Stimulus-spectrum irregularity and the generation of evoked and spontaneous otoacoustic emissions: Comments on the model of Nobili et al.

Christopher A. Shera
Eaton-Peabody Laboratory of Auditory Physiology
Massachusetts Eye & Ear Infirmary
243 Charles Street
Boston, MA 02114, USA

Department of Otology & Laryngology
Harvard Medical School
Boston, MA 02115, USA

Arnold Tubis
Department of Physics
Purdue University
West Lafayette, IN 47907, USA

Institute for Nonlinear Science
University of California, San Diego
La Jolla, CA 92093, USA

Carrick L. Talmadge
National Center for Physical Acoustics
University of Mississippi
University, MI 38677, USA

Abstract

Nobili and colleagues (2003, *J. Assoc. Res. Otolaryngol.*, **4**:478–494) propose that transient evoked otoacoustic emissions (TEOAEs) result from spatially complex "residual oscillations" of the basilar membrane that trace their origin to spectral irregularities in the forward middle-ear transfer function. In this paper we comment on Nobili et al.'s model and conclusions while trying to clarify some of the broader issues they raise. Although Nobili et al.'s published OAE simulations are of uncertain reliability, simple arguments that do not depend on solving the model equations establish that their proposed middle-ear filtering mechanism conflicts with basic experimental findings about TEOAEs. Furthermore, the proposed mechanism cannot produce spontaneous (SOAEs) or stimulus-frequency emissions (SFOAEs) at any level of stimulation. Models of TEOAEs, SFOAEs, and SOAEs based on wave reflection due to scattering by impedance perturbations in the mechanics of the cochlear partition suffer none of these deficiencies.

I. Introduction

Ever since their discovery, transient-evoked otoacoustic emissions (TEOAEs) have generally been ascribed to the reflection of cochlear traveling-wave energy by mechanical impedance perturbations arrayed (or induced) in various ways along the cochlear partition (e.g., Kemp, 1978; Manley, 1983; Ruggero et al., 1983; Sutton and Wilson, 1983; Zwicker, 1986; Furst and Lapid, 1988; Strube, 1989; but see Yates and Withnell, 1999). Nobili and colleagues, however, have recently proposed a new mechanism for generating TEOAEs (Nobili, 2000; Nobili et al., 2003a,b). As an alternative to scattering by mechanical perturbations, Nobili et al. suggest that TEOAEs result from prolonged "residual oscillations" of the basilar membrane (BM) that trace their origin to spectral irregularities in middle-ear transmission.

In their model simulations, residual BM oscillations and TEOAEs appear when the evoking stimulus has an intensity sufficient to partially saturate the nonlinear amplification mechanisms within the cochlea and a frequency spectrum irregular enough to produce a complex spatial vibration pattern along the basilar membrane. They suggest that most sounds, no matter how smooth their frequency spectra appear in the ear canal, acquire the necessary spectral irregularity simply by passing through the middle ear, from which they inherit the spectral features characteristic of middle-ear transfer functions. Nobili et al. also find—in agreement with many others before them—that simulated TEOAEs can be produced by introducing small mechanical perturbations into the equations representing the organ of Corti. In their model, however, these mechanical perturbations often create spontaneous emissions (SOAEs), an emission type not invariably associated with TEOAEs. Nobili et al. therefore argue that "when found in the absence of spontaneous emissions, transient evoked OAEs are mainly attributable to the characteristics of forward middle-ear filtering."

Nobili et al. show several computer simulations but provide relatively little comparison between theory and experiment. This paper grew out of our attempt to understand whether the predictions of the proposed mechanism were truly "in impressive accord with experimental data," as claimed. Evaluating the model in this way requires knowing what the model predicts. As we demonstrate below, however, Nobili et al. have left their readers unable to determine which features of the OAE-like oscillations evident in the simulations are actually predicted by the model and which features may result from unrecognized artifacts of the computation. Nevertheless, by identifying qualitative predictions of the proposed middle-ear filtering mechanism that can be deduced without recourse to numerical modeling, we sidestep the uncertainties surrounding Nobili et al.'s simulations and demonstrate that their model fails to reproduce basic empirical properties of actual evoked OAEs.

II. Accuracy of the Model Simulations

Although Nobili et al. solve their time-domain equations of motion using a computational technique better known for its conceptual simplicity than for its numerical accuracy (e.g., Press et al., 1992; Diependaal et al., 1987), they report no checks of the validity of their model simulations. An elementary test—necessary but not sufficient to ensure the integrity of the computation—would be to verify that the purported solution does not change significantly when the integration step size is decreased. We performed such a test using the program Nobili et al. published on the internet to allow readers to simulate emissions using the model (Nobili, 2003). Although limitations of the program precluded a definitive analysis, our preliminary results were not reassuring.

To rebut our analysis, Nobili and Mammano performed an authoritative test as part of their review of a previous version of this manuscript. At one location along the BM they computed the model response to an acoustic click applied at the eardrum and compared the answer obtained using their standard integration step size to a computation performed when the temporal resolution was increased by a factor of 2. As Nobili and Mammano point out in their review, the two oscillatory responses are qualitatively similar in appearance. However, when the responses are overlaid on the same graph, large quantitative differences immediately become apparent. The two curves are clearly distinguishable as early as 3 ms after stimulus onset, and by 10 ms the waveforms are more than 180 degrees out of phase with one another. The presence of such large quantitative discrepancies at times coincident with the appearance of TEOAE-like oscillations in their simulations demonstrates that Nobili et al.'s published otoacoustic responses cannot be reliable solutions of their model equations.

The tests described above explore only the accuracy of the time-domain integration. Nobili et al. also necessarily discretized the spatial coordinate in their model. To approximate the desired spatial integration they divided the cochlear partition into 500 longitudinal segments and summed the results from each section, weighting each response by a numerical approximation to the local hydrodynamic Green's function. Just as with the temporal integration, employing too coarse a grid can lead to spurious results. Unfortunately, Nobili et al. again present no checks on the validity of their procedure.

Since the optimal grid spacing depends on both the numerical algorithms employed and the size of the acceptable error, there are no hard and fast rules for determining the number of required sections. In the context of modeling OAEs, a reasonable lower bound might be the number ($N_{\min}$) necessary to represent the spatial frequencies important for emission generation. The theory of coherent reflection filtering yields the estimate $N_{\min} \sim 8L/\hat{\lambda}$ sections, where L is the cochlear length and $\hat{\lambda}$ is the wavelength at the traveling-wave peak for the frequency of interest. [Note that $8 = 4 \times 2$, where the factor of 4 is needed to encompass the range of spatial frequencies falling within the pass-band of the "spatial-frequency filter" (e.g., Zweig and Shera, 1995, Fig. 6) and the additional factor of 2 arises from Nyquist's sampling theorem.] Using estimates of $\hat{\lambda}$ for the human cochlea obtained from measurements of SFOAEs (Shera and Guinan, 2003, Table II) yields $N_{\min} \sim 650$ sections for a model that matches human SFOAE group delays over the full range of human hearing. Nobili et al. use a nonuniform grid spacing with a number of sections roughly equivalent to this lower bound. Whether any particular grid spacing suffices in practice can only be determined by detailed numerical analysis. As with the time-domain integration, decreasing the grid spacing until the solution no longer varies on scales relevant to the issues at hand often provides a useful empirical assay. Our experience modeling OAEs indicates that the necessary number of sections can sometimes be significantly greater than the lower bound estimated above. Talmadge et al. (1998), for example, found that obtaining reliable solutions can require as many as 4000 sections, considerably more than the number used by Nobili et al.

The computational dangers are especially acute when simulating OAEs in active cochlear models. Active models propagate and amplify numerical errors much as they do actual responses to the stimulus. Once they appear, small errors can grow rapidly and thereafter masquerade as genuine otoacoustic responses. Since relative OAE amplitudes are often quite small—human TEOAEs and SFOAEs are typically 10–100 times smaller than the stimulus—computational procedures that suffice when solving solely for the primary response to the stimulus may fail completely when calculating OAEs.

III. Qualitative Tests of the Model

The unresolved computational issues discussed above imply that the characteristics of any OAEs predicted by Nobili et al.'s proposed mechanism remain largely unknown. It is therefore difficult to compare the model's predictions with measured OAEs and determine whether the responses are indeed "strikingly similar to those well known to audiologists." For example, measured TEOAEs are generally rather dispersive, meaning that their waveforms exhibit a decrease over time in their instantaneous frequency of oscillation (e.g., Kemp, 1978). The TEOAE waveforms computed from the model, however, show little evidence of this phenomenon (see Nobili et al. 2003b, Fig. 1e′). Unfortunately, the reader cannot easily determine whether the apparent absence of frequency dispersion reflects a shortcoming of the model or whether it arises as an artifact of inaccurate numerical methods.

Despite limitations such as these on any quantitative comparison between theory and experiment, several important qualitative predictions of the proposed middle-ear filtering mechanism can be deduced without solving the model equations. It is easy to demonstrate, for example, that the mechanism cannot account for stimulus-frequency OAEs (SFOAEs): Although TEOAEs are evoked by transient stimuli containing many frequency components, and are therefore potentially sensitive to frequency variations in middle-ear transmission as proposed, SFOAEs are evoked by pure (single-frequency) tones and, *ipso facto*, cannot originate via any mechanism that operates *across* frequency. Since SOAE bandwidths are much smaller than any significant variation in middle-ear transfer functions, similar remarks apply to the generation of spontaneous emissions.

Because their proposed mechanism can produce neither SFOAEs nor SOAEs, Nobili et al. are forced to introduce mechanical perturbations along the BM, a model for which there is longstanding precedent. Compelled to invoke two different mechanisms to explain the appearance of TEOAEs, SFOAEs, and SOAEs, Nobili et al. conclude that "there are at least two main sources of OAEs in the cochlea: one related to CA [cochlear-amplifier] gain irregularities [i.e., mechanical perturbations] and the other to middle-ear characteristics." But since mechanical perturbations, by themselves, can give rise to all three emission types (e.g., Shera and Zweig, 1993b; Talmadge and Tubis, 1993; Zweig and Shera, 1995; Talmadge et al., 1998), the proposed middle-ear filtering mechanism may be entirely superfluous.

If it operates at all, the proposed mechanism is clearly limited to high sound pressure levels. Nobili et al. remark that a linearized version of their model, identical to the nonlinear model at low sound pressure levels, yields unmeasurable TEOAEs. But any emission mechanism that relies on driving the cochlear amplifier into saturation is inconsistent with some of the key phenomenology of evoked OAEs. Both TEOAEs and SFOAEs can be measured at sound-pressure levels near threshold (i.e., far below saturation levels and in the regime of near-linear BM amplification). Indeed, it is at near-threshold stimulus levels that the amplitudes of both TEOAEs and SFOAEs are *largest* relative to the evoking stimulus (e.g., Kemp, 1978; Wilson, 1980; Zwicker and Schloth, 1984; Shera and Zweig, 1993a). Rather than rapidly disappearing to zero, as Nobili et al.'s model predicts, TEOAE amplitudes actually *grow* relative to the stimulus as sound-pressure levels decrease below 30–40 dB SPL.

Nobili et al.'s proposed mechanism cannot reproduce this basic empirical finding. If the mechanism worked in the linear regime near the threshold of hearing, then the otoacoustic response to any transient (i.e., TEOAEs) could be synthesized by superposition from responses to pure tones (i.e., SFOAEs). But, as discussed above, the proposed mechanism cannot produce SFOAEs at any level of stimulation; as a result, the model cannot produce TEOAEs at low sound levels, in clear contradiction with experimental data.

Although this is hardly the place to rehearse the arguments in detail (for which see, e.g., Zweig and Shera, 1995; Talmadge et al., 1998; Shera and Guinan, 1999; Talmadge et al., 2000; Shera, 2003), we note that models that trace the origin of reflection-source OAEs to scattering by mechanical perturbations suffer none of these deficiencies. Indeed, these models provide both a unified framework for exploring the generation of TEOAEs, SFOAEs, and SOAEs (as well as DPOAE fine-structure) and a successful, quantitative account of what Zwicker and Schloth (1984) once characterized as the manifold "interrelations of different otoacoustic emissions."

IV. Summary

By circumventing uncertainties about the numerical accuracy of their published simulations, we have demonstrated that Nobili et al.'s proposed middle-ear filtering mechanism fails to reproduce basic empirical properties of actual evoked OAEs (e.g., persistence of TEOAEs at low stimulation levels, relations between TEOAEs and SFOAEs, etc). In addition, their model cannot produce both TEOAEs and SFOAEs (or SOAEs) unless the model is supplemented with mechanical perturbations along the BM; in this case, Nobili et al.'s results corroborate findings well established in the literature (e.g, Sutton and Wilson, 1983; Zwicker, 1986; Furst and Lapid, 1988; Strube, 1989; Shera and Zweig, 1993b; Zweig and Shera, 1995; Talmadge et al., 1998).

Acknowledgments

We thank Egbert de Boer, Paul F. Fahey, John J. Guinan, Jr., Radha Kalluri, and Robert H. Withnell for valuable discussions and suggestions. Renato Nobili, Fabio Mammano, and a third, anonymous reviewer provided helpful comments on the manuscript. This work was supported by grants R01 DC03687 and R29 DC03094 from the NIDCD, National Institutes of Health.

References

Diependaal, R. J., H. Duifhuis, H. W. Hoogstraten, and M. A. Viergever (1987). Numerical methods for solving one-dimensional cochlear models in the time domain. *J. Acoust. Soc. Am. 82*, 1655–1666.

Furst, M. and M. Lapid (1988). A cochlear model for acoustic emissions. *J. Acoust. Soc. Am. 84*, 222–229.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *J. Acoust. Soc. Am. 64*, 1386–1391.

Manley, G. A. (1983). Frequency spacing of acoustic emissions: A possible explanation. In W. R. Webster and L. M. Aitkin (Eds.), *Mechanisms of Hearing*, pp. 36–39. Clayton, Australia: Monash University Press.

Nobili, R., A. Vetešník, L. Turicchia, and F. Mammano (2003b). Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model. *J. Assoc. Res. Otolaryngol. 4*, 478–494.

Nobili, R., A. Vetešník, L. Turicchia, and F. Mammano (2003a). Otoacoustic emissions simulated in the time domain by a hydrodynamic model of the human cochlea. In A. W. Gummer (Ed.), *Biophysics of the Cochlea: From Molecules to Models*, pp. 524–530. Singapore: World Scientific.

Nobili, R. (2000). Otoacoustic emissions simulated by a realistic cochlear model. In H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (Eds.), *Recent Developments in Auditory Mechanics*, pp. 402–408. Singapore: World Scientific.

Nobili, R. (2003). Simoae at http://www.pd.infn.it/∼rnobili/simoae/simoae.htm (as of 29 July 2003).

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge: Cambridge University Press.

Ruggero, M. A., N. C. Rich, and R. Freyman (1983). Spontaneous and impulsively evoked otoacoustic emissions: Indicators of cochlear pathology? *Hear. Res. 10*, 283–300.

Shera, C. A. and J. J. Guinan (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. *J. Acoust. Soc. Am. 105*, 782–798.

Shera, C. A. and J. J. Guinan (2003). Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. *J. Acoust. Soc. Am. 113*, 2762–2772.

Shera, C. A. and G. Zweig (1993a). Noninvasive measurement of the cochlear traveling-wave ratio. *J. Acoust. Soc. Am. 93*, 3333–3352.

Shera, C. A. and G. Zweig (1993b). Order from chaos: Resolving the paradox of periodicity in evoked otoacoustic emission. In H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (Eds.), *Biophysics of Hair Cell Sensory Systems*, pp. 54–63. Singapore: World Scientific.

Shera, C. A. (2003). Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves. *J. Acoust. Soc. Am. 114*, 244–262.

Strube, H. W. (1989). Evoked otoacoustic emissions as cochlear Bragg reflections. *Hear. Res. 38*, 35–45.

Sutton, G. T. and J. P. Wilson (1983). Modelling cochlear echoes: The influence of irregularities in frequency mapping on summed cochlear activity. In E. Boer and M. A. Viergever (Eds.), *Mechanics of Hearing*, pp. 83–90. The Hague: Martinus Nijhoff.

Talmadge, C. L., A. Tubis, G. R. Long, and P. Piskorski (1998). Modeling otoacoustic emission and hearing threshold fine structures. *J. Acoust. Soc. Am. 104*, 1517–1543.

Talmadge, C. L., A. Tubis, G. R. Long, and C. Tong (2000). Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure. *J. Acoust. Soc. Am. 108*, 2911–2932.

Talmadge, C. L. and A. Tubis (1993). On modeling the connection between spontaneous and evoked otoacoustic emissions. In H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (Eds.), *Biophysics of Hair Cell Sensory Systems*, pp. 25–32. Singapore: World Scientific.

Wilson, J. P. (1980). Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus. *Hear. Res. 2*, 233–252.

Yates, G. K. and R. H. Withnell (1999). The role of intermodulation distortion in transient-evoked otoacoustic emissions. *Hear. Res. 136*, 49–64.

Zweig, G. and C. A. Shera (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J. Acoust. Soc. Am. 98*, 2018–2047.

Zwicker, E. and E. Schloth (1984). Interrelation of different oto-acoustic emissions. *J. Acoust. Soc. Am. 75*, 1148–1154.

Zwicker, E. (1986). ‘Otoacoustic’ emissions in a nonlinear cochlear hardware model with feedback. *J. Acoust. Soc. Am. 75*, 154–162.