JARO
Journal of the Association for Research in Otolaryngology

CrossMark

Research Article

Modeling the dependence of the distortion product otoacoustic emission response on primary frequency ratio

RENATA SISTO,[1] UZMA SHAHEEN WILSON,[2] SUMITRAJIT DHAR,[2] AND ARTURO MOLETI[3]

[1]*INAIL Research, 00044, Monteporzio Catone (RM), Italy*

[2]*Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208, USA*

[3]*Physics Department, University of Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133, Rome, Italy*

ABSTRACT

When measured as a function of primary frequency ratio $r = f_2/f_1$, using a constant f_2, distortion product otoacoustic emission (DPOAE) response demonstrates a bandpass shape, previously interpreted as the evidence for a cochlear "second filter." In this study, an alternate, interference-based explanation, previously advanced in variants, is forwarded on the basis of experimental data along with numerical and analytical solutions of nonlinear and linear cochlear models. The decrease of the DPOAE response with increasing and decreasing ratios is explained by a diminishing "overlap" generation region and the onset of negative interference among wavelets of different phase, respectively. In this paper, the additional quantitative hypothesis is made that negative interference becomes the dominant effect when the spatial width of the generation (overlap) region exceeds half a wavelength of the DPOAE wavelets. Therefore, r is predicted to be optimal when this condition is matched. Additionally, the minimum on the low-ratio side of the DPOAE curve is predicted to occur as the overlap region width equals one wavelength. As the width of the overlap region depends on both tuning and ratio, while wavelength depends on tuning only, an experimental method for estimating tuning from either the width of the pass band or the optimal ratio of the DPOAE vs. ratio curve has been theoretically formulated and evaluated using numerical simulations. A linear model without the possibility of nonlinear suppression is shown to reasonably approximate data from human subjects at low ratios reinforcing the relevance of the proposed negative interference effect. The different dependence of the distortion and reflection DPOAE components on r as well as the nonmonotonic behavior of the distortion component observed at very low ratios are also in agreement with this interpretation.

Keywords: cochlear mechanics, nonlinear models, cochlear tuning, time-frequency analysis

INTRODUCTION

According to a widely accepted taxonomy (Shera and Guinan 1999), DPOAE components of different phase-gradient delays arise from two different generation mechanisms. The main component (distortion or D) of the $2f_1 - f_2 = f_{DP}$ DPOAE is nonlinearly generated by cubic intermodulation distortion of the f_2 and f_1 tones. A backward distortion product (DP) wave is generated in the region around the tonotopic characteristic frequency place for f_2 (x_2), where the overlap of the mechanical response to the two primary tones is maximal. The forward wave generated by the same distortion mechanism acts as an intracochlear stimulus for the generation of the second component (reflection or R) from the f_{DP} tonotopic region through a linear coherent reflection mechanism due to the presence of randomly distributed irregularities (roughness) along the basilar membrane (BM). The R component is therefore approximately equivalent to a stimulus-frequency OAE (SFOAE) or transient-evoked OAE (TEOAE) evoked by a stimulus

Correspondence to: Arturo Moleti · Physics Department · University of Roma Tor Vergata · Via della Ricerca Scientifica 1, 00133, Rome, Italy. Telephone: +39-0672594517; email: moleti@roma2.infn.it

generated within the cochlea. The two DPOAE components are characterized by very different phase-gradient delays, so they can be effectively separated by filtering in the time-frequency domain and separately studied either experimentally or using cochlear modeling tools. Indeed, in DPOAE acquisition paradigms where the ratio r is kept constant, the components D and R generated by wave- and place-fixed mechanisms are, respectively, characterized by constant and rapidly rotating phase in a scaling-symmetric cochlea. In this study, a fixed-f_2 paradigm is considered instead, which slightly affects this prediction, but a separation of the two components is still generally possible in the time-frequency domain.

The two DPOAE components are also differently sensitive to r. For the nonlinear distortion component D, the overlap between the f_2 and f_1 excitation patterns increases with decreasing r, so the DPOAE level would tend to increase with the recruitment of a greater number of DPOAE generators. However, this effect is progressively balanced by the loss of coherence of the intra-cochlear backward DP wave as its constituent wavelets have increasingly different phases. The reflection component R remains unaffected as coherence is regained along the forward path to the place-fixed reflection place(s) (e.g., Shera and Guinan 2007). The typically dominant D component, at commonly used stimulus levels, determines the basic bandpass shape of the overall DPOAE response as a function of r (or f_{DP}) for recordings using a fixed f_2. This bandpass function has been proposed to be a manifestation of cochlear tuning by some (Brown et al. 1992; Brown et al. 1993), and second cochlear filter by others (Allen 1980). The second filter explanation found some credence as it could also account for the presumed discrepancy between the tuning measured on the BM and that derived from neural thresholds in early animal experiments. Although more recent experiments have shown nearly identical BM and neural tuning, at least in the basal part of the cochlea (as reviewed in Robles and Ruggero 2001), the existence of a second filter related to the mechanical action of other cochlear elements is still in consideration as a possible reconciliation between BM and neural measurements as well as a possible explanation for the results of fixed-f_2 DPOAE experiments.

Interference-based explanations of the bandpass shape of the DPOAE response, in which no second filter is involved, have been proposed by several authors. Van Hengel (1996) was able to reproduce the bandpass DPOAE response with a nonlinear model numerically solved without a second filter in the time domain. Similar results were also obtained by Liu and Neely (2010) using their nonlinear cochlear model. Talmadge et al. (1998) used a perturbative approach to generate DPOAEs in a linear delayed-stiffness model, demonstrating that the bandpass shape was generated by interference by removing the phase dependence of the basis functions in the generation integral of the backward DP wave. The same mechanism highlighting the importance of the relation between the spatial width of the OAE generation region and the local wavelength of the traveling wave was also forwarded by Shera and colleagues (Shera 2003; Shera and Guinan 2008). Interference between elementary wavelets coming from extended cochlear regions is also relevant for interpreting the Allen-Fahey experiments (Allen and Fahey 1992), where DPOAE levels are reported as a function of ratio while holding the intra-cochlear DP level in the x_2 region constant, as a method for estimating the cochlear amplifier gain function. Also in that case, the variable size of the DP generation region plays a relevant role (e.g., Shera 2003).

In this report, experimental data and numerical simulations of DPOAEs are presented as functions of r, with the explicit purpose of studying the behavior of the unmixed D and R components. An interpretation of the observed phenomenology is proposed, based on the increasing loss of coherence of the backward DP waves as the overlap region size increases. Theoretical quantitative relationships are found between cochlear tuning and the optimal value of r, where the DPOAE response peaks, based on a heuristic model of the DPOAE response. Equivalently, the relationship between tuning and the width of the DPOAE vs. ratio function is also examined. These relationships are successfully tested by analytical and numerical model simulations and subsequently applied to sample experimental data to get tuning estimates at different frequencies and stimulus levels.

METHODS

Model

A linear transmission line cochlear model was used, based on that proposed by Zweig in 1991, in which the cochlea is schematized by a one-dimensional box model, with a tonotopic basilar membrane (BM) of surface density σ_{bm} separating two volumes filled by a homogeneous incompressible fluid. A nonlinear version (Sisto et al. 2015) of the same delayed-stiffness cochlear model was also implemented and solved in time domain with an iterative technique. Both models were discretized using $N = 2000$ partitions and tuned with an exponential scale-invariant map $\omega_{bm}(x) = \omega_0 e^{-k_\omega x}$, with $\omega_0 = 2\pi \cdot 20{,}600$ rad/s and $k_\omega = 1.38$ cm^{-1}, representing the tonotopic structure of the human cochlea.

The equations for the differential pressure p and the transverse BM displacement ξ are those of a one-dimensional tonotopic transmission line:

$$\begin{cases} \dfrac{\partial^2 p(x,t)}{\partial x^2} = k_0^2 \sigma_{bm} \ddot{\xi}(x,t) \\ \dfrac{p(x,t)}{\sigma_{bm}} = \ddot{\xi}(x,t) + \omega_{bm}(x)\delta\dot{\xi}(x,t) \\ \qquad + \omega_{bm}^2(x)\xi + \rho\omega_{bm}^2(x)\xi(t-\tau) \end{cases} \quad (1)$$

where δ, ρ, and τ are the parameters of the local BM admittance; $k_0 = 31$ cm^{-1} (Talmadge et al. 1998) is a constant related to average geometrical and density parameters of the cochlear cavity; and σ_{bm} is the BM surface density. The linear version of the model (1) was solved in the frequency domain for different values of cochlear tuning. A double-pole structure near the peak of the admittance associated with the linear model (1) was assumed, in which the cochlear partition impedance has the form:

$$Z \propto (s-\hat{s})^2, \quad (2)$$

where s is the Laplace complex variable $j\omega + \Gamma$ and $\hat{s}$ is the pole position in the complex plane. Following Shera (2001), Eq. (2) implies that both Z and its first derivative should be zero at the pole $\hat{s}$.

$$Z = 0; \quad \left.\frac{dZ}{ds}\right|_{s=\hat{s}} = 0 \quad (3)$$

From this condition, four independent equations can be derived setting four parameters, the three model parameters δ, ρ, τ, and the distance of the pole from the imaginary axis α^*. Following Shera's strategy, imposing the conditions (Eq. (3)) and setting the tuning value Q, it is possible to set the model parameters in such a way that the poles move horizontally in a direction that is approximately parallel to that of the real axis.

In the linear version of the model, which is solved in the frequency domain, Q is set as a constant value along the BM, and the four parameters of the model are computed accordingly, using an empirical relation (Sisto et al. 2016) between tuning and α^*. The generation of DP waves, eventually producing DPOAEs at the cochlear base, was introduced as a perturbation, along the lines of the scheme outlined by Talmadge et al. (1998). In this scheme, the perturbative DP source function is proportional to the local value of the product of the displacement at frequency f_1 and that at frequency f_2 squared. An integral over the BM length is performed to compute

how the wavelets associated with these sources propagate back to the cochlear base and add as complex vectors (i.e., in a phase-dependent way) to generate the total DPOAE signal. As the DP source term depends on the local displacement at both primary frequencies, the generation region coincides with the so-called overlap region around $x(f_2)$. No roughness perturbation was introduced, to focus on the D component of the response.

The nonlinear model was numerically solved in the time domain, using a fast algorithm described elsewhere (e.g., Sisto et al. 2015). The dynamic nonlinearity was introduced assuming a dependence of tuning on the instantaneous and local value of the BM velocity:

$$Q_{nl}(\xi) = Q_{pass} + \left(Q_{act} - Q_{pass}\right)\left(1 - \tanh\left(\frac{\dot{\xi}^2(x,t)}{\dot{\xi}_{sat}^2}\right)\right), \quad (4)$$

where Q_{pass} and Q_{act} are the constant tuning factors of the two asymptotic linear regimes approached by the nonlinear model, respectively, at high and low stimulus levels, and $\dot{\xi}_{sat}$ is the BM velocity threshold value for the onset of nonlinear saturation phenomena.

Extracting Tuning Estimates From DPOAE Level vs. Ratio Curves

From an experimental viewpoint, one may distinguish between "direct" tuning estimates, in which the experimental output is the ratio between frequency and bandwidth of some bandpass response, and "indirect" estimates, in which tuning is estimated by measuring another physical quantity (e.g., the OAE group delay), and assuming the validity of some theoretical relation predicted by cochlear mechanics. Several methods have been proposed so far for measuring cochlear tuning, based on OAE measurements. The ultimate goal of these methods needs to be specified, in order to evaluate their performance. Behavioral tuning estimates, obtained with masking techniques, might be considered as the "gold standard" as they directly interrogate perception. However, estimates of such tuning capture any and all filtering that may occur in the auditory system. On the other hand, mechanical tuning in the cochlea is best defined by the ratio between frequency and bandwidth of the BM response to a pure tone, "directly" measured as a function of level and frequency. Behavioral tuning estimates obtained in humans are related to but not quantitatively identical to the underlying tuning properties of the BM—the differences between them arising from critical differ-

ences in measurement methods that evoke additional phenomena such as nonlinear suppression as well as the possible influence of additional filters in the auditory system.

Cochlear models provide a theoretical link between BM tuning and the properties of the OAE complex response (both amplitude and phase). Measurement of OAE suppression tuning curves (e.g., Gorga et al. 2011) perhaps most closely mimic the conditions under which behavioral tuning is measured with simultaneous masking. While nonlinear suppression shapes the outcomes of both experiments, the tuning estimates from behavioral and OAE suppression cannot be readily compared due to the lack of a valid theoretical framework connecting these measurements at different frequencies and stimulus levels. Indeed, the estimated tuning value depends on the level (relative to the maximum of the bandpass response) at which the width is measured, and the dependent variables (perceptual threshold and OAE levels) are not necessarily equivalent or proportional to each other. Other OAE-based tuning estimates (Shera et al. 2002, 2010; Moleti and Sisto 2003, 2016; Sisto and Moleti 2007; Sisto et al. 2013) are based on measurements of the phase-gradient delay, $\tau(f)$, of the OAE response by linear coherent reflection (SFOAE, TEOAE, or reflection component of DPOAE). In this case, cochlear models provide a relation between τ and BM tuning only, so one should consider that these indirect methods do not even attempt to reproduce behavioral tuning estimates, unless one assumes that BM tuning and behavioral tuning are the same quantity. Moreover, as a dependence of OAE delays (and therefore, tuning) on stimulus level has been observed (e.g., Sisto et al. 2013; Moleti and Sisto 2016), one must compare τ-derived estimates of tuning with behavioral tuning estimates at the same level of BM displacement.

As the DPOAE level vs. ratio curve shows a bandpass shape as a function of f_{DP}, it is possible to get a direct estimate of tuning from the width of this curve, e.g., by measuring the equivalent rectangular bandwidth (ERB) and computing the ratio between frequency and bandwidth. This procedure may be applied to any curve with a bandpass shape, but what one estimates this way is just the "tuning" of the bandpass curve itself, whose relation with the tuning properties of the underlying resonant system, if any, is necessarily model dependent. In the case of the DPOAE level vs. ratio curve, whether the bandpass shape reflects any filtering properties of the auditory periphery has been questioned (e.g., Van Hengel 1996; Talmadge et al. 1998; Shera 2003; Shera and

Guinan 2008). Nevertheless, a theoretical relation between the parameters of the bandpass response and the BM tuning may be used to get indirect estimates of BM tuning, either from the value of the optimal ratio of the response or from the width of the DPOAE level vs. ratio curve and BM tuning. A simple model was developed to get a heuristic relation between the parameters of the DPOAE level vs. ratio curve and BM tuning. The proposed schematization is a very simple one, chosen to highlight the physical explanation of the observed dependence on tuning of the optimal ratio and of the width of the DPOAE vs. r curve.

Assuming, for simplicity, equal amplitude (and, therefore, BM tuning Q) for the two primary tones, the spatial width at half maximum of each resonant response peak is

$$\Delta x_Q = \frac{1}{k_\omega Q} \qquad (5)$$

The width at half maximum of the superposition region between the two primaries can be found with elementary geometrical considerations (see Fig. 1) as

$$\Delta x_{over} = \Delta x_Q - \frac{\Delta x_r}{2}, \qquad (6)$$

where Δx_r represents the spatial distance between the tonotopic places of the primary tones, related to r:

$$\Delta x_r = x(f_1) - x(f_2) = \frac{r-1}{r k_\omega} = \frac{\beta}{k_\omega}, \qquad (7)$$

and

$$\beta = \frac{\Delta f}{f_2} = \frac{r-1}{r} \qquad (8)$$

With decreasing ratio, significant negative interference starts to occur as the width of the DPOAE generation region exceeds half the f_{DP} wavelength at the peak, as wavelets in phase opposition begin to cancel each other. Therefore, we predict that the maximum of the DPOAE vs. r curve should occur (approximately) for r where the width of the DPOAE generation region is equal to half the f_{DP} wavelength at the peak. The f_{DP} wavelength at the peak, for a second-order filter function, is

$$\hat{\lambda} = \frac{2\pi}{\hat{k}} = \frac{2\pi}{k_0 \sqrt{Q}} \qquad (9)$$

From Eqs. (5), (6), (7), and (9), one may estimate the ratio corresponding to the maximum DPOAE level:

$$\frac{\hat{\lambda}}{2} = \Delta x_{over} \Rightarrow \frac{\pi}{k_0 \sqrt{Q}} = \frac{1}{k_\omega Q} - \frac{\beta_{max}}{2k_\omega},$$
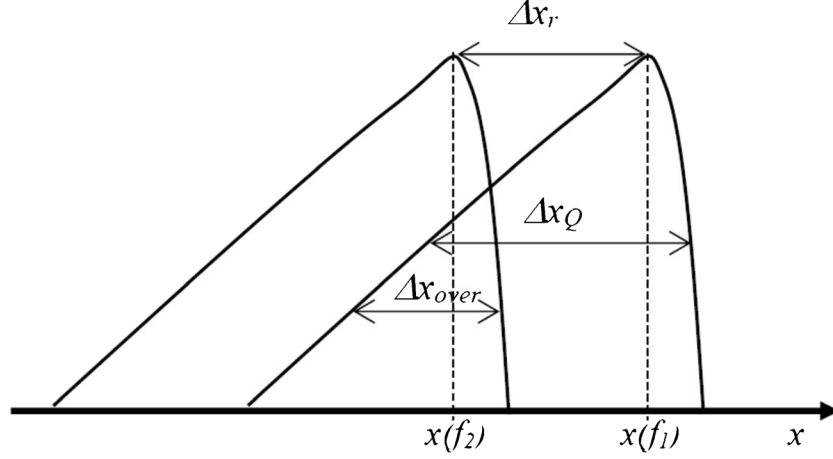
FIG. 1. Schematic view, on a linear scale, of the dependence on ratio and tuning of the width of the overlap region. Because the wavelength of the $2f_1$-f_2 traveling wave and the spatial extent of the peak region are determined by the mechanical tuning of the cochlea, whereas the width of the overlap region also depends on ratio, the main features (optimal ratio and the width) of DPOAE level vs. r curve are dependent on tuning, in a theoretically (model-dependent) predictable form.

hence,

$$\beta_{max} = \frac{2}{Q} - 2\pi \frac{k_\omega}{k_0 \sqrt{Q}} = \frac{2}{Q} - \frac{b_0}{\sqrt{Q}}, \tag{10}$$

where (using the cochlear parameters of Talmadge et al. (1998))

$$b_0 = 2\pi \frac{k_\omega}{k_0} \cong 0.28 \tag{11}$$

Solving the algebraic Eq. (10) for the square root of tuning, and keeping the positive solution, one gets a tuning estimate based on the optimal ratio:

$$Q_{peak} = \left(-\frac{b_0}{2\beta_{max}} + \sqrt{\left(\frac{b_0}{2\beta_{max}}\right)^2 + \frac{2}{\beta_{max}}} \right)^2 \tag{12}$$

A similar relationship may be found between the width of the DPOAE level vs. ratio curve and BM tuning, assuming that the minimum of the experimental curve on the low-ratio side corresponds to the condition in which the width of the generation region equals one wavelength:

$$\beta_{max} - \beta_{min} = \frac{b_0}{\sqrt{Q}}, \tag{13}$$

hence,

$$Q_{width} = \left(\frac{b_0}{\beta_{max} - \beta_{min}}\right)^2. \tag{14}$$

In the analysis of experimental curves, a Gaussian fit to the data may be useful to minimize fluctuations of the estimated tuning. In this case, the ratio corresponding to the minimum on the low-ratio side may be approximated as that at two standard deviations from the peak.

It must be evaluated here how the crudeness of the schematization of Fig. 1 could affect the quantitative estimates of tuning one obtains using Eqs. (12) and (14): Several effects should be considered indeed:

1) The f_{DP} wavelength in the generation region is longer than its value at the peak, by a factor dependent on Q and r, of order 2 at moderate ratios (e.g., for $Q = 6$ and $r = 1.2$);

2) The variation of the DPOAE phase coming from the spatial dependence of the primary tone phase within the overlap region is not negligible. This correction affects the evaluation of the "optimal" width of the overlap region.

3) The slope of the BM profile is not constant, as implicitly assumed in Fig. 1, and, more important, the effective width of the generation region is narrower than that obtained by the geometrical approach, because the DPOAE source is a cubic function of the local amplitudes at the primary frequencies, therefore much steeper than the overlap region plotted in Fig. 1. The latter effect would require correcting the width by a factor of order 2;

4) In Eq. (10), we use the same variable Q to denote tuning in the x_{DP} region (related to the wavelength) and tuning in the overlap region (related to BM response spatial widths). Moreover, if the level of the f_1 stimulus is higher than that of the f_2

stimulus, as it generally happens in experiments, Eq. (6) still holds, but with Δx_Q relative to f_1. Therefore, a frequency shift (which is not negligible for large ratios) occurs between the two terms involving tuning in Eq. (10).

Taking into account all the above uncertainties, the proposed method should be considered just as a "first-order" improvement, relative to the direct method, in which one just divides the frequency by the experimental width of the DPOAE vs. r curve. Any refinement of this method would be strongly model dependent, and would lead to much more complicated and entangled relations; therefore, we chose to use Eqs. (12) and (14) in their simplest form, without introducing any correction. The justification for this choice comes from (a) the results of the numerical simulations of the full cochlear model (see Fig. 5 in the "Results" section) and (b) the results of a semi-analytical computation, which both account for all the above-mentioned effects, to a different degree of approximation.

The proposed semi-analytical computation assumes that, in the frequency domain, the analytical form of the local DPOAE source is proportional to $\xi_1^2 \xi_2$, and that the BM response ξ is proportional to $k^{3/2}$. The wave vector function is approximated as $k(\omega, x) = \frac{k_0 \omega}{\sqrt{\omega^2(x) - \omega^2 + i\omega\gamma}}$, with $\gamma = \omega/Q$. The DPOAE phase shift between x_2 and at $x_2 - \Delta x_{\mathrm{over}}$ is computed for each frequency as the spatial integral of $Re(k)$, including also the contribution $2\phi_1 - \phi_2$, and the "optimal ratio" condition is reformulated by requesting that between x_2 and at $x_2 - \Delta x_{\mathrm{over}}$:

a) The amplitude of the DPOAE source drops by a factor of 2 and
b) The phase of the DPOAE wave changes by π.

This request yields an implicit relation between tuning and optimal ratio, which has to be inverted numerically. One eventually finds a monotonic dependence of Q on r_{max}, which is not much different from that predicted by Eq. (12) (see Fig. 2). Despite the crudeness of the scheme of Fig. 1, the different effects seem therefore to partly compensate each other, because the difference between the two curves in Fig. 2 is smaller than 20 % for r between 1.1 and 1.4. Note that the physical explanation for the pseudo-resonant curve is the same in both schematizations, and that no adjustment of the cochlear parameters was made in both cases.

Data Acquisition and Subjects

Sample DPOAE measurements were performed on six young normal hearing subjects for primary levels of 65/55- and 55/40-dB SPL (L_1/L_2), using a fixed-f_2 paradigm where f_2 was fixed at 1, 2, 4, 6, 8, 10, and 12.5 kHz while f_1 was linearly swept over 4 s to get f_2/f_1 ratios between 1 and 1.5. Responses were averaged for eight sets of measurements for each f_2 frequency. Magnitude and phase of the total ear canal DPOAEs were derived using the least squares fit method. The DPOAE response may be represented as a function of r or as a function of f_{DP}. In either case, it generally shows a bandpass pattern, with a broad peak around $r = 1.2$–1.3.

The same experiment was simulated numerically for $f_2 = 2$ kHz, using both the nonlinear model and the analytical perturbative model, to further validate Eqs. (12) and (14).

Wavelet Filtering Technique

A time-frequency filtering method based on the wavelet transform (Moleti et al. 2012; Sisto et al. 2013) was applied for separating the different DPOAE components generated by different emission mechanisms, both for the experimental data and for the model simulations. In the fixed-ratio paradigm that is generally used for the DPOAE acquisition, the component D has nearly zero phase-gradient delay while the component R has rapidly rotating phase, as a function of the DP frequency. The filtering regions for the different DPOAE components are therefore delimited by hyperbolic curves in the time-frequency domain.

RESULTS

The time-frequency representation of the DPOAE vs. f_{DP} "spectra" show a distortion component with approximately zero delay, along with a second component with greater delay. The delay of this second component decreases approximately as $1/f_2$. This pattern is visible in the experimental data displayed in Fig. 3 (top), for the same subject, for two different values of f_2 (2 and 6 kHz).

After time-frequency filtering, the D component of the DPOAE response is plotted as a function of r in Fig. 3 (bottom) for the same data. The bandpass shape of the response is evident, sufficiently well-fit by a Gaussian function of frequency (dashed line). The data from this subject exhibit two general features: both the optimal ratio and the width of the curve decrease with increasing frequency, consistent with an increase in sharpness of BM tuning (Eqs. (12) and (14)).

The D component of the DPOAE response is also plotted against r for two subjects in Fig. 4 (top) for $f_2 =$
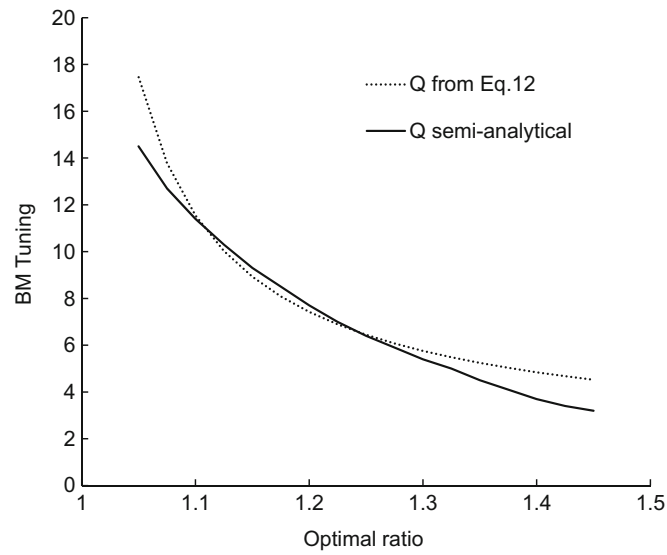
FIG. 2. Comparison between the relation between the optimal ratio and tuning predicted by the rough schematization of Fig. 1 (Eq. (12)), and by a more accurate implicit relation between the two physical quantities, based on a semi-analytical computation

2 kHz, for two different stimulus levels. One may note that, as the stimulus level increases, both the optimal

ratio and the curve width increase, suggestive of decreasing sharpness of tuning. The D component
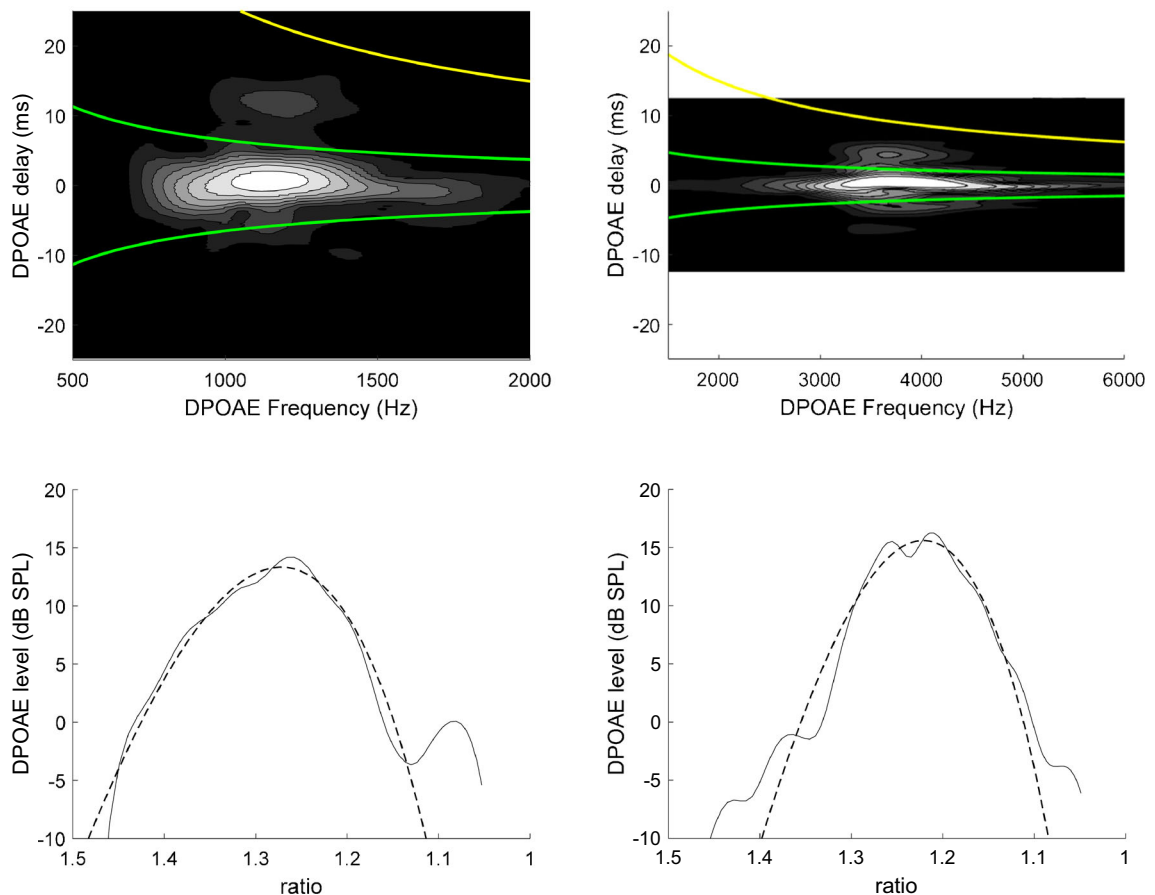
FIG. 3. *top* Time-frequency representation of a human DPOAE vs. ratio response, for $f_2 = 2$ kHz (*left*) and 6 kHz (*right*). The stimulus level is 65- and 55-dB SPL for f_1 for f_2, respectively. The green lines delimit the filtering region of the D component, while the yellow line is the upper delay limit for single-reflection R components. *bottom*

Corresponding D components plotted as a function of ratio. With increasing frequency, both the optimal ratio and the width of the curve decrease, which, according to Eqs. (12) and (14), mean increasing cochlear tuning

obtained from the nonlinear model at $f_2 = 2$ kHz is also shown in Fig. 4 (bottom left) as a function of ratio, for several stimulus levels. As in the experimental data, the maximum is shifted to larger ratios and the bandwidth increases as the stimulus level increases, probably related to lowering of effective tuning by nonlinear saturation. A very similar behavior is obtained (Fig. 4, bottom right) with the analytical linear model, in which tuning is explicitly changed as an external parameter.

The theoretical relations between cochlear tuning, optimal ratio, and width of the DPOAE vs. ratio curve were tested on simulations obtained with the linear model. Cochlear tuning was directly estimated from the width of the computed BM response to a pure tone, and compared in Fig. 5 to the theoretical predictions of Eq. (12) (squares) and to that of Eq. (14) (circles). The overall agreement is quite satisfactory, considering the crudeness of the geometrical schematization leading to Eqs. (12) and (14), and the fact that no parameter has been tuned

in the model. This agreement suggests that the tuning of the experimental DP vs. ratio curves can be directly used to estimate the sharpness of BM tuning.

In Fig. 6, we report the average of the tuning estimates obtained for six subjects as a function of frequency from the DPOAE vs. ratio curves at 65–55-dB SPL stimulus level, using three different methods. Open circles represent tuning computed from the optimal ratio, using Eq. (12); open diamonds represent tuning computed from the optimal and minimal ratio, using Eq. (14); and full squares represent tuning computed with a direct method, by dividing the frequency of the maximum by the bandwidth of the DPOAE vs. f_{DP} curves. In any case, the method was applied to a Gaussian fit to the actual experimental curve. The three methods yield similar results, the second one being affected by larger numerical instability (due to the presence, in the denominator of Eq. (14), of the difference between two terms of the same order). Tuning
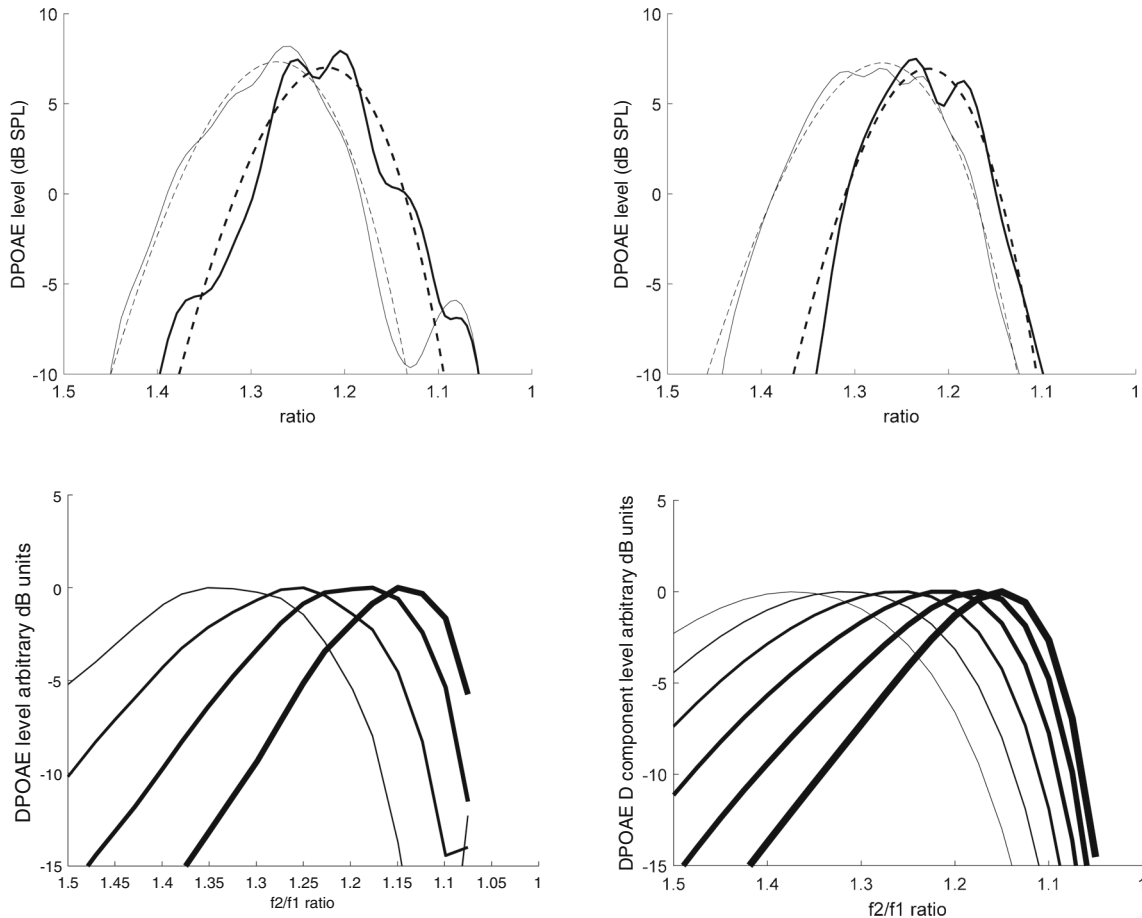
FIG. 4. *top* Filtered DPOAE D component as a function of *r*, fitted to a Gaussian profile, for $f_2 = 2$ kHz for two different stimulus levels (65–55- and 55–40-dB SPL, thin and thick lines, respectively) from two different ears. *bottom left* Dependence on *r* of the DPOAE component D in the nonlinear model solved in the time domain, for $f_2 = 2$ kHz and for four different stimulus levels of L_2 between 40- and 55-dB SPL in 5-dB steps (lines of decreasing thickness). In the analytical linear model (*bottom right*), the overall tuning was varied between 3 and 8 with unit steps (lines of increasing thickness)
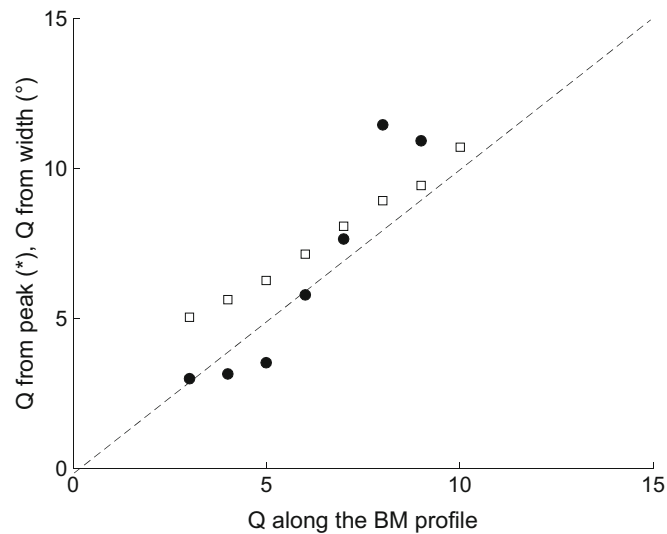
FIG. 5. Correlation between Q measured on the BM in the linear model and that estimated either from the width (Eq. (14), filled circles) or from the peak (Eq. (12), open squares) of the DPOAE vs. *r* response, for $f_2 = 2$ kHz. The model tuning was varied between 4 and 10

generally increases with frequency, reaching a flat maximum plateau around 6–10 kHz. The fact that the absolute value of the direct tuning estimate is almost coincident with that obtained using Eq. (12) was not expected, because no adjustment had been applied to the parameters of Eq. (12), directly taken from Talmadge et al. (1998). In the "Methods" section, the difficulty of comparing tuning estimates obtained with different techniques had been highlighted. In the present case, the crudeness of the theoretical assumptions underlying Eqs. (12) and (14) suggests an even more cautious interpretation of such comparisons. Nevertheless, it may be noted

that the behavioral estimates by Glasberg and Moore (1990), obtained with simultaneous masking, approximately represented by the dashed-dotted line, are not inconsistent with the results of the present study. The OAE-based estimates by Shera et al. (2002), obtained from SFOAE phase-gradient delay data, have similar frequency dependence and sharper tuning, in agreement with behavioral tuning estimates by the same authors (Oxenham and Shera 2003; not reported in Fig. 6), obtained with forward masking. On the other hand, the tuning estimates by Moleti and Sisto (2016), obtained from the phase-gradient delay of the DPOAE reflection
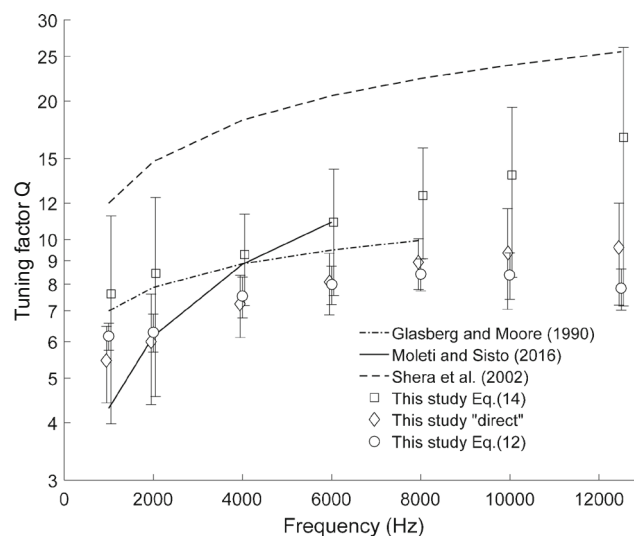
FIG. 6. Comparison between three different average tuning estimates obtained for six subjects from the DPOAE vs. ratio curve using Eqs. (12) and (14), and directly dividing the frequency of the peak by the bandwidth of the DPOAE vs. ratio curve. Error bars represent one standard deviation. For reference, two OAE-based tuning estimates, and a behavioral tuning estimate, are also shown. A small frequency shift, not present in the data, was added to enhance visibility of the error bars

components at the same stimulus level (55–65-dB SPL), are comparable in sharpness but show a steeper slope as a function of frequency.

DISCUSSION

Overall, the DPOAE level vs. r functions obtained experimentally, and shown in Figs. 3 and 4, are more symmetrical than those predicted by both numerical and analytical models, which also tend to overestimate the shift of the maximum and the change in width with stimulus level. Nevertheless, both the shift of the maximum and the variation of the width are qualitatively well predicted by both models. The agreement between the experimental results and the analytical linear model, in which two-tone suppression effects are obviously not accounted for, suggests that interference phenomena play indeed a very crucial role in determining the behavior of the DPOAE vs. r curves. This behavior could be mostly attributed to a compromise between the positive effect of increasing the width of the overlap region and the negative interference among DP wavelets of different phases within the generation region. These observations further suggest that the shape of the DPOAE vs. r curves depends on a single parameter, i.e., the mechanical tuning of the BM, which is explicitly changed in the linear model, and changes with stimulus level in the nonlinear model. Therefore, the results of this study strongly support the possibility of using DPOAE vs. r curves for objective estimates of cochlear tuning. Interesting, in Liu and Neely (2010), bandpass DPOAE vs. r curves were obtained using a nonlinear cochlear model, showing the same dependence on stimulus level as demonstrated here, which we attribute to the dynamical change of tuning of the cochlear amplifier—an explanation applicable to their results as well.

A few studies have explored the DPOAE stimulus parameter range in detail, varying the f_2 frequency, r, and both primary levels, L_1 and L_2 (Kummer et al. 1998, 2000; Johnson et al. 2006). In particular, Johnson et al. (2006) performed extensive measurements of the DPOAE response at selected frequencies $f_2 = 1, 2, 4, 8$ kHz as a function of three parameters: L_1, L_2, and r. In such studies, the nonmonotonic dependence of the DPOAE level on L_1, at fixed L_2 and r, was interpreted as the result of nonlinear suppression of the f_2 response by the f_1 tone. Although one cannot rule out a significant role of suppression phenomena, a possible alternative explanation of this phenomenology is that the optimal ratio condition $\frac{\hat{\lambda}}{2} = \Delta x_{over}$ is

fulfilled, at a fixed ratio, for a specific level L_1, according to a variation of our simple "geometrical" model in which the condition $L_1 = L_2$ is relaxed.

The crudeness of the geometrical model used to estimate the relationship between BM tuning and the parameters of the DPOAE vs. r curve suggests that the tuning estimates obtained this way could be affected by rather large systematic errors. Another uncertainty comes from that on the numerical values of the cochlear parameters k_0 and k_ω, which, in this study, have been taken from Talmadge et al. (1998). On the other hand, the proposed methods have been shown to be in rather good agreement with numerical simulations in which tuning was a controlled parameter, based on more realistic shapes of the BM response. The proposed method could be further refined and validated by animal experiments, in which a direct comparison between such tuning estimates and those obtained from direct measurements of the BM response width would be possible.

Other OAE-based techniques for estimating BM tuning, such as those based on SFOAE or TEOAE group delay (e.g., Shera et al. 2002; Sisto et al. 2013; Moleti and Sisto 2016), have been proposed. The comparison shown in Fig. 6 between the tuning dependence on frequency obtained with behavioral techniques, with OAE-delay-based techniques, and with those based on our analysis of the DPOAE vs. ratio curves, shows that the latter are in rather satisfactory agreement with behavioral tuning estimates obtained with simultaneous masking. On the other hand, Fig. 6 also shows how different theoretical assumptions significantly affect the steepness of the estimated dependence of tuning on frequency. It may be profitable to focus on the frequency dependence estimates, as comparison of absolute tuning values is difficult, for the previously discussed reasons. For example, Moleti and Sisto (2016) seem to overestimate this steepness, probably because the assumption that the wavelength in the peak region scales as the square root of tuning is strictly valid in a long-wave limit only. This assumption is present also in the present model (Eq. (9)), which could slightly affect the steepness of the tuning dependence on frequency estimated by Eq. (14) also, and to a lesser extent, that estimated following Eq. (12). Generally speaking, the methods based on the DPOAE vs. r curve are less sensitive to this assumption because they measure the DPOAE D component, whose generation source amplitude and phase is also sensitive to the f_1 primary tone, for which the long-wave approximation is more accurate in the DP generation region. Equation (12) is even less sensitive to this assumption, because tuning enters in two terms of Eq. (10), and only one of them contains the long-wave assumption. This brief discussion was meant to

emphasize that, in addition to opening the possibility of using different independent techniques to get objective estimates of cochlear tuning, such comparisons are also important to validate or falsify the models of the cochlea that provide the theoretical basis for the proposed methods.

The fact that the DPOAE R component typically shows a less sharply peaked dependence on ratio, or even a monotonic dependence (see, e.g., Botti et al. 2016), further supports the above interpretation of the behavior of the D component. Indeed, the beamforming mechanism (Shera and Guinan 2007) predicts that the phase shifts producing negative interference for the D response would be canceled in the residual path of the forward DP wave that is partially reflected in the $x(f_{DP})$ region to generate the R component.

Although the theoretical explanation for the bandpass shape of the DPOAE vs. ratio curves that is provided in this study is consistent with theoretical models and experimental data, an alternative (or complementary) explanation, based on the existence of a mechanical second filter, cannot be ruled out, also because its predictions are not clearly stated. It would be useful to design experiments capable of discriminating between the two hypotheses.

CONCLUSIONS

The peaked dependence of the DPOAE (D component) level on the primary frequency ratio, considered as evidence for the existence of a second cochlear filter, could alternatively be explained by a simple linear interference phenomenon: with decreasing ratio, as the width of the overlap region increases, destructive interference causes the decrease of the DPOAE D component response. A simple model was developed that assumes that destructive interference occurs as the overlap region starts exceeding half the local wavelength of the distortion product. On the basis of this simple geometrical model, a relation was derived between the ratio corresponding to the maximum of the DPOAE vs. ratio curve and the tuning factor along the BM. A similar relation holds between tuning and the width of the same curve. The width of the DP vs. ratio curves was estimated using a perturbative linear model and a nonlinear model solved in time domain, confirming the relation with the width of the BM profile obtained with the heuristic geometrical model. These findings also support the idea of using measurements of the experimental DPOAE vs. ratio curve for objective cochlear tuning estimates in humans. The application of the method to human subjects yields tuning values

(and dependence on frequency and stimulus level) that are in reasonable agreement with previous tuning estimates.

ACKNOWLEDGMENTS

REFERENCES

Allen JB (1980) Cochlear micromechanics–a physical model of transduction. J Acoust Soc Am 68:1660–1670

Allen JB, Fahey PF (1992) Using acoustic distortion products to measure the cochlear amplifier gain on the basilar membrane. J Acoust Soc Am 92:178–188

Botti T, Sisto R, Sanjust F, Moleti A, D'Amato L (2016) Distortion product otoacoustic emission generation mechanisms and their dependence on stimulus level and primary frequency ratio. J Acoust Soc Am 139:658–673

Brown AM, Gaskill SA, Williams DM (1992) Mechanical filtering of sound in the inner ear. Proc Biol Sci 250:29–34

Brown AM, Gaskill SA, Carlyon RP, Williams DM (1993) Acoustic distortion as a measure of frequency selectivity: relation to psychophysical equivalent rectangular bandwidth. J Acoust Soc Am 93:3291–3297

Glasberg BR, Moore BC (1990) Derivation of auditory filter shapes from notched-noise data. Hear Res 47:103–138

Gorga MP, Neely ST, Kopun J, Tan H (2011) Distortion-product otoacoustic emission suppression tuning curves in humans. J Acoust Soc Am 129:817–827

Johnson TA, Neely ST, Garner CA, Gorga MP (2006) Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions. J Acoust Soc Am 119:418–428

Kummer P, Janssen T, Arnold W (1998) The level and growth behavior of the 2 f1-f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss. J Acoust Soc Am 103:3431–3444

Kummer PJ, Janssen T, Hulin P, Arnold W (2000) Optimal L(1)-L(2) primary tone level separation remains independent of test frequency in humans. Hear Res 146:47–56

Liu Y, Neely ST (2010) Distortion product emissions from a cochlear model with nonlinear mechanoelectrical transduction in outer hair cells. J Acoust Soc Am 127:2420–2432

Moleti A, Longo F, Sisto R (2012) Time-frequency domain filtering of evoked otoacoustic emissions. J Acoust Soc Am 132:2455–2467

Moleti A, Sisto R (2016) Estimating cochlear tuning dependence on stimulus level and frequency from the delay of otoacoustic emissions. J Acoust Soc Am 140:945–959

Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J Assoc Res Otolaryngol 4:541–554

Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. Physiol Rev 81:1305–1352

Shera CA (2001) Intensity-invariance of fine time structure in basilar-membrane click responses: implications for cochlear mechanics. J Acoust Soc Am 110:332–348

Shera CA (2003) Wave interference in the generation of reflection- and distortion-source emissions. In: Gummer AW (ed) Biophysics of the cochlea: from molecules to models. World Scientific, Singapore, pp 439–453

SHERA CA, GUINAN JJJ (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am 105:782–798

SHERA CA, GUINAN JJJ (2007) Cochlear traveling-wave amplification, suppression, and beamforming probed using noninvasive calibration of intracochlear distortion sources. J Acoust Soc Am 121:1003–1016

SHERA CA, GUINAN JJ (2008) "Mechanisms of mammalian otoacoustic emission," in: Manley GA, Fay RR, Popper AN (eds), Active Processes and Otoacoustic Emissions. New York: Springer, 305–342

SHERA CA, GUINAN JJ JR, OXENHAM AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci U S A 99:3318–3323

SISTO R, MOLETI A, ALTOÈ A (2016) Decoupling the level dependence of the basilar membrane gain and phase in nonlinear cochlea models. J Acoust Soc Am 138:EL155–EL160. https://doi.org/10.1121/1.4928291

SISTO R, MOLETI A, SHERA CA (2015) On the spatial distribution of the reflection sources of different latency components of otoacoustic emissions. J Acoust Soc Am 137:768–776

SISTO R, SANJUST F, MOLETI A (2013) Input/output functions of different-latency components of transient-evoked and stimulus-frequency otoacoustic emissions. J Acoust Soc Am 133:2240–2253

TALMADGE CL, TUBIS A, LONG GR, PISKORSKI P (1998) Modeling otoacoustic emission and hearing threshold fine structures. J Acoust Soc Am 104:1517–1543

VAN HENGEL PWJ (1996) Emissions from cochlear modelling. In: PhD thesis, Rijksuniversiteit Groningen

ZWEIG G (1991) Finding the impedance of the organ of Corti. J Acoust Soc Am 89:1229–1254