Available online at www.sciencedirect.com
SCIENCE DIRECT®
Hearing Research 207 (2005) 76–86
HEARING RESEARCH
www.elsevier.com/locate/heares

A comparison of OAEs arising from different generation mechanisms in guinea pig

Robert H. Withnell *, Sumitrajit Dhar [1], Angela Thomsen

Department of Speech and Hearing Sciences, Indiana University, 200 South Jordan Avenue, Bloomington, IN 47405, United States

Received 30 September 2004; received in revised form 14 April 2005; accepted 14 April 2005
Available online 1 June 2005

Abstract

Otoacoustic emissions provide unambiguous evidence that the cochlea supports energy propagation both towards, and away from, the stapes. The standard wave model for energy transport and cochlear mechanical amplification provides for compressional and inertial waves to transport this energy, the compressional wave through the fluids and the inertial wave along the basilar membrane via fluid coupling. It is generally accepted that energy propagation away from the stapes is dominated by a traveling wave mechanism along the basilar membrane. The mechanism by which energy is predominantly transported back to the stapes remains controversial. Here, we compared signal onset delay measurements and rise/steady-state/fall times for SFOAEs and $2f_1 - f_2$ OAEs ($f_2/f_1 = 1.2$) obtained using a pulsed-tone paradigm in guinea pig. Comparison of $2f_1 - f_2$ OAE signal onset delay for the OAE arising from the f_2 region with SFOAE signal onset delay (matched to the f_2 stimulus frequency) based on signal onset occurring at 10% of the peak signal amplitude was suggestive of a bi-directional traveling wave mechanism. However, significant variability in signal onset delay and signal rise, steady-state duration, and fall times for both the $2f_1 - f_2$ OAE and SFOAE was found, qualifying this interpretation. Such variability requires explanation, awaiting further studies.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Otoacoustic emission; Guinea pig; Traveling wave; Signal onset delay

1. Introduction

The hydrodynamics of the cochlea provides for energy transport away from the stapes in one of two forms: an acoustic compression wave (fast wave) and via an inertially mediated bulk fluid flow (slow wave) that produces a pressure difference across the basilar membrane (Lighthill, 1991; Yates, 1995). Sensory transduction is mediated by the latter or slow wave, the graded mechanical properties of the basilar membrane (BM) providing a spatial tuning with each frequency mapped to a different cochlear location. It was originally believed that there was a preferential direction for the propagation of energy along the cochlear partition toward the helicotrema (Zwislocki, 1953; Bekesy, 1960). However, with the discovery of otoacoustic emissions (OAEs) (Kemp, 1978) it was clear that the cochlea supported the propagation of energy in both directions. Essentially all models of cochlear function assume that this propagation of energy in both directions involves a traveling wave mechanism, i.e., spatial filtering associated with the mechanical properties of the BM (stiffness and mass) will influence energy propagation equally in both directions. However, as early as 1980, a reverse traveling wave mechanism was questioned (Wilson, 1980), at issue being the delay time or lack thereof for the reverse propagation of energy. Recently, Ren

* Corresponding author. Tel.: +1 812 855 9339; fax: +1 812 855 5531.
E-mail addresses: rwithnel@indiana.edu (R.H. Withnell), s-dhar@northwestern.edu (S. Dhar), athomsen@indiana.edu (A. Thomsen).
[1] Present address: Hugh Knowles Center, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, United States.

0378-5955/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2005.04.005

(2004) reported no evidence for reverse traveling waves from BM measurements, and stapes vibration that preceded BM vibration. In contrast to all previous studies of BM vibration, Ren (2004) was able to measure from not one cochlear location but rather longitudinally over about 1 mm. Ren measured the intermodulation distortion product $2f_1 - f_2$ phase for stimulus frequency ratios (f_2/f_1) ranging from 1.05 to 1.2 and found in all cases a negative slope, i.e., the data indicated only forward traveling waves for $2f_1 - f_2$. Ren concluded, based on the lack of evidence for a reverse traveling wave, that energy propagates back to the stapes nearly instantaneously as an acoustic compression wave. However, this interpretation, while persuasive, is not unequivocal. Ren's findings do not exclude the possibility that the $2f_1 - f_2$ forward traveling wave dominance is a construct of stimulus frequency ratio and that at higher stimulus frequency ratios, a $2f_1 - f_2$ reverse traveling wave would dominate (Shera, 2003).

Bell and Fletcher (2004) recently proposed a model for acoustic wave generation within the cochlea. This model utilizes a local resonance phenomenon. Acoustic waves would also be produced with the standard wave model for energy transport and cochlear mechanical amplification – additional energy added to BM vibration would produce both compressional and inertial waves that propagate back to the stapes. Regardless of whether fluid coupling is achieved by a local resonance or wave model, an acoustic compressional wave must be present in reverse energy transport. But is it significant?

Investigation of the existence of a reverse traveling wave can also be done non-invasively using OAEs, specifically, comparing DPOAE travel times with SFOAE travel times. DPOAEs arise from the nonlinear interaction on the BM of the responses to two pure tone stimuli with frequencies that are arithmetically related to the stimulus frequencies. DPOAEs in humans have been shown to arise from the complex interaction of components coming from two different locations on the BM (Heitmann et al., 1998), with different generating mechanisms dominating the production of each component (Talmadge et al., 1999; Kalluri and Shera, 2001). This is illustrated schematically in Fig. 1 for a cochlea that has been unfurled with the membranous labyrinth (Scala media) represented by the BM. Subsequent to an input of stimulus tones of frequencies f_1 and f_2, nonlinear interaction on the BM results in $2f_1 - f_2$ being generated in the region of overlap of the f_1 and f_2 excitation patterns (mostly near f_2), this $2f_1 - f_2$ component then propagating both in the forward and reverse directions along the BM. The component propagating in the forward direction propagates as a traveling wave (Kirk and Yates, 1994; Ren, 2004) to its own characteristic frequency place from where it has been suggested that it is reflected back to the stapes (Talmadge et al., 1999; Kalluri and Shera, 2001).

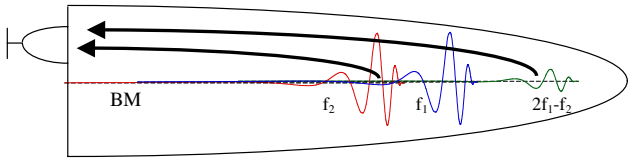

Fig. 1. Schematic of $2f_1 - f_2$ generation with the cochlea unfurled and the membranous labyrinth (Scala media) represented by the BM. Subsequent to an input of stimulus tones of frequencies f_1 and f_2, nonlinear interaction on the BM results in $2f_1 - f_2$ being generated in the region of overlap of the f_1 and f_2 excitation patterns (mostly near f_2), this $2f_1 - f_2$ component then propagating both in the forward and reverse directions along the BM. The component propagating in the forward direction then propagates as a traveling wave that is then thought to be reflected from its characteristic frequency region back to the stapes.

The physical delay for the nonlinear distortion component of the $2f_1 - f_2$ DPOAE can be expressed in terms of the travel times of the transpartition pressure waves of frequencies f_2 and $2f_1 - f_2$ (Talmadge et al., 1999):

$$\tau_{\mathrm{dpoae}} = \tau_2 + \tau_{\mathrm{dp}}, \tag{1}$$

where τ_{dpoae} is the physical delay for the $2f_1 - f_2$ DPOAE; τ_2 is the physical delay of the forward traveling wave of frequency f_2 to propagate to the region of f_2 plus the delay for the f_2 signal to propagate from the point of measurement in the ear canal to the stapes; τ_{dp} is the physical delay for the reverse traveling wave of frequency $2f_1 - f_2$ to propagate from the f_2 region back to the stapes plus the delay for the $2f_1 - f_2$ energy to propagate back through the middle ear and then to the microphone in the ear canal.

The above physical delays refer only to delay measurements made in the time domain (Tubis et al., 2000) for a pulsed f_2 paradigm.

Based on a bi-directional traveling wave mechanism, at a stimulus frequency ratio of 1.2 it is expected that the $2f_1 - f_2$ energy that arises from the f_2 region should propagate back to the stapes along the BM with a velocity that is not significantly altered by the graded mechanical properties of the BM. This is evident from inspection of Fig. 2 which shows BM phase versus frequency recorded from the first turn of a guinea pig cochlea (40 dB SPL stimulus tones) (data courtesy of De Boer and Nuttall). Fig. 2 suggests that group delay ($\mathrm{d}\phi/\mathrm{d}\omega$), measured at the 18 kHz CF location, is relatively constant up to about 14 kHz in frequency. A constant velocity of propagation along the BM back to the stapes means that that the $2f_1 - f_2$ DPOAE is arising in the long-wave region.[2] τ_{dp} (the reverse travel time for the $2f_1 - f_2$ DPOAE component arising from the f_2 CF

[2] Wave velocity need not be constant for a wave to be arising in the long-wave region, but a wave that propagates with constant velocity has to be in the long-wave region (wave velocity cannot be constant in the short wave region).

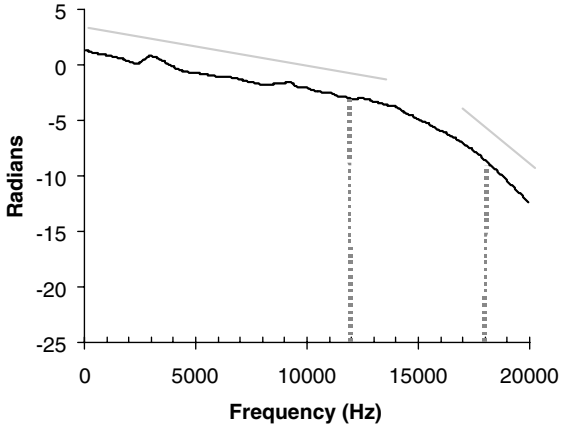

Fig. 2. Basilar membrane phase versus frequency recorded from the first turn of a guinea pig cochlea (40 dB SPL stimulus tones) (data courtesy of De Boer and Nuttall). Group delay ($d\phi/d\omega$) is relatively constant up to about 14 kHz, with a delay of 32 μs. At the characteristic frequency (CF) location of 18 kHz the group delay is 280 μs.

region) should be of the order of 10% of the forward travel time (for f_2 = 18 kHz, $2f_1 - f_2$ = 12 kHz, when f_2/f_1 = 1.2) based on a bi-directional traveling wave mechanism. τ_{dp} would be negligible for $2f_1 - f_2$ propagating back to the stapes via an acoustic compression wave mechanism.

τ_2 is assumed to represent the delay for the forward traveling wave of frequency f_2 to reach its CF place. This assumption is based on the notion that the region of origin that contributes most to the generation of $2f_1 - f_2$ will be that region where the amplitude and phase of the f_1 and f_2 traveling waves combine to produce the maximum $2f_1 - f_2$. Assuming that the dominant cochlear outer hair cell nonlinearity produces amplitude but not phase distortion, this region will be where $A_1^2 * A_2$ is maximized (where A_1 and A_2 are the BM displacement amplitudes for the f_1 and f_2 traveling waves), i.e., at and perhaps slightly apical to the f_2 place. While the phase of f_2 rotates rapidly near the f_2 place, the phase of $2f_1 - f_2$, determined by the phase rotation of f_1 and f_2, does not, and so $2f_1 - f_2$ being generated over this localized region will add in-phase (Yates, 1998). The round-trip travel time for the nonlinear distortion component of the $2f_1 - f_2$ DPOAE measured at a frequency ratio of 1.2, based on a bi-directional traveling wave mechanism, is expected to be dominated by the forward travel time (τ_2) of the f_2 stimulus to its CF region.

At low stimulus levels, SFOAEs are thought to arise from reflection from randomly distributed spatial inhomogeneities (Zweig and Shera, 1995), i.e., a linear place-fixed reflection mechanism, with reflection arising predominantly from the peak or tip of the BM excitation pattern. For such a bi-directional traveling wave mechanism, SFOAEs, in contrast to the $2f_1 - f_2$ DPOAE (at a frequency ratio of 1.2), represent a round trip travel time with both the forward and reverse-going waves being equally affected by spatial filtering. This OAE would have a cochlear delay that is approximately equivalent to twice the delay of the forward travel time of the f_2 stimulus to its CF place (where the SFOAE is evoked by a stimulus of frequency f_2). Small random fluctuations in the density of these irregularities are proposed to significantly affect cochlear reflectance (Zweig and Shera, 1995) and so amplitude and phase of the SFOAE. Experimental findings consistent with this hypothesis have been reported by Shera and Guinan (2003) and Goodman et al. (2003). Significantly, such variations in cochlear reflectance will produce considerable variability in cochlear delay times. However, based on average SFOAE delay data, it is expected, for a bi-directional traveling wave mechanism for the propagation of energy along the BM, that SFOAE delay will be approximately equal to twice the forward travel time (τ_2) of the f_2 stimulus

$$\text{i.e., } \tau_{sfoae} \approx 2 * \tau_2 \approx 2(\tau_{dpoae} - \tau_{dp}), \quad \text{where } \tau_{dp} \approx 0.1 * \tau_2. \tag{2}$$

An acoustic compression wave mechanism for the reverse propagation of energy would mean that the reverse travel time would be negligible and

$$\tau_{sfoae} \approx \tau_2 \approx \tau_{dpoae} - \tau_{dp}, \tag{3}$$

where f_2/f_1 = 1.2 for the DPOAE.

Cochlear travel times for DPOAEs based on a phase-gradient estimate of delay are confounded by there not being a clear understanding of the relationship of the phase-gradient (for either of the f_1 or f_2 stimulus tones held constant with the other tone swept in frequency) with physical cochlear delay (Shera et al., 2000). In contrast, a pulsed-tone paradigm with a continuous f_1 tone and a pulsed f_2 tone provides a direct time-domain measure of cochlear delay (Tubis et al., 2000), i.e., signal onset delay. In this study, signal onset delay for the $2f_1 - f_2$ DPOAE arising from the f_2 region (f_2/f_1 = 1.2) is examined in terms of its relationship with SFOAE signal onset delay. Guinea pigs are used as the experimental subject. Unlike in humans, the $2f_1 - f_2$ OAE measured in the guinea pig at a stimulus frequency ratio of 1.2 arises primarily from the f_2 region, the DPOAE component from the DP place not being significant (Withnell et al., 2003; Schneider et al., 2003).

2. Method

2.1. Animal surgery

Albino guinea pigs (300–650 g) were anesthetized with Nembutal (30–35 mg/kg i.p.) and Atropine (0.06–0.09 mg i.p.), followed approximately 15 min later by Hypnorm (0.1–0.15 ml i.m.). Anesthesia was maintained

with supplemental doses of Nembutal and Hypnorm. In a number of animals, Pancuronium (0.15 ml i.m.) was administered to reduce physiological noise associated with spontaneous muscle contractions. Guinea pigs were tracheotomized and mechanically ventilated on Carbogen (5% CO_2 in O_2) with body rectal temperature maintained at approximately 38.5 °C. The head was positioned using a custom-made head holder that could be rotated for access to the ear canal. Heart rate was monitored throughout each experiment. The bulla was opened dorso-laterally and a silver wire electrode placed on the round window niche for the recording and monitoring of the compound action potential (CAP). A plastic tube was placed in the bulla opening to ensure that the bulla was adequately ventilated, although no attempt was made to seal the bulla.

$2f_1 - f_2$ OAE data reported here was obtained from ten animals while SFOAE data was obtained from five animals (a different set of animals from the 10 used for DPOAE experiments). Experimentation on animals used in this study was approved by Indiana University Bloomington Animal Care and Use Committee.

2.2. Signal generation and data acquisition

OAEs were recorded with stimulus delivery and response acquisition computer-controlled using custom-software with either a MOTU 828, 24 bit, 44.1 kHz digital audio workstation (DPOAEs) or a Card Deluxe sound card (SFOAEs). The acoustic stimuli were delivered by a Beyer DT48 dynamic loudspeaker placed approximately 4 cm from the entrance to the ear canal. Ear canal sound pressure recordings were made by a Sennheiser MKE 2–5 electrostatic microphone fitted with a metal probe tube (1.2 mm long, 1.3 mm i.d., 1500 Ω acoustic resistor) positioned approximately 2 mm into the ear canal. The microphone and probe tube combination was calibrated against a Bruel and Kjaer 1/8 in. microphone. The output from the probe tube microphone was amplified 20 dB, high-pass filtered, and transmitted as a balanced input to one of the analogue input channels of the Card Deluxe sound card or MOTU workstation. It was subsequently digitized in 250 ms epochs at a rate of 44.1 kHz (DPOAEs) or in 42.6 ms epochs at a rate of 96 kHz (SFOAEs).

2.2.1. DPOAEs

The stimulus complex consisted of two tones generated using a pulsed-tone paradigm (Talmadge et al., 1999) with one of the stimulus tones (f_1) presented continuously while the other tone (f_2) was pulsed on for 100 ms in every 250 ms with ~2.5 ms rise and fall times and 98.5 ms duration. Stimuli were digitally generated and output separately on two different output channels, mixed without amplification and buffered by a Tucker-Davis Technologies HB6 loudspeaker buffer-amplifier. Stimulus frequency ratio was fixed at 1.2, with f_2 = 5400–11,340 Hz (540 Hz step-size) and stimulus level ~74 dB pSPL – stimulus level was based on a constant voltage delivered to the loudspeaker with probe tube corrections made post hoc; the data set was restricted to that with L_2 ranging from 70 to 78 dB pSPL with $0 \leqslant L_1/L_2 \leqslant 10$ dB.

The f_2 stimulus tone and $2f_1 - f_2$ OAE were extracted from the averaged time domain ear canal sound pressure recording using narrow band filtering of the time domain signal about the center-frequency of each component of interest. Using Matlab, each ear canal sound pressure recording was filtered using the filtfilt function (performs zero-phase digital filtering by processing the input data in both the forward and reverse directions). The filter used was a band-pass, linear phase FIR digital filter (Hamming window) with a bandwidth of 200 Hz and a filter order of 200.

2.2.2. SFOAEs

The stimulus was an amplitude modulated (188 Hz) tone burst with ~3 ms rise and fall times and 27.5 ms duration. Stimulus frequency varied between 5400 and 10150 Hz. Stimulus level was ~70 dB pSPL (65–72 dB pSPL). The OAE was extracted from averaged ear canal sound pressure recording using the nonlinear differential extraction technique (see Kemp et al., 1990) with a stimulus level ratio of 6 dB. The ear canal sound pressure and nonlinear derived OAE were narrow-band filtered as per DPOAEs.

2.2.3. Definition of signal onset

Signal onset was defined in two ways:

i. the point in time at which the absolute value of the Hilbert transform of the signal equals 10% of the peak amplitude;
ii. the point in time at which the absolute value of the Hilbert transform of the signal equals −3 dB relative to the peak amplitude.

OAE delay was quantified in terms of signal onset delay, i.e., OAE onset minus signal onset. Definition (i) is intended to be a good approximation to signal onset. (ii) was included for comparison purposes as this is the definition most commonly used in the literature. Both of these definitions are independent of the noise floor. An alternative to definitions of signal onset referenced to the peak amplitude of the signal would be a signal onset based on the noise floor but such a definition does not provide a comparable estimate of signal onset for signals with differing signal to noise ratios.

Fig. 3 provides an example of the f_2 stimulus and $2f_1 - f_2$

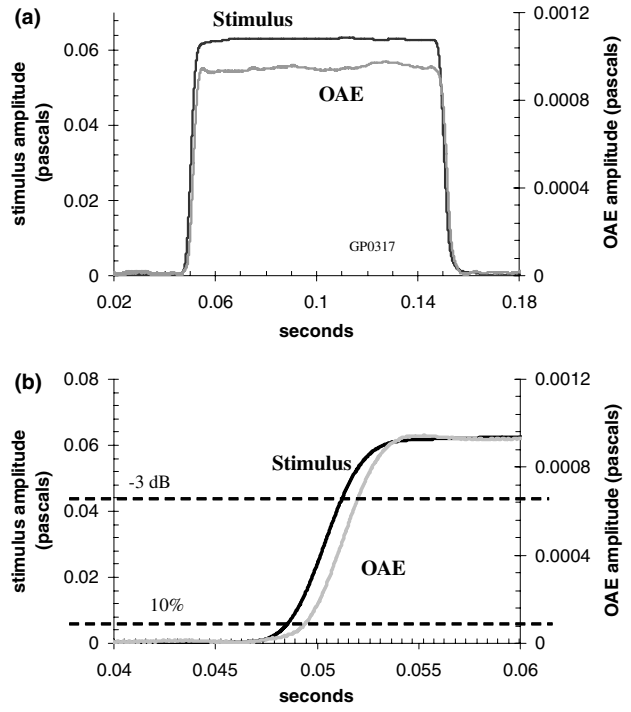

Fig. 3. An example of the f_2 stimulus and $2f_1 - f_2$ OAE obtained (magnitude of the Hilbert transform of the time domain waveform) in panel (a) and illustration of the signal onset for each of the two definitions described in the text in panel (b) where the time axis has been abbreviated to concentrate on the region of stimulus and OAE onset.

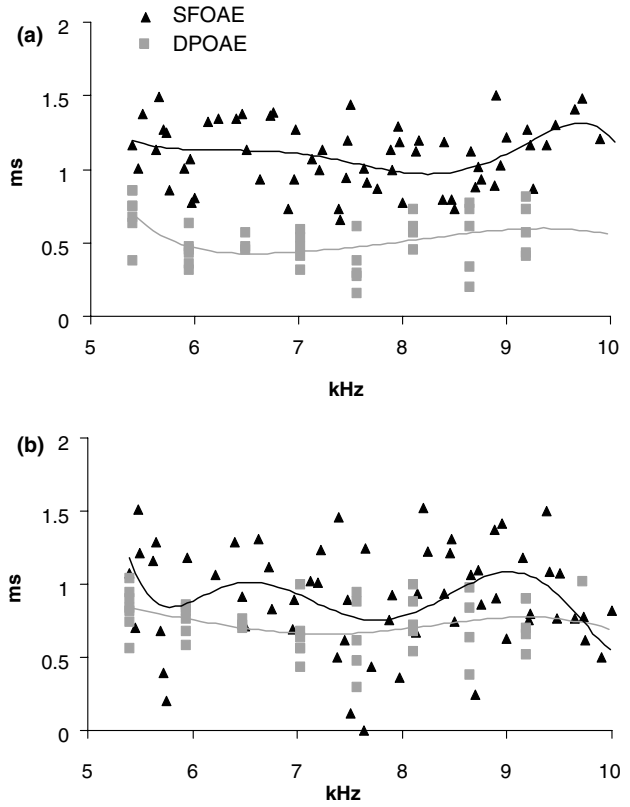

Fig. 4. $2f_1 - f_2$ OAE and SFOAE signal onset delay versus frequency calculated from 10% of the peak amplitude in panel (a) and −3 dB of the peak amplitude in panel (b). In both panels, signal onset delay data (10th to 90th percentile) for each OAE was fit with a sixth order polynomial. The difference in the fitted curves in panel (a) is suggestive of a bi-directional traveling wave mechanism. Panel (b) shows the $2f_1 - f_2$ OAE and SFOAE signal onset delay data based on −3 dB re the peak amplitude to overlap considerably – see text for further discussion.

in panel (b) where the time axis has been abbreviated to concentrate on the region of stimulus and OAE onset.

2.2.4. Choice of stimulus level

The choice of stimulus level for the DPOAE and SFOAE (~70 dB pSPL) was made based on the goal of obtaining OAEs with a good signal to noise ratio, to reduce the effect of noise on the measurement of delay. This does however introduce a stimulus-level dependent confounding effect of a possible mixing of components arising from different mechanisms (Talmadge et al., 2000; Goodman et al., 2003). However, the signal onset definition of 10% of the peak amplitude corresponds to an effective stimulus level of approximately 50 dB pSPL, a stimulus level that should produce an SFOAE dominated by a linear coherent reflection mechanism (Goodman et al., 2003). The potential problem of a stimulus level-dependent mixing of components arising from different mechanisms will be considered further in Section 4.

3. Results

3.1. Signal onset delay

Signal onset delay for the $2f_1 - f_2$ OAE (versus f_2 frequency) and SFOAE (stimulus frequency = f_2) calculated from the signal onset of the f_2 stimulus and OAE using each of the two definitions is shown in Fig. 4. For both the $2f_1 - f_2$ OAE, generated by a nonlinear distortion mechanism, and the SFOAE, generated by a place-fixed reflection mechanism, considerable scatter was present in the data for both definitions of signal onset delay, confounding obtaining a fit to the data. To facilitate curve-fitting while removing the effect of data outliers, data is shown from the 10th to 90th percentile of signal onset delay values obtained for each of the OAEs, with curves fitted to this data set. Each data set was fit with a sixth order polynomial.[3] It was expected that a power law ($\tau \propto f^b$) would best describe cochlear delay versus frequency (Shera and Guinan, 2003), but the scatter in the data and the data being measured over only a one octave range presumably obfuscates such a relationship. The scatter in the data for SFOAEs has

[3] Microsoft Excel provides for curve fitting of data sets using six different curve options: linear, log, polynomial, exponential, power, and moving average. For all four data sets, a sixth order polynomial provided the best fit to the data (maximized R^2).

previously been argued to not be a construct of noise but rather, it has been suggested, comes from "intrinsic variations in emission phase that are correlated with variations in emission amplitude across frequency" (Shera and Guinan, 2003, p. 2764). The delay values obtained by Shera and Guinan (2003) were calculated from the first derivative of the phase with respect to frequency and not directly from the time domain. Here, the scatter is also evident in time domain measurements, illustrating that variations in cochlear reflectance directly affect SFOAE travel times.[4] The two definitions of signal onset provide for differing effective stimulus levels. For SFOAEs, stimulus level was approximately 70 dB pSPL. For the signal onset definition of 10% of the peak amplitude, this corresponds to an effective stimulus level of approximately 50 dB pSPL, a stimulus level that should produce an SFOAE dominated by a linear coherent reflection mechanism and so produce scatter in SFOAE delays. The signal onset definition of −3 dB, re the peak amplitude corresponds to an effective stimulus level of 67 dB pSPL, a stimulus level that produces an SFOAE that is a mixture of OAE arising from linear coherent reflection and nonlinear distortion (Talmadge et al., 2000; Goodman et al., 2003) – it is not clear whether the linear coherent reflection model of Zweig and Shera (1995) provides for place-fixed reflection at higher stimulus levels. Further discussion on the source of the scatter in the data for the SFOAE and the $2f_1 - f_2$ OAE generated by a nonlinear distortion mechanism is left for Section 4.

Comparison of the sixth order polynomial fitted to each of the data sets reveals that for the signal onset definition of 10% re the peak amplitude (panel (a)) there is a difference between the fitted curves that would be consistent with Eq. (2) and a bi-directional traveling wave mechanism. Panel (b) shows the $2f_1 - f_2$ OAE and SFOAE signal onset delay data based on −3 dB re the peak amplitude to overlap considerably with the fitted curves, more suggestive of Eq. (3) and an acoustic compression wave mechanism for energy to propagate back to the stapes. Neither figure provides curve fits that provide for relationships that are well described by Eq. (2) or (3), presumably due to the scatter or variability in the data. The source of the discrepancy in the relationship between $2f_1 - f_2$ OAE and SFOAE signal onset delay based on the two definitions of signal onset is considered in Section 4.

3.2. Rise-time, steady-state duration, and fall-time

Fig. 5 shows the relationship between stimulus and OAE rise time, steady-state duration, and fall-time for the $2f_1 - f_2$ OAE (panels (a)–(c)). Rise-time was calculated as

$$\text{Rise-time} = \text{Signal onset } (-3\text{dB point of the peak amplitude}) - \text{Signal onset } (10\% \text{ of peak amplitude})$$

It is evident in Fig. 5(a) that the rise time of the $2f_1 - f_2$ OAE is typically longer than the f_2 stimulus. The line at 45° with a slope of one represents equal rise times. Most of the OAE rise times fall above this line with rise times longer than the stimulus that evoked them. OAE rise time also shows a much greater variation than stimulus rise time, OAE rise time varying from 1.6 to 4.9 ms, while stimulus rise time varies from 2.2 to 2.6 ms. An obvious possible source for the extended OAE rise time is a secondary emission arising from the $2f_1 - f_2$ place.

Fig. 5(b) shows the stimulus steady-state duration, i.e., the time interval between −3 dB re peaks at onset and offset of the f_2 stimulus, plotted versus the steady-state duration of the $2f_1 - f_2$ OAE (the time interval between −3 dB re peaks at onset and offset of the $2f_1 - f_2$ OAE). The steady-state duration of the stimulus, as expected, is constant at 98.6 ± 0.1 ms (2 s.d.). The steady-state duration of the $2f_1 - f_2$ OAE, in contrast, is 99.0 ± 2.2 ms, i.e., the duration of the OAE in the steady-state region is not the same as the stimulus, although the histogram inset in panel (b) of OAE steady-state duration minus stimulus steady-state duration (Δ_{ss}) shows that 50% of all OAE steady-state durations fall within ±0.3 ms of the stimulus steady-state duration. The histogram suggests that the difference in steady-state duration (Δ_{ss}) is random.[5] Differences in steady-state duration of ±0.1 ms can be explained by measurement precision (while no error analysis was performed the steady-state duration of the stimulus has a measurement error of ±0.1 ms (2 s.d.), providing a value for measurement error based on the fact that this duration should be constant). Differences exceeding this value are presumably not due to measurement error. An OAE steady-state duration that deviates from the stimulus steady-state duration is not readily explainable in terms of an OAE that arises out of a cochlear mechanical amplification process that responds on a cycle-by-cycle basis. The difference, when present, suggests some hitherto unknown contribution to the generation of the nonlinear component of the $2f_1 - f_2$ OAE that appears to randomly slightly alter the steady-state duration of the OAE.

Fig. 5(c) shows the fall time of the stimulus plotted versus the fall time of the $2f_1 - f_2$ OAE. Significant variation in SFOAE fall time is observed relative to the stimulus fall time. Panel (c) shows more variation in f_2 stimulus fall

[4] We wish to thank Christopher Shera for pointing out that the scatter in physical SFOAE delays could be attributed to variation in cochlear reflectance.

[5] A normal probability plot shows Δ_{ss} to be normally distributed except for the two most negative and two most positive Δ_{ss}'s.

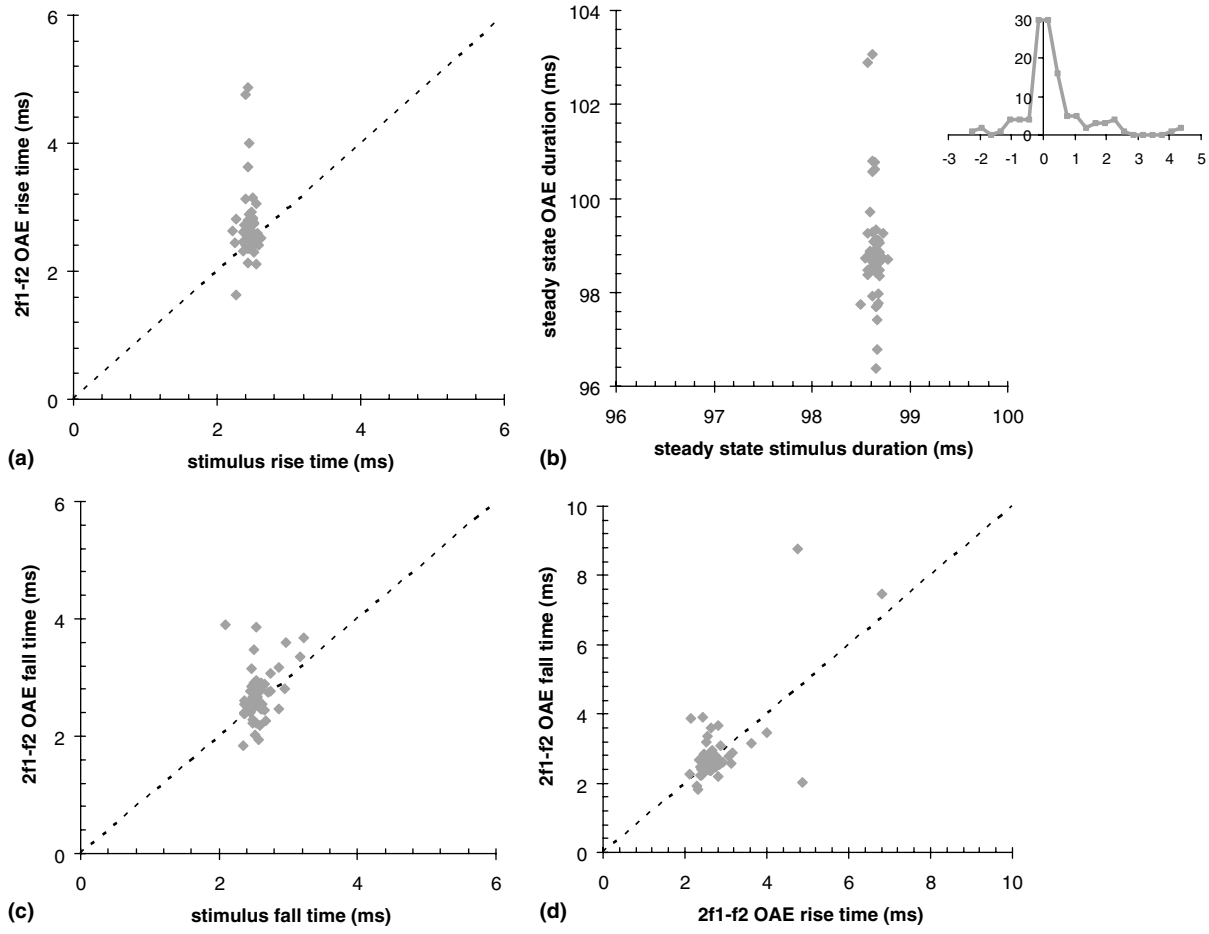

Fig. 5. Stimulus and OAE rise time, steady-state duration, and fall time for the $2f_1 - f_2$ OAE. Panel (a) shows stimulus rise time versus OAE rise time, panel (b) shows stimulus steady-state duration versus OAE steady-state duration, panel (c) shows stimulus fall time versus OAE fall time, and panel (d) shows OAE rise versus fall time. See text for details.

time than was found for the f_2 stimulus rise time in Fig. 5(a), longer stimulus offsets being due to stimulus ringing. $2f_1 - f_2$ OAE rise time in panel (a) had a mean of 2.4 ms versus a mean OAE fall time of 2.8 ms in panel (c), there being a reasonable correlation between the two (0.7). Fig. 5(d) illustrates this correlation, showing $2f_1 - f_2$ OAE rise versus fall time. A fall time similar in value to the rise time suggests that the majority of $2f_1 - f_2$ OAEs had little or no secondary emission arising from the $2f_1 - f_2$ place. As such, the extended OAE rise times in Fig. 5(a) is not explained by a secondary emission.

Fig. 6(a) shows the rise time of the stimulus plotted versus the rise time of the SFOAE. SFOAE rise time shows considerable variation either side of the mean of 3.0 ms, stimulus rise time varying between only 3.1 and 3.3 ms. This differs from Fig. 5(a) where DPOAE rise time was typically longer than the f_2 stimulus rise time but in both cases stimulus rise time varied by only a small amount (0.2–0.3 ms) while OAE rise time varied over a range of about 3 ms.

Fig. 6(b) shows the stimulus steady-state duration (the time interval between −3 dB re peaks at onset and offset). The steady-state duration of the stimulus was 27.3 ± 0.3 ms (2 s.d.). The steady-state duration of the SFOAE was 26.9 ± 4.8 ms, i.e., as was found for the $2f_1 - f_2$ OAE, the average duration of the stimulus and OAE are similar but the OAE steady-state duration varies considerably.

Fig. 6(c) shows the fall time of the stimulus plotted versus the fall time of the SFOAE. Significant variation in SFOAE fall time is observed relative to the stimulus fall time (the variation being greater than that observed for the $2f_1 - f_2$ OAE). As was found for the $2f_1 - f_2$ OAE, the fall time of the stimulus shows greater variation than the rise time. SFOAE fall time has a mean of 3.5 ms compared with a mean rise time (panel (a)) of 3 ms. Fig. 6(d) shows SFOAE rise time versus fall time – unlike for the $2f_1 - f_2$ OAE, there is no correlation between SFOAE rise and fall time ($r = 0.2$).

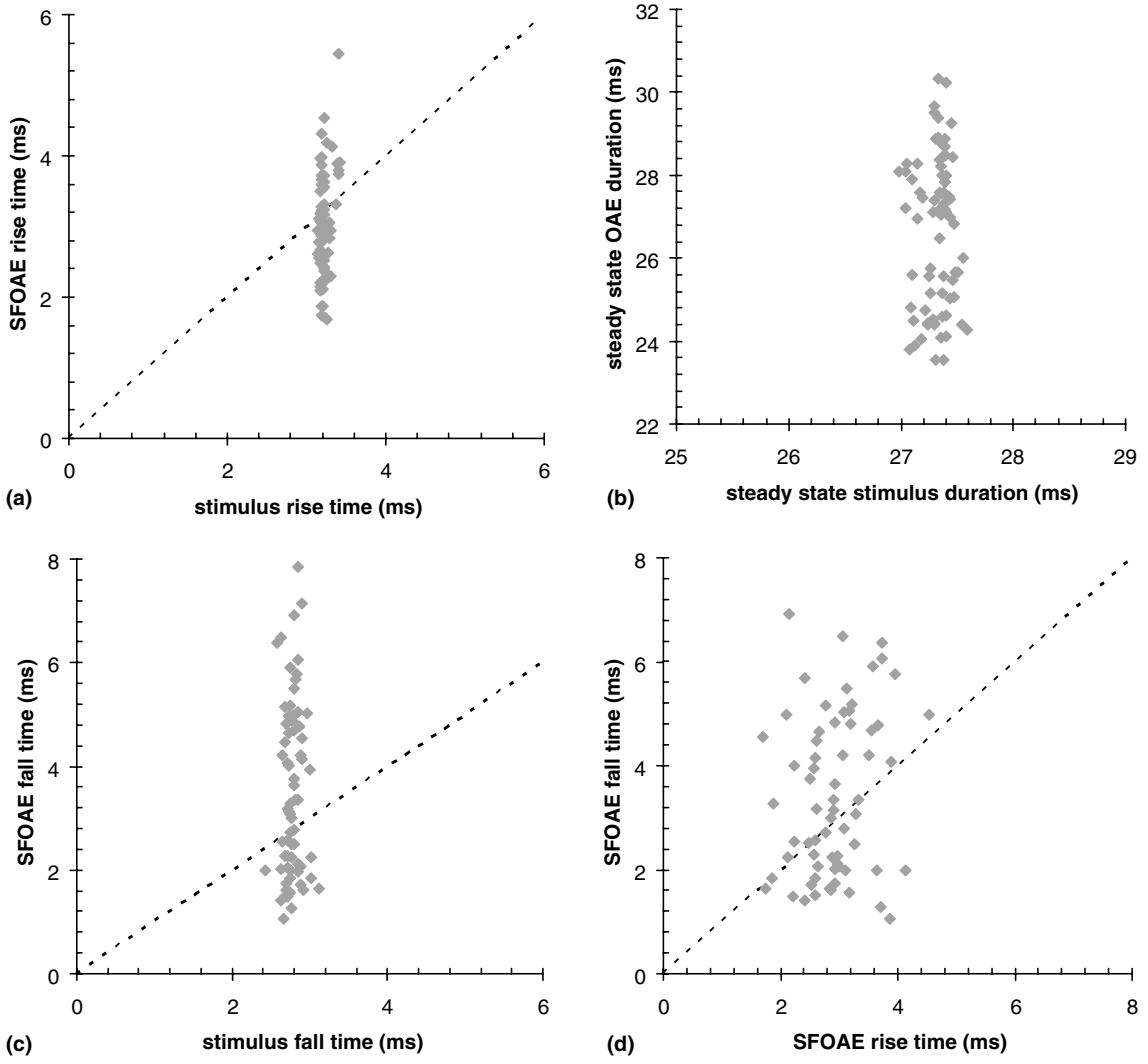

Fig. 6. Stimulus and OAE rise time, steady-state duration, and fall time for the SFOAE. Panel (a) shows stimulus rise time versus OAE rise time, (b) shows stimulus steady-state duration versus OAE steady-state duration, (c) shows stimulus fall time versus OAE fall time, and (d) shows OAE rise versus fall time. See text for details.

4. Discussion

Reports of signal onset delay as a measure of OAE delay include Whitehead et al. (1996) and Konrad-Martin and Keefe (2003). Both of these studies included an objective estimate of signal onset of the −3 dB point re the peak amplitude. Comparison of signal onset delay for 10% versus −3 dB of the peak amplitude for the SFOAE and the $2f_1 - f_2$ OAE reveals that

i. $2f_1 - f_2$ OAE signal onset delay for the −3 dB definition is greater than the 10% definition, but the variance is similar, and the two estimates are correlated ($r = 0.68$; $r = 0.83$ with outliers (5%) not included in analysis). Unlike for SFOAEs, the $2f_1 - f_2$ OAE rise time is consistently longer than the stimulus rise time (see Fig. 5).

ii. SFOAE signal onset delay for the −3 dB definition is, on average, slightly less that the 10% definition, the variance is greater, and the two estimates are not correlated ($r = 0.27$). This is not surprising given the variance in SFOAE rise time (see Fig. 6) that suggests a −3 dB definition is not representative of signal onset delay.

It is evident that signal onset based on a −3 dB definition will be a reasonable estimate of the 10% definition of signal onset for the $2f_1 - f_2$ OAE (subject to a correction factor) but that for SFOAEs it will not provide a good estimate of the 10% estimate of signal onset due to the fact that the OAE rise time is unrelated to stimulus rise time. For the −3 dB signal onset definition, secondary emission from the $2f_1 - f_2$ place may extend the estimate of the $2f_1 - f_2$ OAE signal onset, but the

correlation between stimulus rise time and OAE rise time and between OAE rise and fall time suggests such a secondary emission to not be significant in the majority of cases, as has previously been reported in the literature for a stimulus frequency ratio of 1.2 (Schneider et al., 2003; Withnell et al., 2003). Note that contamination by a secondary emission does not occur for the 10% definition of signal onset because of insufficient time for the $2f_1 - f_2$ from the DP place to contribute to the OAE. A -3 dB definition of signal onset assumes that stimulus rise time and OAE rise time are correlated. While this assumption appears to be valid for the $2f_1 - f_2$ OAE, it is not valid for the SFOAE and so a -3 dB definition is not representative of signal onset delay for SFOAEs and so suggests that Fig. 4(b) is not suggestive of an acoustic compression wave mechanism for energy to propagate back to the stapes.

Comparison of $2f_1 - f_2$ OAE signal onset delay for the OAE arising from the f_2 region with SFOAE signal onset delay (matched to the f_2 stimulus frequency) based on signal onset occurring at 10% of the peak signal amplitude was suggestive of a bi-directional traveling wave mechanism. Significant variability in signal onset delays obtained for both the $2f_1 - f_2$ OAE and SFOAE introduce a complexity that precludes establishing definitively that this is the case. However, Fig. 4(a) argues persuasively against an acoustic compression wave mechanism dominating energy transmission back to the stapes in guinea pig, such a mechanism requiring $\tau_{\text{sfoae}} \approx \tau_{\text{dpoae}}$. A bi-directional cochlear traveling wave mechanism posits that energy is predominantly propagated to and from the stapes through a continuous exchange of potential and kinetic energy along the BM and so is influenced in both directions by the graded mechanical properties of the BM. For stimulus frequency otoacoustic emissions arising from a place fixed mechanism, the delay is, on average, expected to be twice the forward propagation time (Shera and Guinan, 2003). For $2f_1 - f_2$ OAEs, the OAE arising from the f_2 region is arising basal to the place on the BM tuned to that frequency and so OAE delay is expected to be substantially less than twice the forward propagation time.

4.1. Variability in signal onset delay and rise/steady-state duration/fall time

Considerable variability was found for $2f_1 - f_2$ OAE and SFOAE signal onset delays obtained using both a signal onset of 10% of the peak amplitude and -3 dB of the peak amplitude. Variable times was also a feature of OAE rise time, steady-state duration, and fall time, the variability being present for OAEs purportedly generated by different mechanisms. Variability in signal onset delay for the nonlinear distortion $2f_1 - f_2$ OAE (see Fig. 4) i.e., the delay was not a smoothly varying function of stimulus frequency, was a somewhat surprising finding given the purported mechanism for the generation of this OAE being one that is directly related to traveling wave amplitude, i.e., a wave-fixed mechanism (Kemp, 1986; Shera and Guinan, 1999; Talmadge et al., 1999). Such variability might be explained by a localized breaking of scaling symmetry whereby place-fixed irregularities distort an otherwise regular spatial tuning and so produce small variations in the traveling wave envelope. Scaling symmetry means that

$$\tau * f_{\text{CF}} = \text{a constant}.$$

This can be violated in a smooth cochlea by changes in cochlear tuning from base to apex that vary in a smooth manner (Shera, 2002, personal communication) but such a smooth transition in tuning should not produce the irregular variability in delay seen in Fig. 4. Nonlinear distortion $2f_1 - f_2$ OAE signal onset delay could also vary in a quasi-irregular fashion if the amplitude of the emission were not solely determined by traveling wave amplitude. However this would be analogous to a mixing of wave and place fixed mechanisms and such a role for a place-fixed mechanism for the nonlinear distortion $2f_1 - f_2$ OAE is inconsistent with previous findings for an OAE arising from a place-fixed mechanism (Zweig and Shera, 1995; Shera and Guinan, 2003), i.e., the emission arising basal to the peak of the traveling wave ($2f_1 - f_2$ arising from f_2 region) appears to undergo little place-fixed reflection.

The $2f_1 - f_2$ OAE was obtained with the stimulus level of f_1 fixed (f_1 on continuously) while pulsing the f_2 stimulus (2.5 ms rise/fall time, 98.5 ms steady state duration), with $f_2/f_1 = 1.2$. The rise time of the f_2 stimulus should, as a result, produce a $2f_1 - f_2$ OAE that grows as per the basilar membrane displacement amplitude produced by f_2, subject to two-tone suppression effects (Withnell and Yates, 1998). The $2f_1 - f_2$ OAE may not grow linearly over the entire rise time of the stimulus and so this could be a source of variability in the signal onset delay and rise and fall times.[6] Nonlinear growth of the OAE means that referencing to the peak amplitude does not provide an equivalent onset in time for the OAE versus the stimulus, i.e., saturation of the OAE results in the 10% definition of signal onset not corresponding to the 10% signal onset definition for the stimulus. Nonlinear growth of the $2f_1 - f_2$ OAE could also produce jitter in the $2f_1 - f_2$ OAE steady-state duration, this being dependent on if, and at what stimulus level, the $2f_1 - f_2$ OAE growth saturates.

The $2f_1 - f_2$ OAE does not appear to be contaminated by a mixing of components arising from different mechanisms, the data being consistent with only one

[6] Withnell and Yates (1998) found $2f_1 - f_2$ OAE growth to be linear up to $L_2 = 42$ dB pSPL with $L_1 = 58$ dB pSPL in one guinea pig; an L_1 of approximately 77 dB pSPL as used in this study presumably extends the linear growth to higher values of L_2 due to two-tone suppression.

mechanism contributing to this OAE. This is not, however, the case for the SFOAE. A stimulus level of approximately 70 dB pSPL may be a potential source of jitter in time domain measurements of SFOAE steady-state duration due to the mixing of two components arising from different mechanisms (Goodman et al., 2003). The potential complication of a mixing of mechanisms should not be significant for the signal onset definition of 10% of the peak amplitude, this definition corresponding to an effective stimulus level of approximately 50 dB pSPL, a stimulus level that should produce an SFOAE dominated by a linear coherent reflection mechanism (Goodman et al., 2003). However, the SFOAE in guinea pig has previously been found to grow nonlinearly (Souter, 1995), such nonlinear growth[7] perhaps producing variability in signal onset delay measures and SFOAE rise and fall times.

Another possible source of variability in delay would be variation in stimulus levels. The stimulus levels of the f_2 tone varied between 70 and 78 dB pSPL, while the stimulus level of the f_1 tone varied between 74 and 82 dB pSPL with 0 dB $\leqslant L_2/L_1 \leqslant$ 10 dB. However, no significant correlation was found between L_1 and signal onset delay, L_2 and signal onset delay, or L_2/L_1 and signal onset delay.

The common feature of variability of OAE rise, steady-state duration, and fall times for OAEs purportedly generated by different mechanisms and signal onset delays may have a common "effective" origin. Such variability may arise from place-fixed irregularities that, acting through the cochlear mechanical amplifier feedback loop (e.g., shift in operating point), produce a localized breaking of scaling symmetry and perturb an otherwise regular spatial tuning, providing for the origin of the variability in OAEs arising from a nonlinear distortion mechanism, with the irregularities themselves providing the origin of OAEs arising from a linear coherent reflection mechanism. However, there is no evidence for amplitude microstructure in the frequency domain for the wave-fixed DPOAE that would corroborate such an origin.

4.2. Equating stimulus level

In this paper, we have equated SFOAE stimulus level to the f_2 tone stimulus level, a comparison that assumes that the spatial extent of the region of the cochlea that predominantly contributes to the generation of the DPOAE is determined solely by the f_2 stimulus and that both emissions arise from the same part of the traveling wave. For a two-tone complex, determining the effective level of stimulation relative to a single tone stimulus is complicated by both stimulus tones contributing to the generation of DPOAEs and two-tone suppression. Additionally, the SFOAE, unlike the DPOAE, cannot be extracted using Fourier analysis, the method of extraction (i.e., suppression, nonlinear derived technique) not necessarily isolating the whole emission which may have implications for the delay. Given these limitations, and the variability in signal onset delays, definitive proof of the existence of a reverse traveling wave mechanism for energy to propagate back to the stapes awaits further studies.

[7] A qualification to this nonlinear growth of the SFOAE is that the SFOAE in this study was extracted using the nonlinear derived extraction method with a stimulus ratio of 6 dB, this extraction method linearizing the otherwise nonlinear growth observed using suppression methods and vector subtraction.

Acknowledgments

We are indebted to Drs. Christopher Shera and Mark Chertoff and one anonymous reviewer for providing valuable feedback on earlier versions of this manuscript.

References

Bell, A., Fletcher, N.H., 2004. The cochlear amplifier as a standing wave: squirting waves between rows of outer hair cells? J. Acoust. Soc. Am. 116, 1016–1024.

Bekesy, G., 1960. Experiments in Hearing. McGraw-Hill Book Company, USA.

Goodman, S.S., Withnell, R.H., Shera, C.A., 2003. The origin of SFOAE microstructure in guinea pig. Hear. Res. 183, 7–17.

Heitmann, J., Waldmann, B., Schnitzler, H., Plinkert, P.K., Zenner, H., 1998. Suppression of distortion product otoacoustic emissions (DPOAE) near $2f_1 - f_2$ removes DP-gram fine structure – evidence for a secondary generator. J. Acoust. Soc. Am. 103, 1527–1531.

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

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

Kemp, D.T., 1986. Otoacoustic emissions, travelling waves and cochlear mechanisms. Hear. Res. 22, 95–104.

Kemp, D.T., Ryan, S., Bray, P., 1990. A guide to the effective use of otoacoustic emissions. Ear Hear. 11, 93–105.

Kirk, D.L., Yates, G.K., 1994. Evidence for electrically evoked travelling waves in the guinea pig cochlea. Hear. Res. 74, 38–50.

Konrad-Martin, D., Keefe, D.H., 2003. Time-frequency analyses of transient-evoked stimulus-frequency and distortion-product otoacoustic emissions: testing cochlear model predictions. J. Acoust. Soc. Am. 114, 2021–2043.

Lighthill, J., 1991. Biomechanics of hearing sensitivity. J. Vibr. Acoust. 113, 1–13.

Ren, T., 2004. Reverse propagation of sound in the gerbil cochlea. Nature Neurosci. 7, 333–334.

Schneider, S., Prijs, V.F., Schoonhoven, R., 2003. Amplitude and phase of distortion product otoacoustic emissions in the guinea pig in an (f1,f2) area study. J. Acoust. Soc. Am. 113, 3285–3296.

Shera, C.A., 2003. Wave interference in the generation of reflection- and distortion-source OAEs. Biophysics of the Cochlea: from Molecule to Model. World Scientific, Singapore, pp. 439–449.

Shera, C.A., Talmadge, C.L., Tubis, A., 2000. Interrelations among distortion-product phase-gradient delays: their connection to

scaling symmetry and its breaking. J. Acoust. Soc. Am. 108, 2933–2948.

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

Souter, M., 1995. Stimulus frequency otoacoustic emissions from guinea pig and human subjects. Hear. Res. 90, 1–11.

Talmadge, C.L., Long, G.R., Tubis, A., Dhar, S., 1999. Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. J. Acoust. Soc. Am. 105, 275–292.

Talmadge, C.L., Tubis, A., Long, G.R., 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.

Tubis, A., Talmadge, C.L., Tong, C., Dhar, S., 2000. On the relationship between the fixed-f1, fixed-f2, and fixed ratio phase derivatives of the $2f_1 - f_2$ distortion product otoacoustic emission. J. Acoust. Soc. Am. 108, 1772–1785.

Whitehead, M.L., Stagner, B.B., Martin, G.K., Lonsbury-Martin, B.L., 1996. Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency. J. Acoust. Soc. Am. 100, 1663–1679.

Wilson, J.P., 1980. Model for cochlear echoes and tinnitus based on an observed electrical correlate. Hear. Res. 2, 527–532.

Withnell, R.H., Shaffer, L.A., Talmadge, C.L., 2003. Generation of DPOAEs in the guinea pig. Hear. Res. 178, 106–117.

Withnell, R.H., Yates, G.K., 1998. Onset of basilar membrane nonlinearity reflected in cubic distortion tone input–output functions. Hear. Res. 123, 87–96.

Yates, G.K., 1995. Cochlear structure and function. In: Moore, B.C.J. (Ed.), Hearing. Academic Press, New York.

Yates, G.K., 1998. Mechanical signal conditioning in the cochlea: nonlinear processes. In: From Structure to Information in Sensory SystemsBiophysics, vol. 5. World Scientific Publishing, Singapore.

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

Zwislocki, J., 1953. Wave motion in the cochlea caused by bone conduction. J. Acoust. Soc. Am. 25, 986–989.