JARO 16: 679–694 (2015)
DOI: 10.1007/s10162-015-0543-7

JARO
Journal of the Association for Research in Otolaryngology

Research Article

Stimulus Frequency Otoacoustic Emission Delays and Generating Mechanisms in Guinea Pigs, Chinchillas, and Simulations

Maria A. Berezina-Greene[1,2] and John J. Guinan Jr.,[1,2,3]

[1]*Eaton-Peabody Lab, Mass. Eye and Ear Infirmary, 243 Charles St, Boston, MA 02114, USA*
[2]*Harvard-MIT HST Speech and Hearing Bioscience and Technology Program, Cambridge, MA, USA*
[3]*Harvard Medical School, Boston, MA, USA*

Received: 12 January 2015; Accepted: 30 August 2015; Online publication: 15 September 2015

ABSTRACT

According to coherent reflection theory (CRT), stimulus frequency otoacoustic emissions (SFOAEs) arise from cochlear irregularities coherently reflecting energy from basilar membrane motion within the traveling-wave peak. This reflected energy arrives in the ear canal predominantly with a single delay at each frequency. However, data from humans and animals indicate that (1) SFOAEs can have multiple delay components, (2) low-frequency SFOAE delays are too short to be accounted for by CRT, and (3) "SFOAEs" obtained with a 2nd ("suppressor") tone ≥2 octaves above the probe tone have been interpreted as arising from the area basal to the region of cochlear amplification. To explore these issues, we collected SFOAEs by the suppression method in guinea pigs and time-frequency analyzed these data, simulated SFOAEs, and published chinchilla SFOAEs. Time-frequency analysis revealed that most frequencies showed only one SFOAE delay component while other frequencies had multiple components including some with short delays. We found no systematic patterns in the occurrence of multiple delay components. Using a cochlear model that had significant basilar membrane motion only in the peak region of the traveling wave, simulated SFOAEs had single and multiple delay components similar to the animal SFOAEs. This result indicates that multiple components (including ones with short delays) can originate from cochlear mechanical irregularities in the SFOAE peak region and are not necessarily indicative of SFOAE sources in regions ≥2 octaves basal of the SFOAE peak region. We conclude that SFOAEs obtained with suppressors close to the probe frequency provide information primarily about the mechanical response in the region that receives amplification, and we attribute the too-short SFOAE delays at low frequencies to distortion-source SFOAEs and coherent reflection from multiple cochlear motions. Our findings suggest that CRT needs revision to include reflections from multiple motions in the cochlear apex.

Keywords: Coherent reflection, SFOAE, TEOAE, Cochlear mechanics

INTRODUCTION

Otoacoustic emissions (OAEs) are cochlear responses in the form of sounds that can be recorded in the ear canal. Most OAEs are the result of cochlear active processes returning some of the energy in the cochlear response back into the ear canal. A single tone evokes an emission at the tone frequency, known as a stimulus frequency otoacoustic emission (SFOAE). The amplitudes and phases of SFOAEs vary with tone frequency and provide a non-invasive way to assess the cochlea and its mechanics.

Correspondence to: Maria A. Berezina-Greene · Eaton-Peabody Lab, Mass. Eye and Ear Infirmary · 243 Charles St, Boston, MA 02114, USA. Telephone: 603 387 6118; email: maria.berezina@gmail.com

According to classic coherent reflection theory, SFOAEs are generated by cochlear irregularities that coherently reflect energy within the broad, tall peak of the traveling wave (the "SFOAE CF region") (Zweig and Shera 1995). In the simplest view, this coherently reflected energy arrives in the ear canal with a single delay that is approximately twice the travel time of the basilar membrane (BM) traveling wave (Zweig and Shera 1995). Because cochlear irregularities are random, they sometimes produce cancellations of the energy reflected from the peak of the SFOAE CF region. Such cancellations result in deep notches in SFOAE amplitude vs. frequency plots, and can result in SFOAE phase vs. frequency patterns that are complicated and show phase reversals at SFOAE notches.

Recent work shows that cancellations of wavelets that originate at different places along the cochlea can produce multiple SFOAE delay components, i.e., multiple peaks of SFOAE energy that arrive in the ear canal with different delays (Sisto et al. 2015). When cochlear responses are sharply tuned (have high Q), multiple SFOAE components arise within the traveling-wave peak region and SFOAE components arising basal of the amplified peak region are negligibly small (Zweig and Shera 1995). However, when the cochlear responses have low Q, components that arise basal of the amplified peak region can be a significant fraction of the SFOAE (Sisto et al. 2015).

A major issue in measuring SFOAEs is separating the SFOAE from the probe tone that evokes the SFOAE. SFOAEs are usually obtained by the suppression method (Guinan 1990; Shera and Guinan 1999). In this method, the SFOAE is the vector difference between two measurements, one with only the probe tone and one with the probe tone plus a 2nd tone. The difference is termed as "residual." Most often, the 2nd tone is close in frequency and higher in level than the probe tone and is considered to suppress the generation of the SFOAE. If the SFOAE is fully suppressed and the addition of the suppressor does not generate additional probe-frequency sources, then the residual equals the SFOAE. However, a 2nd tone interacting in the cochlea with the probe tone may also create new (additional) probe-frequency energy that was not present with the probe tone alone, and this new source of energy at the probe frequency can contribute to, or produce, the residual (Shera et al. 2004). The residual (which is the difference between two measurements) is sometimes said to be the SFOAE, which can make the term "SFOAE" confusing. We will use the terminology "true SFOAE" when we want to specify the probe-frequency OAE from a probe tone with no other sound present. What fraction of the measured residual is the true SFOAE vs. a new source can depend on the frequency of the 2nd tone relative to the probe tone (Shera et al. 2004). Consequently, how well the residual obtained with a suppressor represent the true SFOAE may vary with 2nd tone frequency.

The 2nd tone suppresses the SFOAE by reducing cochlear amplification and the mechanical response to the probe tone. Suppression is caused by the 2nd tone response deflecting outer-hair-cell (OHC) stereocilia into low-slope "saturation" regions of their receptor-current/deflection response (Geisler et al. 1990). With this mechanism, suppression at the probe tone frequency by a 2nd tone only takes place when the 2nd tone produces a response in cochlear regions where the response to the probe frequency is actively amplified. SFOAE components that originate within the amplified peak region we term "near basal" SFOAE components, and ones that originate basal to the probe tone amplification region we term "far-basal" SFOAE components. Near-basal SFOAE components can be suppressed by a 2nd tone that produces a response in the amplification region; however, a 2nd tone whose response is basal to the probe-frequency amplification region cannot suppress the SFOAE. A far-basal SFOAE component is not amplified and cannot be suppressed by a 2nd tone.

SFOAEs have been usually measured by presenting a 2nd tone with a frequency near the probe-tone frequency. We will term SFOAEs obtained this way as "near-suppressor SFOAEs" or "nsSFOAEs." Ideally, a 2nd tone near the probe tone frequency at a high enough level would suppress all of the near-basal SFOAE components. However, a 2nd tone with a frequency near the probe-tone frequency cannot suppress far-basal SFOAE components and thus an nsSFOAE will not equal the true SFOAE if the true SFOAE has significant far-basal components (e.g., if it is from a low-Q cochlear region). For reference, the four new terms used in this paper are defined in Table 1.

Although coherent reflection theory (CRT) adequately describes much of the production of SFOAEs, the validity of CRT has been questioned based on two kinds of evidence. A 2nd tone several octaves above the SFOAE frequency sometimes produces a substantial SFOAE-frequency residual (Guinan 1990; Siegel et al. 2003; Charaziak et al. 2013; Charaziak and Siegel 2014, 2015a, b). These, and other experiments, have been interpreted as evidence against CRT because they mean that some SFOAE sources are too basal to be accounted for by CRT (Siegel et al. 2005; Choi et al. 2008; Charaziak et al. 2013; Charaziak and Siegel 2014, 2015a, b). The second challenge to classic CRT comes from low-frequency SFOAEs. SFOAE group delays at mid- to high-frequencies fit well with classic CRT (Shera and Guinan 2003; Shera et al. 2008), but low-frequency SFOAE delays are shorter than predict-

TABLE 1

Definitions of new terms

Term	*Definition*
Near suppressor SFOAE (nsSFOAE)	SFOAEs measured with a suppression method where the 2nd tone (suppressor) has frequency near the probe-tone frequency
True SFOAE	Probe-frequency OAE from a probe tone with no other sound present (e.g., suppressor tone)
Near basal SFOAE	SFOAE energy that originates within the amplification region of the traveling wave and as such can be suppressed by a 2nd tone
Far basal SFOAE	SFOAE energy that originates basal to the amplification region of the traveling wave and cannot be suppressed by a 2nd tone

ed in cats, guinea pigs (Shera and Guinan 2003), and chinchillas (Siegel et al. 2005).

SFOAEs are useful scientifically and may be useful clinically. For instance, their delay has been used to predict human cochlear tuning (Shera et al. 2002, 2010). However, for SFOAEs to be used successfully, we must understand where along the cochlea they are generated, how well their latency reflects cochlear motions, and whether coherent reflection theory adequately accounts for SFOAE generation.

We concentrate here on nsSFOAEs because they are a metric of high practical importance even though they do not include or measure any far-basal SFOAE components. To shed light on the mechanisms that produce nsSFOAEs, where they originate and why their group delays are shorter than expected at low frequencies, we (1) measured nsSFOAEs from low-level tones at closely spaced frequencies in guinea pigs and did time-frequency analyses of the results looking for evidence of multiple SFOAE components that might be distinguished by their delay (as was done by Sisto et al. 2013), (2) did equivalent time-frequency analyses on simulated SFOAEs arising from coherent reflection in a simple cochlear model, and (3) also conducted this analysis on chinchilla nsSFOAE data provided by the Siegel laboratory (Siegel et al. 2005). A previous time-frequency analysis of guinea pig SFOAEs was done by Goodman et al. (2003); however, they used a nonlinear SFOAE extraction paradigm and high sound levels, so their paradigm will not reveal SFOAE components that grow linearly and may generate excessive distortion-source SFOAEs due to high sound levels. Our suppression paradigm should more effectively reveal SFOAE components found at low sound levels.

METHODS

General Methods

The nsSFOAEs were collected by suppression as in Guinan et al. (2003) over a frequency range 500 to 10,000–12,000 Hz in 83 Hz steps. Suppressor tones (presented at 60 dB sound pressure level (SPL) for 50 ms every 100 ms) at 50 Hz above the probe frequency were presented simultaneously with continuous, 40 dB SPL probe tones. At each probe frequency, responses were averaged over 6–8 s and nsSFOAEs were calculated from the vector difference between the probe-frequency parts of the probe-alone and the probe-plus-suppressor responses. The probe-frequency part of each response was obtained using a fast Fourier transform (FFT) of the time waveform. For each probe frequency, the noise floor was estimated as the average of FFT amplitudes at eight frequencies below and eight frequencies above the probe frequency. nsSFOAE vs. probe-tone frequency functions were measured in 17 animals, mostly from one ear in each. Ears that had thresholds ≥40 dB SPL were not included. To allow equal-data comparisons with the 10 chinchilla and 10 simulated data sets, in the "Results" we considered data only from the first 10 guinea pig ears.

Albino guinea pigs were anesthetized with Nembutal, followed by fentanyl and haloperidol (initial doses: 25, 0.2, and 10 mg/kg, respectively). Additional doses were given as necessary as judged by change in heart rate or reflex withdrawal from a toe pinch. Animals were tracheotomized and mechanically ventilated. The heart rate, breath rate, expired CO_2, rectal temperature, and electroencephalogram were continuously monitored. The rectal temperature was maintained at approximately 37–38 °C. The soft tissue around the skull was removed and a head holder bar was cemented to the skull. The ear canals were truncated and acoustic assemblies were inserted. The bulla was opened dorso-laterally and a silver wire electrode was placed near the round window to monitor the auditory nerve compound action potential (CAP). Thresholds were estimated from CAPs in response to tone pips (5 ms duration, 0.5 ms raised-cosign-shaped rise/fall) by an automated up/down procedure which determined the sound level needed to produce 10 μV pp CAPs at frequencies 1000–32,000 Hz in octave, or less, steps. Some guinea pigs also underwent a craniotomy and cerebellar aspiration to expose the floor of the 4th ventricle for use in another experiment. Experimental protocols were

approved by the Massachusetts Eye and Ear Infirmary Animal Care Committee.

Dr. Christopher Shera kindly provided us with 10 sets of simulated SFOAEs vs. frequency functions that were derived from the simple cochlear model described in Shera and Bergevin (2012). In this linear model, the amplitude of the traveling wave was approximated by a Gaussian envelope with a bandwidth set to match the Q_{10} of guinea pig auditory nerve fiber tuning curves (Tsuji and Liberman 1997) from the corresponding frequency region. At each cochlear place, the model response had a straight-line phase vs. frequency function with a slope at each frequency set to the delay determined by the straight-line power-law fit to the guinea pig SFOAE data above 2000 Hz of Shera and Guinan (2003; the dashed line in their Fig, 5). Ten different sets of random cochlear irregularities were fed to the model and resulted in 10 sets of simulated emissions. The model did not include reflections at the stapes which removed a source of long-latency energy. However, our guinea pig data showed little energy at latencies of twice the expected SFOAE peak latency which is consistent with reflections at the stapes being negligible.

Chinchilla nsSFOAE data from Siegel et al. (2005) were kindly provided by Dr. Jonathan Siegel. They were collected using the suppression method with probe and suppressor tones of 30 and 55 dB SPL, respectively, and the suppressor frequency 43 Hz below the probe frequency. The probe frequency spacing between measurements was 21.83 Hz for frequencies from 300 to 2000 Hz and 43.06 Hz for higher frequencies.

Time-Frequency Analysis

The time-frequency analysis used short-term inverse Fourier transforms (IFFTs) to obtain nsSFOAE energy distributions as a function of time from a series of narrow frequency regions, thereby yielding each region's "Impulse Response." The analysis was implemented using custom software written in MATLAB (version 7.6.0 R2008a, The MathWorks, Inc.).

A central methodological issue is how the analysis-window bandwidth affects the results. We can increase time resolution by increasing the bandwidth of the analysis window but this reduces frequency resolution. The spectra from window bandwidths of 500, 1000, and 2000 Hz in two guinea pigs, one with many frequency-domain complexities and one with few, are shown in Figure 1. The wide 2000-Hz window shows more peaks at each frequency than the narrow 500-Hz window, but it is not clear to what extent these peaks represent multiple SFOAE components from narrow frequency regions or result from the energy latency pattern varying across frequency within the 2000-Hz window and producing the multiple peaks and valleys by interference (i.e., combining out-of-phase sources can produce dips and peaks even when no single frequency region shows such dips and peaks). The peaks are narrower with the 2000-Hz window than with the 500-Hz window, but where in the 2000-Hz window, this energy originates from is unknown. For instance, consider 1900 Hz in Figure 1, left. With the 2000-Hz window, there is a peak at a latency of ~1.5 ms, but with the 1000-Hz window, there is a dip at 1.5 ms and energy at 1.5 ms is found from 2600 to 2900 Hz. Thus, the 2000-Hz window shows a peak at

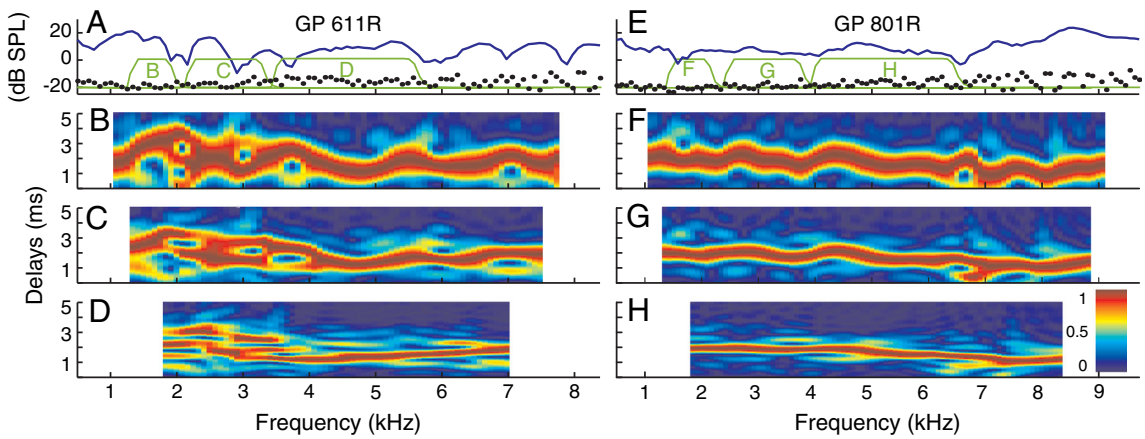

FIG. 1. Time frequency analysis of near-suppressor stimulus frequency otoacoustic emission (nsSFOAE) data from two guinea pigs (*left* and *right*) using different analysis bandwidths. *Top row*, measured nsSFOAE magnitudes vs. frequency (*solid lines*), noise floors (*dots*), and example outlines of the analysis windows. *Rows 2–4*, nsSFOAE energy as a function of frequency and delay for analysis windows with bandwidths of 500, 1000, and 2000 Hz (*rows 2*, *3* and *4*, respectively). The panels labeled *B*, *C*, *D*, *F*, *G*, and *H* correspond to the window-bandwidth examples in the top row. In rows 2–4, nsSFOAE energy at each frequency is normalized by the peak energy at that frequency. The *key at bottom right* shows the correspondence between color and the fraction of the maximum energy at each frequency.

1.5 ms and 1900 Hz because of energy from CF places >1900 Hz. Overall, the better time resolution of the 2000-Hz window does not help in determining whether an nsSFOAE in a particular frequency region has short-latency components. On the other hand, to determine whether nsSFOAE group delays at low frequencies (<2 kHz) are too short, a narrow window is necessary.

If the SFOAE phase slope changes a lot with frequency, then varying the window bandwidth across frequency might help. To obtain a relevant metric, nsSFOAE phase accumulation across frequency was analyzed for all 17 guinea pig ears. In each ear, unwrapped nsSFOAE phase was plotted from low to high frequency, and for each successive one-cycle phase change, the center frequency and bandwidth of the segment in Hertz were determined. From 500 to 9000 Hz, the width of a single cycle of nsSFOAE phase accumulation changed roughly by a factor of 2 (Fig. 2) but the scatter was much larger. These data present no strong reason for varying the bandwidth of the window with frequency. Considering this, and the need to evaluate changes in latency with frequency at 2000 Hz and below, we chose a 500-Hz frequency-window bandwidth.

Analysis of the Frequency-Window Impulse Response

The impulse response is strongly affected by the shape of the frequency-selection window. Our window used the *magnitude* of a band-pass Butterworth filter with an effective pass-band of 500 Hz bounded by 600 Hz attenuation bands reaching 30 dB attenuation (e.g., thin lines in Fig. 1A) and with amplitudes >850 Hz from the center frequency set to zero over the range 0–16,000 Hz. The windowing changed only magnitudes not phases. The window was moved in 83 Hz steps centered on each data frequency (windows overlapped) and each result was attributed to that window's center frequency. For chinchilla and model data, the step size was adjusted to match the data while keeping a bandwidth of ~500 Hz.

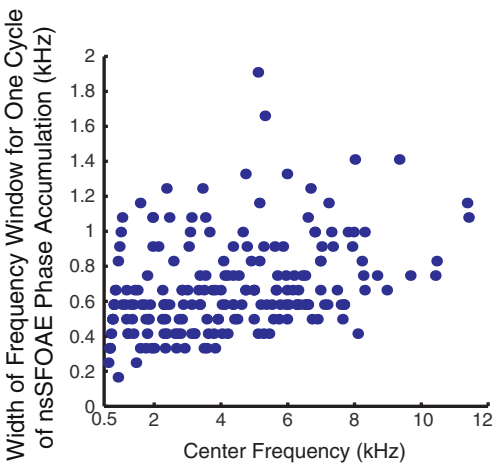

FIG. 2. The dots show the widths of the frequency segments over which one cycle of nsSFOAE phase change was accumulated, vs. the center frequency of that segment. Data from 17 guinea pig ears.

Each set of windowed data was IFFTed using Matlab's *ifft* function. The magnitude of the complex IFFT output was divided by the square root of two to obtain the magnitude (in rms Pascals) vs. delay function or impulse response (e.g., Fig. 3, D1–D4). The impulse response represents the arrival of nsSFOAE energy over time in the ear canal from cochlear regions in the frequency window. Energy with delays over 5 ms was mostly noise and was removed with a recursive exponential filter (Kalluri and Shera 2001). Using a purpose-written Matlab program, peaks were detected by zero crossings of the impulse-response derivative and were identified using three criteria. Peaks must have at least ¼ the amplitude of the largest peak and a signal-to-noise ratio (SNR) ≥15 dB (noise impulse responses were obtained from the noise at each frequency using the Butterworth window followed by an IFFT). The SPLs of identified peaks were obtained by separating peaks (using recursive exponential filters with cutoffs at the dips between peaks) and FFTing these time-domain peaks back to the frequency domain. If a frequency-domain peak had a SNR <6 dB, then the peak was no longer considered as separate, and SPL calculations of the other peaks were repeated without separation of the no-longer-considered peak.

The delays of the impulse response peaks that passed the above criteria are shown for one ear in Figure 3C. The equivalent decibel SPLs of these peaks are shown in Figure 3B. Also in Figure 3C is the phase-gradient delay calculated from the slope of the phase-vs.-frequency function. In contrast to an individual time-frequency component, the phase-gradient delay tracks how the phase of the overall SFOAE varies with frequency. In particular, in each frequency range, the slope of the phase is mostly determined by the delay of the dominant component. The phase-gradient delay was obtained at each frequency from a regression line fit to the phases in a window equal to the pass-band (500 Hz) of the Butterworth window. This window was used for calculating the phase-gradient delay because it best corresponds to the delay from the time-frequency analysis (The phase-gradient delay from adjacent points produces a result that is too noisy to be useful, especially near nsSFOAE amplitude dips).

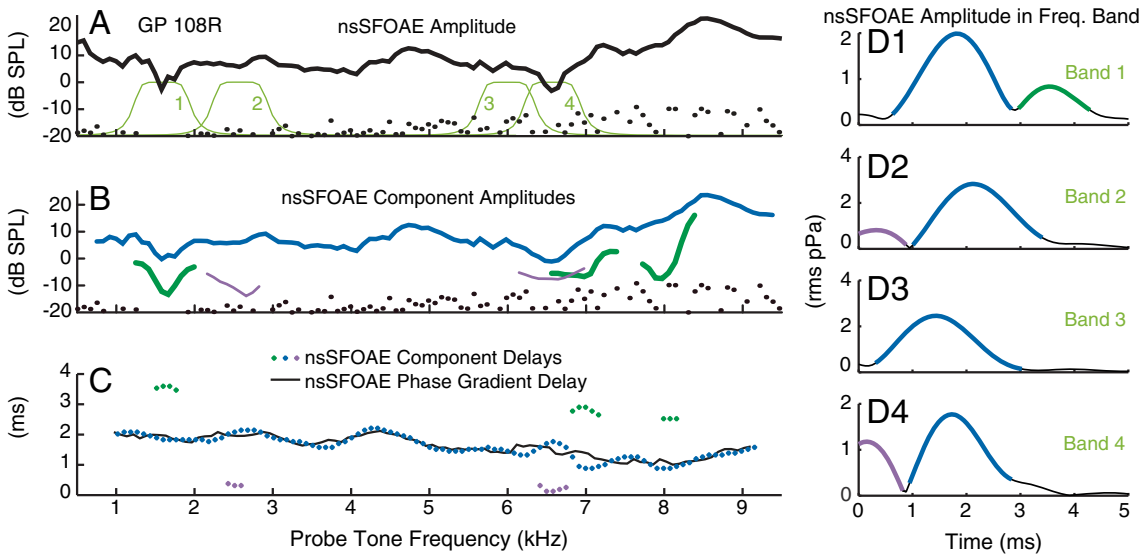

Fig. 3. Time-frequency component analysis of nsSFOAEs from one guinea pig ear (GP108R). (*A*) nsSFOAE magnitude vs. frequency (*thick line*), noise levels (*dots*), and four examples (*thin lines 1–4*) of windows that selected nsSFOAE frequency regions to be inverse Fourier transformed (IFFTed). The scale for windows 1–4 is decibel of gain (which is negative), not re SPL. (*B*) Amplitudes of the nsSFOAE components obtained from time-frequency analysis. *Blue line*, the component with a delay corresponding to the phase-gradient delay. *Thin purple* and *thick green lines*, components with delays shorter and longer, respectively, than the phase-gradient delay. (*C*) nsSFOAE-component delays (*diamonds*) vs. center frequency of the IFFT selection window. *Blue diamonds*, the component with a delay corresponding to the phase-gradient delay. *Purple* and *green diamonds*, components with delays shorter and longer, respectively, than the phase-gradient delay. *Solid black line*, nsSFOAE phase-gradient group delay. (*D1–D4*) nsSFOAE energy vs. delay plots (impulse response functions) from IFFTs of windows 1–4 of panel *A*. Each impulse response was attributed to the center frequency of its window.

RESULTS

Overall Results From Time-Frequency Analysis of Guinea Pig nsSFOAEs

We used data at frequencies from 500 to 8000–12,000 Hz from 10 ears with normal hearing. The nsSFOAE magnitudes had broad peaks and sharp dips (Figs. 1A, 3A), as previously reported (Goodman et al. 2003). Many of the dips were accompanied by near half-cycle phase changes (not shown). Every ear had frequency regions in which the nsSFOAE energy as a function of time (the region's impulse response) had energy at a single delay, and regions with two nsSFOAE delay peaks. A few ears had small regions with three delay peaks. Across all frequencies and ears, ~33 % of frequencies had two peaks and only 1 % of frequencies showed three peaks. When there was a single peak, its delay was almost always similar to the phase-gradient group delay (see Figs. 3, 4 (left), and 5 and the next section). When there were two peaks, the delay of one was usually similar to the phase-gradient delay, while the delay of the other could be longer or shorter than the phase-gradient delay (Figs. 3, 4 (left), and 5). Two peaks were almost always present at frequencies near dips in the nsSFOAE magnitude. Occasionally, the analysis did not show two peaks near a dip because one of the peaks failed the peak selection criteria. We will refer to the peak of an impulse response function that passed the selection criteria as an nsSFOAE "component." Each component was characterized by its delay (the delay of the peak) and energy (the energy within the delay region demarcated by the dips or inflection points on either side of the peak—see "Methods").

Classifying nsSFOAE Components by Their Delays

We attempted to determine whether the multiple nsSFOAE components showed patterns that would allow them to be classified by their delays. In particular, Sisto et al. (2013) reported a finding that SFOAE delay components in humans could be classified by their delays and we sought to determine if such a pattern was present in guinea pigs.

To classify nsSFOAE delay components, we considered their relationship to three delay metrics (phase gradient, power law, and neural). To choose a dividing line between components that agree with the phase gradient delay and those that are longer or shorter, we plotted all nsSFOAE component delays at all frequencies from the 10 guinea pig ears relative to the corresponding phase-gradient delay (Fig. 5). From Figure 5, we chose ±0.6 ms as the dividing line. Similar

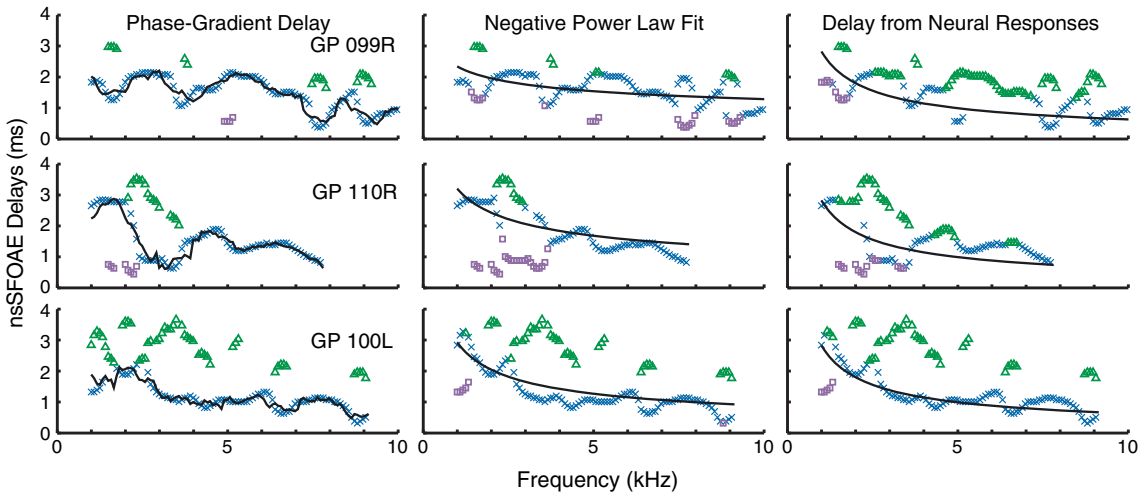

FIG. 4. nsSFOAE delay components from three ears (one per row, guinea pig numbers are in the *left column*) classified by their relation to the phase-gradient delay (*left*), to the negative-power-law fit to the data (*center*), or to neural delays (*right*). Delays were classified as agreeing with the solid-line delay if they were within ±0.6 ms (*xs*); otherwise, they were shorter (*squares*) or longer (*triangles*). *Left solid lines*=phase-gradient delays. *Center solid lines*=negative power-law fit to the distribution of energy in the time-frequency plane. *Right solid lines*=delays (doubled) obtained from the one-line fit to the high-level click responses of guinea pig auditory-nerve fibers from Versnel et al. (1990).

plots for power law and neural delays were diffuse so the ±0.6-ms dividing line was also used with these metrics. nsSFOAE component relative to phase gradient delays is shown in Figures 3 and 4, left. Unfortunately, this simple and intuitively appealing method did not produce a dependable classification. Although the phase-gradient delay usually followed the component with the largest energy, when two components of nearly equal amplitude were present, it appeared random which of the two components agreed with the phase-gradient delay.

The second method followed Sisto et al. (2013) and determined the power-law fit for an individual ear from the negative-power-law curve that yielded the maximum energy integrated across frequency in a narrow band around the curve. This method worked in humans; however, our results show that it is less effective in guinea pigs. The coefficient of determina-

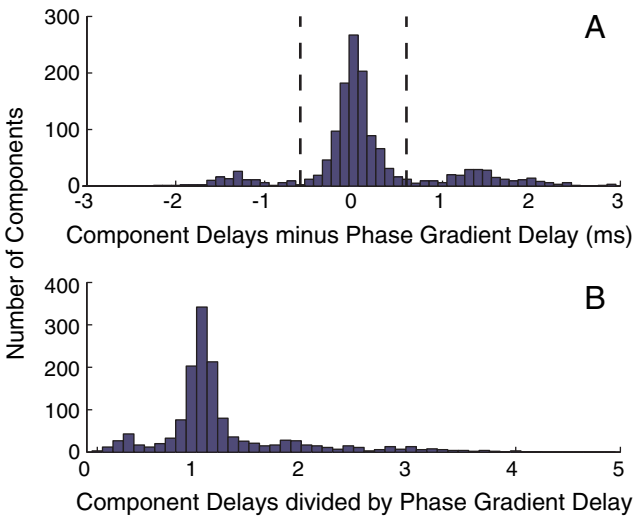

FIG. 5. The distribution of component delays relative to the phase-gradient delay from 10 guinea pig ears. *Top*, the difference in millisecond. *Bottom*, the difference expressed as a ratio. *Dashed lines* show the ±0.6-ms dividing lines between components whose delay agreed with the phase gradient delay and those that had shorter or longer delays. These distributions were influenced in complex ways by the parameters of the time-frequency analysis, e.g., doubling the window bandwidth shifts the shape of the function but does not compress it by a factor of two in time.

tion, R^2, for the guinea pig delay component that best fit the power-law curve averaged only R^2=0.38±0.22 for the 10 guinea pig ears (R^2 is the proportion of the data variance explained by the fit and varies from 0 to 1). The low R^2 indicates that the change across frequency in guinea pig nsSFOAE delays is not well described by a power law. Overall, using the negative-power-law curve (Fig. 4, center column) delay components were inconsistently classified just as they were using phase-gradient delays.

The third method used neural latencies as functions of frequency obtained from post-stimulus-time histogram from guinea pig auditory-nerve-fiber responses to 92 dB peSPL clicks (Versnel et al. 1990). Traveling-wave delays were calculated as the neural delays minus 1.01 ms as in Versnel et al. We classified nsSFOAE components relative to SFOAE delays estimated by doubling the traveling-wave delays from Versnel et al. one-line fit to their data. There was no consistent match between the nsSFOAE component delays and the delays predicted by the neural latencies (Fig. 4, right), especially when there were two nsSFOAE components.

Overall, none of the methods provided a consistent way to classify cases in which there were multiple delay components. Thus, the delay patterns did not provide evidence for there being systematically different origins for nsSFOAE delay components in guinea pigs, in contrast to the consistent delay patterns reported for human SFOAEs by Sisto et al. (2013).

The overall relationships of the nsSFOAE components divided into groups by the 0.6-ms dividing lines can be seen in Figure 6. Also in Figure 6 are dashed lines showing the SFOAE delays predicted from Versnel et al. (1990) fit to their data. Figure 6 also shows the fit to guinea pig SFOAE delays above 2000 Hz from Shera and Guinan (2003, their Fig. 6 and Table 1).

Time-Frequency Analysis of Simulated SFOAEs

Model guinea pig SFOAEs from Dr. Christopher Shera (see "Methods") were analyzed in the same way as our real guinea pig data. Color-coded, time-frequency plots of simulated guinea pig SFOAEs (like those in Fig. 1) from the largest-bandwidth window showed multiple narrow peaks and valleys spread over a wide latency range, similar to the plots from real guinea pig SFOAEs. Thus, as for the real guinea pig data, the simulated guinea pig data were analyzed using narrow, 500 Hz windows. SFOAE component delay vs. frequency plots for three representative SFOAE simulations are shown in Figure 7, middle. For comparison, Figure 7 (left) shows three more examples of real guinea pig nsSFOAE component delays. In both simulated and real ears, there are multiple SFOAE components at irregular intervals along the frequency axis, and the patterns of the SFOAE-component latencies look very similar, i.e., frequency regions around SFOAE-magnitude dips show two delay peaks. One difference is that the simulated SFOAEs have longer delays at low frequencies than the delays of real nsSFOAEs. The SFOAE simulation model was constructed to have the delays from Shera and Guinan (2003) fit to their guinea

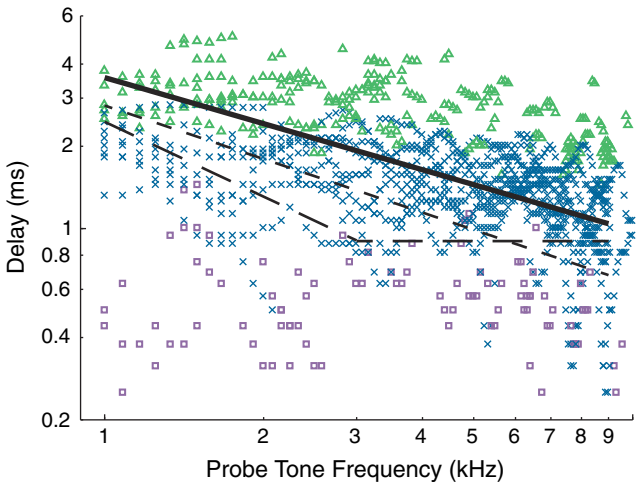

FIG. 6. Guinea pig nsSFOAE component delays obtained by time-frequency analysis, compared to doubled auditory-nerve delays (*dashed lines*) and nsSFOAE delays from guinea pig data above 2000 Hz of Shera and Guinan (2003) (*solid line*). nsSFOAE-component delays are from 10 guinea pig ears and were classified as agreeing with the phase-gradient delay if they were within ±0.6 ms (*xs*); otherwise, they were shorter (*squares*) or longer (*triangles*). The dashed-line estimates of nsSFOAE delays were derived from the single-line (*short dashes*) and two-line (*long dashes*) fits to guinea pig single-auditory-nerve-fiber delays to high level clicks (Versnel et al. 1990).

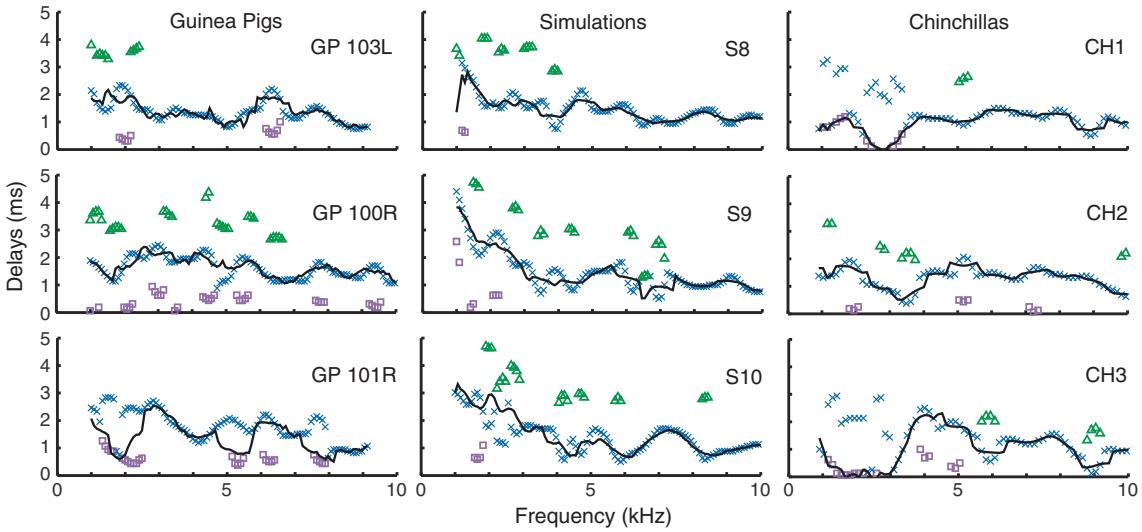

FIG. 7. Time-frequency delay components in three representative ears (*top to bottom*) of guinea pigs (*left*), models (*center*), and chinchillas (*right*). Each panel shows the nsSFOAE (or simulated SFOAE) component delays (*xs*) of different ears, their phase-gradient delays (*solid lines*), and the animal or simulation number. Delays were classified as agreeing with the phase-gradient delay if they were within ±0.6 ms (*xs*); otherwise, they were shorter (*squares*) or longer (*triangles*).

pig data above 2000 Hz. They also noted that this fit was a reasonable approximation to the predicted SFOAE delay based on available guinea pig mechanical measurements. Thus, the greater upward bends in the simulation delays at low frequencies compared to the real guinea pig delays (Fig. 7) provide a way to appreciate the difference between the measured and predicted delays at low frequencies.

Time-Frequency Analysis of Chinchilla nsSFOAEs

The nsSFOAE data from 10 chinchillas provided by Dr. Jon Siegel (see "Methods") were analyzed in the same way as was done for the guinea pig and model data. nsSFOAE component delay vs. frequency plots from three representative chinchilla ears are shown in Figure 7, right. These chinchilla nsSFOAEs showed delay vs. frequency patterns similar to those in guinea pigs, namely multiple nsSFOAE components especially around frequencies that have sharp dips in nsSFOAE magnitude. The chinchilla data also show another pattern, regions of near-zero delays (Fig. 7, right). Regions like these were also observed in guinea pigs, but they were smaller and less frequent. They were almost never seen in the model data.

Comparison Across Species/Model and Classification Schemes

Component delays from guinea pig and chinchilla nsSFOAEs, as well as from guinea pig SFOAE simulations, all referenced to their respective phase-gradient delays, power-law fits, and neural delays are shown in Figure 8. Several patterns can be seen in these data. With the phase-gradient-delay and power-law-fit references, the main components (the ones that best match the reference—blue xs) are centered on the reference. In contrast, with the neural-delay references, the main component delay is too long at mid- to high frequencies and more closely centered, or too short, at low frequencies. For both guinea pigs and chinchillas, the neural delay that we used is a delay that tracks the first neural response in that species, and at high frequencies, this delay is shorter than the BM group delay. This accounts for the SFOAE main component delays being too long at high frequencies. In the "Discussion," we present hypotheses for why the SFOAE delays at low frequencies match, or are shorter than, the SFOAE delay predicted by the first neural response. Another pattern in Figure 8 is that multiple delay components occur more frequently at low frequencies than at high frequencies. This is due, in part, to the linear frequency scale in Figure 8 which shows an octave as wider at high frequencies. However, other factors may also be involved such as the sharper tuning found at high frequencies or that there may be more irregularities at low frequencies which may produce more cochlear reflections. It can also be seen in Figure 8 that very short delays (those less than 1/10 the reference delay—red circles) are more prevalent in chinchillas than in guinea pigs and are the least common in the simulated SFOAEs.

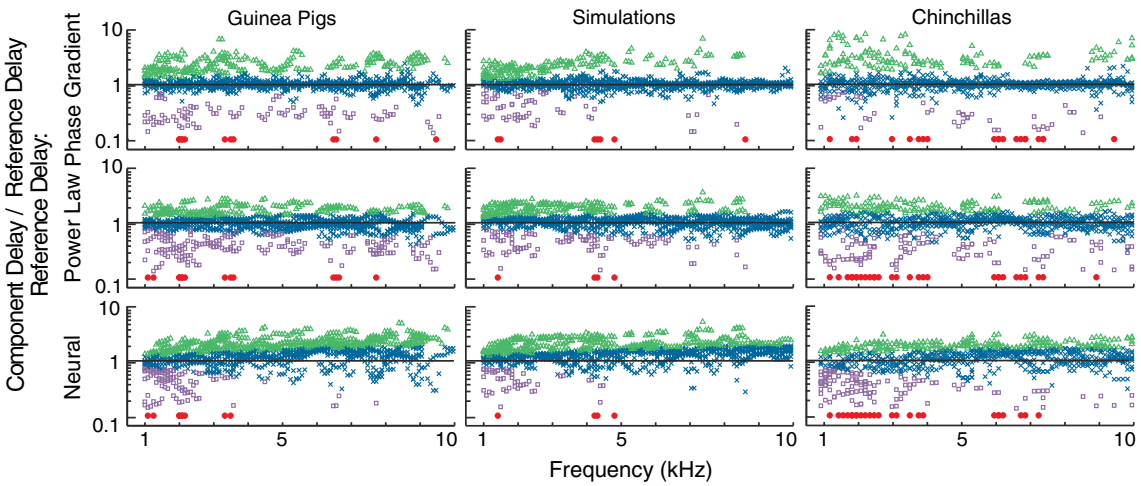

FIG. 8. nsSFOAE component delays relative to the phase-gradient delay (*top row*), the power-law fit (*middle row*), and twice the neural delay (*bottom row*). Each column shows data from the same 10 ears from the species/model labeled at *top*. Delays were classified as agreeing with the reference delay if they were within ±0.6 ms (*blue xs*); otherwise, they were shorter (*purple squares*) or longer (*green triangles*). Values less than 0.1 are shown at 0.1 as *red circles*. Phase-gradient and power-law references were derived separately for each ear/simulation. Reference neural delays for guinea pigs and simulations were from Versnel et al. (1990). Reference neural delays for chinchilla were from Temchin et al. (2005; Fig. 13, τ_{BFGD}).

DISCUSSION

SFOAE Components From Near the Peak of the SFOAE CF Region

The time-frequency analysis of guinea pig and chinchilla nsSFOAEs and simulated SFOAEs usually found one delay component in 2/3 of frequencies measured in guinea pigs, a component that presumably originated from the peak of the traveling wave, which we termed the SFOAE CF region. In frequency regions with one delay component, the lack of a short-delay component is evidence that far-basal sources are not important contributors to nsSFOAEs. However, our methods do not rule out the presence of small short-delay nsSFOAE components. In addition, nsSFOAEs do not include any SFOAE components that would not have been suppressed.

At some frequencies, most often near dips, multiple SFOAE components were found. The presence of two components near an SFOAE dip is explained well by classic coherent reflection theory which predicts that sometimes the pattern of irregularities will be such that the SFOAE energy originating from the peak of the SFOAE CF region cancels. Such cancellations are what produce SFOAE dips in SFOAE amplitude vs. frequency plots. If the energy from the peak region is cancelled, then what remains is energy from the edges of the SFOAE CF region and this energy results in two delay components in the time-frequency analysis. Wavelets from the basal and apical edges may then form the largest SFOAE components, and these have delays shorter and longer than the CF peak would have had. Thus, the presence of short and long delay components near a dip does not necessarily mean that the short component is a far-basal component. A recent model simulation of SFOAEs produced by coherent reflection shows results that agree with the above scenario and further supports the hypothesis that SFOAE wavelet cancellation can produce more than two delay components Sisto et al. (2015).

We were unable to find a consistent method to classify SFOAE delay components when two components were present. We hypothesize that this is because the two components did not originate from different mechanisms (i.e., the shorter-delay component was not from a far-basal source); instead, both components had a common origin in the coherent reflection of energy near the SFOAE CF region. Supporting this hypothesis is the similarity of the multiple SFOAE components in real guinea pig ears and the multiple SFOAE components from the simple guinea pig model (Figs. 7 and 8). In the model, the *only* origin of SFOAEs was from coherent reflection of energy in the peak region of the traveling wave. The amplitude of the Gaussian that approximated the envelope of the traveling wave was diminishingly small far basal to the SFOAE CF region and could not produce substantial SFOAE components. Nonetheless, the model showed multiple delay peaks in the resulting SFOAEs, some with short latencies. This again shows that multiple delay peaks such as those produced by the model do not mean that the delay components originated far-basal to the SFOAE CF region. The multiple delay components produced by the simple model had to arise as a result of the pattern of irregularities producing a cancellation of energy from the peak region of the traveling wave and

causing edge-region SFOAE wavelets to add coherently with long and short delays.

From a time-frequency analysis (based on wavelets) applied to human SFOAEs and transient evoked otoacoustic emissions (TEOAEs), the Sisto-Moleti lab has suggested that there is a long-delay OAE component that originates near the peak of the SFOAE CF region and a shorter-delay OAE component that originates 2 mm (~1/3 octave) basal to the CF peak of the traveling wave, i.e., it is a "near-basal" component (Moleti et al. 2013; Sisto et al. 2013). Consistent with this interpretation, the longer-delay SFOAE component grew compressively and the shorter-delay SFOAE component grew more linearly (Sisto et al. 2013—note that just because a component is highly compressive does not prove it originates at the peak of the traveling wave, the apical edge of the traveling wave is also very compressive; Ruggero et al. 1997). It is not clear to what extent the two components distinguished by Sisto and co-workers in human SFOAEs correspond to the two components we found in guinea pig SFOAEs. Because guinea pig SFOAE delays are much shorter than those in humans, it is difficult to resolve SFOAE components in guinea pigs with the close delay ratios that Sisto and coworkers found in humans. It seems likely that the human and guinea pig SFOAE components arise from similar coherent reflection cancellations, but the human SFOAE components have much smaller delay ratios than those in guinea pigs because humans have much sharper cochlear tuning than guinea pigs (Shera et al. 2002, 2010; Joris et al. 2011).

Interpreting Proposed Evidence for SFOAE Components with a Far-Basal Origin

The hypothesis that significant SFOAE components have "far basal" origins (basal to the probe-response cochlear amplification and suppression region) comes from the finding of probe-frequency residuals produced by a 2nd tone that is two octaves or more above the SFOAE tone frequency (Guinan 1990; Siegel et al. 2003; Charaziak et al. 2013; Charaziak and Siegel 2014, 2015a). In particular, Charaziak and Siegel (2014) showed that, in chinchillas, the tuning curves produced by varying the level of a 2nd tone to produce a criterion residual at the probe frequency, which they called "SF-STCs," were similar in shape to CAP suppression tuning curves for probe frequencies >3000 Hz, but for probes <3000 Hz, the SF-STCs extended to frequencies several octaves above the CAP-STCs. At the 2nd tone frequencies where the SF-STCs extended above the CAP-STCs, the residuals must have received no cochlear amplification, i.e., they were far basal. Charaziak and Siegel (2015b) provide a compelling demonstration that such far-basal residuals do indeed originate from the response region of the 2nd tone because these residuals are abolished in the same way by a nearby 3rd tone or by acoustic trauma. Charaziak and Siegel interpreted these far-basal residuals as showing suppression of SFOAE components that originated from these far-basal regions. However, since there is no cochlear amplification at the probe frequency in these far-basal regions, there is no probe-frequency cochlear amplification for the 2nd tone to suppress. Indeed, they show that cochlear amplification was not suppressed by showing that the far-basal 2nd tones did not change CAP thresholds. A high-frequency 2nd tone can still push OHC stereocilia into their low-slope saturation region, but if this happens far basal of the probe-tone amplification region, it does not change probe-frequency cochlear amplification. Pushing OHC stereocilia into their saturation regions does, however, produce a place where local nonlinearities created by the 2nd tone might interact with the local response to the probe tone to produce new-source (i.e., not part of the true SFOAE—explained further in the "Introduction" and Shera et al. 2004) probe-frequency wavelets that give rise to the SFOAE residual measured in the ear canal. This mechanism is our working hypothesis for the generation of SFOAE residuals by far-basal 2nd tones.

The model analysis of Sisto et al. (2015) shows that far-basal SFOAE components can arise by coherent reflection for probe tones with CFs in low-Q cochlear regions. However, if there are far-basal SFOAE components, these components will not be suppressed and will be missed in nsSFOAE measurements. It is possible that a far-basal 2nd tone somehow interferes with the production of passive far-basal SFOAE components. If so, this would not be by suppression. Also note that the presence of passive far-basal SFOAE components or their possible interference by a 2nd tone does not preempt that the 2nd tone creates new SFOAE residual components (Shera et al. 2004). It may be difficult to experimentally determine the extent to which a far-basal SFOAE component is from a passive true SFOAE component or from a 2nd-tone-generated new residual component. However, it should be kept in mind that the theory that predicts the possibility that there are passive far-basal SFOAE components also predicts that such components will be small (Sisto et al. 2015). In cats, Guinan (1990) found that 2nd tones at frequencies five times higher than a 1,500-Hz probe tone produced residuals similar in amplitude to those from near-probe 2nd tones, and even higher residuals were produced by 2nd tones at some intermediate frequencies. Such high-amplitude, far-basal residuals provide strong evidence that high-frequency 2nd tones can produce substantial new probe-frequency

residual components. The only known mechanism to account for such high-amplitude residuals is the production of a new probe-frequency source (Shera et al. 2004). Our assessment is that far basal residuals are entirely, or almost entirely, from new probe-frequency sources.

Far-basal, passive SFOAE components are not suppressed and therefore will not be included in nsSFOAE measurements. How much do far-basal passive SFOAE actually contribute to true SFOAEs? The finding that the compression, suppression, and spectral smoothing methods produce very similar SFOAEs in humans (Kalluri and Shera 2007) suggests far-basal passive SFOAE components are generally small, at least in humans. This result also suggests that any new-source SFOAE components produced by the near-frequency suppressor used to obtain the nsSFOAE are also small. The amplitude of new-source SFOAE components produced by a second tone may be species dependent. High-frequency 2nd tones appear to create lower-amplitude residuals in humans than they do in cats or chinchillas (e.g., Guinan 1990; Backus and Guinan 2006; Keefe et al. 2008; Boothalingam and Lineton (2012); Charaziak et al. 2013; Charaziak and Siegel 2014). More work is needed to determine just how well nsSFOAEs equal true SFOAEs.

Charaziak and Siegel (2015a) measured ear canal sound pressures from a single tone stepped across a wide range of frequencies in chinchillas, and from an analysis of these data concluded that SFOAEs have components with short delays that arise more than an octave basal to the probe-frequency cochlear place. Since no 2nd tone was used, these data are not contaminated by new SFOAE components from the 2nd tone. Charaziak and Siegel's analysis involved transforming the data to the pseudo-time domain (as in Shera et al. 2008) where the different SFOAE components were separated by their pseudo-time delays, followed by a transformation of the components back to the frequency domain. The analysis procedure attempted to compensate for cochlear frequency dispersion but it is unclear how successful this was and, in particular, whether it adequately accounted for the lack of scaling symmetry in the cochlear apex. It seems possible that the wider tuning in the apex was transferred into the base and produced the apparent short delays for basal CFs. Noteworthy in Charaziak and Siegel (2015a) results is that the analysis shows short-latency components even in the 9000–10,000-Hz region of the chinchilla where BM motion shows nonlinearity primarily within ½ octave of the probe tone and where Charaziak et al. (2013) measurements showed no far-basal SFOAE components. The possible problems in this technique due to the lack of scaling might be avoided if the procedure was applied only to frequencies above 4000 Hz where the chinchilla cochlea is scaling symmetric (Temchin et al. 2008). Even if Charaziak and Siegel's analysis were correct, it only indicates that there are short latency SFOAE components but does not say where relative to the probe CF place they originate. One possibility is that the short-delay components revealed by Charaziak and Siegel's analysis are due to distortion from the peak region of the traveling wave, not reflections from far-basal regions. Charaziak and Siegel say that if short and long latency components are separated in the time domain and the short latency component is transformed to the frequency domain, the short-latency component is similar in form to the residual from a high-frequency suppressor. However, the match they show was adequate in one example and poor in the other, especially considering that they arbitrarily selected the time cutoff (with different cutoffs for the two cases) to get the best match. Overall, this evidence that a single probe tone evokes far-basal SFOAE components is not convincing.

Large SFOAEs components that originated far basal to the SFOAE CF region were produced in a model in which the cochlear irregularities were spatially low-pass filtered (Choi et al. 2008). This model filtered out cochlear irregularities that were on the scale of individual OHCs, which seems unrealistic (variation in motility from one OHC to the next may be the most important cochlear irregularity). The presence of small SFOAE components that originate far basal to the SFOAE CF region is expected from classic coherent reflection theory (Zweig and Shera 1995; Sisto et al. 2015). In essence, the Choi et al. model exaggerates the importance of these basal sources by filtering out SFOAE contributions from the peak of the SFOAE CF region. This filtering was done to mimic the high-amplitude frequency fine structure Choi et al. found in human SFOAEs using swept-tones and analyzed by a digital heterodyne method. However, Kalluri and Shera (2013), in a more comprehensive study using swept tones, did not find such a SFOAE fine structure. Furthermore, Kalluri and Shera's analysis showed that a fine structure can be artificially produced by using a frequency analysis window that is too short. Thus, the Choi et al. model may have mimicked a fine structure that was artificially produced. Overall, we do not consider this model demonstration that SFOAEs originate from the far-basal region as a good indication of what happens in a normal cochlea.

Implications for Transient Evoked Otoacoustic Emissions

TEOAEs are generated by the same mechanisms as SFOAEs (Shera and Guinan 1999; Kalluri and Shera 2007), so TEOAE data can shed light on SFOAE

generation and SFOAE conclusions can apply to TEOAEs. Lewis and Goodman (2014b) found multiple TEOAE components in humans with the shortest-delay peaks suppressed by tones about an octave above the nominal probe frequency. Since humans have sharper tuning than animals (Shera et al. 2002, 2010), one octave might be too far from the probe frequency to be within the cochlear-amplified region of the response to the tone pip. However, Lewis and Goodman's three-cycle tone pip had a very broad spectrum: at 1/2 octave above the pip center frequency, the energy was only about 10 dB below the energy at the peak (Lewis and Goodman 2014a). So a suppressor tone one octave above the peak frequency would be ½ octave above the upper shoulder of the stimulus bandwidth and perhaps within the suppression region for this energy. Overall, the multiple short-delay TEOAE peaks found by Lewis and Goodman can be reasonably interpreted as being near-basal OAE components.

TEOAEs, like SFOAEs, may have far basal components that are not cochlear amplified, grow linearly, and are not suppressed (Moleti et al. 2012; Moleti et al. 2013; Sisto et al. 2013; Sisto et al. 2015). Far-basal OAE components might be easier to demonstrate in TEOAEs than in SFOAEs because they are separated in time from the major response component. However, typical TEOAE data processing ignores short-latency responses because they can be contaminated by ringing of the sound stimulus. Stimulus ringing can be removed by various "nonlinear TEOAE" methods, but these methods also remove far-basal TEOAE components that grow linearly. The removal of far-basal TEOAE components might be an advantage in that the TEOAE components that remain provide more frequency-specific information about cochlear amplification than the whole TEOAE. Measuring TEOAEs by the nonlinear method and measuring SFOAEs by nsSFOAEs are theoretically deficient because they miss linearly growing or far-basal parts of the OAEs. However, these methods provide the practical advantage of capturing the part of the OAE that contributes the most information about cochlear amplification (unless the TEOAE measurements are at levels so low that the cochlea is operating linearly—see Kalluri and Shera 2007).

Why Do Low-Frequency SFOAEs Have Overall Delays That Are Too Short for Classic Coherent Reflection Theory?

First, we note that Siegel et al. (2005) concluded that chinchilla SFOAE group delays at *all* frequencies were too short to be accounted for by classic coherent reflection theory. However, a detailed analysis of chinchilla data that took into account the correction terms in the comparison of SFOAE group delays and BM group delays (theory indicates that they are not exactly a factor of two apart) showed that above 4000 Hz, SFOAE delays are accurately predicted from BM delays using classic coherent reflection theory (Shera et al. 2008; Shera et al. 2010). However, below 4000 Hz, the SFOAE group delays were too short—which is the issue we address here (note that this transition frequency may vary across species).

Our guinea pig SFOAE latencies are very similar to those measured by Shera and Guinan (2003) using the phase-gradient method, and our conclusion that below 3000 Hz overall SFOAE latencies are too short to be accounted for by classic coherent reflection theory concurs with theirs. What might account for this and the similar discrepancies in other species? One possibility is that low frequency BM motion cannot be extrapolated from high-frequency measurements and/or that existing apical mechanical measurements are inaccurate at low frequencies. Apical mechanical measurements have been done without a good test for cochlear sensitivity at low frequencies (see Lichtenhan et al. 2013) so they may not show normal cochlear motion. These possibilities cannot be ruled out, but several other hypotheses may better explain the too-short SFOAE delays.

Low-frequency regions where SFOAE components had extremely short latencies were sometimes found in chinchillas (Fig. 7). These regions have delays that are so uniformly short that it seems likely that these are regions where distortion produced the SFOAEs. Models predict that cochlear nonlinearity can create wave-fixed distortion with near-zero group delays at the SFOAE frequency (Talmadge et al. 2000). Furthermore, the low-frequency SFOAEs from chinchillas sometimes show a fine frequency structure as if there were two beating SFOAE components (Siegel et al. 2005; Shera et al. 2008) (In contrast, SFOAE dips in guinea pigs did not show regular patterns). In addition, there were frequency regions in chinchillas where the average SFOAE group delay was shorter than a single forward traveling-wave delay (Siegel et al. 2005). Both of these would be accounted for by the chinchilla SFOAE group delays originating from a combination of distortion and coherent reflection. The chinchilla data were obtained from 30 dB SPL tones and it might seem surprising to find distortion producing significant SFOAEs at such low sound levels. However, that explanation fits well with the chinchilla data (Fig. 7), and at 30 dB SPL, chinchilla BM motion (in the cochlear base) shows highly nonlinear growth (e.g., Ruggero et al. 1997; Rhode 2007). Note that the nsSFOAE components from both distortion and coherent reflection originate from the traveling-wave peak region, are a consequence of cochlear amplification at the SFOAE

frequency, and can be revealed by a close-frequency suppressor.

Revising Coherent Reflection Theory for Multiple Cochlear Motions

Another hypothesis to explain low-frequency SFOAEs with group delays that are too short to be from classic coherent reflection of BM motion is that these SFOAEs arise by coherent reflection from cochlear motion(s) that have a shorter group delay than BM motion. Strong evidence for short-delay motion in the cochlear apex comes from single auditory-nerve-fiber responses that have phase vs. frequency plots showing frequency regions with different group delays in cats, guinea pigs, gerbils, and chinchillas (Pfeiffer and Molnar 1970; Kiang 1984; Ronken 1986; Palmer and Shackleton 2009; Temchin and Ruggero 2010). These data imply that an individual inner hair cell receives mechanical drives that have different group delays. In these data, the delay at the characteristic frequency (CF) (which is usually presumed to show the delay of BM motion) was typically longer than the delays at frequencies away from CF. Thus, the delays in the shallower-slope regions away from CF would be shorter than the group delay of BM motion.

The auditory-nerve-initial-peak (ANIP) response to high-level clicks has a shorter delay than the bulk of the neural response and may come from a cochlear motion that has a shorter group delay than BM motion (Guinan et al. 2005; Guinan and Cooper 2008). The ANIP response is strongly and selectively inhibited by MOC stimulation, which shows that the ANIP response behaves differently than the bulk of the response. The ANIP response was produced by 90 dB SPL clicks in cats (i.e., it is a high-level response, but note that the multiple auditory-nerve-fiber group delays cited above were found at low sound levels). A neural response in guinea pigs that was elicited by 92 dB SPL clicks was reported by Versnel et al. (1990; see "Results" for a further description). For fibers with low best frequencies, it seems likely that Versnel et al. largest PST peak in guinea pigs is equivalent to the ANIP peak in cats. Thus, for low frequencies, the dashed lines in Figure 6 would show the expected SFOAE delays that result from the motion that produces the ANIP response. Below 2000 Hz, the SFOAE latencies estimated from the auditory nerve data fit well with the main delay of the SFOAE data (Figs. 6 and 8, bottom left). Based on these data, we hypothesize that at low frequencies, coherent reflections arise primarily from cochlear motion(s) that have shorter group delays than BM motion. The ANIP motion is one such motion, but there may be others (Guinan 2012).

When classic coherent reflection theory was formulated, it was thought that BM motion and other motions of the organ of Corti were locked together. However, direct measurements in living, sensitive cochleae show that the motion of the reticular lamina can be twice as large as BM motion, that these motions have different phases, and that the nonlinearities are different in the reticular lamina vs. the BM (Chen et al. 2011; Zha et al. 2012). In guinea pigs, for motions at frequencies of 3000 Hz or less, the reticular lamina tilts about the top of the pillar cells allowing fluid flow in and out of the reticular-lamina tectorial-membrane gap (Nowotny and Gummer 2006; Guinan 2012). Thus, other cochlear motions can be very different from BM motion, particularly at low frequencies. The theory of coherent reflection needs to be revised to include coherent reflections from the motion of structures other than the BM. Which of these motions is the dominant SFOAE source may vary with stimulus frequency, stimulus level, cochlear location, species, and the pattern of cochlear irregularities. Furthermore, reflections from the motions of structures that are vibrating out of phase can interfere so that phase variations across cochlear structures will also shape SFOAEs. With these things in mind, we hypothesize that "too-short" low-frequency SFOAEs are produced, at least in part, by coherent reflection from motions of cochlear structures that have group delays that are shorter than the group delays of BM motion.

CONCLUSIONS

At many frequencies, the time-frequency analysis of nsSFOAEs found only one component, which supports the hypotheses that nsSFOAEs predominantly provide information about the peak region of the traveling wave (the SFOAE CF region). Multiple SFOAE components can be a result of the pattern of cochlear irregularities producing cancellation of the energy reflected in the SFOAE CF region and they do not necessarily indicate that there are SFOAE sources that come from outside of the SFOAE amplification region. Our analysis of existing data indicate that SFOAE-frequency residuals produced by a 2nd tone that is two octaves or more above the SFOAE frequency are most likely created by nonlinearity at the peak of the 2nd tone and are not present without the 2nd tone. SFOAEs at low frequencies that have delays too short to be reflections from BM motion may be due to a combination of distortion-produced SFOAE components and coherent reflections from apical cochlear motion that has a shorter group delay than BM-motion group delay. Finally, our results

indicate that coherent reflection theory needs to be revised to include reflections from the motions of multiple cochlear structures, not just from the basilar membrane.

ACKNOWLEDGMENTS

We thank Drs. C.A. Shera, M.C. Brown, D.C. Mountain, and K. K. Charaziak for comments on earlier versions of the manuscript. This work was supported by RO1 DC000235, R01 DC003687, T32 DC00038, P30 DC005209, and a NSF GRFP, and was part M. B-G's PhD thesis.

Conflict of Interest

The authors declare that they have no conflict of interest.

REFERENCES

Backus BC, Guinan JJ Jr (2006) Time course of the human medial olivocochlear reflex. J Acoust Soc Am 119:2889–2904

Boothalingam S, Lineton B (2012) Effect of contralateral acoustic stimulation on cochlear tuning measured using stimulus frequency and distortion product OAEs. Int J Audiol 51:892–9

Charaziak KK, Siegel JH (2014) Estimating cochlear frequency selectivity with stimulus-frequency otoacoustic emissions in chinchillas. J Assoc Res Otolaryngol 15:883–896

Charaziak KC, Siegel JH (2015a) Short-latency SFOAE components extracted with different methods. J Assoc Res Otolaryngol 38:169, **#282**

Charaziak KK, Siegel JH (2015b) Tuning of SFOAEs evoked by low-frequency tones is not compatible with localized emission generation. J Assoc Res Otolaryngol 16:317–29

Charaziak KK, Souza P, Siegel JH (2013) Stimulus-frequency otoacoustic emission suppression tuning in humans: comparison to behavioral tuning. J Assoc Res Otolaryngol 14:843–862

Chen F, Zha D, Fridberger A, Zheng J, Choudhury N, Jacques SL, Wang RK, Shi X, Nuttall AL (2011) A differentially amplified motion in the ear for near-threshold sound detection. Nat Neurosci 14:770–774

Choi YS, Lee SY, Parham K, Neely ST, Kim DO (2008) Stimulus-frequency otoacoustic emission: measurements in humans and simulations with an active cochlear model. J Acoust Soc Am 123:2651–2669

Geisler CD, Yates GK, Patuzzi RB, Johnston BM (1990) Saturation of outer hair cell receptor currents causes two-tone suppression. Hear Res 44:241–256

Goodman SS, Withnell RH, Shera CA (2003) The origin of SFOAE microstructure in the guinea pig. Hear Res 183:7–17

Guinan JJ Jr (1990) "Changes in stimulus frequency otoacoustic emissions produced by two-tone suppression and efferent stimulation in cats". In: Dallos P, Geisler CD, Matthews JW, Steele CR (eds) Mechanics and biophysics of hearing. Springer, Madison, WI, pp 170–177

Guinan JJ Jr (2012) How are inner hair cells stimulated? Evidence for multiple mechanical drives. Hear Res 292:35–50

Guinan JJ Jr, Cooper NP (2008) Medial olivocochlear efferent inhibition of basilar-membrane responses to clicks: evidence for two modes of cochlear mechanical excitation. J Acoust Soc Am 124:1080–1092

Guinan JJ, Backus BC, Lilaonitkul W, Aharonson V (2003) Medial olivocochlear efferent reflex in humans: otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs. J Assoc Res Otolaryngol 4:521–540

Guinan JJ Jr, Lin T, Cheng H (2005) Medial-olivocochlear-efferent inhibition of the first peak of auditory-nerve responses: evidence for a new motion within the cochlea. J Acoust Soc Am 118:2421–2433

Joris PX, Bergevin C, Kalluri R, Mc Laughlin M, Michelet P, van der Heijden M, Shera CA (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. Proc Natl Acad Sci U S A 108:17516–17520

Kalluri R, Shera CA (2001) Distortion-product source unmixing: a test of the two-mechanism model for DPOAE generation. J Acoust Soc Am 109:622–637

Kalluri R, Shera CA (2007) Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions. J Acoust Soc Am 121:2097–2110

Kalluri R, Shera CA (2013) Measuring stimulus-frequency otoacoustic emissions using swept tones. J Acoust Soc Am 134:356–368

Keefe DH, Ellison JC, Fitzpatrick DF, Gorga MP (2008) Two-tone suppression of stimulus frequency otoacoustic emissions. J Acoust Soc Am 123:1479–94

Kiang, NYS (1984) Peripheral neural processing of auditory information. In: Handbook of physiology, section 1: the nervous system, Vol. 3 (Sensory processes). American Physiological Society, Bethesda, MD, pp. 639–674

Lewis, J D, Goodman, S S (2014a). Basal contributions to short-latency transient-evoked otoacoustic emission components. J Assoc Res Otolaryngol. 1-17

Lewis JD, Goodman SS (2014b) The effect of stimulus bandwidth on the nonlinear-derived tone-burst-evoked otoacoustic emission. J Assoc Res Otolaryngol 15:915–931

Lichtenhan JT, Cooper NP, Guinan JJ Jr (2013) A new auditory threshold estimation technique for low frequencies: proof of concept. Ear Hear 34:42–51

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

Moleti A, Mohsin Al-Maamury A, Bertaccini D, Botti T, Sisto R (2013) Generation place of the long- and short-latency components of transient-evoked otoacoustic emissions in a nonlinear cochlear model. J Acoust Soc Am 133:4098–4108

Nowotny M, Gummer AW (2006) Nanomechanics of the subtectorial space caused by electromechanics of cochlear outer hair cells. Proc Natl Acad Sci U S A 103:2120–2125

Palmer AR, Shackleton TM (2009) Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency. J Assoc Res Otolaryngol 10:233–250

Pfeiffer RR, Molnar CE (1970) Cochlear nerve fiber discharge patterns: relationship to the cochlear microphonic. Science 167:1614–1616

Rhode WS (2007) Basilar membrane mechanics in the 6–9 kHz region of sensitive chinchilla cochleae. J Acoust Soc Am 121:2792–2804

Ronken DA (1986) Anomalous phase relations in threshold-level responses from gerbil auditory nerve fibers. J Acoust Soc Am 79:417–425

Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L (1997) Basilar-membrane responses to tones at the base of the chinchilla cochlea. J Acoust Soc Am 101:2151–2163

SHERA CA, BERGEVIN C (2012) Obtaining reliable phase-gradient delays from otoacoustic emission data. J Acoust Soc Am 132:927–943

SHERA CA, GUINAN JJ JR (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 JJ JR (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 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

SHERA C A, TUBIS A, TALMADGE C L, GUINAN J J, JR. (2004). THE DUAL EFFECT OF "SUPPRESSOR" TONES ON STIMULUS-FREQUENCY OTOACOUSTIC EMISSIONS. ASSOC RES OTOLARYNGOL ABSTR 27, ABS. 776

SHERA CA, TUBIS A, TALMADGE CL (2008) Testing coherent reflection in chinchilla. J Acoust Soc Am 123:3851

SHERA CA, GUINAN JJ JR, OXENHAM AJ (2010) Otoacoustic estimation of cochlear tuning: validation in the chinchilla. J Assoc Res Otolaryngol 11:343–365

SIEGEL JH, TEMCHIN AN, RUGGERO M (2003) Empirical estimates of the spatial origin of stimulus-frequency otoacoustic emissions. Assoc Res Otolaryngol Abstr 26:172 **(#679)**

SIEGEL JH, CERKA AJ, RECIO-SPINOSO A, TEMCHIN AN, VAN DIJK P, RUGGERO M (2005) Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering. J Acoust Soc Am 118:2434–2443

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

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

TALMADGE CL, TUBIS A, LONG GR, TONG C (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

TEMCHIN AN, RUGGERO MA (2010) Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics. J Assoc Res Otolaryngol 11:297–318

TEMCHIN AN, RECIO-SPINOSO A, VAN DIJK P, RUGGERO MA (2005) Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks and noise and comparison with basilar-membrane vibrations. J Neurophysiol 93:3635–48

TEMCHIN AN, RICH NC, RUGGERO MA (2008) Threshold tuning curves of chinchilla auditory-nerve fibers. I. Dependence on characteristic frequency and relation to the magnitudes of cochlear vibrations. J Neurophysiol 100:2889–98

TSUJI J, LIBERMAN MC (1997) Intracellular labeling of auditory nerve fibers in guinea pig: central and peripheral projections. J Comp Neurol 381:188–202

VERSNEL H, PRIJS VF, SCHOONHOVEN R (1990) Single-fibre responses to clicks in relationship to the compound action potential in the guinea pig. Hear Res 46:147–160

ZHA D, CHEN F, RAMAMOORTHY S, FRIDBERGER A, CHOUDHURY N, JACQUES SL, WANG RK, NUTTALL AL (2012) In vivo outer hair cell length changes expose the active process in the cochlea. PLoS One 7, e32757

ZWEIG G, SHERA CA (1995) The origin of periodicity in the spectrum of evoked otoacoustic emissions. J Acoust Soc Am 98:2018–2047