Reconciling the origin of the transient evoked ototacoustic emission in humans

Robert H. Withnell,[a] Chantel Hazlewood, and Amber Knowlton
Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405, USA

(Received 29 June 2007; revised 4 October 2007; accepted 4 October 2007)

A pervasive theme in the literature for the transient evoked otoacoustic emission (TEOAE) measured from the human ear canal has been one of the emission arising solely (or largely) from a single, place-fixed mechanism. Here TEOAEs are reported measured in the absence of significant stimulus contamination at stimulus onset, providing for the identification of a TEOAE response beginning within the time window that is typically removed by windowing. Contrary to previous studies, it was found that in humans, as has previously been found in guinea pig, the TEOAE appears to arise from two generation mechanisms, the relative contributions of these two mechanisms being time and stimulus-level dependent. The method of windowing the earliest part of the ear canal measurement to remove stimulus artifact removes part of the TEOAE i.e., much of the component arising from a nonlinear generation mechanism. This reconciliation of TEOAE origin is consistent with all OAEs in mammals arising in a stimulus-level dependent manner from two mechanisms of generation, one linear, one nonlinear, as suggested by Shera and Guinan [J. Acoust. Soc. Am. **105**, 782–798 (1999)]. © *2008 Acoustical Society of America.* [DOI: 10.1121/1.2804635]

PACS number(s): 43.64.Jb, 43.64.Kc [BLM] Pages: 212–221

I. INTRODUCTION

The relationship between the stimulus spectrum and the otoacoustic emission response spectrum for transient-evoked otoacoustic emissions (TEOAEs) has repeatedly been found to have a one-to-one correspondence, consistent with a linear generation mechanism (e.g., Kemp and Chum, 1980; Kemp, 1986; Probst *et al.*, 1986; Hauser *et al.*, 1991, Xu *et al.*, 1994; Prieve *et al.*, 1996; Killan and Kapadia, 2006; Kalluri and Shera, 2007). Discordant with these studies using exclusively human subjects have been studies using guinea pigs, which have suggested a contribution from a nonlinear mechanism of generation (Avan *et al.*, 1995; Withnell and Yates, 1998; Yates and Withnell, 1999). And there have been a few studies with human subjects that suggest other than a solely linear mechanism of generation for the TEOAE (Avan *et al.*, 1993, 1997; Carvalho *et al.*, 2003). Insight into this disparity was provided by Withnell and McKinley (2005), who reported that, in guinea pigs, the dominant mechanism of generation of the TEOAE appeared to be time dependent, shifting from a wave-fixed generation mechanism to a place-fixed generation mechanism over the time course of the response.

Wave and place fixed, first described by Kemp (e.g., see Kemp, 1986), argues for two mechanisms of OAE generation, distinguishable by their phase characteristic, with the region of generation shifting with the traveling wave for the wave-fixed and being spatially fixed for the place-fixed mechanism. For an exactly scaling-symmetric cochlea, OAEs would be produced solely by a wave-fixed mechanism. A place-fixed mechanism involves random variations in the impedance of the basilar membrane versus length (Zweig and Shera, 1995), producing, it seems, a breaking of scaling symmetry near threshold. Acting through the cochlear amplifier feedback loop, these small variations in basilar membrane impedance act as a reflection source from which intracochlear standing waves can develop between the reflected wave and the incident wave. At or near threshold, a variation in hearing threshold versus stimulus frequency on the order of 20 dB can be observed (e.g., Long, 1984), presumably associated with these intracochlear standing waves (Talmadge *et al.*, 1998).

Studies of the origin of the TEOAE in humans have windowed or filtered the first 2.5–5 ms of the ear canal recording or nonlinear derived response (e.g., Kemp *et al.*, 1990) to remove the ringing of the speaker associated with delivering a click stimulus subsequent to loading it with the ear canal. This windowing removes any part of the TEOAE that is present in the early part of the response. As a result, examinations of TEOAE origin (e.g., Probst *et al.*, 1986; Xu *et al.*, 1994; Killan and Kapadia, 2006) that have found the TEOAE to arise from a linear generation mechanism require that the windowed early part of the response contain no TEOAE or that this early part of the TEOAE have the same generation mechanism. A time dependence to the origin of the TEOAE in guinea pig raises the possibility that a similar relationship may exist in humans and that if the early part of the response can be preserved and not removed with windowing, then this early part of the TEOAE may have a nonlinear generation mechanism. To explore this possibility, a similar approach to measuring the TEOAE was completed to that performed previously in guinea pig (see Withnell *et al.*, 1998), with the earphone not physically coupled to the ear canal and sound delivered to the ear with the recording microphone positioned in the ear canal.

[a]Electronic mail: rwithnel@indiana.edu

 0001-4966/2008/123(1)/212/10/$23.00

II. METHOD

A. Subject

Seven adult females, less than 30 years of age, served as subjects for this study. Data for this study were collected from one ear of each subject, hearing being within clinically normal limits for the ear tested.

B. Signal generation and data acquisition

Signal generation and data acquisition were computer-controlled using custom software and a Card Deluxe soundcard with 96 kHz sampling rate. Sound stimuli were delivered by a Beyer DT48 loudspeaker positioned approximately 4 cm from the entrance to the ear canal. A Sennheiser M series electrostatic microphone (6 mm diameter) was placed in the ear canal to measure ear canal sound pressures, the depth of insertion defined by the length of the microphone capsule (<1 cm). The frequency response of the loudspeaker at the position of the microphone in the ear canal was determined by delivering pseudorandom electrical noise (a sum of sine waves, spectrally flat with random phase) to the loudspeaker; the impulse response of the loudspeaker was then calculated from the frequency response measured at the microphone. The stimulus wave form for evoking a TEOAE was generated using a sinc function $[\sin(\omega t)/(\omega t)]$ deconvolved with the impulse response of the loudspeaker (see Yates and Withnell, 1999, for further details). This provided for a stimulus with a flat amplitude spectrum and linear phase delay at the measurement microphone in the ear canal. TEOAEs were obtained using the nonlinear derived extraction technique (Kemp *et al.*, 1990) with a 21.3 ms time base, a 9 dB stimulus level ratio, and 1000–4000 averages [see Withnell and McKinley (2005) for more details regarding extraction of the TEOAE from ear canal sound pressure recordings]. For the data reported here, TEOAEs were obtained over a range of stimulus levels and bandwidths.

C. Data analysis

Time domain windowing to separate TEOAE components with different phase characteristics was achieved using a recursive exponential filter (see Shera and Zweig, 1993; Kalluri and Shera, 2001)

TEOAE early latency component=TEOAE$\cdot F(t)$,

TEOAE late component=TEOAE$\cdot F(t)$,

where $F(t)=1/\Gamma_n(\tau)$, $\Gamma_n(\tau)$ is defined recursively as $\Gamma_{n+1}(\tau)=e^{\Gamma_n(\tau)-1}$, with $\Gamma_1(\tau)=e^{\tau 2}$,

$\tau=t/\tau_{\text{cut}}$, where t is time and τ_{cut} is the length of the window, filter order $(n)=16$.

The value of τ_{cut} was chosen such that it:

i. Separates the TEOAE into two components with different phase characteristics, one with a shallow slope, the other a steep slope.

ii Matches the amplitudes of the two components with the TEOAE phase. The slope of the TEOAE phase identifies which of the early and late components dominates the TEOAE and the amplitudes of each of the components should be in agreement with this, e.g., when the slope of the TEOAE phase is steep, the late component should be larger in amplitude than the early component.

This study was completed with the approval of the Human Ethics Committee, Indiana University.

III. RESULTS

Figure 1 shows examples of time-averaged, nonlinear derived, ear canal sound pressure recordings from six of the subjects, the acoustic stimulus varying in bandwidth but having a similar peak sound pressure level (~71 dB pSPL) at the measurement microphone. Each of the panels includes the stimulus wave form, scaled to the amplitude of the time-averaged, nonlinear derived, ear canal sound pressure recordings. Inset in each panel is the earliest part of the recording. Each of the first four panels show the earliest part of the response in time to be higher in frequency than the later part of the response. Figure 1(e) shows the earliest part of the response to have a higher frequency than the later part but the response as a whole is delayed relative to the responses in Figs. 1(a)–1(d). Stimulus contamination of the response coincident in time with the stimulus is notable in Fig. 1(a) and may be present in other panels.

A. Quantifying stimulus contamination of the nonlinear derived response

When a short duration electrical pulse is delivered to an earphone/loudspeaker, it rings at its resonant frequency (where the electrical pulse has a frequency spectrum that includes one or more of the resonant frequencies of the loudspeaker). If the loudspeaker is driven with a sufficiently large current that the diaphragm response is no longer linear, then the speaker ringing at the resonant frequency will no longer be linear. This speaker-generated nonlinearity will contaminate the nonlinear derived response recorded from a human ear with normal hearing. Figure 2 shows the nonlinear derived sound pressure level recorded in response to a short duration stimulus[1] over a range of stimulus levels with the Senheisser microphone placed in the ear canal of KEMAR, with the Beyer DT48 loudspeaker used to generate the acoustic stimulus positioned approximately 4 cm from the ear canal entrance. KEMAR provides a passive cavity with an input impedance that is intended to match the input impedance of the average human ear (at higher stimulus levels where the cochlear input impedance is resistive). Each of the panels includes the stimulus wave form, scaled to the amplitude of the time-averaged, nonlinear derived, ear canal sound pressure recordings. Each panel shows the stimulus amplitude at the microphone of the stimulus i.e., 57–77 dB pSPL, but this is not the determinant of stimulus contamination; rather, it is the magnitude of the input voltage to the speaker coupled with the acoustic load. The uppermost panel shows little or no stimulus contamination of the nonlinear derived response. The bottom panel shows stimulus contamination with the earphone ringing. The amplitude spectrum of this ringing shows two resonant modes, one centered on ~3.2 kHz with another mode centered on ~4.5 kHz. Note that the stimulus contamination is largest coincident in time

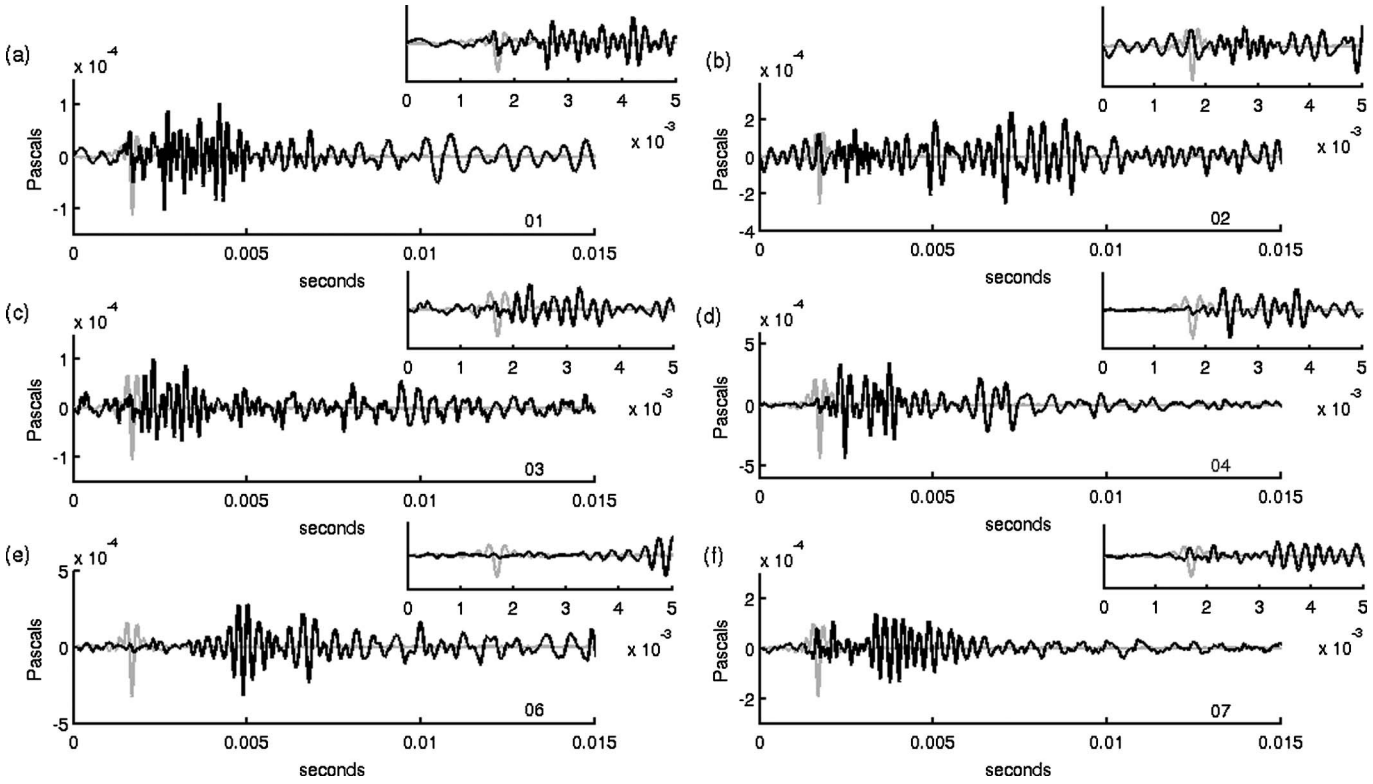

FIG. 1. TEOAE responses from six subjects, the acoustic stimulus varying in bandwidth but having a similar peak sound pressure level (~71 dB pSPL) at the measurement microphone. Each of the panels includes the stimulus wave form, scaled to the amplitude of the time-averaged, nonlinear derived, ear canal sound pressure recordings. Inset in each panel is the earliest part of the recording. Stimulus level ratio in each case was 9 dB. Stimulus bandwidths were (a) 0.5–5.5 kHz, (b) 1.5–7 kHz, (c) 1.5–5 kHz, (d) 1–5 kHz, (e) 1–5 kHz, and (f) 1–5 kHz.

with the stimulus, decaying to the noise floor within ~1.5 ms of the stimulus peak amplitude. The stimulus generated in the bottom panel was for the maximum input voltage provided by the software. At five stimulus levels (three of which are shown), stimulus contamination, if present, was largest coincident in time with the stimulus. It would appear that the upper limit to the magnitude of the stimulus contamination of the nonlinear derived response is defined at stimulus onset, prior to the onset of a physiological response, and that stimulus contamination has a very short duration.

It is apparent in Fig. 1 that the magnitude of the speaker-generated nonlinearity at stimulus onset is insufficient in amplitude for speaker-generated nonlinearity to contaminate significantly these TEOAE responses recorded from six human ears.

B. TEOAE versus stimulus level

Figure 3 shows the nonlinear derived response recorded over a range of stimulus levels from one subject. Figure panels are arranged in descending stimulus level (top to bottom) from 78 to 63 dB pSPL in 5 dB steps. Each panel shows the nonlinear residual and the click stimulus wave form (lighter lines), the click stimulus peak occurring at 1.7 ms. In all four recordings, a robust TEOAE is observed; stimulus contamination coincident in time with the stimulus is small (or not present) relative to the response that follows. As in Fig. 1, the earliest part of the response is higher in frequency than the response later in time. Inspection of the four panels reveals that the early part of the TEOAE (<3.7 ms or <2 ms

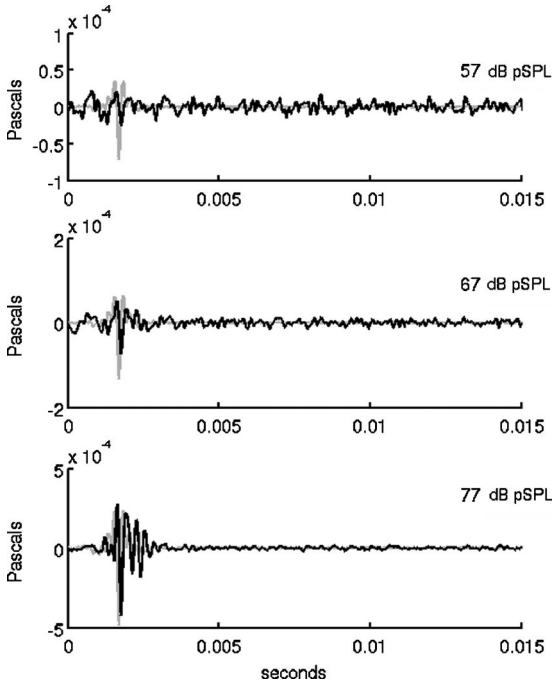

FIG. 2. The nonlinear derived sound pressure level recorded in response to a short duration stimulus over a range of stimulus levels in KEMAR. Each of the panels includes the stimulus wave form, scaled to the amplitude of the time-averaged, nonlinear derived, ear canal sound pressure recordings. Stimulus bandwidth was 1–5 kHz. The uppermost panel shows little or no stimulus contamination of the nonlinear derived response. The bottom panel shows stimulus contamination with the earphone ringing. Stimulus contamination is largest coincident in time with the stimulus, decaying to the noise floor within ~1.5 ms of the stimulus peak amplitude.

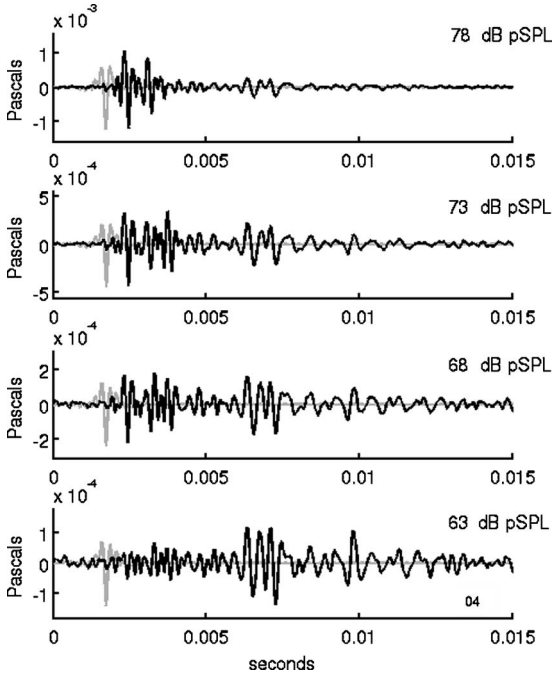

FIG. 3. The TEOAE recorded over a range of stimulus levels from one subject. Panels are arranged in descending stimulus level (top to bottom) from 78 to 63 dB pSPL in 5 dB steps. Each panel shows the nonlinear residual and the click stimulus wave form (lighter lines), the click stimulus peak occurring at 1.7 ms. In all four recordings, a robust TEOAE is observed; stimulus contamination coincident in time with the stimulus is small (or not present) relative to the response that follows. Stimulus bandwidth was 1–5 kHz.

poststimulus peak) decreases in amplitude relative to the later part (>4.7 ms) with decreasing stimulus level. The method of TEOAE extraction used in this study (nonlinear derived extraction technique) however, increasingly underestimates the TEOAE as stimulus level is reduced and the stimulus level ratio is maintained constant due to OAE growth not being saturated (Lonsbury-Martin *et al.*, 1988) and so the relative amplitude relationship of the early and late parts of the TEOAE may be affected.

Examples of TEOAEs obtained across a range of stimulus levels from two other subjects are shown in Fig. 4. As observed in Fig. 3, the early part of the TEOAE decreases in amplitude relative to the later part as stimulus level decreases. In contrast to Fig. 3, stimulus contamination of the response is more evident, particularly for the responses to the highest stimulus levels.

C. Time-domain windowing of the TEOAE

Here we examine the proposal that the TEOAE in humans can be separated into two components, the early part of the TEOAE having a shallow phase slope consistent with a wave-fixed mechanism and the late part having a steep phase slope consistent with a place-fixed mechanism. Equipment used to record TEOAEs in humans typically window some part of the averaged ear canal sound pressure recording poststimulus onset (e.g., 2.5 ms, Kemp *et al.*, 1990) to reduce stimulus contamination of the recording. This windowing would remove some or all of the early part of the TEOAE response in Figs. 3 and 4. Figure 5 shows amplitude and phase spectra corresponding to the TEOAE in the top panel of Fig. 3 and the spectra resulting from time-domain windowing this TEOAE. Also shown are the original TEOAE time wave form, and the early and late TEOAE wave forms extracted by time domain windowing. The black line is the original TEOAE. The phase spectrum of the TEOAE shows the slope of the phase to be rotating rapidly up to about 2.2 kHz with a shallow slope from 2.2 to 5.5 kHz. This indicates that the early component should be larger in amplitude than the late component over the 2.2–5.5 kHz frequency range and vice-versa below 2.2 kHz. For a cut-off value of 3.7 ms (see Sec. III D for how this value was ar-

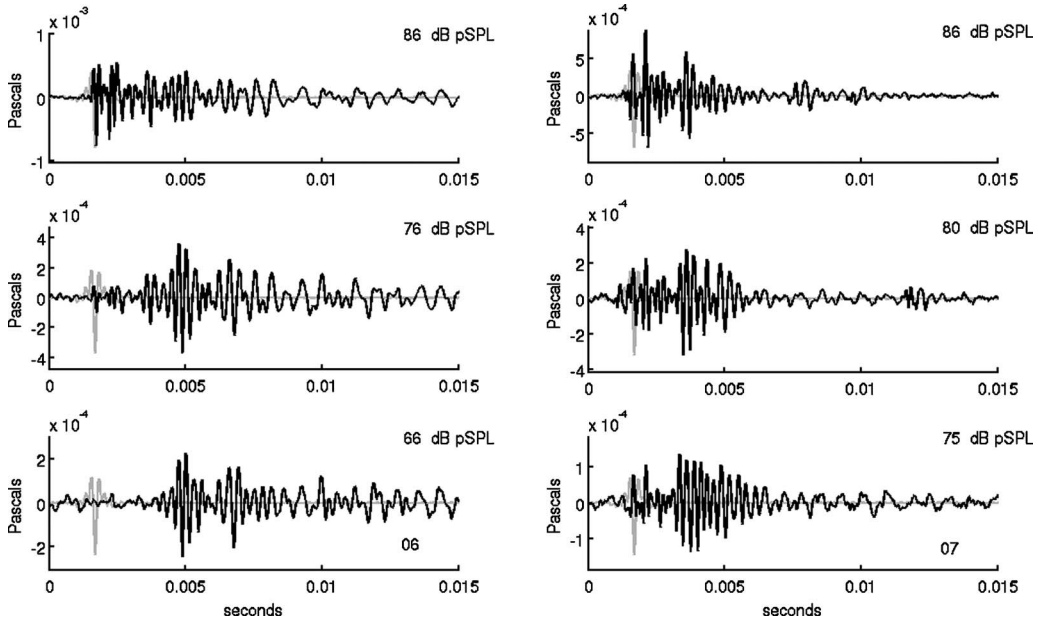

FIG. 4. Examples of TEOAEs obtained across a range of stimulus levels from two other subjects. As observed in Fig. 3, the early part of the TEOAE decreases in amplitude relative to the later part as stimulus level decreases. In contrast to Fig. 3, stimulus contamination of the response is more evident, particularly for the responses to the highest stimulus levels. Stimulus bandwidth was 1–5 kHz.

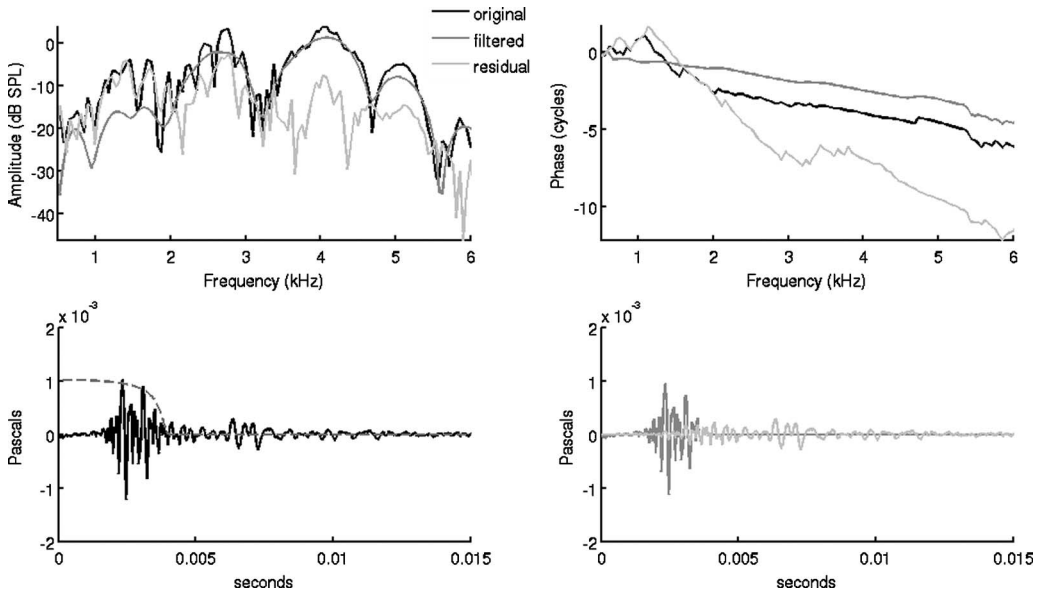

FIG. 5. (a) The amplitude spectra corresponding to the TEOAE in (c) and the spectra resulting from time-domain windowing this TEOAE. (b) The phase spectra corresponding to the TEOAE in (c) and the spectra resulting from windowing this TEOAE. (c) TheTEOAE time wave form (the top panel of Fig. 3) and the windowing function, $F(t)$. (d) TEOAE component wave forms extracted by time domain windowing.

rived at), the early part of the TEOAE rotates ~1.5 cycles over a 2 kHz range, corresponding to a group delay of ~0.8 ms, a value that suggests either a wave-fixed mechanism of generation or the emission arising from the basal region of the cochlea. In contrast, the late component has a group delay of ~6.9 ms centered on 2 kHz and ~3.8 ms centered on 4 kHz, values that suggest forward travel times of >3.5 and 1.9 ms [see Shera *et al.* (2005), Eq. (59)].

Figure 6 is as for Fig. 5, but from another subject corresponding to the TEOAE obtained to a stimulus level of 86 dB pSPL from Fig. 4 (left top panel). Unlike the TEOAE analyzed in Fig. 5 (top panel of Fig. 3), there is evidence of stimulus contamination of the TEOAE. To reduce this stimulus contamination, the response was time-domain windowed with a time cut of 2.38 ms (0.68 ms poststimulus peak) before analysis, the wave form obtained after this windowing then being time-domain windowed to separate the early and late components of the TEOAE response. In contrast to the TEOAE in Fig. 5, this TEOAE has a phase response that rotates rapidly with frequency over most of the frequency range, indicating that the late part of the TEOAE should be larger than the early component at most frequencies. For a cut-off value of 3.6 ms, the early part of the TEOAE rotates ~1.8 cycles over a 2 kHz range, corresponding to a nonphysical group delay for the 2.5–4.5 kHz region of the cochlea of ~0.9 ms. In contrast, the late component has a

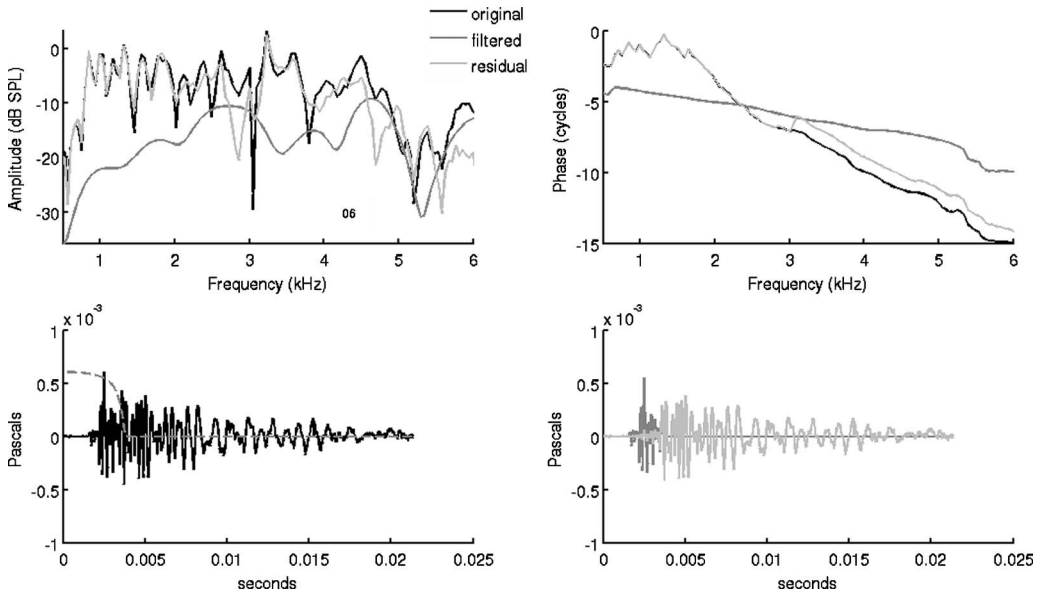

FIG. 6. As for Fig. 5, but from another subject corresponding to the TEOAE obtained to a stimulus level of 86 dB pSPL from Fig. 4 (left top panel). To reduce stimulus contamination, the response was time-domain windowed with a time cut of 2.38 ms (0.68 ms poststimulus peak) before analysis, the wave form obtained after this windowing then being time-domain windowed to separate the early and late components of the TEOAE response.

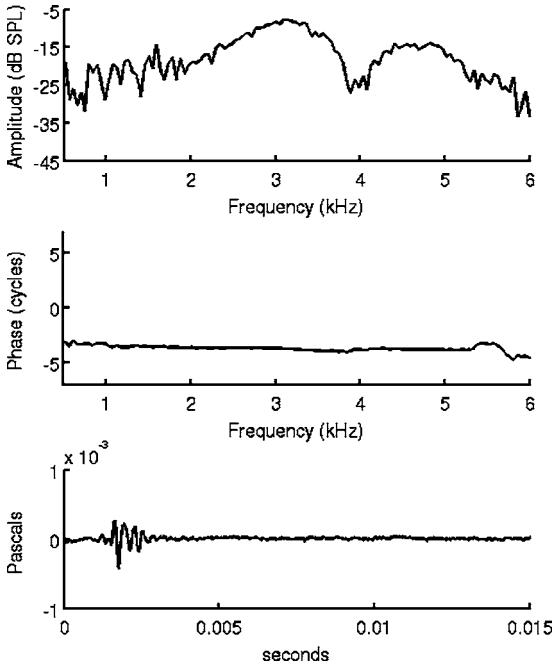

FIG. 7. The stimulus artifact recorded with KEMAR. It is notable that the amplitude and phase spectra for the artifact differs significantly from the TEOAE recorded from a human subject (see Figs. 5 and 6). The amplitude spectrum shows two resonant modes, one centered on ~3.2 kHz, the other centered on ~4.5 kHz. The phase of the artifact has a zero slope, consistent with no delay, the onset of the artifact coincident in time with the stimulus.

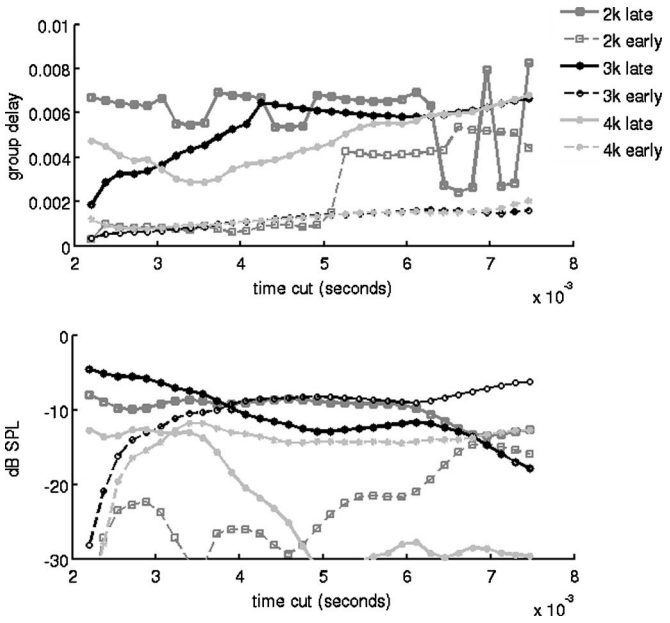

FIG. 8. The slope of the phase and the amplitude for the early and late components of the TEOAE in Fig. 5 at 2, 3, and 4 kHz over a range of values of τ_{cut}.

group delay of ~6.4 ms centered on 2 kHz and ~4.2 ms centered on 4 kHz, values that suggest forward travel times of >3.2 and 2.1 ms. These values are similar to that found for the TEOAE from the subject in Fig. 5.

Figure 7 shows the equivalent representation of Figs. 5 and 6 for the stimulus artifact recorded with KEMAR. It is notable that the amplitude spectra for the artifact differs significantly from the TEOAE recorded from a human subject (see Figs. 5 and 6), and the phase of the artifact has a zero slope consistent with no delay while the early component of the TEOAE has a negative phase slope. The negative slope for the early component of the TEOAE in Figs. 5 and 6 is thought to be due to a breaking of scaling and the slope of the phase of the early component of the TEOAEs clearly distinguish them from stimulus artifact.

D. Determination of τ_{cut}

Determination of the appropriate time cut, τ_{cut}, for the analysis presented in Figs. 5 and 6 was made by evaluating the rate of change of phase versus frequency ($d\Phi/d\omega$) at 2, 2.5, 3, 3.5, and 4 kHz for the early and late components for a range of τ_{cut}. For τ_{cut}'s shorter than the optimum value, some of the component arising from a nonlinear mechanism remains in the late part of the response, reducing the rate of change of phase versus frequency for this component, or the "group delay." Figure 8 shows the slope of the phase and the amplitude for the early and late components of the TEOAE in Fig. 5 at 2, 3, and 4 kHz over a range of values of τ_{cut} The optimal τ_{cut} is between 3.4 and 3.75 ms; for cut-off values beyond 3.75 ms the value of $d\Phi/d\omega$ [Fig. 8(a)] for the 4 kHz late component of the TEOAE starts to increase by virtue of the main energy in this part of the response being removed by windowing. Figure 8(b) shows the amplitudes of the 3 and 4 kHz late components to, in general, be a decreasing function of τ_{cut} these component include a significant early component at low τ_{cut}'s. As τ_{cut} increases, the late component should plateau in amplitude until a cut-off value is reached where some of the late component is removed with time domain windowing; for TEOAE arising from a linear coherent reflection mechanism this cut-off value would be a function of frequency, higher frequency components being removed first, i.e., with the lowest cut-off value. Concomitantly, the early component of the TEOAE will increase in amplitude versus τ_{cut} until it plateaus, any subsequent alteration in amplitude as τ_{cut} increases dependent on the magnitude and phase of the late component versus the early component. The 2 and 3 kHz early component frequencies seem to reach a plateau in amplitude in the 3.4 and 3.75 ms regions. This emphasizes an observation made by Withnell and McKinley (2005) that there is no one cut-off value that optimally separates both components. A cut-off value for the recursive exponential filter (defined as $1/e$) that separates the early and late components of 3.7 ms is within the optimal τ_{cut} range of 3.4–3.75 ms.

E. TEOAE early and late components versus stimulus level

The stimulus-level dependence of the contributions of the early versus late components is illustrated in Fig. 9 with the amplitude and phase spectra of the TEOAE and the early and late components corresponding to the four TEOAEs obtained over a stimulus range of 78 to 63 dB pSPL from Fig. 3. Values of τ_{cut} at each stimulus level with justification for the selection of τ_{cut} are given in Table I. Panel A, Fig. 9 provides the amplitude and phase spectra of the TEOAE and early and late components for a stimulus level of 78 dB

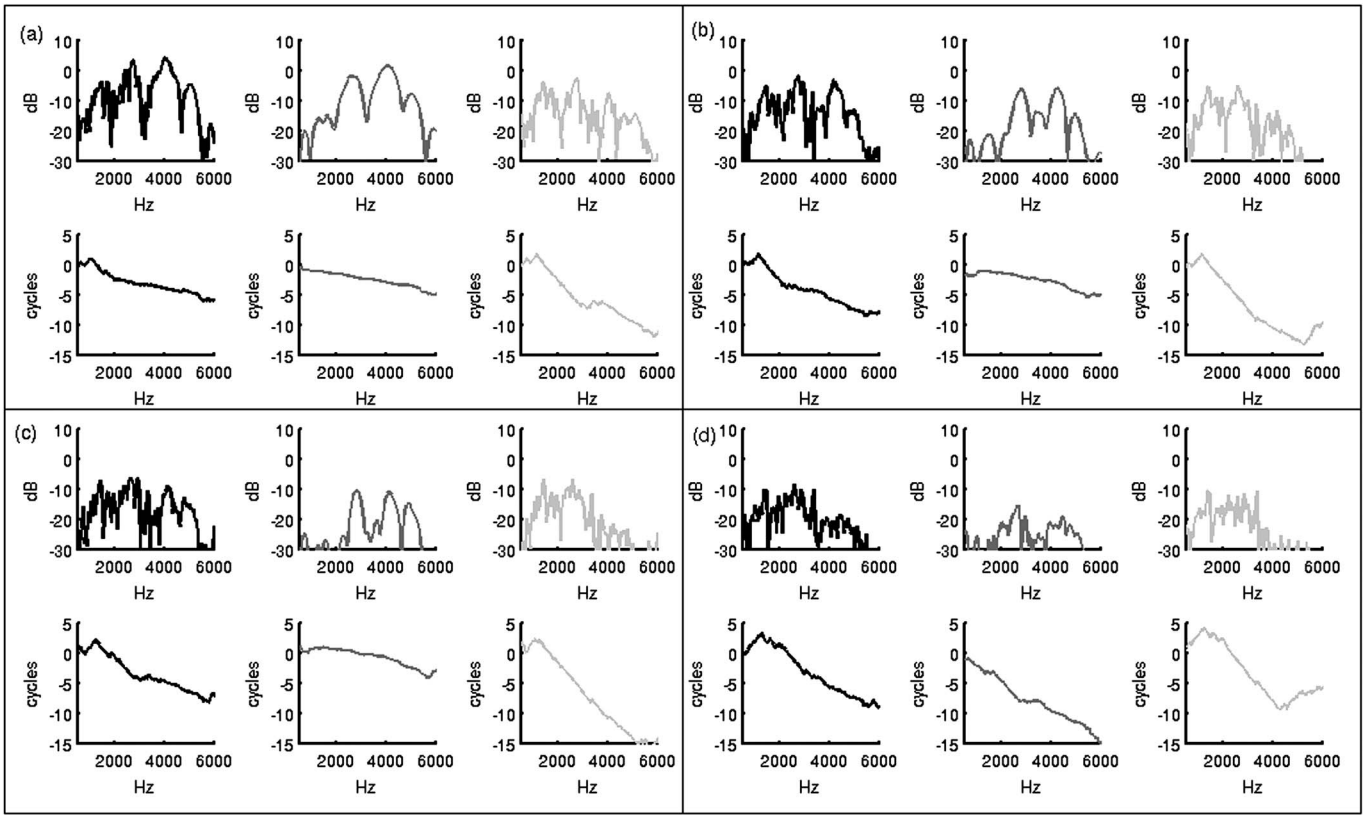

FIG. 9. The stimulus-level dependence of the contributions of the early vs late components is illustrated over a stimulus range of 78 to 63 dB pSPL (TEOAEs from Fig. 3). See the text for further details.

pSPL (see Fig. 5), panels B to D for stimulus levels of 73–63 dB pSPL. TEOAE amplitude spectra decrease in magnitude as stimulus level decreases. Inspection of the phase spectra for the TEOAE reveals that the slope of the phase increases with decreasing stimulus level (steep phase component extends to higher frequencies), arguing for a decrease in the relative contribution of the early component and an increase in the relative contribution of the late component as stimulus level decreases. Time-domain windowing of the TEOAE can isolate an early and a late component with signature phases that argue for distinct mechanisms of generation. The early component amplitude spectrum in Fig. 9 exhibits a microstructure that argues for a more complex origin than can be ascribed to a simple reflection from a nonlinear variation in basilar membrane impedance (Talmadge *et al.*, 2000) where the gain of the cochlear amplifier feedback loop is constant or a slowly varying function of frequency. The early component phase spectra in panels A to C have a slope that does not alter significantly between 1.5 and 4 kHz with decreasing stimulus level; the lack of change in slope of the phase of this component as stimulus level decreases argues against the delay being a construct of the slope of the phase being nonzero may represent a breaking of scaling symmetry (cochlear Q is not constant as a function of frequency). Human cochlear Q's are not thought to be constant as a function of frequency (Shera and Guinan, 2003), although a recent report of cochlear Q's inferred from stimulus frequency otoacoustic emission measurements suggested cochlear Q's to be constant from 1 to 4 kHz (Schairer *et al.*, 2006). The bandwidth of the amplitude spectra of the late component of the TEOAE becomes smaller with decreasing stimulus level, the basal region of the cochlea contributing less to this component. No stimulus-level dependent variation in the slope of the late component phase is evident up to 3 kHz.

TABLE I. Values of τ_{cut} at each stimulus level with justification for the selection of τ_{cut} corresponding to Fig. 9.

Stimulus level	τ_{cut} (ms)	Justification
78	3.73	See the text
73	4.24	TEOAE phase has a steep slope below 2.3 kHz arguing for the late component to be larger than the early component over this frequency range. A saw-tooth pattern is observed from 2.3 to 3 kHz, indicating that the early and late components have similar amplitudes over this frequency range. Above 3 kHz, TEOAE phase has a shallow slope and so the early component must be larger than the late component.
68	4.92	TEOAE phase has a steep slope below 2.7 kHz, a saw-tooth pattern from 2.7 to 3.8 kHz, and a shallow slope above 3.8 kHz
63	6.79	TEOAE phase has a steep slope below 3.0 kHz, a saw-tooth pattern from 3.0 to 3.7 kHz, and a shallow slope above 3.7 kHz

IV. DISCUSSION

The TEOAE in humans has consistently been found to arise predominantly from one mechanism of generation, a

place-fixed mechanism (e.g., Kemp, 1986; Probst *et al.*, 1986; Prieve *et al.*, 1996; Killan and Kapadia, 2006; Kalluri and Shera, 2007). The TEOAE phase spectrum in this case shows a steep slope. In all previous studies that have examined the origin of the TEOAE in humans, the early part of the TEOAE recording was contaminated by stimulus artifact and so was removed prior to analysis, typically the first 2.5 ms (e.g., Kemp *et al.*, 1990), but 6 ms (Prieve *et al.*, 1996) and more (e.g., Killan and Kapadia, 2006) have been extracted from the recording prior to analysis. In all of these studies, stimulus contamination of the TEOAE response was produced by ringing of the earphone centered on the major resonant frequencies of the earphone with the earphone physically coupled to the ear canal.

Here, TEOAEs were recorded with the earphone not physically coupled to the ear canal. Measurements in KEMAR with the earphone not physically coupled to the ear canal suggested that the upper limit to the magnitude of the stimulus contamination of the TEOAE is defined at stimulus onset, prior to the onset of a physiological response. In the absence of significant stimulus contamination at stimulus onset, TEOAEs were recorded with a response beginning within the time window that typically is removed by windowing (see Fig. 1). Time-domain windowing of the TEOAE using a recursive exponential filter suggests that the TEOAE can be separated into two components, one with a shallow phase slope that gives a short group delay, the other having a steep phase slope that gives a group delay suggestive of a round-trip cochlear delay. The shallow phase slope for the early component of the TEOAE differs from stimulus artifact that has zero phase slope and so identifies this early component as a physiological response. Examination of the TEOAE versus stimulus level (Fig. 9) reveals a stimulus-level dependence for the two components identified as contributing to the total TEOAE. As stimulus level increases, the absolute and relative contribution of the early component reduces as stimulus increases while the relative contribution of the late component decreases. It seems then that the TEOAE in humans, in accord with other types of OAE recorded in rodents (TEOAEs: Yates and Withnell, 1999; SFOAEs: Goodman *et al.*, 2003; DPOAEs: Withnell *et al.*, 2003; Schneider *et al.*, 2003) and humans (DPOAEs: Talmadge *et al.*, 1999; Kalluri and Shera, 2001; Knight and Kemp, 2001; SFOAEs: Schairer and Keefe, 2005), has two components with phase spectra that suggest that the TEOAE arises from two distinct mechanisms of generation.

It could be argued that time-domain windowing the TEOAE will produce two components, one with a shallow phase, the other with a steeper phase, simply by virtue of the windowing process. The group delay for the early component (e.g., the early component of the TEOAE in Figs. 5 and 6 had a "group delay" of ~0.8 ms) it could be argued represents contributions from the tails of the traveling waves generated by the stimulus, the response having a short delay by virtue of it arising basal to the active region of the traveling wave for each frequency component. Two testable predictions arise from this explanation: (i) The early component of the TEOAE should have a group delay that is an increasing function of τ_{cut} and (ii) the group delay for the early component of the TEOAE should increase with decreasing stimulus level.

With respect to the first prediction, inspection of the group delay for the early component as a function of τ_{cut} in Fig. 8 at 3 kHz reveals that up to a $\tau_{cut} \sim 7.5$ or 5.8 ms poststimulus peak the group delay is between 0.6 and 1.6 ms. The delay increases from 0.6 to 1.6 ms over the range $2.5 < \tau_{cut} < 6.4$ ms, and then stays relatively constant at a delay of ~1.6 ms from $6.4 < \tau_{cut} < 7.5$ ms. Note that we are considering here the proposition that the TEOAE arises from one generation mechanism so that increasing τ_{cut} increases the OAE within the time window and so eventually, as τ_{cut} increases, it should encompass the whole TEOAE. A $\tau_{cut} \sim 7.5$ or 5.8 ms poststimulus peak provides for a TEOAE that should receive significant contribution from the cochlear region near the characteristic place for 3 kHz (mean SFOAE group delay at 3 kHz to 5.5 ms; Shera and Guinan, 2003). A time dependence for the 3 kHz component represents an increasing contribution from OAE later in time but the group delay peaks at 1.6 ms, a value not in accord with one mechanism of generation, i.e., the TEOAE at 3 kHz encompassed within a time window of 5.8 ms poststimulus peak should have a much longer delay if the OAE arises from only mechanism and the amplitude of the emission is dependent on the displacement amplitude of the basilar membrane. Alternatively, an argument based on one mechanism where the emission is generated in the basal region of the cochlea would not provide for a late component with an OAE group delay commensurate with a round-trip cochlear travel time.

The second prediction can be tested by examining Fig. 9. The phase spectra for the early component of the TEOAE has a slope that does not alter significantly between 1.5 and 4 kHz with stimulus level from 78 to 68 dB pSPL, i.e., the group delay for the early component of the TEOAE does not increase with decreasing stimulus level. It would appear, then, that the group delay for the early component extracted by time-domain windowing is consistent with the emission arising from a wave-fixed mechanism and inconsistent with the delay being a construct of the windowing process.

Zweig and Shera (1995) provided a theoretical description for the generation of OAEs exhibiting a steep phase slope, the coherent reflection filter (CRF) theory. The CRF theory is expected to apply only in the low-level, linear region of operation of the cochlea; with increasing stimulus level, the traveling wave peak broadens, presumably reducing the phase coherence of reflections across this region. In addition to reflection from randomly distributed cochlear irregularities (place-fixed OAE), emissions also are thought to arise as a consequence of cochlear nonlinearity acting through the cochlear amplifier feedback loop (wave-fixed OAE). To date, no theory of cochlear mechanical function provides an adequate description of OAE generation incorporating a stimulus-level dependent cochlear nonlinearity.

Windowing or removing the early part of the TEOAE recording so as to isolate TEOAE that is generated predominantly by one mechanism, a linear, place-fixed mechanism, serendipitously has clinical virtue. This TEOAE component will have a one-to-one correspondence with the stimulus fre-

quencies that generated it subject to variation in cochlear reflectance. The amplitude spectrum microstructure, according to the CRF theory, is produced by random variation in cochlear reflectance (Zweig and Shera, 1995). The SFOAE, according to the CRF theory, arises from reflections from cochlear irregularity with the signal source being the amplitude of displacement of the basilar membrane traveling wave, a source that is largest at the peak of the traveling wave subject to phase coherence across this peak. A number of studies have established that SFOAEs arise predominantly from this tip region by virtue of their group delay (e.g., Shera and Guinan, 2003; Goodman *et al.*, 2004) with TEOAEs being effectively a composite of SFOAEs (Kalluri and Shera, 2007).

In guinea pigs, it has been suggested that the TEOAE arises predominantly from intermodulation distortion energy generated by the cochlear nonlinear response to the stimulus component frequencies (Yates and Withnell, 1999). A nonlinear generation mechanism will be both within channel and between channel, the extent of the between channel contribution (intermodulation distortion) being presumably a consequence of the amount of overlap of the cochlear filters (Withnell and McKinley, 2005). In humans it has been found that cochlear filters are sharper than rodents (Shera *et al.*, 2002) and so the contribution of intermodulation distortion products to the TEOAE in humans is presumably less than in rodents. The within channel nonlinear contribution presumably arises from the outer hair cell nonlinearity/ies acting through the cochlear amplifier feedback loop generating a periodic basilar membrane response to a sinusoidal input that is a sum of the fundamental plus higher order harmonic distortion. Higher order harmonics do not couple well into the basilar membrane and the resultant fundamental response is not linearly related to the change in input stimulus level.

Previous studies suggesting that the TEOAE recorded from a human ear arises solely from one mechanism is presumably a result of windowing the earliest part of the TEOAE response, removing much of the component arising from a nonlinear generation mechanism. The TEOAE, as has been found in guinea pig, appears to arise from two distinct mechanisms, the relative contributions of these two mechanisms being time dependent and stimulus level dependent. It seems, then, that all OAEs in mammals arise in a stimulus level dependent manner from two mechanisms of generation, one linear, one nonlinear, as suggested by Shera and Guinan (1999).

ACKNOWLEDGMENT

The authors would like to thank Chris Shera and two anonymous reviewers for valuable feedback on this manuscript.

[1]The stimulus wave form was generated using a sinc function $[\sin(\omega t)/(\omega t)]$ deconvolved with the impulse response of the loudspeaker, providing a stimulus with a flat amplitude spectrum and linear phase delay at the measurement microphone in the ear canal.

Avan, P., Bonfils, P., Loth, D., Elbez, M., and Erminy, M. (**1995**). "Transient-evoked otoacoustic emissions and high-frequency acoustic trauma in the guinea pig," J. Acoust. Soc. Am. **97**, 3012–3020.

Avan, P., Bonfils, P., Loth, D., and Wit, H. P. (**1993**). "Temporal patterns of transient-evoked otoacoustic emissions in normal and impaired cochleae," Hear. Res. **70**, 109–120.

Avan, P., Elbez, M., and Bonfils, P. (**1997**). "Click-evoked otoacoustic emissions and the influence of high-frequency hearing losses in humans," J. Acoust. Soc. Am. **101**, 2771–2777.

Carvalho, S., Buki, B., Bonfils, P., and Avan, P. (**2003**). "Effect of click intensity on click-evoked otoacoustic emission wave forms: Implications for the origin of emissions," Hear. Res. **175**, 215–225.

Goodman, S. S., Withnell, R. H., De Boer, E., Lilly, D. J., and Nuttall, A. L. (**2004**). "Cochlear delays measured with amplitude-modulated tone-burst evoked OAEs," Hear. Res. **188**, 57–69.

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

Hauser, R., Probst, R., and Lohle, E. (**1991**). "Click- and tone-burst-evoked otoacoustic emissions in normally hearing ears and in ears with high-frequency sensorineural hearing loss," Eur. Arch. Otorhinolaryngol. **248**, 345–352.

Kalluri, R., and 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.

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

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

Kemp, D. T., and Chum, R. (**1980**). "Properties of the generator of stimulated acoustic emissions," Hear. Res. **2**, 213–232.

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

Killan, E. C., and Kapadia, S. (**2006**). "Simultaneous suppression of tone burst-evoked otoacoustic emissions—Effect of level and presentation paradigm," Hear. Res. **212**, 65–73.

Knight, R. D., and Kemp, D. T. (**2001**). "Wave and place fixed DPOAE maps of the human ear," J. Acoust. Soc. Am. **109**, 1513–1525.

Long, G. R. (**1984**). "The microstructure of quiet and masked thresholds," Hear. Res. **15**, 73–87.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R., and Coats, A. C. (**1988**). "Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy," Hear. Res. **33**, 69–93.

Prieve, B. A., Gorga, M. P., and Neely, S. T. (**1996**). "Click- and tone-burst-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears," J. Acoust. Soc. Am. **99**, 3077–3086.

Probst, R., Coats, A. C., Martin, G. K., and Lonsbury-Martin, B. L. (**1986**). "Spontaneous, click-, and toneburst-evoked otoacoustic emissions from normal ears," Hear. Res. **21**, 261–275.

Schairer, K. S., Ellison, J. C., Fitzpatrick, D., and Keefe, D. H. (**2006**). "Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears," J. Acoust. Soc. Am. **120**, 901–914.

Schairer, K. S., and Keefe, D. H. (**2005**). "Simultaneous recording of stimulus-frequency and distortion-product otoacoustic emission input-output functions in human ears," J. Acoust. Soc. Am. **117**, 818–832.

Schneider, S., Prijs, V. F., and 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., and Guinan, J. J., Jr. (**1999**). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," J. Acoust. Soc. Am. **105**, 782–798.

Shera, C. A., and Guinan, J. J., 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, C. A., Guinan, J. J., and Oxenham, A. J. (**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., and Talmadge, C. L. (**2005**). "Coherent reflection in a two-dimensional cochlea: Short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions," J. Acoust. Soc. Am. **118**, 287–313.

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

Talmadge, C. L., Long, G. R., Tubis, A., and 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., and Piskorski, P. (**1998**). "Modeling otoacoustic emission and hearing threshold fine structures," J. Acoust. Soc. Am. **104**, 1517–1543.
Talmadge, C. L., Tubis, A., Long, G. R., and 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.
Withnell, R. H., Kirk, D. L., and Yates, G. K. (**1998**). "Otoacoustic emissions measured with a physically open recording system," J. Acoust. Soc. Am. **104**, 350–355.
Withnell, R. H., and McKinley, S. (**2005**). "Delay dependence for the origin of the nonlinear derived transient evoked otoacoustic emission," J. Acoust. Soc. Am. **117**, 281–291.
Withnell, R. H., Shaffer, L. A., and Talmadge, C. L. (**2003**). "Generation of DPOAEs in the guinea pig," Hear. Res. **178**, 106–117.
Withnell, R. H., and Yates, G. K. (**1998**). "Enhancement of the transient-evoked otoacoustic emission produced by the addition of a pure tone in the guinea pig," J. Acoust. Soc. Am. **104**, 344–349.
Xu, L., Probst, R., Harris, F. P., and Roede, J. (**1994**). "Peripheral analysis of frequency in human ears revealed by tone burst evoked otoacoustic emissions," Hear. Res. **74**, 173–180.
Yates, G. K., and Withnell, R. H. (**1999**). "The role of intermodulation distortion in transient-evoked otoacoustic emissions," Hear. Res. **136**, 49–64.
Zweig, G., and Shera, C. A. (**1995**). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am. **98**, 2018–2047.