SDT methods based on the three types of AN fibres, multiple CFs and rate and synchrony cues could not explain the behaviourally derived "BM I/O functions". Encoding of the probe is dominated by high-SR fibres due to the use of low intensity probes. BM nonlinearities are not reflected in the high-SR fibres (Yates *et al.*, 1990), which are limited by their saturation at high intensities and their SR at lower stimulus levels. This suggests that behaviourally derived "BM I/O

functions" with the current stimulation paradigm do not reflect the compressive growth of the BM, even though this occurs previous to the AN. Furthermore, physiological data suggest that contributions from higher neural stages are needed to account for the behavioural data (Nelson *et al.*, 2009). And even more, in the case of mild threshold elevation due to OHC damage, CFs outside the narrow frequency range where the healthy BM velocity was found to grow compressively with input level carry probe information, and may contribute to the perceptual outcome.

ACKNOWLEDGEMENTS

This work was funded by the Novo Nordisk Foundation grant NNF17OC0027872.

REFERENCES

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. Otolaryngol., **20** (4),363–382. doi: 10.1007/s10162-019-00721-7.

Jepsen, M. L., and Dau, T. **(2011)**. "Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss", J. Acoust. Soc. Am., **129** (1),262–281. doi: 10.1121/1.3518768.

Jones, P. R. **(2016)**. "A tutorial on cue combination and Signal Detection Theory: Using changes in sensitivity to evaluate how observers integrate sensory information", J. Math. Psychol., **73**,117–139. doi: 10.1016/j.jmp.2016.04.006.

Liberman, M. C., and Dodds, L. W. **(1984)** "Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves", Hear. Res., **16**(1),55–74. doi: 10.1016/0378-5955(84)90025-X.

Nelson,D. A., Schroder, A. C., and Wojtczak, M. **(2001)** "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners", J. Acoust. Soc. Am., **110**(4),2045–2064, 2001. doi: 10.1121/1.1404439.

Nelson, P. C., Smith, Z. M., and Young, E. D. **(2009)** "Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking", J. Neurosci., **29**(8),2553–2562. doi: 10.1523/JNEUROSCI.5359-08.2009.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. **(1997)** "Basilar-membrane responses to tones at the base of the chinchilla cochlea", J. Acoust. Soc. Am., **101**(4),2151–2163. doi: 10.1121/1.418265.

Yates, G. K., Winter, I. M., and Robertson, D. **(1990)** "Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range", Hear. Res., **45**(3),203 – 219. doi: 10.1016/0378-5955(90)90121-5.

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. **(2014)** "Updated parameters and expanded simulation options for a model of the auditory periphery", J. Acoust. Soc. Am., **135**(1),283–286. doi: 10.1121/1.4837815.