

Comparison of Behavioral and Physiological Measures of the Status of the Cochlear Nonlinearity

Trends in Hearing
Volume 25: 1–11
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165211016155
journals.sagepub.com/home/tia



Michał Fereczkowski^{1,2,3,4} , Torsten Dau¹, and Ewen N. MacDonald¹

Abstract

While an audiogram is a useful method of characterizing hearing loss, it has been suggested that including a complementary, suprathreshold measure, for example, a measure of the status of the cochlear active mechanism, could lead to improved diagnostics and improved hearing-aid fitting in individual listeners. While several behavioral and physiological methods have been proposed to measure the cochlear-nonlinearity characteristics, evidence of a good correspondence between them is lacking, at least in the case of hearing-impaired listeners. If this lack of correspondence is due to, for example, limited reliability of one of such measures, it might be a reason for limited evidence of the benefit of measuring peripheral compression. The aim of this study was to investigate the relation between measures of the peripheral-nonlinearity status estimated using two psychoacoustical methods (based on the notched-noise and temporal-masking curve methods) and otoacoustic emissions, on a large sample of hearing-impaired listeners. While the relation between the estimates from the notched-noise and the otoacoustic emissions experiments was found to be stronger than predicted by the audiogram alone, the relations between the two measures and the temporal-masking based measure did not show the same pattern, that is, the variance shared by any of the two measures with the temporal-masking curve-based measure was also shared with the audiogram.

Keywords

peripheral compression, notched-noise test, otoacoustic emissions, temporal-masking curve test

Received 27 September 2020; Revised 19 April 2021; accepted 20 April 2021

Currently, the main method used to characterize hearing loss is the audiogram, reflecting sensitivity to pure tones. While useful, it is not sufficient to predict suprathreshold perception and performance of individual hearing-impaired (HI) listeners. In other words, two individuals with similar audiograms can differ widely on perceptual and hearing-aid-outcome measures. Therefore, there is a need for additional suprathreshold measures to better characterize HI listeners.

The active mechanism in the cochlea depends on the outer hair cells' (OHCs) operation (e.g., Ruggero & Rich, 1991), and it is the main source of the nonlinear response of the basilar membrane (BM) to tonal stimuli, that is, the compressive input–output characteristic (BM I/O function). Simple models of BM I/O employ a broken-stick nonlinearity comprising a linear low-level region, a compressive mid-level region, and a linear

high-level region (e.g., Lopez-Poveda & Johannesen, 2012; Plack et al., 2004). The two most important parameters of this model are the compression threshold, or knee-point (KP), between the low-level and mid-level regions, and the slope of the mid-level region (i.e., the

¹Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Lyngby, Denmark

²Faculty of Health Sciences, Institute of Clinical Research, University of Southern Denmark, Odense, Denmark

³Research Unit for ORL – Head & Neck Surgery and Audiology, Odense University Hospital, Odense, Denmark

⁴Institute of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark

Corresponding Author:

Michał Fereczkowski, Technical University of Denmark, Bygning 352, Ørstedes Plads Kongens, Lyngby 2800, Denmark.
Email: mfer@dtu.dk



compression exponent [CE] in dB/dB units). Importantly, Lopez-Poveda and Johannesen (2012) underlined that this model of the cochlear nonlinearity is functional in nature and does not attempt to tap into physiological state and structural damage of the OHCs. Behavioral experiments show that while the KP estimates are highly correlated with the audiometric threshold (Lopez-Poveda & Johannesen, 2012), CE estimates are not, at least for listeners with mild-to-moderate hearing loss (e.g., Jepsen & Dau, 2011; Lopez-Poveda & Johannesen, 2012; Plack et al., 2004 or Johannesen et al., 2016).

Several researchers have suggested that individual estimates of cochlear compression, such as CE, are likely to improve predictions of suprathreshold performance and hearing-aid fitting procedures (Johannesen & Lopez-Poveda, 2008; Mills et al., 2007; Mueller & Janssen, 2004). However, studies along these lines have reported mixed findings (Johannesen et al., 2016; Kortlang et al., 2016; Lopez-Poveda et al., 2017). For instance, while behavioral CE estimates, obtained from temporal-masking curves (TMCs, Nelson et al., 2001), were found to be important predictors of speech intelligibility in speech-shaped noise, hearing thresholds did not predict the performance in this case. However, the same estimates could not predict performance in time-reversed two-talker babble, nor subjective hearing-aid benefit (Johannesen et al., 2016; Lopez-Poveda et al., 2017). In Kortlang et al. (2016), a model-based compressor, designed to restore BM I/O functions in HI listeners, performed similarly to linear gain and a solution based on National Acoustic Laboratories' nonlinear fitting procedure, version 1 (NAL-NL1) in terms of speech intelligibility in noise across several conditions, thus casting doubt on the utility of estimating the BM I/O characteristics in individual listeners. However, the BM I/O estimation and restoration in that study were based on an audiogram and categorical loudness scaling (ACALOS, Brand & Hohmann, 2002) data from Jürgens et al. (2011), who compared the ACALOS parameter estimates to cochlear compression estimates obtained from the TMC method. However, while ACALOS data could predict OHC gain loss (i.e., loss of cochlear gain provided by OHCs, as defined in Plack et al., 2004) estimated from the TMC experiments, there was no correlation between the ACALOS data and the TMC-based CE estimates. While there are several potential sources of this lack of correlation (of which the measurement variability of either method is the simplest), the TMC-based CE remains a potentially interesting beyond-audiogram metric that cannot be predicted from the audiogram or ACALOS data.

One of the challenges faced by the studies is estimating CE in a sufficiently large sample size. TMC-based experiments are very time consuming, require extensive training

before listeners reach stable performance, and the estimated thresholds are often characterized by large within-subject variability (Rosengard et al., 2005). While physiological methods are available, such as those based on distortion-product otoacoustic emissions (DPOAEs; e.g., Johannesen & Lopez-Poveda, 2010; Neely et al., 2009), they return unreliable estimates in individual listeners, and in others, they do not return an estimate at all.

Johannesen and Lopez-Poveda (2010) observed that the CE estimates measured with TMC and DPOAE methods were correlated in normal-hearing (NH) listeners at 4 kHz but not at lower frequencies. The latter was an unexpected finding given that it is assumed that both methods probe the state of the OHCs in the cochlea. Anyfantakis et al. (2017) hypothesized that the lack of correlation was due to the fine-structure pattern in the DPOAE response magnitude as a function of test frequency, which can vary across input level, leading to a high variance in the estimated distortion-product input/output (DP I/O) function slopes. When using a presentation method designed to minimize the fine-structure effects, high correlation was reported between the TMC- and DPOAE-based estimates measured at 1 and 2 kHz in HI listeners but not in NH listeners. However, only a small listener sample was tested in that study and a larger scale effort was considered necessary to confirm the findings.

Estimates of auditory filter characteristics, and thus frequency selectivity, can be derived behaviorally by examining how threshold for a pure tone presented in notched noise (NN) varies as a function of notch bandwidth (Patterson, 1976). As OHCs influence both the nonlinear compression and frequency selectivity in the cochlea, some characteristics of the auditory filter and estimates of cochlear compression should be related. Specifically, the filters in listeners with active OHCs and thus strong cochlear compression are believed to be more sharply tuned than in listeners with damaged OHCs (Oxenham & Wojtczak, 2010). Moore et al. (1999) found a very strong correlation between the ratio of slopes of the growth-of-masking (GOM, Oxenham & Plack, 1997) curves and the corresponding equivalent rectangular bandwidth (ERB) estimates ($r = .92$) in HI listeners at frequencies at and above 2 kHz, and this correlation was shown to exceed the correlation between the estimates from the two measures and the OHC gain estimates obtained from a loudness-matching experiment. As GOM is an alternative to TMC method of estimating CE and both methods rely on forward masking, it can be expected that auditory-filter properties measured with the NN method should be strongly correlated to the CEs measured with other methods.

The aim of this study was to investigate the relation between the two above-mentioned measures of peripheral compression in a large group of listeners. In addition,

sharpness of the auditory filters was estimated with the NN method, as it should also reflect the state of the active mechanism of the cochlea, while NN being a perceptually simpler task than TMCs. Here, we analyze the results of the psychoacoustic and physiological experiments to determine whether the TMC, NN, and otoacoustic emission (OAE)-based estimates are consistent and could complement each other. Furthermore, if we assume that the audiogram is affected by factors beyond the state of the cochlear active mechanism (e.g., the inner-hair cell function), it is interesting to investigate whether the TMC-, NN-, and OAE-based estimates can provide more information about the state of OHCs than the audiogram alone. This beyond-audiogram information may then be useful for better characterization of the hearing loss in individual listeners.

Methods

Participants

Forty-five HI listeners were recruited and tested at the audiological departments of the Bispebjerg Hospital and the State Hospital (Rigshospitalet), located in Copenhagen, Denmark. The existing audiological data (air- and bone-conduction thresholds measured at octave frequencies between 125 and 8000 Hz) were used to assess two recruitment criteria. First, the listeners were selected to ensure that they had sensorineural hearing loss (based on air-bone gaps not greater than 10 dB at any audiological frequency). Second, the listeners' hearing thresholds at 1 kHz could not exceed 45 dB hearing level (HL) in the ears with the better four-frequency pure-tone average. In addition, five NH listeners were tested using the same experimental paradigm. All listeners provided informed consent prior to participation in the experiments. The experimental procedure was approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391).

Setup and Procedure

The experiments were performed over three, 2-to-3-h sessions, on 3 separate days, in a double-walled listening booth. The OAE measurement was performed using the ER10-X probe system. The behavioral tasks were performed on a Windows PC equipped with Matlab with stimuli presented via calibrated Sennheiser HDA300 headphones connected directly to a Fireface UCX soundcard. All tests focused on investigation of the status of the cochlear active mechanism at two frequencies: 1 and 2 kHz. These frequencies were chosen, as a compromise between their importance for speech intelligibility (e.g., American National Standards Institute,

ANSI S3.5 1997), DPOAE noise floor levels (relatively high at lower frequencies, Gorga et al., 1993), and the degree of hearing loss at different frequencies (i.e., expected strength of the cochlear active mechanism).

DPOAEs were measured on the first day, to make sure that the probe tip could be well fitted (i.e., a good seal could be achieved) to the better ear. If a stable fit could not be obtained, the other ear was used for all further tests. Two OAE measurement paradigms were employed—one used swept tones and the other pure tones. This was done to enable future comparisons between the two methods. The calibration (including estimation of the forward-pressure level), swept-tone OAE measurements, and analysis (including source unmixing) were very similar to the scissors-rule-based paradigm used in Anyfantakis et al. (2017), but with two major changes. First, the presentation levels of the second primary (L2) in this study were set to six values that uniformly span the range from 40 to 80 dB SPL. Second, in an attempt to reduce the measurement noise, and thus improve the signal-to-noise ratio, participants were seated (Driscoll et al., 2004) in the test booth. In the pure-tone paradigm, the same presentation levels, level rule, and primary frequency ratio (1.22) were used, as in the swept-tone paradigm. The pure tones presented had a duration of 1.1 s, including 50 ms cosine ramps. Nine recordings of each combination of test-frequency and presentation level were performed.

The classic NN paradigm was used to investigate the sharpness of the auditory filters, and it was performed on Day 2. The reason for not randomizing the order of the NN and the TMC tasks was the higher perceptual difficulty of the TMC task. It was assumed that listeners' overall performance would be better if they started with simpler tasks and the difficulty was gradually increased. (For the same reason, the test conditions within the NN and TMC task were not randomized.) In the NN task, the broadband, 300 ms (including 8 ms cosine ramps) masker had a constant spectral density of 40 dB/Hz and the participants' task was to detect a 200 ms pure tone (which also included 8 ms cosine ramps), henceforth referred to as the target tone. The stimuli were presented in a 2-alternative forced-choice (2-AFC) task, with a one-up-three-down version of the Grid method (Fereczkowski, 2015). There was a 0.7 s silence interval between the two presented maskers and the target tone was temporally centered in a randomly chosen masker. The minimum notch width was set to 0 (which corresponds to the tone-in-noise threshold), and the maximum was set to 0.85 (as a proportion of the test frequency). Only symmetrical notches were tested, and they were obtained by setting the corresponding Fourier coefficients to 0 in a randomly generated noise portion. The step sizes were 3 dB in the target-tone-level and 0.05 in the notch-width dimensions. Prior to the test, each participant was trained in the 2-AFC task, and the target-

tone-detection thresholds were measured using the standard 2-AFC, one-up-three-down paradigm. Next, a procedure similar to guided-stepwise training (Hietkamp et al., 2009) was used to familiarize the listener with the NN task and a warm-up run with the Grid method was administered for the 1 kHz target tone. This training procedure typically took 10 to 15 min. Subsequently, three test runs were executed at 1 kHz. If the standard error of the estimated thresholds (averaged across the tested notch widths) exceeded 3 dB, a fourth run was administered, and all runs were averaged. Finally, a warm-up run was performed for the 2 kHz target and three or four test runs followed. Visual feedback was provided after participant's responses during all training and test runs.

On the last day, the TMC experiments were performed. In this task, a 200-ms pure-tone masker was followed by a 16-ms pure-tone target (presented at 12 dB above the threshold measured using the 2-AFC one-up-three-down procedure). As 8-ms cosine ramps were used in both (masker and target) cases, the target tone had no steady state. As in the NN task, a 2-AFC one-up-three-down procedure was used and the two maskers were separated by 0.7 s. The minimum and maximum masker-target temporal gaps (measured between the zero-voltage points) were 10 and 200 ms. The maximum masker level was 95 dB SPL in the on-frequency condition and 100 dB SPL in the off-frequency condition. The step sizes were set to 3 dB for the tone level and 5 ms for the temporal gap. Unlike in the NN case, three conditions were tested (2 kHz off- and on-frequency and 1 kHz on frequency, in this order). In the on-frequency condition, the target and the masker tones had the same frequency and in the off-frequency condition, the masker frequency was set to 55% of the target frequency. Again, a procedure similar to guided-stepwise training was employed before each test condition (unlike the NN case) and within each condition one warm-up and two test runs were performed. If the average standard error exceeded 3 dB, up to three extra runs could be administered, and all runs were averaged. As in the NN case, visual feedback was provided after each response given by a participant, during all training and test runs.

Data Analysis

The DPOAE data from the swept-tone paradigm were processed using the same procedure as used by Anyfantakis et al. (2017), which included source unmixing. The distortion-product (DP) response at a given input level was considered valid if the estimated signal-to-noise ratio exceeded 5 dB. For the case of the pure-tone paradigm, the eight recordings with the lowest root mean square (RMS) were selected and high-pass filtered at 500 Hz, to reduce artifacts and excessive noise. The eight recordings were then averaged, and a Fast Fourier Transform (FFT)

analysis was performed on the 1 s long fragment that did not contain the ramps. The strength of the 2f₁-f₂ component was recorded as the final DP response.

The DP input/output (I/O) curves, resulting from both presentation paradigms, were fitted independently with a broken-stick (one, two, or three sections, joined at KPs) function using a similar constrained fitting procedure as used by Fereczkowski et al. (2017). The constraints were set on the fitted slopes but not on the KPs. In the three-section case, the slopes of the first and the third section were limited to values between 1/1.05 and 2, and the slope of the second (mid-) section was limited to values between 0 and 1/1.05. This was done to ensure that the fitted model could approximate a function with characteristics typical for BM I/O curves, with linear behavior at the lowest and the highest input levels and a compressive section at the mid-levels. The lower and upper bounds were chosen based on the range of values found in the literature (Plack et al., 2004, Rosengard et al., 2005).

In the case of the two-section model, both fitted slopes were constrained to values between 0 and 2. For consistency, in the one-section case (i.e., linear regression), the fitted slopes were also constrained to values between 0 and 2. To limit the risk of overfitting, the Akaike Information Criterion was used to determine the best fit. As in Plack et al. (2004), if any of the fitted sections was beyond the range of the data, the data were effectively fit using a reduced number of parameters and this effective number of parameters was used in Akaike Information Criterion calculations. Here, the slope of the most compressive (shallow) portion of the best-fitting I/O function (in dB/dB coordinates), henceforth referred to as CE, was recorded. As an effect, the CE value could vary between 0 and 2 (e.g., in cases when few data points were available for fitting and the best fit was provided by a single section with a steep slope). While CE values exceeding 1 are not physiologically plausible, they have nonetheless been reported in other studies (e.g., Rosengard et al., 2005) and are kept unchanged here to avoid artificial limitation of the variability of the estimates obtained from different paradigms.

To reduce the variance of the OAE-based CE estimates, and facilitate the comparison with the psychophysical methods, the CE values obtained from the pure-tone and swept-tone DPOAE paradigms were averaged to obtain the final OAE CE estimate. This was done as the CE values from both methods were found not to be significantly different, according to a rank-sum test ($p = .54$ at 1 kHz and $p = .55$ at 2 kHz), but significantly correlated in Spearman's sense, $r_s(17) = .52$, $p < .05$ at 1 kHz and $r_s(27) = .76$, $p < .0001$ at 2 kHz.

The TMC I/O curves were obtained from the on- and off-frequency data, averaged across all test-runs.

Following the procedure from Nelson et al. (2001), for each temporal-gap value, the corresponding on-frequency threshold is taken as the input-level coordinate and the off-frequency threshold is taken as the output-level coordinate. As an effect, the off-frequency thresholds are plotted against the on-frequency thresholds. To better illustrate the procedure, one can assume that the off-frequency thresholds follow a straight line with a relatively shallow slope and the on-frequency thresholds follow a three section with a shallow segment at low gap values followed by a steep segment at moderate gap values and again a shallow segment at the highest gaps. If the shallow segments of the off- and on-frequency curves have similar slopes, the resulting TMC I/O curve will have a three section form, starting with a linear segment at low input levels, followed by a shallow (compressive) segment at moderate levels and with another steep, linear segment at the highest levels. Finally, the TMC CE estimate was obtained in the same manner as for the DP I/O paradigm described earlier.

For the auditory filter estimates, fitting one- and two-parameter rounded exponential functions ($roex(p)$ and $roex(p, r)$; Patterson et al., 1982), using the Akaike Information Criterion for model selection, but without allowing for off-frequency listening, occasionally led to unstable (i.e., very high) estimates of the rounding parameter, p . Thus, the sharpness of tuning in the NN task was estimated in a simpler way. Through interpolation, the notch width that resulted in a threshold 10 dB lower than the tone-in-noise threshold was found and termed the NN10 threshold.

To avoid potential distribution-related issues, Spearman correlation coefficients are reported. However, to estimate potential *beyond-audiogram* information common between the NN, TMC, and OAE estimates, linear models are used.

Results

The top left panel of Figure 1 presents the distribution of hearing thresholds across the HI listeners.

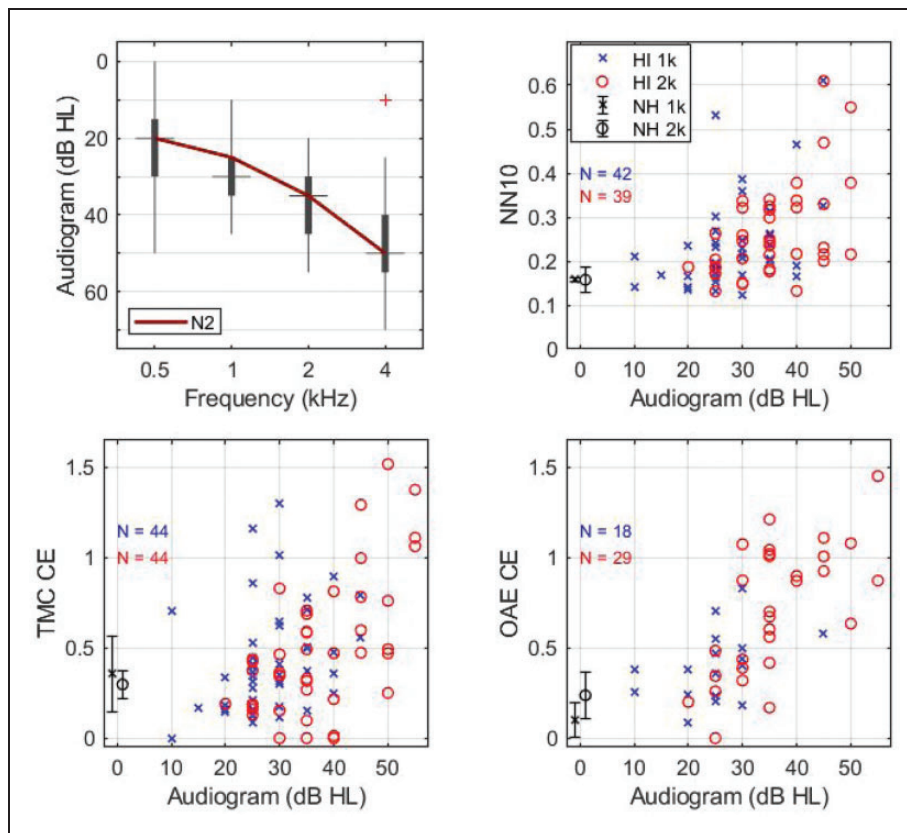


Figure 1. Hearing Thresholds and The Relations with NN10, TMC CE, and OAE CE. The top-left panel presents the distribution of hearing thresholds across the listeners. The median threshold is the same as the N2 audiogram (red, solid line) from Bisgaard et al. (2010) at 0.5, 2 and 4 kHz and is 5 dB higher at 1 kHz. The remaining three panels show scatterplots of the estimates of NN10, TMC CE, and OAE CE with the corresponding hearing threshold at 1 kHz (blue crosses) and 2 kHz (red circles). In the top-right panel, the error bars located near 0 dB HL show mean and standard deviation of the NN10 estimates obtained with five NH listeners. In case of all nonaudiometric estimates, lower values indicate stronger cochlear nonlinearity. The numeric insets indicate the number data points at 1 kHz (blue, top) and at 2 kHz (red, bottom). TMC = temporal-masking curves; OAE = otoacoustic emission; CE = compression exponent; NN = notched noise; NH = normal-hearing.

The distribution of hearing thresholds at four frequencies (500, 1000, 2000, and 4000 Hz) and median values correspond well with N2 standard audiogram from Bisgaard et al. (2010), with only a small deviation of 5 dB at 1000 Hz, which suggests that the listeners exhibited mild-to-moderate losses. The remaining panels present scatterplots between the pure-tone threshold at 1 and 2 kHz and the NN10 or the CE estimates from TMC and DPOAE methods for those frequencies (remaining three panels). In those panels, blue crosses represent the data collected for the HI participants at 1 kHz and red circles represent the 2-kHz data as a function of the hearing threshold at that frequency. The black error bars show data collected for five NH listeners, except for the TMC task, where data from only four NH listeners were collected.

The top-right panel presents NN10 estimates. In the NH listeners, the mean values are 0.158 for 1 kHz (black cross) and 0.157 at 2 kHz (black circle). The corresponding standard deviations are 0.006 and 0.028. In addition, NN10 was estimated from the data in Rosen and Baker (1994), collected for the 40 dB/Hz condition in three NH listeners at 2 kHz. The estimate was 0.148 and thus was within 2 standard deviations from both NH NN10 estimates obtained from the current data. Therefore, it is henceforth assumed that the NH reference value of the NN10 is between 0.15 and 0.16 for the two tested frequencies. As could be expected, the estimates in the HI listeners are usually larger, indicating broader auditory filters. The NN10 estimates from the HI listeners vary between 0.12 and 0.61. In some listeners, the NN10 was not possible to estimate due to the differences between the thresholds estimated at one frequency not reaching 10 dB. This suggests broad auditory filters; however, a decision was made to treat these cases as missing values.

From the scatterplots in Figure 1, it appears that the results from all three measures are correlated with audiometric data. Indeed, NN10 estimates (top-right panel) were found to be significantly correlated with hearing threshold (right-upper panel) both at 1 kHz, $r_s(df=40) = .47$, $p < .01$, and at 2 kHz, $r_s(37) = .51$, $p < .001$. Similarly, the correlations between TMC CE estimates and hearing threshold (bottom-left panel) were significant both at 1 kHz, $r_s(42) = .41$, $p < .01$, and at 2 kHz, $r_s(42) = .57$, $p < .0001$. Finally, the correlations between OAE CE estimates and hearing threshold (bottom-right panel) were trending toward significance at 1 kHz, $r_s(16) = .47$, $p = .051$, and significant at 2 kHz, $r_s(27) = .64$, $p < .001$.

The OAE experiment returned a particularly low number of estimates at 1 kHz. This is due to the averaging of the results from the pure-tone and the swept-tone paradigms, as it was only possible to generate estimates from 20 listeners using the pure-tone paradigm at 1 kHz, even if the swept-tone paradigm returned 38 estimates.

While the OAE strength was higher at 1 kHz than at 2 kHz when using the pure-tone paradigm, the noise-floor levels were higher as well. The average of the median noise floor levels for the listeners for whom CE could be estimated was -16.5 dB SPL ± 2.7 dB, while the average for those for whom it was not possible to estimate a CE was -12.4 dB SPL ± 6.2 dB. This suggests that increased noise floor in some listeners reduced the ability to estimate CE at 1 kHz using the DPOAE pure-tone paradigm. In addition, while more CE estimates were obtained from the swept-tone paradigm, they correlated more poorly than the averaged OAE CE with non-OAE estimates.

The correlation with hearing thresholds seems consistent across the three discussed measures, which suggests a common underlying mechanism. To further investigate this, Figure 2 shows scatterplots of estimates from the three non audiogram measures. The left-upper panel of Figure 2 presents a scatterplot of the TMC CE and the NN10 estimates, that is, the two psychophysical measures. These measures were significantly correlated at 1 kHz, $r_s(40) = .41$, $p < .01$, but not at 2 kHz $r_s(36) = .13$, $p = .42$. Spearman's partial correlation between NN10 and TMC CE was estimated for the 1 kHz case. The air-conduction threshold from the audiogram was used here (and elsewhere) to control for the effect of hearing threshold at that frequency. The correlation was insignificant, $r_{ps}(38) = .29$, $p = .07$, which suggests that most of the variance shared by the NN10 and TMC CE is shared by the audiogram.

The left-lower panel of Figure 2 presents the scatterplot of TMC CE and OAE CE estimates. The Spearman correlations were not significant at 1 kHz, $r_s(15) = .40$, $p = .12$, or at 2 kHz, $r_s(N=27) = .32$, $p = .09$. The right-upper panel of Figure 2 presents a scatterplot of the NN10 and OAE CE estimates. The scatterplot suggests a stronger relationship than in the two previous cases, at least for the 1 kHz data, and the correlation analysis supports this (1 kHz: $r_s(16) = .68$, $p < .01$; 2 kHz: $r_s(26) = .50$, $p < .05$). When considering the audiometric thresholds, the partial correlations between NN10 and OAE CE estimates were moderate and significant (1 kHz): $r_{ps}(15) = .60$, $p < .05$; 2 kHz: $r_{ps}(25) = .49$, $p < .05$).

The relationship between OAE CE, NN10, and HL was further investigated using linear models with OAE CE as the dependent variable. Specifically, it was tested whether the introduction of the NN10 estimate (NN10 + HL model of OAE CE) leads to improved predictions, when compared with the hearing-threshold-based model (HL-model). Table 1 provides the summary of the modeling results.

At 1 kHz, the adjusted R^2 was .11 for the HL-model (M1, insignificant at $p = .09$) and increased to .52 for the NN10 + HL model. An F test confirmed that this increase was significant, $F(1, 15) = 14.7$, $p < .01$, and HL was not a

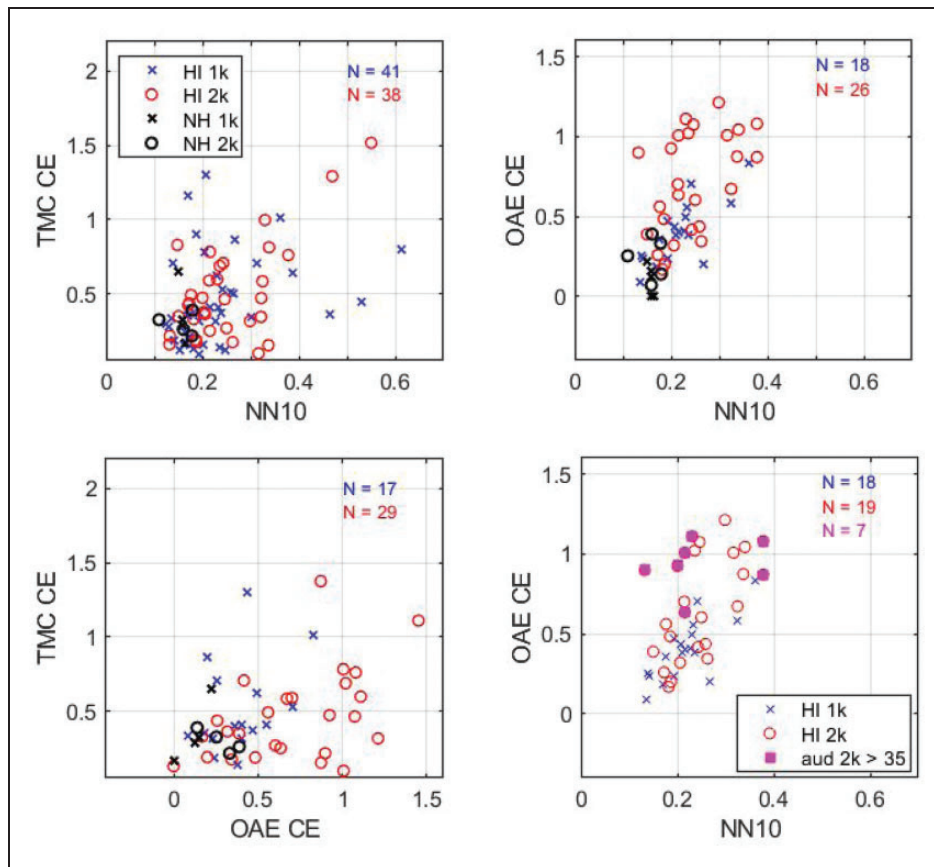


Figure 2. Relations Between the Estimates of NN10, TMC CE, and OAE CE. Large variability is evident in the two panels on the left, the panels comparing NN10 and OAE CE reveal more structure in the data. TMC = temporal masking curves; OAE = otoacoustic emission; CE = compression exponent; NN = notched noise.

Table 1. Linear-Model Analysis of the Relationship Between Pure-Tone Thresholds (HL), NN10, and OAE CE.

Frequency	1 kHz		2 kHz		
	M1	M2	M1	M2	M3
Model					
Variables	HL	+ NN10	HL	+ NN10	HL < 40
No. data points	18		27		20
HL <i>p</i> val	.098	.782	.0001**	.004**	.129
NN10 <i>p</i> val	–	.002**	–	.021*	.008**
Model <i>F</i> stat.	3.09	10.2	20.6	10.5	9.1
Model <i>p</i> val	.098	.002**	.0001**	.0006*	.0023**
<i>R</i> ²	.16	.58	.43	.48	.53
Adjusted <i>R</i> ²	.11	.52**	.41	.43**	.47**

Note. The statistical significance of models M2 and M3 (see “Discussion” section) was determined by means of model comparison against M1. *p* value codes: * < .05. ** < .01. *** < .0001. NN = notched noise; HL = Hearing Level.

significant factor in the NN10 + HL model (*p* = .78). At 2 kHz, the adjusted *R*² increased from .41 for the HL model to .43 for the NN10 + HL model. An *F* test confirmed that this increase was significant, *F*(1, 23) = 12.4, *p* < .01. However, here, HL was a significant predictor in the NN10 + HL model of OAE CE (*p* < .05).

The Spearman correlation and partial correlation as well as the linear models for the NN10 and OAE CE data suggest that the correspondence is stronger at 1 kHz than at 2 kHz and the linear model of OAE CE involves audiogram as a predictor only at 2 kHz. One of the main differences between data collected at 1 and 2 kHz is the

distribution of hearing thresholds (see e.g., Figure 1). Kowalewski (2014) investigated auditory filters in NH listeners at very high frequencies (>10 kHz), where hearing thresholds exhibit a steep slope. He reported that tuning sharpness could be overestimated in such cases, due to the limited audibility of the high-frequency portion of the NN masker. This could be corrected for by suitably amplifying the inaudible portion of the masker. It is possible that the high-frequency portion of the 40 dB/Hz masker was not sufficiently audible for some of the listeners tested in this study. The bottom-right panel of Figure 2 presents the scatterplot of NN10 versus OAE CE estimates separated based on whether the listener's hearing threshold was greater than 35 dB HL. At 1 kHz, only one listener had a threshold greater than 35 dB HL. However, at 2 kHz, seven listeners had thresholds greater than 35 dB HL. Furthermore, the NN10 estimates corresponding to those listeners span the entire range of the estimates presented in the panel. Thus, for the listeners that exhibited both relatively low NN10 and relatively high OAE CE estimates, the masker may have been insufficiently audible leading to a low NN10, that is, an overestimate of tuning sharpness.

If we assume that the NN10 estimates obtained here are unreliable for listeners whose audiometric thresholds exceed 35 dB HL, the analyses can be recomputed with these data omitted. The subsequent Spearman correlation of the NN10 and OAE CE data at 2 kHz was significant, $r_s(17) = .63$, $p < .01$, partial correlation $r_{ps}(16) = .58$, $p < .05$. In the subsequent linear-model analysis, the adjusted R^2 of the HL + NN10 model is .47 and was significantly ($p < .01$) better than the simple HL-based model (adjusted $R^2 = .28$). In addition, HL was not a significant predictor in the HL + NN10 model ($p = .13$). These results are similar to the results obtained at 1 kHz, which suggests that the correspondence between NN10 and OAE CE data is similar across the two frequencies, as long as the investigated hearing threshold range is matched. This suggests that there is a significant relation between the NN10 and OAE CE estimates at both tested frequencies and part of this relationship is not related to the underlying hearing threshold.

Discussion

All three measures of the state of the cochlear active mechanism showed a similar correlation with hearing thresholds, with Spearman correlation coefficients ranging from .41 to .47 at 1 kHz and from .51 to .64 at 2 kHz. These correlations are not unexpected, at least in the case of the NN- and OAE-based estimates, as previous studies have observed a broader auditory filter bandwidths (Moore et al., 1999), and higher slopes of OAE

I/O functions (Neely et al., 2009) with higher hearing thresholds.

The significant correlation between the TMC CE and the hearing threshold is inconsistent with findings of Plack et al. (2004), Jepsen and Dau (2011), Lopez-Poveda and Johannesen (2012), and Johannesen et al. (2016), who found no such correlation for similar groups of HI listeners with mild-to-moderate hearing loss. One reason for the discrepancy could be a lack of statistical power given the smaller number of participants in the first three studies. However, the last study involved 68 HI listeners and significant effort was made (each listener received 2 h of training and up to 3 h per frequency of testing time) to ensure high data quality. Therefore, differences between this study and the other studies in terms of data analysis are a more likely source of the lack of agreement. For instance, while in this study and the study of Plack et al. (2004), the models used to fit the I/O estimates are reasonably similar, with both studies using broken-stick functions with identical minimum limit on CE, the upper limit on the CE slope, and therefore the range of CE estimates, was different. To be more specific, in this study, the slopes of the fitted I/O function were allowed to exceed 1 (i.e., exhibit expansive behavior) and ranged from 0 to 1.52, while the maximum theoretical CE in Plack et al. (2004) was 1 and the reported CEs ranged between 0 and 0.46. However, the range of CEs reported in this study is consistent with Rosengard et al. (2005), where the values fell between -0.07 and 1.70.

For each experimental method tested here, it is assumed that the resulting estimate is influenced by the status of the cochlear active mechanism, and thus the outcomes are related. Because the OHC activity is the main source of the human-hearing sensitivity at low levels, it is not necessarily surprising that the outcome metrics from all three measures were correlated with audiometric thresholds. Quantitatively, the OHC gain loss has been estimated to correspond to roughly 65% of the hearing threshold (Plack et al., 2004), and the current data show a very similar pattern (not shown). However, the main hypothesis of this study was that if the three discussed measures are closely related to the OHC status, the audiogram should not be enough to fully predict the relations between them, due to audiometric threshold being influenced by, for example, inner-hair cell loss, or subclinical conductive losses. In other words, the three measures were expected to return estimates that would provide beyond-audiogram information about the state of the cochlear active mechanism. Moreover, as OHC gain can be modeled as a product of KP and $1 - CE$ (Equation 2 in Lopez-Poveda & Johannesen, 2012), and KP is highly correlated to the hearing threshold, it could be expected that the estimates which are highly correlated with the slope of the BM I/O

function would be interrelated to a stronger degree than they are related with the OHC gain loss estimate. This would be consistent with Moore et al. (1999) who found a trend toward higher correlations of ERB estimates and ratio of slopes of GOM curves than between the same ERB estimates and OHC gain loss estimates, in HI listeners. In that study, the correlation between the ratio of GOM slopes and the ERBs was also found to significantly exceed the correlation between the ratio of slopes and OHC gain loss estimates. The simple and partial correlations between the three variables observed in this study show a similar pattern (with respect to the audiogram instead of the OHC gain loss) for only some of the variable pairs.

For the two psychophysical measures, NN10 and TMC CE, the current results show a significant correlation at 1 kHz but not at 2 kHz. Moreover, the correlation at 1 kHz seems to be accounted for by the hearing sensitivity at this frequency, due to insignificant partial correlation. This is an unexpected finding, given that Moore et al. (1999) found a very strong Pearson correlation between the ratio of slopes of the GOM curves and the corresponding ERB estimates ($r = .92$) in HI listeners at frequencies at and above 2 kHz. Clearly, their results did not involve identical measures as this study, but GOM and TMC estimates have been shown to be correlated, at least for NH listeners (Rosengard et al., 2005) and the NN10 and ERB estimates are expected to be correlated as both describe sharpness of tuning. In fact, the NN10 and ERB are perfectly correlated for the filter shapes following the roex(p) model which was tested in a simple simulation and very highly correlated with the roex(p, r) model: with Pearson's r equal to .99. However, this finding is valid for data perfectly following the roex shape. In case of noisy data, the fitted roex parameters and the NN10 values may differ, as was the case in this study. On the other hand, Rosengard et al. (2005) did not find any correlation between GOM and TMC measures in HI listeners and attributed this to high variability of both measures, and in particular the TMC measure. Therefore, the lack of high correlation between the psychophysical measures tested here may be a result of the high variance of estimates produced using the TMC method.

The comparison of the TMC and OAE CE estimates shows no significant correlation at 1 or 2 kHz. This is consistent with Johannesen and Lopez-Poveda (2010), who found no relation at frequencies below 4 kHz in NH listeners, but inconsistent with Anyfantakis et al. (2017) who reported a correlation between TMC and OAE CEs at 1 and 2 kHz in HI listeners. However, the study by Anyfantakis et al. differed from this study in two important aspects. First, the correlation analyses in that study were limited to eight data points from six

listeners. Second, the procedure used by Anyfantakis et al. to collect the TMC data aimed at minimizing the variability of the estimated thresholds. Each participant had at least 2 h of training with the TMC task before the experiment and six test runs per condition were administered. Due to time limitations, the participants in this study received less training and fewer test runs were administered. While a range of audiometric thresholds of the HI listeners was similar across both studies, the range of CE's reported by Anyfantakis et al. (between 0.25 and 1) was narrower than in this study. This supports the conclusion that the nonsignificant correlation observed here is because of the high variance in the individual CE estimates collected using the TMC method. In addition, the variability of the TMC-based estimates may be increased with respect to other measures, due to the findings of Wojtczak and Oxenham (2009), who reported that the rate of recovery of the off-frequency masker may depend on the masker level.

This potential limitation in terms of the TMC data quality suggests caution when interpreting the fact that only the correlations between NN10 and OAE CE measures are stronger than what could be predicted by the hearing threshold. Given that Moore et al. (1999) found a significant correlation between ERB and GOM measures, it is tempting to speculate that if GOM-based CEs were estimated in the current experiment they would correlate well with the NN10 estimates and, potentially, OAE CEs. Furthermore, this correlation would be expected not to be fully explained by the hearing thresholds or OHC gain, which would make the GOM CE, NN10, and OAE CE, or some combination of those (e.g., based on principal component analysis) an interesting, beyond-audiogram characteristic of individual hearing abilities. In the next step, this characteristic could be compared with speech-in-noise performance, as was done in the case of TMC CE in Johannesen et al. (2016).

However, the variability of the TMC CE estimates is not the only limitation of this study. The second limitation is a relatively low number of OAE CE estimates that could be obtained from the HI listeners. Despite attempts to reduce the effects of DPOAE fine structure, it was only possible to estimate CEs from the OAE data from 19 and 29 listeners at 1 kHz and 2 kHz, respectively. This is most likely due to the increased noise floor at low frequencies in the pure-tone paradigm.

As the NN method exhibits less test-retest variability than TMC and returns valid estimates in larger number of participants than the OAE methods, it seems to be the best candidate for the laboratory or clinical use as a suprathreshold performance measure, especially if high-frequency masker shapes can be efficiently adjusted for the audiogram shape, as in Kowalewski (2014),

where such adjustment led to improved estimates of auditory filter widths. However, in the form employed here, the NN method is rather time consuming. The test sessions involving the NN procedure lasted up to 90 min and returned estimates at only two test frequencies. A separate analysis would be necessary to estimate the minimum amount of training- and test-time that would allow for maintaining the high quality of the NN10 estimates. Another approach could involve testing listeners outside the laboratory, using, for example, consumer-grade mobile audio equipment (Hyvärinen et al., 2019). The NN experiment may be a particularly suitable candidate for this approach due to employing noise-based stimuli and suprathreshold presentation levels. If the NN test is indeed feasible outside a laboratory or a clinic, it might become a useful tool for characterizing individual hearing loss.

Summary

This study investigated the relation between three measures that are expected to tap into the status of the cochlear active mechanism: estimates of slopes of the TMC- and DPOAE-based I/O functions and a measure of sharpness of tuning from the NN experiment. The first finding is that a measure of sharpness of tuning was highly correlated with an OAE-based measure of I/O response growth and that the relation is independent on the hearing sensitivity. The second finding is that the low reliability of the TMC-based estimates may be the main reason for the limited correspondence between this and the two other measures. In addition, a GOM-based measure (GOM slope ratio) is expected to correlate highly with the OAE- and NN-based measures and if this was the case, it would be interesting to investigate the relations between the three measures and, for example, speech-in-noise performance, to investigate the beyond-audiogram contribution of the state of the cochlear mechanism to speech intelligibility. Meanwhile, the NN experiment appears as the most reliable of the three tested measures and if it can be reliably implemented outside the laboratory, it might become a useful tool for characterizing individual hearing loss.


Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Oticon Foundation.

ORCID iD

Michał Fereczkowski  <https://orcid.org/0000-0002-7960-1188>

References

- Anyfantakis, K., MacDonald, E., Epp, B., & Fereczkowski, M. (2017). Comparison of objective and subjective measures of cochlear compression in normal-hearing and hearing-impaired listeners. In S. Santurette, T. Dau, J. Christensen-Dalsgaard, L. Tranebjærg, T. Andersen, & T. Poulsen (Eds), *International symposium on auditory and audiological research* (Vol. 6: pp. 167–174). The Danavox Jubilee Foundation. <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2017-20/328>
- Bisgaard, N., Vlaming, M. S., & Dahlquist, M. (2010). Standard audiograms for the IEC 60118-15 measurement procedure. *Trends in Amplification*, *14*(2), 113–120. <https://doi.org/10.1177/1084713810379609>
- Brand, T., & Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. *The Journal of the Acoustical Society of America*, *112*(4), 1597–1604. <https://doi.org/10.1121/1.1502902>
- Driscoll, C., Kei, J., Shyu, J., & Fukai, N. (2004). The effects of body position on distortion-product otoacoustic emission testing. *Journal of the American Academy of Audiology*, *15*(8), 566–573. <https://doi.org/10.3766/JAAA.15.8.4>
- Fereczkowski, M. (2015). *Time-efficient behavioral estimates of cochlear compression* [PhD thesis]. Technical University of Denmark. <https://orbit.dtu.dk/en/publications/time-efficient-behavioural-estimates-of-cochlear-compression>
- Fereczkowski, M., Jepsen, M. L., Dau, T., & MacDonald, E. N. (2017). Investigating time-efficiency of forward masking paradigms for estimating basilar membrane input-output characteristics. *PLoS One*, *12*(3), pp. 1–20. <https://doi.org/10.1371/journal.pone.0174776>
- Gorga, M. P., Neely, S. T., Bergman, B. M., Beauchaine, K. L., Kaminski, J. R., Peters, J., . . . , Jesteadt, W. (1993). A comparison of transient-evoked and distortion product otoacoustic emissions in normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *94*(5), 2639–2648. <https://doi.org/10.1121/1.407348>
- Hietkamp, R. K., Andersen, M. R., & Lunner, T. (2009, December). Perceptual audio evaluation by hearing-impaired listeners—some considerations on task training. In J. M. Buchholz, T. Dau, J. Christensen-Dalsgaard, T. Poulsen (Eds), *Proceedings of the international symposium on auditory and audiological research* (Vol. 2, pp. 487–496). The Danavox Jubilee Foundation. <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2009-50/222>
- Hyvärinen, P., Fereczkowski, M., & MacDonald, E. N. (2019). Evaluation of a notched-noise test on a mobile phone. In A. A. Kressner, J. Regev, J. Christensen-Dalsgaard, L. Tranebjærg, S. Santurette, & T. Dau (Eds.), *Proceedings of the international symposium on auditory and audiological research* (Vol. 7, pp. 125–132). The Danavox Jubilee Foundation. proceedings.isaar.eu/index.php/isaarproc/article/view/2019-15

- Jepsen, M. L., & Dau, T. (2011). Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *129*(1), 262–281. <https://doi.org/10.1121/1.3518768>
- Johannesen, P. T., & Lopez-Poveda, E. A. (2008). Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions. *Journal of the Acoustical Society of America*, *124*, 2149–2163. <https://doi.org/10.1121/1.2968692>
- Johannesen, P. T., & Lopez-Poveda, E. A. (2010). Correspondence between behavioral and individually “optimized” otoacoustic emission estimates of human cochlear input/output curves. *The Journal of the Acoustical Society of America*, *127*(6), 3602–3613. <https://doi.org/10.1121/1.3377087>
- Johannesen, P. T., Pérez-González, P., Kalluri, S., Blanco, J. L., & Lopez-Poveda, E. A. (2016). The influence of cochlear mechanical dysfunction, temporal processing deficits, and age on the intelligibility of audible speech in noise for hearing-impaired listeners. *Trends in Hearing*, *20*: 1–14. <https://doi.org/10.1177/2331216516641055>
- Jürgens, T., Kollmeier, B., Brand, T., & Ewert, S. D. (2011). Assessment of auditory nonlinearity for listeners with different hearing losses using temporal masking and categorical loudness scaling. *Hearing Research*, *280*(1–2), 177–191. <https://doi.org/10.1016/j.heares.2011.05.016>
- Kortlang, S., Grimm, G., Hohmann, V., Kollmeier, B., & Ewert, S. D. (2016). Auditory model-based dynamic compression controlled by subband instantaneous frequency and speech presence probability estimates. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *24*(10), 1759–1772. <https://doi.org/10.1109/TASLP.2016.2584705>
- Kowalewski, B. (2014). Modified notched-noise method for investigation of auditory filter shapes at high frequencies. In Forum acusticum 2014, Retrieved from: https://www.researchgate.net/publication/266208217_Modified_notched_noise_method_for_investigation_of_auditory_filters_at_high_frequencies
- Lopez-Poveda, E. A., & Johannesen, P. T. (2012). Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss. *Journal of the Association for Research in Otolaryngology*, *13*(4), 485–504. <https://doi.org/10.1007/s10162-012-0327-2>
- Lopez-Poveda, E. A., Johannesen, P. T., Perez-González, P., Blanco, J. L., Kalluri, S., & Edwards, B. (2017). Predictors of hearing-aid outcomes. *Trends in Hearing*, *21*: 1–28. <https://doi.org/10.1177/2331216517730526>
- Mills, D. M., Feeney, M. P., & Gates, G. A. (2007). Evaluation of cochlear hearing disorders: Normative distortion product otoacoustic emission measurements. *Ear and Hearing*, *28*, 778–792. <https://doi.org/10.1097/AUD.0b013e3181576755>
- Moore, B. C., Vickers, D. A., Plack, C. J., & Oxenham, A. J. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *The Journal of the Acoustical Society of America*, *106*(5), 2761–2778. <https://doi.org/10.1121/1.428133>
- Mueller, J., & Janssen, T. (2004). Similarity in loudness and distortion product otoacoustic emission input/output functions: Implications for an objective hearing aid adjustment. *Journal of the Acoustical Society of America*, *115*, 3081–3091. <https://doi.org/10.1121/1.1736292>
- Neely, S. T., Johnson, T. A., Kopun, J., Dierking, D. M., & Gorga, M. P. (2009). Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears. *The Journal of the Acoustical Society of America*, *126*(2), 728–738. <https://doi.org/10.1121/1.3158859>
- Nelson, D. A., Schroder, A. C., & Wojtczak, M. (2001). A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *110*(4), 2045–2064. <https://doi.org/10.1121/1.1404439>
- Oxenham, A. J., & Plack, C. J. (1997). A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America*, *101*(6), 3666–3675. <https://doi.org/10.1121/1.418327>
- Oxenham, A. J., & Wojtczak, M. (2010). Frequency selectivity and masking. In C. J. Plack (Ed.), *The Oxford handbook of auditory science: Hearing* (pp. 5–44). <https://doi.org/10.1093/oxfordhb/9780199233557.001.0001>
- Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *The Journal of the Acoustical Society of America*, *59*(3), 640–654. <https://doi.org/10.1121/1.380914>
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., & Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *The Journal of the Acoustical Society of America*, *72*(6), 1788–1803. <https://doi.org/10.1121/1.388652>
- Plack, C. J., Drga, V., & Lopez-Poveda, E. A. (2004). Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *115*(4), 1684–1695. <https://doi.org/10.1121/1.1675812>
- Rosen, S., & Baker, R. J. (1994). Characterising auditory filter nonlinearity. *Hearing Research*, *73*(2), 231–243. <https://doi.org/10.1.1.453.3443>
- Rosengard, P. S., Oxenham, A. J., & Braida, L. D. (2005). Comparing different estimates of cochlear compression in listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America*, *117*(5), 3028–3041. <https://doi.org/10.1121/1.1883367>
- Ruggero, M. A., & Rich, N. C. (1991). Furosemide alters organ of corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane. *Journal of Neuroscience*, *11*(4), 1057–1067. <https://doi.org/10.1523/JNEUROSCI.11-04-01057.1991>
- Wojtczak, M., & Oxenham, A. J. (2009). Pitfalls in behavioral estimates of basilar-membrane compression in humans. *The Journal of the Acoustical Society of America*, *125*(1), 270–281. <https://doi.org/10.1121/1.3023063>