Near equivalence of human click-evoked and stimulus-frequency otoacoustic emissions

Radha Kalluri
Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye & Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114 and Speech and Hearing figureBioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02138

Christopher A. Shera[a)]
Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye & Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114; Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02138; and Department of Otology & Laryngology, Harvard Medical School, Boston, Massachusetts 02115

(Received 15 August 2006; revised 29 December 2006; accepted 29 December 2006)

Otoacoustic emissions (OAEs) evoked by broadband clicks and by single tones are widely regarded as originating via different mechanisms within the cochlea. Whereas the properties of stimulus-frequency OAEs (SFOAEs) evoked by tones are consistent with an origin via linear mechanisms involving coherent wave scattering by preexisting perturbations in the mechanics, OAEs evoked by broadband clicks (CEOAEs) have been suggested to originate via nonlinear interactions among the different frequency components of the stimulus (e.g., intermodulation distortion). The experiments reported here test for bandwidth-dependent differences in mechanisms of OAE generation. Click-evoked and stimulus-frequency OAE input/output transfer functions were obtained and compared as a function of stimulus frequency and intensity. At low and moderate intensities human CEOAE and SFOAE transfer functions are nearly identical. When stimulus intensity is measured in "bandwidth-compensated" sound-pressure level (cSPL), CEOAE and SFOAE transfer functions have equivalent growth functions at fixed frequency and equivalent spectral characteristics at fixed intensity. This equivalence suggests that CEOAEs and SFOAEs are generated by the same mechanism. Although CEOAEs and SFOAEs are known by different names because of the different stimuli used to evoke them, the two OAE "types" are evidently best understood as members of the same emission family. [DOI: 10.1121/1.2435981]

PACS number(s): 43.64.Jb, 43.64.Bt, 43.64.Kc, 43.58.Ry [BLM] Pages: 2097–2110

I. INTRODUCTION

The stimuli used to evoke otoacoustic emissions (OAEs) range from the spectrally dense (e.g., the broadband clicks used to elicit click-evoked OAEs) to the spectrally sparse (e.g., the single pure tones used to evoke stimulus-frequency OAEs). Although linear reflection models of OAE generation (e.g., Zweig and Shera, 1995; Talmadge *et al.*, 1998) predict that both click-evoked and stimulus-frequency otoacoustic emissions (CEOAEs and SFOAEs) originate via essentially linear mechanisms (i.e., wave reflection off preexisting mechanical perturbations), other models imply that differences in the spectrum of the evoking stimulus result in differences in the mechanisms of OAE generation. Nobili *et al.* (2003a), for example, use model simulations to argue that the mechanisms responsible for CEOAE generation are both inherently nonlinear and fundamentally different from those responsible for generating SFOAEs. In Nobili *et al.*'s simulations, CEOAEs result from spatially complex "residual oscillations" of the basilar membrane that trace their origin to spectral irregularities in middle-ear transmission (see also Nobili, 2000; Nobili *et al.*, 2003b). Based on OAE measurements in guinea pig, Yates and Withnell (1999b) also posit a distinction between OAEs evoked by narrow- and broadband stimuli. They argue that although SFOAEs may originate from the independent "emission channels" predicted by linear reflection models, CEOAEs are essentially broadband distortion-product emissions (broadband DPOAEs). In this view, CEOAEs arise not from independent channels but from intermodulation distortion sources induced as a consequence of nonlinear interactions among the multiple frequency components of the broadband click stimulus (see also Withnell and Yates, 1998; Yates and Withnell, 1999a; Carvalho *et al.*, 2003).

The work reported here was motivated by these basic disagreements about the influence of stimulus spectrum on mechanisms of OAE generation. Our goal was to determine the relationship between the OAEs evoked by stimuli with the most dissimilar temporal and spectral structure (i.e., CEOAEs and SFOAEs). Interpretation of the experiments assumes that differences in OAE spectral characteristics im-

a)Electronic mail: shera@epl.meei.harvard.edu

ply differences in OAE generating mechanisms, and conversely. Similar logic has been used to distinguish "reflection-" and "distortion-source" OAEs. Whereas reflection-source OAEs (e.g., SFOAEs at low levels) have a rapidly varying phase and a slowly varying amplitude occasionally punctuated by sharp notches, distortion source OAEs (e.g., DPOAEs evoked at fixed, near-optimal primary-frequency ratios) have an almost constant amplitude and phase. These differences in OAE spectral characteristics are taken as indicative of fundamental differences in their mechanisms of generation (e.g., Kemp and Brown, 1983; Shera and Guinan, 1999).

Despite its fundamental importance, only a handful of studies have addressed the comparison between CEOAEs and SFOAEs. Although Zwicker and Schloth (1984) measured tone- and click-evoked frequency responses in the same human subject, the uncertain reliability of their tone-evoked data precludes any compelling comparison. Unlike the tone-evoked responses observed in subsequent studies, the "synchronous-evoked" OAEs reported by Zwicker and Schloth appear inconsistent with an origin in a causal system (Shera and Zweig, 1993). Furthermore, the emission data for the two stimulus types are presented in different ways: Although the CEOAE data represent the emission alone, the tone-evoked data represent the combined pressure of the stimulus and the emission. In what appears to be the only other study to explicitly address the issue, Prieve *et al.* (1996) found that the OAEs evoked by clicks and by tone bursts have similar intensity dependence, consistent with a common mechanism of generation. Unfortunately, the bandwidths of their tone bursts were not all that narrow (they typically spanned an octave or more), and their data do not allow a comparison of spectral structure or phase.

The experiments reported here examine the effect of stimulus bandwidth on OAE generation mechanisms by measuring and appropriately comparing the emissions evoked by wideband clicks (CEOAEs) with those evoked by tones (SFOAEs). Comparisons are made across stimulus frequency and intensity in the same human subjects.

II. MEASUREMENT METHODS

Measurements were made in one ear of each of four ($n=4$) normal-hearing human subjects who were comfortably seated in a sound-isolated chamber. All procedures were approved by human studies committees at the Massachusetts Eye and Ear Infirmary and the Massachusetts Institute of Technology.

Stimuli were digitally generated and recorded using implementations of standard OAE measurement protocols on the Mimosa Acoustics measurement system. The measurement system consists of a DSP-board (CAC Bullet) installed in a laptop computer, an Etymōtic Research ER10c probe system, and two software programs—one for measuring CEOAEs (T2001 v3.1.3) and another for measuring SFOAEs (SF2003 v2.1.18).

Signals were delivered and recorded in the ear canal. In-the-ear calibrations were made before each measurement. Stimuli were digitally generated using a fixed sampling rate of 48 kHz and data buffer lengths of 4096 samples, resulting in a frequency resolution of approximately 12 Hz. Potential artifacts were detected in real time by computing the difference between the current data buffer and an artifact-free reference buffer. The current data buffer was discarded whenever the rms value of the difference waveform exceeded a subject-specific criterion. Accepted data buffers were added to the averaging buffer. Continual replacement of the reference buffer minimized the effects of slowly varying drifts in the baseline.

We briefly outline the procedures for measuring each type of OAE below. Interested readers can consult Mimosa Acoustics technical documentation for more detailed descriptions of the measurement system (see also Lapsley-Miller *et al.*, 2004a,b).

A. Measuring CEOAEs

CEOAEs were evoked using broadband clicks (0.5–5 kHz) ranging in intensity from 35 to 80 dB pSPL (peak-equivalent SPL). To enable comparisons with SFOAEs, which are evoked using iso-intensity pure tones, the click waveform was adjusted using the in-the-ear calibration data to produce a flat-spectrum microphone signal. Responses were averaged across 500–4000 repetitions, depending on the stimulus level. Noise floors for the measurements typically ranged from −25 to −33 dB SPL.

A typical response waveform is shown in Fig. 1. The large pulse is the acoustic click, the smaller, more temporally dispersed portion of the waveform is the CEOAE. CEOAEs were extracted from the ear-canal pressure waveform by using either the linear-windowing or the nonlinear-residual method. The following sections describe each method in turn. Our standard-protocol used a click repetition period T_a of approximately 26 ms (1253 samples). As a check for possible efferent effects (e.g., Guinan *et al.*, 2003), we varied the interstimulus time from roughly 20 ms up to 100 ms but found no significant dependence on repetition period.

1. The linear windowing method

In the linear windowing paradigm (e.g., Kemp, 1978) the stimulus and emission, $p_{ST}(t)$ and $p_{CE}(t)$, are extracted from the total ear-canal pressure, $p(t)$, by applying stimulus and emission windows, $w_s(t)$ and $w_e(t)$:

$$p_{ST}(t) = w_s(t)p(t), \tag{1}$$

$$p_{CE}(t) = w_e(t)p(t). \tag{2}$$

The stimulus and emission spectra are then computed by taking the 4096-point discrete Fourier transform $\mathcal{F}\{\cdot\}$ of zero-padded waveforms $p_{CE}(t)$ and $p_{ST}(t)$:

$$P_{ST}(f) = \mathcal{F}\{p_{ST}(t)\}, \tag{3}$$

$$P_{CE}(f) = \mathcal{F}\{p_{CE}(t)\}. \tag{4}$$

The input/output CEOAE transfer function $T_{CE}(f;A)$ is defined as the ratio of the two spectra:

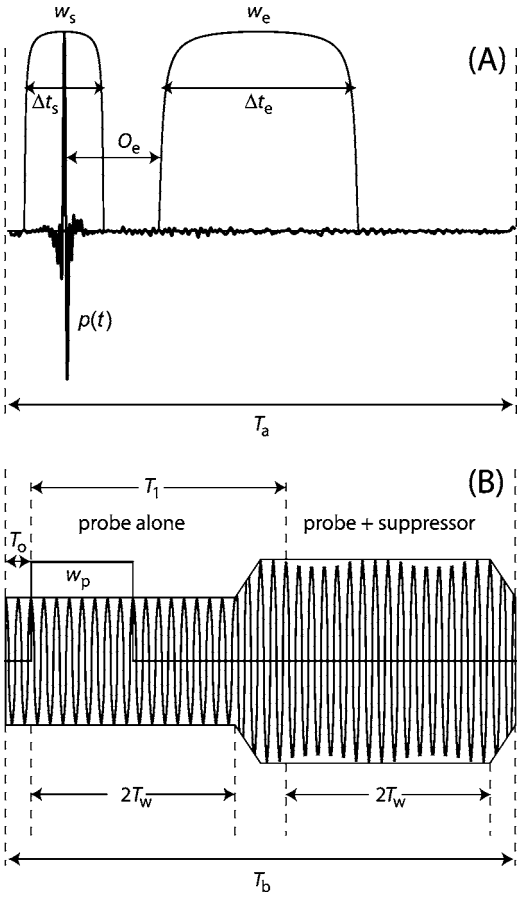

FIG. 1. Schematic diagrams of the measurement paradigms. (A) For CEOAEs, the stimulus and emission are measured using the linear windowing technique by applying recursive-exponential windows w_s and w_e to the ear-canal pressure, $p(t)$. The windows' center positions, t_s and t_e, and widths, Δt_s and Δt_e, are chosen to optimize the separation between the stimulus and emission. (B) For SFOAEs, the stimulus and emission are measured using the interleaved suppression technique. The emission is computed as the Fourier component at the probe frequency by taking the complex difference between the probe-alone and probe-suppressor segments of the ear-canal pressure. The probe-alone and probe-suppressor waveforms are extracted using rectangular windows $w_p^{(n)}$ and $w_{ps}^{(n)}$; only $w_p^{(0)}$ is shown in the figure. The Fourier analysis buffer (duration T_w) contains an integral number of cycles of both probe and suppressor.

$$T_{CE}(f;A) = \frac{P_{CE}(f;A)}{P_{ST}(f)}, \tag{5}$$

where $A \equiv |P_{ST}|$ is the stimulus amplitude. Although we refer to the ratio as a transfer function, $T_{CE}(f;A)$ depends on the stimulus amplitude and is therefore more correctly known as a "describing" function (e.g., Krylov and Bogolyubov, 1947; Gelb and Vander Velde, 1968).

For the windowing technique to work, the stimulus click must be sufficiently localized in time so that the end of the stimulus does not significantly overlap with the early components of the emission. Unless otherwise noted, the clicks used in these experiments were bandlimited from 0.5 to 5 kHz—the broadest flat spectrum click without notches that the measurement system was able to generate. Interference between stimulus and emission can be further reduced by the proper choice of windows. We used tenth-order recursive-exponential windows $w_{rex}(t;\Delta t)$ (Shera and Zweig, 1993;

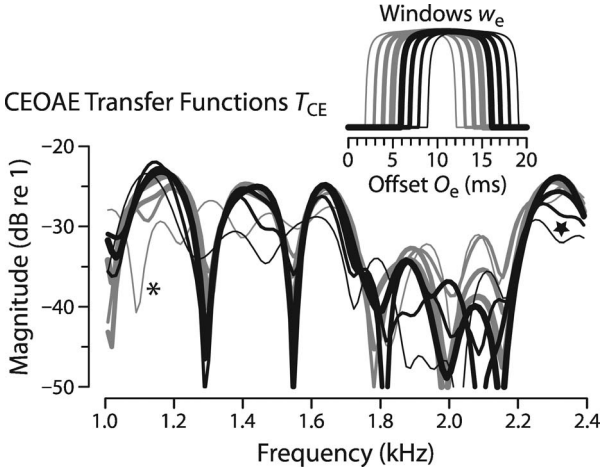

FIG. 2. Dependence of CEOAE spectra on window offset, O_e. The inset illustrates the windows $w_e(t-t_e;\Delta t_e)$ with poststimulus offsets ranging from 2 to 8 ms. The spectral structure of the CEOAE transfer function varies with window offset. For window offsets between 5 and 7 ms, CEOAE transfer functions are almost independent of O_e (thick lines). For shorter offsets (<5 ms) the transfer functions manifest additional spectral structure (*), presumably due to interference-like interactions between the stimulus and the emission. For offsets greater than 7 ms the short latency, high-frequency components of the CEOAE degrade (⋆).

Kalluri and Shera, 2001) with time offsets and widths chosen to reduce interactions between the stimulus and emission. Thus,

$$w_s(t) = w_{rex}(t - t_s;\Delta t_s), \tag{6}$$

$$w_e(t) = w_{rex}(t - t_e;\Delta t_e), \tag{7}$$

with standard offsets $\{t_s, t_e\}=\{0, 10\}$ ms and widths $\{\Delta t_s, \Delta t_e\}=\{5, 10\}$ ms. All offsets are relative to the center of the stimulus click at $t=0$. The recursive-exponential window is defined in footnote 10 of Kalluri and Shera (2001).

The location of the emission analysis window must be carefully chosen to reduce interference caused by interactions between the stimulus and the early components of the emission. The window $w_e(t)$ begins at time $O_e=t_e-\Delta t_e/2$ after the click (see Fig. 1). To determine the "optimal" window offset, we varied O_e until small shifts had negligible effects on the transfer function within the frequency range of interest (1–4 kHz). Offsets smaller than about 5 ms or larger than 7 ms produced significant changes in the magnitude of the transfer function (see Fig. 2). Except where noted, we adopted the value $O_e=5$ ms for all the results shown here. Because CEOAEs are dispersed in time, with high frequency components arriving before the low frequency components, the optimal window for the 1–4-kHz region will not be optimal for emissions in other frequency bands.

2. The nonlinear residual method

The nonlinear residual method is an alternate and generally more popular procedure for measuring CEOAEs. In this method CEOAEs are extracted by exploiting the nonlinear compressive growth of the emissions in conjunction with the linear growth of the stimulus. Three identical clicks are

followed by a fourth click that is three times larger but of the opposite polarity. The CEOAE estimate is the average of the four responses.

Unlike the linear windowing technique, in which short-latency components of the emission are typically eliminated, the nonlinear residual method separates the emission from the stimulus without removing the early arriving components of the emission. However, to avoid confusion, reduce potential artifacts due to system distortion, and enable direct comparison between the two CEOAE techniques, we apply the standard emission window, $w_e(t)$, to the nonlinear-derived emission as well. Therefore, just as in the linear technique, early arriving components of the emission are eliminated.

B. Measuring SFOAEs

We measured the SFOAE pressure, $P_{SF}(f)$, using a variant of the suppression method (Shera and Guinan, 1999; Kalluri and Shera, 2001). As illustrated in Fig. 1, the emission is obtained as the complex difference between the ear-canal pressure at the probe frequency (f) measured first with the probe tone alone and then in the presence of a more intense (55 dB SPL) suppressor tone at a nearby frequency, f_s, roughly 47 Hz below the probe frequency (Fig. 1). The suppressor was presented in interleaved time segments to minimize possible artifactual contamination from time-varying drifts in the base signal. To reduce spurious contamination by earphone distortion, the probe and suppressor were generated using separate sounds sources.

The probe-alone waveform, $p_p(t)$, and probe-suppressor waveform, $p_{ps}(t)$, are obtained from the measured ear-canal pressure, $p(t)$, by averaging over two subsegments extracted using windows $w_p(t)$ and $w_{ps}(t)$:

$$p_p(t) = \frac{1}{2}\sum_{n=0}^{1} w_p^{(n)}(t)p(t), \tag{8}$$

$$p_{ps}(t) = \frac{1}{2}\sum_{n=0}^{1} w_{ps}^{(n)}(t)p(t). \tag{9}$$

In this case, the windows are rectangular boxcars of width T_w:

$$w_p^{(n)}(t) = w_{box}(t - T_0 - nT_w; T_w); \tag{10}$$

$$w_{ps}^{(n)}(t) = w_{box}(t - T_0 - T_1 - nT_w; T_w), \tag{11}$$

where

$$w_{box}(t;\Delta t) = \begin{cases} 1 & \text{for } 0 \le t \le \Delta t, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

The window offsets T_0 and T_1 were chosen to allow the system time to return to steady state after switching the suppressor tone on or off. The window duration T_w equals that of the Fourier analysis buffer. Stimulus frequencies were chosen so that the analysis buffer (of duration $T_w = N\Delta t$, where N is the buffer size and Δt is the reciprocal of the sampling rate) always contained an integral number of cycles of both probe and suppressor. For the measurements reported here $\{T_0, T_1, T_b\} = \{\frac{1}{4}, \frac{5}{2}, 5\}T_w$. The SFOAE pressure is computed as

$$P_{SF}(f) = \mathcal{F}\{p_p(t)\} - \mathcal{F}\{p_{ps}(t)\}e^{i2\pi f T_1}, \tag{13}$$

where $\mathcal{F}\{\cdot\}$ indicates the 4096-point discrete Fourier transform at the probe frequency, f. The stimulus pressure is extracted from the probe-suppressor segment:

$$P_{ST}(f) = \mathcal{F}\{p_{ps}(t)\}e^{i2\pi f T_1}. \tag{14}$$

By analogy with Eq. (5) for $T_{CE}(f;A)$, the transfer function $T_{SF}(f;A)$ is defined as the ratio of probe-frequency spectral components

$$T_{SF}(f;A) = \frac{P_{SF}(f;A)}{P_{ST}(f)}, \tag{15}$$

where we have now explicitly indicated the dependence on stimulus amplitude ($A \equiv |P_{ST}|$). We measured $T_{SF}(f;A)$ with a frequency resolution of approximately 23 Hz using probe-tone levels ranging from approximately 10 to 40 dB SPL. We typically employed 32 averages at the highest probe level and 128 averages at the lowest.

III. EXPERIMENTAL COMPLICATIONS

Before describing our main results, we first address two measurement issues that complicate the comparison between $T_{CE}(f;A)$ and $T_{SF}(f;A)$. The first pertains to differences between the two different CEOAE measurement methods. The second deals with complications arising from synchronized spontaneous otoacoustic emissions (SSOAE).

A. CEOAE transfer functions from linear and nonlinear methods

Figure 3 compares the CEOAE transfer functions $T_{CE}(f;A)$ measured using the linear-windowing and nonlinear-residual methods. We denote transfer functions measured using the two methods by $T_{CE}(f;A)$ and $T_{CE}^{NL}(f;A)$, respectively. For brevity, we show measurements from one subject; similar results were obtained in all.

Although both the linear-windowing and nonlinear-residual techniques yield qualitatively similar values of $T_{CE}(f;A)$ at high stimulus levels, CEOAEs at low levels can only be extracted using the linear technique. As stimulus levels are decreased from 80 to 60 dB pSPL, the magnitudes of both $T_{CE}(f;A)$ and $T_{CE}^{NL}(f;A)$ increase. At these levels $T_{CE}(f;A)$ and $T_{CE}^{NL}(f;A)$ have similar peaks, notches, and phase behaviors. This similarity in behavior does not carry through to the lowest levels. As stimulus levels are further reduced, the magnitude of $T_{CE}(f;A)$ continues to grow and eventually becomes nearly independent of level. By contrast, $T_{CE}^{NL}(f;A)$ reaches a maximum value and then falls quickly into the noise floor. The combination of results—near level independence of $T_{CE}(f;A)$ and the rapid fall of $T_{CE}^{NL}(f;A)$ at low stimulus levels—suggests that CEOAEs grow almost linearly at the lowest stimulus levels. Note, however, that by using the nonlinear-derived method, Withnell and McKinley (2005) found short-latency CEOAE components in guinea pigs that appear to result from nonlinear mechanisms within the cochlea. When measured using the linear-windowing

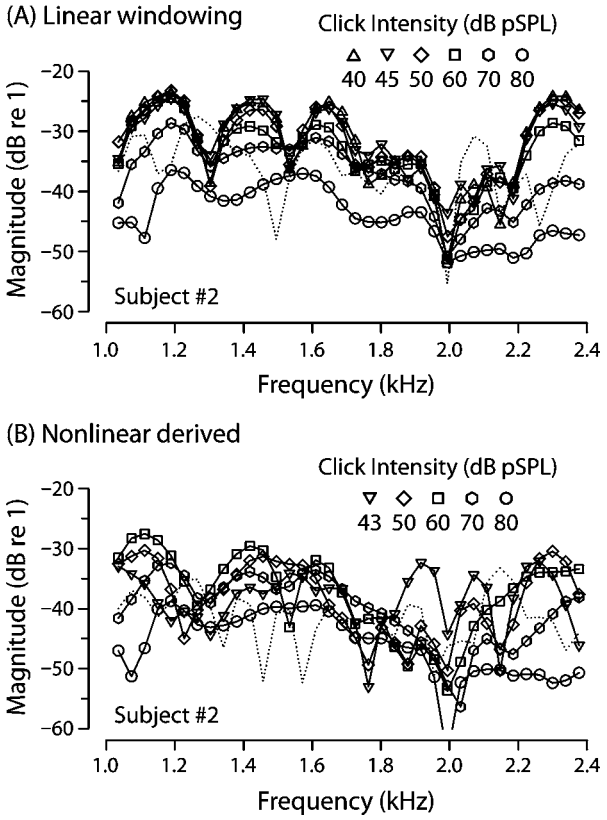

FIG. 3. Linear-windowed and nonlinear-derived CEOAE transfer functions. Panels (A) and (B) show $T_{\mathrm{CE}}(f;A)$ and $T_{\mathrm{CE}}^{\mathrm{NL}}(f;A)$, respectively, at click intensities ranging from 40 to 80 dB pSPL. The two techniques yield qualitatively similar results at high stimulus levels (70–80 dB pSPL). At lower intensities, however, they diverge: Whereas the nonlinear-derived $T_{\mathrm{CE}}^{\mathrm{NL}}(f;A)$ ultimately falls into the noise, the linear-windowed $T_{\mathrm{CE}}(f;A)$ continues to increase until it becomes independent of intensity. The transfer-function noise floors shown here (dotted lines) were measured at the lowest click intensities. Because they are scaled by the stimulus spectrum, transfer-function noise floors are much lower at higher stimulus levels.

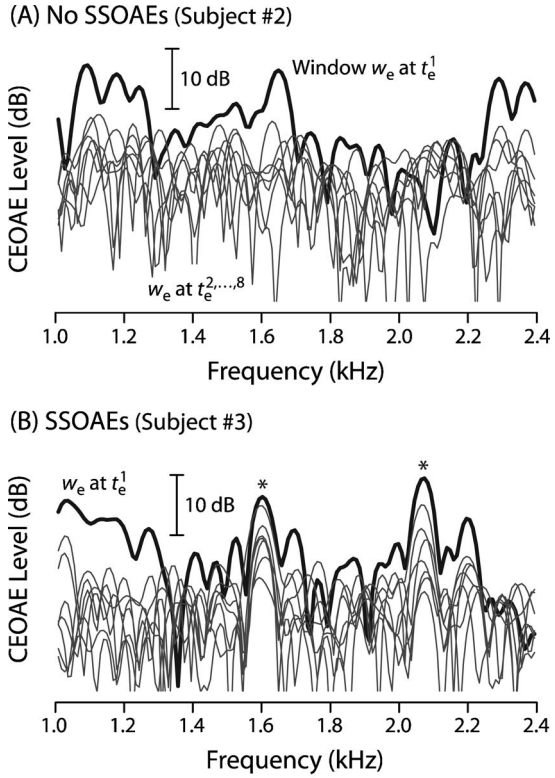

FIG. 4. Sustained activity in the CEOAE spectrum identifies synchronized SOAEs. SSOAEs were identified using the linear-windowing technique with an interstimulus time of 100 ms. Eight 20-ms analysis windows (with center offsets t_e^1 through t_e^8 ranging from 15 to 85 ms after the stimulus) were applied to the measured ear-canal pressure. Panel (A) shows the corresponding CEOAE spectra in a subject without measurable SSOAEs. Only the spectrum of the response during the first window (t_e^1, thick line) contains significant emission energy; responses from all subsequent windows (t_e^2 through t_e^8; thin lines) are small by comparison. Panel (B) shows the spectra in a subject with SSOAEs. A response at SSOAE frequencies appears as a spectral peak in all windows (*).

protocol, these short latency components would typically be obscured by the stimulus. Since our measurements had a residual short-latency stimulus artifact due to earphone nonlinearities (e.g., Kapadia *et al.*, 2005), we cannot rule out the possibility that human CEOAEs also contain small short-latency nonlinear components buried beneath the stimulus artifact. Because the nonlinear technique cannot be used to measure $T_{\mathrm{CE}}(f;A)$ at the lowest stimulus levels, all subsequent CEOAE measurements presented in this paper were made using the linear protocol.

B. Synchronized spontaneous otoacoustic emissions

Some of our subjects had synchronized spontaneous otoacoustic emissions (SSOAEs). SSOAEs are long-lasting transient responses that are not always identifiable by conventional SOAE searches, in which no external stimulus is presented. SSOAEs can, however, be detected when they are evoked by or synchronized to an applied stimulus, in this case the click used to evoke CEOAEs.

We measured SSOAEs using a variant of the standard linear-windowing technique for measuring CEOAEs. The variant employed an interclick time of 100 ms rather than the standard 20 ms used in the CEOAE measurements. To detect SSOAEs we then computed and compared the response spectra, $P_{\mathrm{CE}}^i(f)$, in eight partially overlapping analysis windows centered at different post-stimulus times ($t_e^i = 5 + 10i$ ms):

$$P_{\mathrm{CE}}^i(f) = \mathcal{F}\{p(t)w(t - t_e^i;\Delta t)\}, \quad i = 1,2,\ldots,8, \tag{16}$$

where the nominal window duration Δt is 20 ms. Figure 4 shows the spectra for the eight windowed segments in two subjects. The dark thick line gives the spectrum of the response during the first window, centered at 15 ms after the click. The narrow lines show the spectra measured during subsequent windows. In two of the four subjects, significant response energy occurs only within the first 20 ms, and we considered these subjects to have unmeasurable SSOAEs. In the remaining two subjects, some peaks in the spectrum [e.g., those identified by asterisks (*) in Fig. 4] disappear more slowly over the eight windows. We identified these long-lasting transient responses to the click as SSOAEs.

The existence of SSOAEs complicates the measurement of $T_{\mathrm{CE}}(f;A)$, and to a lesser extent $T_{\mathrm{SF}}(f;A)$, in at least two ways. First, SSOAEs make it more difficult to determine the stimulus spectrum [i.e., the denominator in Eq. (5)]. In sub-

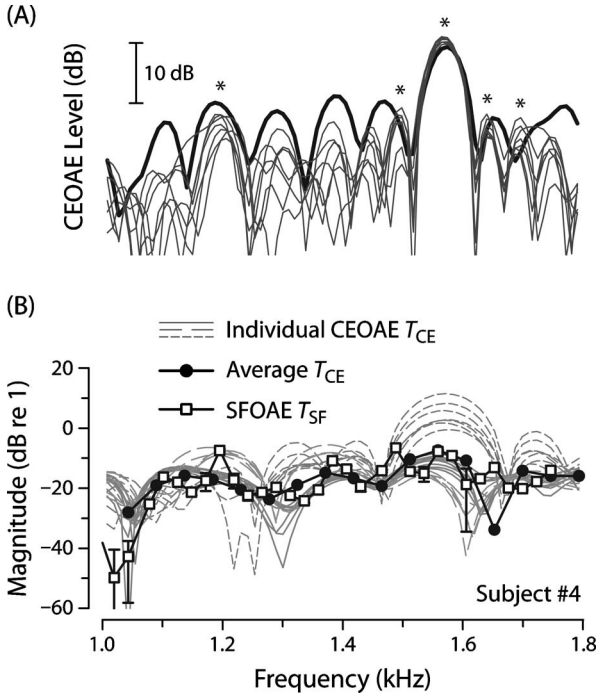

FIG. 5. Averaging out variability due to SSOAEs. Panel (A) shows the results of the SSOAE identification procedure (see Fig. 4) in a subject with strong SSOAEs. The frequencies of significant SSOAEs are marked by asterisks (*). Panel (B) shows the magnitude of individual CEOAE transfer functions (thin gray lines) made at 40 dB pSPL. Gray lines of different styles show transfer functions measured during different days. The solid circles (•) give the magnitude of the complex ensemble average of the multiple measurements. The open squares (□) show the SFOAE transfer function at a comparable stimulus level in the low-level linear regime (20 dB SPL).

jects with long-lasting SSOAEs, the response in the stimulus window is contaminated by responses that have not fully decayed by the time the next stimulus presentation occurs. Contamination by SSOAEs typically creates spurious ripples in the measured stimulus spectrum. To reduce errors in the computation of the transfer function, we estimate the stimulus spectrum at low stimulus levels (i.e., at 40–70 dB pSPL) by appropriately rescaling the stimulus spectrum measured at high levels (i.e., at 80 dB pSPL). This rescaling procedure reduces the error because the relative influence of SSOAE ripples is smallest at high levels.

Second, SSOAEs increase variability in the measurements, as shown in Fig. 5. The gray lines in the bottom panel show individual measurements of $T_{CE}(f;A)$ made during several sessions on two different days. Note the increased variability near SSOAE frequencies, indicated by asterisks in the top panel. The variability presumably reflects an instability in the relative phase with which the stimulus initiates or synchronizes the SSOAE. For example, sometimes the SSOAE seems to add to the CEOAE; at other times it appears to subtract. To reduce this variability in subjects with SSOAE we calculate the (complex) average of measurements made during multiple sessions at each frequency (black circle in the figure) before comparing with measurements of $T_{SF}(f;A)$. We find that matches between $T_{CE}(f;A)$ and $T_{SF}(f;A)$ are generally improved significantly by this ensemble averaging.

IV. COMPARISON OF SFOAE AND CEOAE TRANSFER FUNCTIONS

Figure 6 shows measurements of $T_{CE}(f;A)$ and $T_{SF}(f;A)$ versus frequency in a subject lacking the complications introduced by the existence of SSOAEs. Error bars on the magnitude represent the standard deviation of the mean.[1] The figure demonstrates that $T_{CE}(f;A)$ and $T_{SF}(f;A)$ have qualitatively similar spectral structure, including a rapidly varying phase and magnitude peaks and notches that occur at approximately the same frequencies in both transfer functions. Both transfer functions also share a qualitatively similar dependence on stimulus intensity. At the lowest levels, transfer function magnitudes appear nearly independent of level, consistent with a region of approximate linearity near threshold. At higher intensities, the transfer-function magnitudes generally decrease, consistent with compressive nonlinear growth in emission amplitudes. Although the strong qualitative similarity between the CEOAE and SFOAE transfer functions suggests that clicks and tones evoke emissions via similar mechanisms, definitive conclusions require more careful comparisons.

A. Matching click and tone intensities

At a minimum, the comparisons need to take into account that both CEOAEs and SFOAEs depend on stimulus intensity. Even if the responses are comparable in principle, comparisons made at different effective intensities may amount to comparing apples and oranges. Complicating the situation is the fact that click and tone intensities are conventionally specified in different ways. Whereas pure-tone intensities are measured in SPL, click intensities are measured in peak-equivalent SPL (pSPL), defined as the SPL of a pure tone with the same peak pressure as the click waveform. At what click intensity (in dB pSPL) should one measure CEOAEs in order best to compare them with SFOAEs measured at a given probe level (in dB SPL)? Whether or not this question has a meaningful answer depends on the nature of the nonlinearities involved in OAE generation.

To address this issue we investigated the dependence of CEOAE transfer functions on stimulus intensity and bandwidth. The black symbols in Fig. 7 show the magnitude of $T_{CE}(f_0;A)$ in a narrow frequency range (near $f_0 \cong 1.2$ kHz) measured using clicks of various intensities (pSPL) and bandwidths (data from subject 2, also without SSOAEs). Although $T_{CE}(f_0;A)$ appears nearly independent of intensity and bandwidth at the lowest intensities ($A < 50$ dB pSPL), a systematic dependence on both emerges at higher levels ($A > 70$ dB pSPL). Overall, the trends in the CEOAE data appear well described by a function with the approximate form

$$T_{CE} = T_{CE}(f_0;A_{pk},\mathrm{BW}) \cong T_0 \Big/ \left(1 + \frac{A_{pk}/A_0}{1+\mathrm{BW}/\Delta F}\right)^{\alpha}, \qquad (17)$$

where A_{pk} and BW are, respectively, the peak-equivalent stimulus pressure[2] and equivalent rectangular bandwidth of the stimulus.[3] The corresponding reference values, A_0 and ΔF, have units of pressure and frequency, respectively; the dimensionless constant T_0 sets the overall scale, and the exponent α determines asymptotic growth rates.

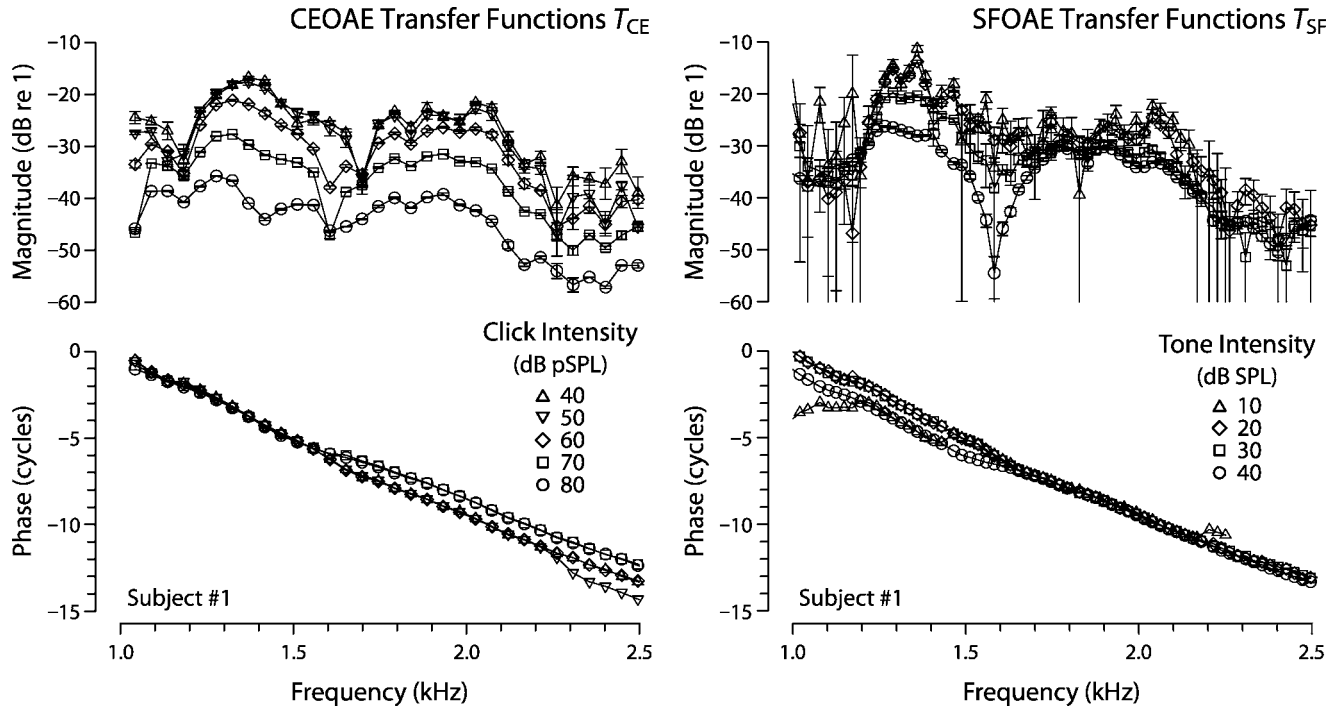

FIG. 6. CEOAE and SFOAE transfer functions.The two columns show the magnitude (top) and phase (bottom) of $T_{CE}(f;A)$ and $T_{SF}(f;A)$, respectively, for a subject without SSOAEs. The symbols identify the stimulus intensity, which ranged from 40 to 80 dB pSPL for clicks and from 10 to 40 dB SPL for tones. Error bars on the magnitude represent the standard error of the mean.

At large stimulus bandwidths and high intensities, the transfer function given by Eq. (17) increases with bandwidth (by about 6α dB/oct at fixed intensity) and decreases with intensity (at a rate approaching $-\alpha$ dB/dB at fixed bandwidth). In the opposite limits, Eq. (17) becomes independent of bandwidth when $\mathrm{BW} \ll \Delta F$ and independent of intensity when $A_{pk} \ll A_0$.

Equation (17) can be simplified for analysis and plotting by rewriting it as

$$T_{CE} = T_{CE}(f_0;A_{eff}) \cong \frac{T_0}{(1 + A_{eff}/A_0)^{\alpha}}, \tag{18}$$

a form obtained by combining the intensity and bandwidth dependence into an effective stimulus pressure, A_{eff}, defined by

$$A_{eff}(A_{pk}, r) \equiv A_{pk}/(1 + r), \tag{19}$$

where $r \equiv \mathrm{BW}/\Delta F$. To the extent that Eq. (17) approximates the data, Eq. (18) predicts that CEOAE transfer functions measured using different stimulus bandwidths can be made to fall along a single curve if plotted against the variable

$$L_{bwc} \equiv L_{pk} - 20 \log (1 + r), \tag{20}$$

where $L_{pk} = 20 \log (A_{pk}/20\ \mu\mathrm{Pa})$ is the click intensity in peak-equivalent SPL (pSPL) and L_{bwc} is what we call "bandwidth-compensated" sound-pressure level (cSPL). The black symbols in Fig. 8 show the CEOAE data replotted versus L_{bwc} with $\Delta F \cong 74$ Hz; as predicted, the data fall approximately along a single curve.

Equation (18) can be used to extrapolate the CEOAE growth function to stimulus bandwidths other than those measured in Fig. 7. As an extreme case, the equation can be used to predict the intensity dependence of SFOAE transfer functions by finding the limit $r \to 0$. This extrapolation regards the SFOAE stimulus as a "click" in which the bandwidth has been reduced so much that the stimulus comprises nothing but a single pure tone. If the extrapolation remains valid across this large reduction in stimulus bandwidth, then

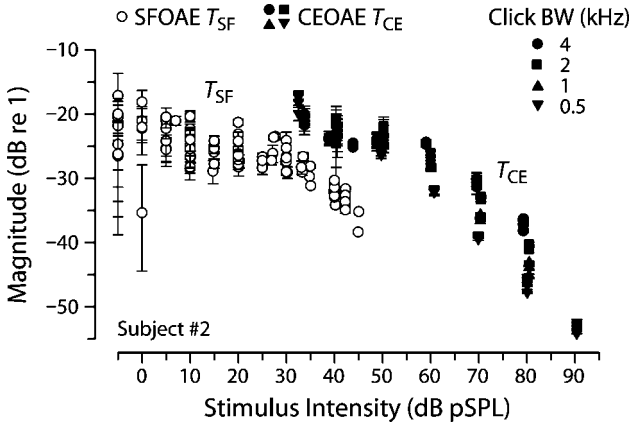

FIG. 7. Intensity and bandwidth dependence of CEOAE and SFOAE transfer functions. Black symbols show the intensity dependence of $T_{CE}(f_0;A)$ measured using clicks with nominal bandwidths of $\{4,2,1,0.5\}$ kHz, corresponding to spectral edges of $\{1-5,1-3,1-2,1-1.5\}$ kHz, respectively. Equivalent rectangular bandwidths are given in footnote 3. For comparison, filled symbols show the intensity dependence of the SFOAE transfer function, $T_{SF}(f_0;A)$. Both CEOAE and SFOAE data were measured at four frequencies in a narrow band near a peak in $T_{CE}(f;A)$ magnitude ($f_0 \in [1.15,1.2]$ kHz). Multiple points at the same intensity and bandwidth represent values at the four different f_0 frequencies in the range. All stimulus intensities are expressed in peak-equivalent SPL (pSPL) or its equivalent (SPL) for the pure tones used to measure SFOAEs.

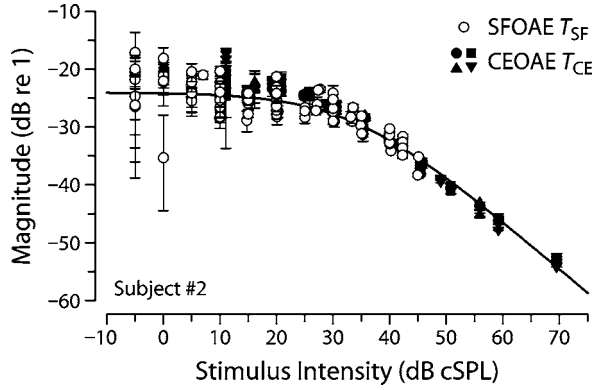

FIG. 8. Unification of CEOAE and SFOAE growth functions. CEOAE transfer functions $T_{CE}(f_0;A)$ from Fig. 7 are shown here with black symbols; SFOAE transfer functions $T_{SF}(f_0;A)$ are shown using open circles. Stimulus levels are expressed in "bandwidth-compensated" sound-pressure level (dB cSPL), defined by Eq. (20). The solid line shows Eq. (17) evaluated with the best-fit parameters given in the text.

$$T_{SF}(f_0;A) \cong \lim_{r\to 0} T_{CE}(f_0;A_{eff}) = \frac{T_0}{(1+A/A_0)^{\alpha}}, \tag{21}$$

where A (the rms pressure of the SFOAE stimulus) and A_{eff} are equivalent for pure tones. In the pure-tone limit, $L_{bwc}\to L_{pk}\to L$, where L is the stimulus sound-pressure level (in dB SPL). The extrapolation thus predicts that SFOAE growth function data lie along the same unified curve found for CEOAEs.

Figure 8 demonstrates that, remarkably, the extrapolation from click to pure-tone bandwidths appears valid: The single growth function shown to characterize the intensity dependence of $T_{CE}(f;A)$ applies also to SFOAEs at the same frequency. Although the measurements of $T_{SF}(f;A)$ are limited to stimulus intensities less than 45 dB SPL, the agreement between the two growth functions (plotted versus cSPL) is excellent throughout the measured range. Although variability in the measurements increases at the lowest intensities (especially for SFOAEs), the common growth function manifests a region of approximately linear growth below about 10–15 dB cSPL. The growth function then gradually transitions into a compressive regime whose slope is somewhat smaller than −1 dB/dB at higher intensities. Fitting Eq. (17) to the pooled CEOAE and SFOAE data yields the parameter estimates $T_0 \cong -24\pm 2$ dB *re* 1, $A_0 \cong 900\pm 400$ μPa (or 33 ± 4 dB SPL), $\Delta F \cong 74\pm 20$ Hz, and $\alpha \cong 0.8\pm 0.08$. (Essentially the same values are obtained when the CEOAE data are fit independently.) The uncertainties represent approximate 95% confidence intervals obtained by bootstrap resampling and are not independent of one another. The solid line in Fig. 8 shows Eq. (17) evaluated using the best-fit parameters.

B. Comparisons at matched stimulus intensities

Figure 8 suggests that CEOAE and SFOAE transfer functions might best be compared at stimulus intensities matched by expressing them in bandwidth-compensated SPL. If the unification achieved in Fig. 8 generalizes across frequency and subject, the transfer functions will then have very similar magnitudes. We test this suggestion in Fig. 9, which compares CEOAE and SFOAE transfer functions across frequency at two matched stimulus intensities in two subjects. The lower of the two intensities (20 dB cSPL) falls within or just above the low-level linear regime; the higher intensity (40 dB cSPL) evokes responses in the region of compressive OAE growth. The two columns shows transfer functions from different subjects, neither of whom had identifiable SSOAEs in the measured frequency range (1–2.4 kHz).

Figure 9 demonstrates that the magnitudes and phases of the transfer functions $T_{CE}(f;A)$ and $T_{SF}(f;A)$ are almost identical at matched intensities. The agreement extends even to the spectral notches, regions where one might expect the responses to be especially sensitive to small changes. Since details of the phase are obscured by the large delay and the phase unwrapping, Fig. 10 replots the transfer-function phases after subtracting out smooth trend lines that capture the secular variation of the phase. The resulting comparisons show that the agreement between $\angle T_{CE}(f;A)$ and $\angle T_{SF}(f;A)$ is generally excellent, even in microstructural detail.

Figure 11 extends the comparison to subjects with SSOAE. In these subjects, CEOAE transfer functions were obtained by averaging across measurement sessions, as described in Sec. III B. Although small differences between $T_{CE}(f;A)$ and $T_{SF}(f;A)$ are found, especially at the lower intensity [e.g., in subject 4, where sharp peaks in $|T_{SF}(f;A)|$ can be seen at SSOAE frequencies], the overall agreement is still excellent.

1. Discrepancies and their possible origin

Despite the compelling overall agreement between $T_{CE}(f;A)$ and $T_{SF}(f;A)$ at matched intensities, the two transfer functions are not always identical to within the measurement error. For example, in subject 2 at 40 dB cSPL (Fig. 9, right-middle panel), the spectral peaks of $T_{CE}(f;A)$ near approximately 1.2 and 1.4 kHz occur at slightly higher frequencies than do the corresponding peaks of $T_{SF}(f;A)$. Do these small differences result from normal session-to-session variability in the emission measurements? Or from methodological differences in the measurement of the two emission types? Or might the discrepancies be more interesting, perhaps reflecting some (presumably subtle) difference in the mechanisms of emission generation?

To characterize the session-to-session variability, we remeasured the two transfer functions multiple times in 1 day (removing, reinserting, and recalibrating the measurement probe each time) and on 9 different days over a 3-month period in a single subject (subject 2). To make the relatively large number of required measurements feasible, we limited the measurements of $T_{SF}(f;A)$ to five frequency points between 1 and 1.2 kHz. In this subject the variability of the measurements across different days was not significantly different from the variability of the measurements made within 1 day.[4] The solid curve with flanking gray region in Fig. 12 shows the mean of 62 measurements of $T_{CE}(f;A)$; the open circles show the mean and deviation of 30 measurements of $T_{SF}(f;A)$. The gray shaded area and error bars extend one standard deviation above and below the mean. The results show that the spectral offset between peaks of $T_{CE}(f;A)$ and

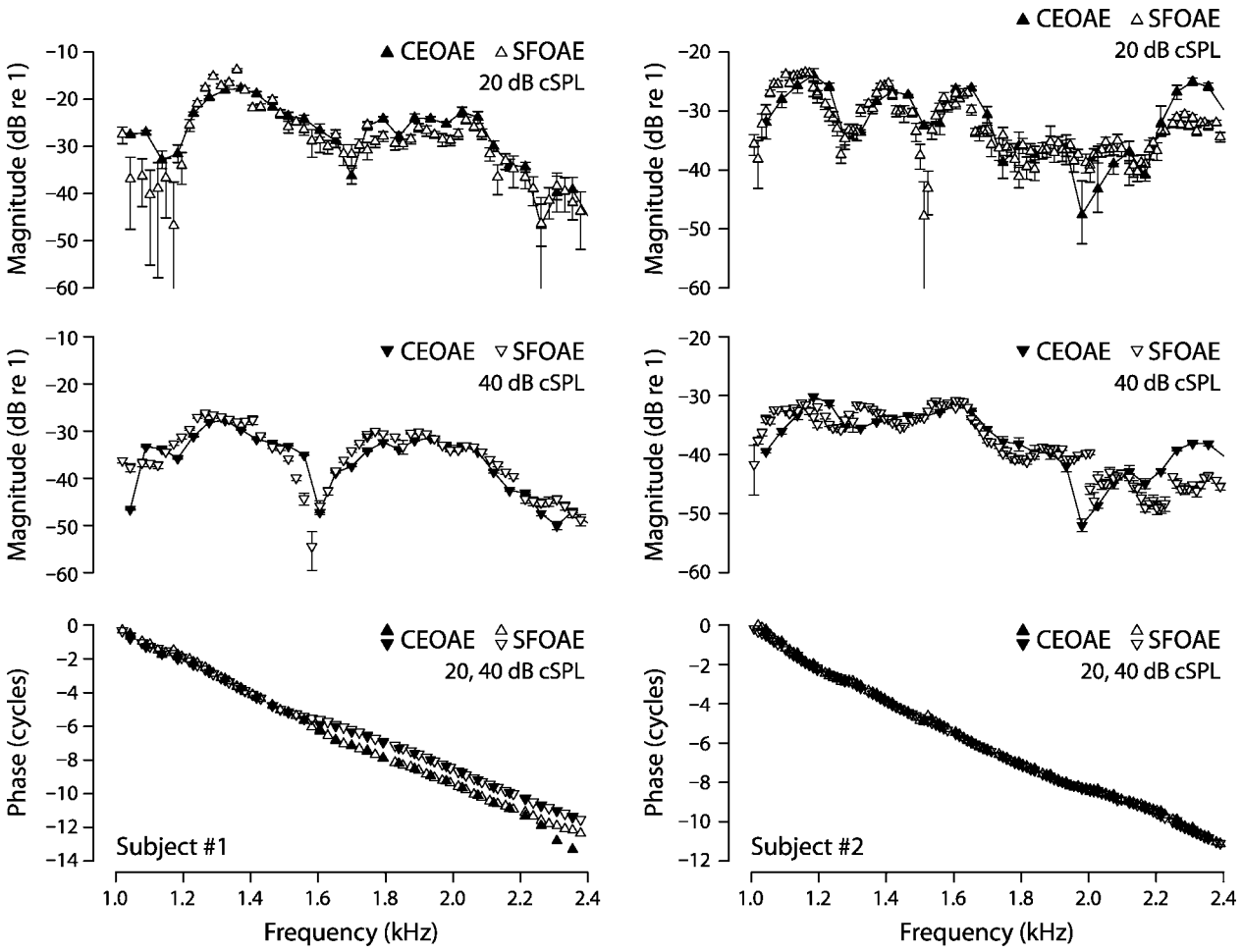

FIG. 9. CEOAE and SFOAE transfer functions at matched intensities in subjects without SSOAEs. Columns show measurements in two different subjects. The top row shows transfer-function magnitudes measured at 20 dB cSPL; the middle row shows magnitudes at 40 dB cSPL. The bottom row shows unwrapped phases at both intensities. In all panels, $T_{CE}(f;A)$ and $T_{SF}(f;A)$ are shown with filled and open symbols, respectively.

$T_{SF}(f;A)$ is larger than the measurement variability, indicating a statistically significant difference between the two measurements.

We doubt, however, that these discrepancies provide evidence for differences in the underlying mechanism of emission generation. Rather, we suspect that the differences are largely if not entirely methodological. As demonstrated in Fig. 2, overlap between the ringing portion of the stimulus and the CEOAE can affect details of the $T_{CE}(f;A)$ spectral shape. The solid curve with hatch marks in Fig. 12 illustrates how a small (ΔO_e=0.5 ms) increase in the offset of the OAE analysis window changes the value of $T_{CE}(f;A)$ extracted from the same measured waveforms. Although larger window offsets reduce interference artifacts due to ringing of the stimulus, they do so at the cost of eliminating short-latency components of the emission. Comparing the values of $T_{CE}(f;A)$ obtained from the same time waveforms using two different window offsets (O_e=5 and 5.5 ms) demonstrates that relatively small changes in the CEOAE measurement paradigm can produce variations in $T_{CE}(f;A)$ comparable to those observed between $T_{CE}(f;A)$ and $T_{SF}(f;A)$. Note, for example, how the peaks in $T_{CE}(f;A)$ near 1.2 and 1.4 kHz shift toward slightly lower frequencies when using the 5.5-ms offset, resulting in an improved match between the $T_{CE}(f;A)$ and $T_{SF}(f;A)$ at these frequencies.

V. DISCUSSION

At low and moderate stimulus intensities human CEOAE and SFOAE input/output transfer functions are nearly identical. When stimulus intensity is measured in bandwidth-compensated SPL (cSPL), we found that CEOAE and SFOAE transfer functions have equivalent growth functions at fixed frequency and equivalent spectral characteristics at fixed intensity. This strong similarity suggests that the OAEs evoked by broad- and narrow-band stimuli (clicks and tones) are generated by the same mechanism.

A. Possible limits of application

Although our conclusions may apply more widely, we summarize below the known limitations of our study. (1) Our comparisons between CEOAEs and SFOAEs were time consuming, and were therefore performed in a relatively small number of subjects (n=4). Nevertheless, since the subjects were selected only for having measurable emissions, and because we found similar results in all, it seems unlikely that the near equivalence we report is merely a statistical fluke. (2) All of our subjects had normal hearing. Although addi-

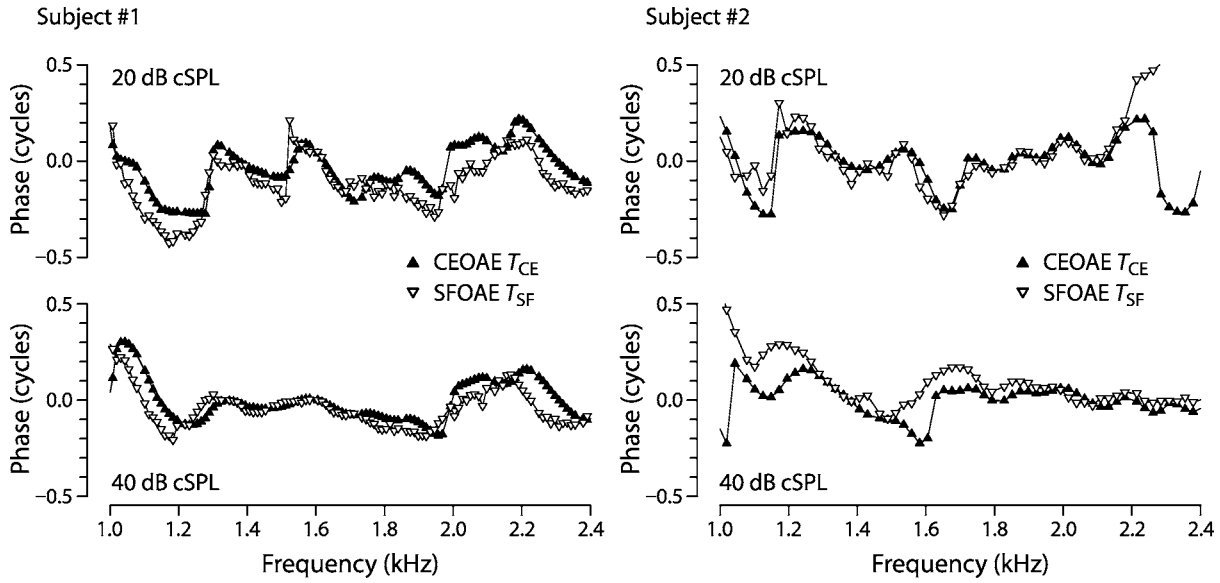

FIG. 10. Details of CEOAE and SFOAE transfer-function phase at matched intensities. The figure shows $\angle T_{CE}(f;A)$ and $\angle T_{SF}(f;A)$ reproduced from Fig. 9 after subtracting out smooth curves that capture the secular variation of the phase. Columns show measurements in two different subjects. The top row shows detrended phases measured at 20 dB cSPL; the bottom row shows detrended phase at 40 dB cSPL. At each level the same trend curves were subtracted from both $\angle T_{CE}(f;A)$ and $\angle T_{SF}(f;A)$.

tional studies are needed to determine whether our findings generalize, we have no reason to suspect that similar conclusions will not apply to impaired ears, so long as their emissions remain measurable. (3) We used low to moderate stimulus levels (35–80 dB pSPL for broadband clicks, 10–40 dB SPL for tones) and cannot rule out the possibility that SFOAE and CEOAE transfer functions differ more significantly at higher stimulus levels. (4) In order to reduce interference between the stimulus and the emission, we used time windows to eliminate CEOAE components arriving earlier than about 5 ms after the stimulus peak. In addition to removing high-frequency components of the response, this windowing may also have removed possible short-latency low-frequency components generated in the base of the cochlea. Although accurate estimates of the magnitudes of these components were compromised by system nonlinearities (see below), measurements in test cavities imply that any such short-latency components must be small relative to the long-latency components. (5) Our comparisons are limited to the frequency range of 1 to 3 kHz. In particular, we did not explore the behavior in more apical regions of the cochlea, where emission mechanisms may differ from those in the base (Shera and Guinan, 2003; Siegel *et al.*, 2005). (6) We did not systematically explore a wide range of stimulus presentation rates (e.g., for the click stimuli) in every subject. Since high-rate clicks are generally much more effective elicitors of efferent activity than the stimuli used to measure SFOAEs (Guinan *et al.*, 2003), we checked for differences related to efferent effects by varying the click-repetition period in two subjects. Although we found no obvious effects of click-repetition period in these subjects, the strength of otoacoustic efferent effects varies from individual to individual, and we may simply have "gotten lucky." It remains possible, even likely, that differences in the strength of efferent feedback elicited by the two stimuli can produce differences in $T_{CE}(f;A)$ and $T_{SF}(f;A)$ in some subjects, at least when $T_{CE}(f;A)$ is measured using high-rate clicks. (7) Finally, our measurements are in humans, a species whose OAE characteristics differ in some respects from those of many laboratory animals (e.g., humans have longer OAE latencies and smaller distortion-source emissions). The near-equivalence we find between $T_{CE}(f;A)$ and $T_{SF}(f;A)$ remains to be examined in other species.

B. Quasilinearity of the transfer functions

If the mechanisms of OAE generation and propagation were completely linear, the near-equivalence we find between CEOAE and SFOAE transfer functions would be entirely expected. The response of a linear system can be equivalently characterized either in the time domain using broadband stimuli (such as the click stimulus used to evoke CEOAEs) or in the frequency domain using narrow-band stimuli (such as the pure tone used to evoke SFOAEs). If the cochlea were a linear system, the principle of superposition would require that transfer functions measured in the time and frequency domains be identical, regardless of the details of emission generation.

Our data support the notion that cochlear responses are nearly linear at levels approaching the threshold of hearing. For example, we find that the transfer functions $T_{CE}(f;A)$ and $T_{SF}(f;A)$ are almost identical and independent of stimulus intensity and bandwidth at low levels. Furthermore, CEOAE transfer functions obtained using the nonlinear-residual method, a method that relies on nonlinear OAE growth to extract the emission, fall into the noise floor at low intensities. These results are consistent with previous OAE measurements (e.g., Zwicker and Schloth, 1984; Shera and Zweig, 1993), including those demonstrating approximate linear superposition among OAEs evoked by various low-level stimuli (Zwicker, 1983; Probst *et al.*, 1986; Xu *et al.*, 1994). Linearity of CEOAE and SFOAE responses at low

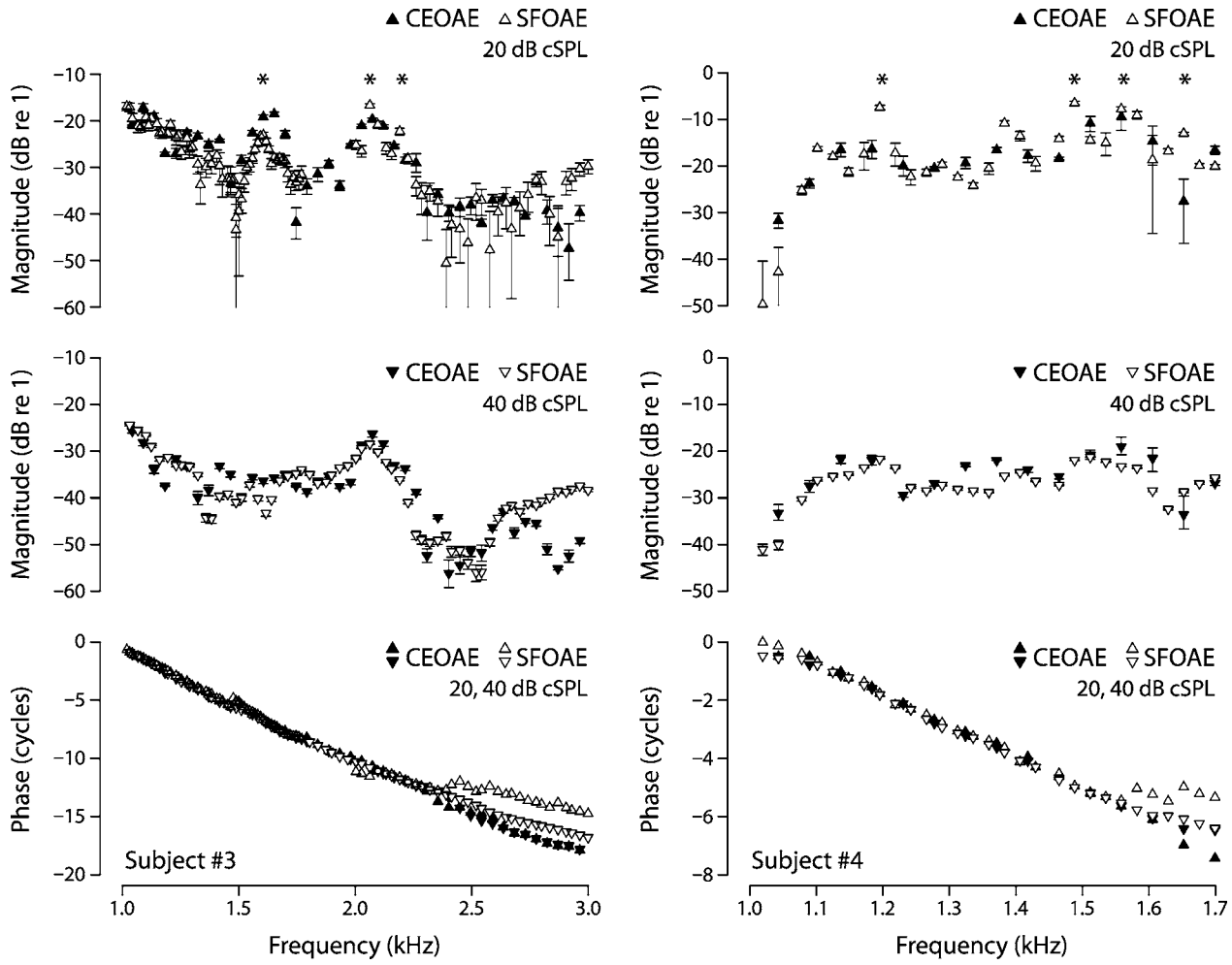

FIG. 11. CEOAE and SFOAE transfer functions at matched intensities in subjects with SSOAE. Columns show measurements in two different subjects. The top row shows transfer-function magnitudes measured at 20 dB cSPL; the middle row shows magnitudes at 40 dB cSPL. The bottom row shows unwrapped phases at both intensities. In all panels, $T_{CE}(f;A)$ and $T_{SF}(f;A)$ are shown with filled and open symbols, respectively. SSOAE frequencies are identified by asterisks (*) in the top panels.

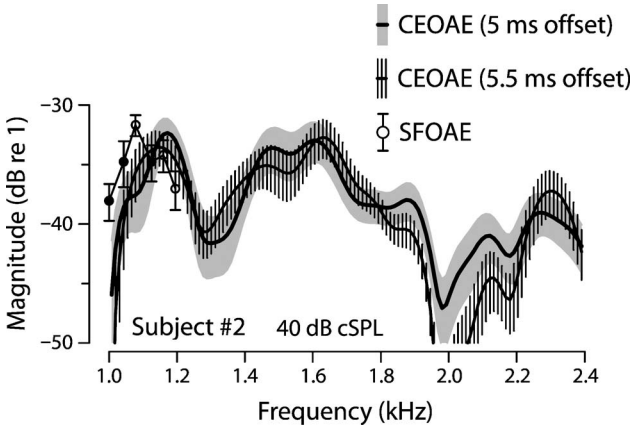

FIG. 12. Session-to-session and methodological variability. Solid curves show mean magnitudes of $T_{CE}(f;A)$ obtained from the same emission measurements (subject 2, 62 measurements distributed over 3 months) using different time offsets for the analysis window (O_e=5 and 5.5 ms). The flanking gray and hatched regions extend one standard deviation above and below the mean. The open circles with error bars show mean magnitudes of $T_{SF}(f;A)$ (30 measurements, same time frame). All stimulus intensities were 40 dB cSPL.

levels is also consistent with basilar-membrane mechanical responses (reviewed in Robles and Ruggero, 2001), which manifest approximate linearity at levels approaching threshold.

Since the operation of the cochlea is certainly nonlinear at intensities not far above threshold, our finding that CEOAE and SFOAE transfer functions continue to match even at moderate levels is more unexpected. The continuing match suggests that as stimulus intensities rise the cochlea emerges gracefully from the low-level linear regime. In particular, the observed match between the spectral structure of CEOAE and SFOAE transfer functions suggests that the reverse-traveling waves that combine to form CEOAEs arise by an approximately linear mechanism (e.g., "scattering") in which interactions among the various frequency components of the stimulus (e.g., intermodulation distortion) play only a secondary role. Although nonlinear interactions such as self- and mutual suppression affect the overall emission magnitudes (e.g., by influencing the gain of the cochlear amplifier), intermodulation distortion does not appear to be primarily responsible for generating the reverse-traveling waves themselves. Our results are consistent with studies of the OAEs evoked by broadband noise (Maat *et al.*, 2000), where

Wiener-kernel analysis indicates that although the overall emission amplitude varies with stimulus intensity, the cochlear response appears approximately linear at each level. Analogous results, including strong if imperfect matches between responses evoked by broad- and narrow-band stimuli, are found in measurements of basilar-membrane motion (e.g., Recio and Rhode, 2000; de Boer and Nuttall, 2002).

Our findings are also consistent with those of Prieve *et al.* (1996), who found that CEOAEs and tone-burst evoked OAEs (TBOAEs) have similar growth functions. They concluded that emissions evoked by the two stimuli share common mechanisms of generation and, in particular, that both are generated by mechanisms acting in independent frequency channels. This conclusion was questioned by Yates and Withnell (1999b), who pointed out that although the tone-burst bandwidths (which generally spanned an octave or more) were narrower than those of the clicks, they were still broad enough to excite the same complex cross-frequency interactions as the click. They therefore argued that the growth functions matched not because of OAE generation via independent frequency channels, but precisely the opposite: because both stimuli produce nonlinear interactions among the different spectral components of the stimulus. Our data do not support Yates and Withnell's suggestion: We measured SFOAEs using continuous narrow-band pure tones (rather than the relatively broadband TBOAEs used by Prieve *et al.*) and still found the reported match between wide- and narrow-band growth functions.

C. Interpreting the unification

The unification between CEOAE and SFOAE growth functions predicted by extrapolation from Eq. (17) and demonstrated experimentally in Fig. 8 was obtained using OAE data from a single individual measured in a narrow range of frequencies. Despite this limitation, the quantitative agreement between $T_{CE}(f;A)$ and $T_{SF}(f;A)$ apparent at matched (bandwidth-compensated) intensities demonstrated in Figs. 9 and 11 suggests that relations similar to Eq. (17) but with frequency- and perhaps subject-dependent parameters (e.g., T_0, A_0, ΔF, α) may apply more widely. [Equation (17) clearly breaks down at frequencies where monotonic emission growth is interrupted by "interference notches;" the frequencies f_0 represented by the data in Fig. 7 were chosen to avoid this behavior.]

Although we caution against overinterpretation of Eq. (17) until its empirical foundation and region of validity are more firmly understood, the frequency scale ΔF merits further comment. Its appearance in the CEOAE growth function presumably reflects an effective integration bandwidth for CEOAE generation, analogous to the equivalent rectangular band (ERB) of an auditory filter. In this regard we note that the best-fit value $\Delta F \cong 74 \pm 20$ Hz obtained from the OAE growth functions at $f_0 \cong 1.2$ kHz is comparable to the auditory-filter ERB [$\mathrm{ERB}(f_0) \cong 90 \pm 10$ Hz] obtained from independent otoacoustic and psychophysical measurements (Shera *et al.*, 2002; Oxenham and Shera, 2003). A more systematic study could determine whether Eq. (17) and the approximate equality of bandwidths observed here hold at other frequencies. Interpreted as an effective integration bandwidth, a nonzero ΔF implies that the "channels" associated with CEOAE generation are not truly independent (e.g., Prieve *et al.* 1996). Evidently, the CEOAE at any given frequency is affected by stimulus energy at nearby frequencies, presumably through suppression.

D. Consistency with the DP-place component of DPOAEs

The match we find between $T_{CE}(f;A)$ and $T_{SF}(f;A)$ is consistent with emission measurements that report strong similarities between CEOAEs and certain DPOAEs (Knight and Kemp, 1999), in particular upper-sideband DPOAEs and lower-sideband DPOAEs measured at f_2/f_1 ratios close to 1. DPOAEs are typically mixtures of emissions originating from at least two different regions of the cochlea, namely the region where the responses to the primaries overlap and the region tuned to the distortion-product frequency (e.g., Kim, 1980; Kemp and Brown, 1983; Gaskill and Brown, 1990; Brown *et al.*, 1996; Engdahl and Kemp, 1996; Brown and Beveridge, 1997; Heitmann *et al.*, 1998; Talmadge *et al.*, 1999; Kalluri and Shera, 2001; Knight and Kemp, 2001). Theory and experiment both indicate that the relative contribution of the components from these two locations varies systematically with stimulus parameters (e.g., Fahey and Allen, 1997; Knight and Kemp, 2001; Shera and Guinan, 2007). In particular, upper-sideband DPOAEs and lower-sideband DPOAEs measured with f_2/f_1 ratios close to 1 are generally dominated by emissions from the distortion-product place, oil whose characteristics are very similar to SFOAEs. Indeed, Kalluri and Shera (2001) showed by direct comparison that the DPOAE component originating from the DP place closely matched the SFOAE evoked at the same frequency. Previous results thus establish that (1) CEOAEs resemble the DP-place component of DPOAEs and (2) the DP-place component of DPOAEs matches SFOAEs. Taken together, these results are consistent with the equivalence reported here between CEOAE and SFOAE transfer functions.

E. Implications for emission mechanisms

Our results contradict two proposed models of CEOAE generation, both of which suggest that CEOAEs originate primarily by nonlinear mechanisms within the cochlea. Nobili and colleagues argue that CEOAEs arise from spatially complex, nonlinear "residual oscillations" of the basilar membrane that trace their origin to spectral irregularities in middle-ear transmission (Nobili, 2000; Nobili *et al.*, 2003a,b). Based on their model simulations, Nobili *et al.* conclude that transient evoked OAEs that occur in the absence of spontaneous emissions result from a "spatial imbalance" in cochlear nonlinearity and amplification caused by rapid frequency variations in forward middle-ear filtering. In this view, CEOAEs result from mechanisms that are both inherently nonlinear and fundamentally different from those responsible for generating SFOAEs. We note, for example, that Nobili *et al.*'s proposed middle-ear filtering mechanism for generating CEOAEs cannot produce SFOAEs at any

level of stimulation: Although CEOAEs are evoked by transient stimuli containing many frequency components, and are therefore potentially sensitive to frequency variations in middle-ear transmission as proposed, SFOAEs are evoked by pure (single-frequency) tones and, *ipso facto*, cannot originate via any mechanism that operates *across* frequency. We show here that the characteristics of CEOAEs and SFOAEs are nearly identical (in ears both with and without SSOAEs), in clear contradiction to Nobili *et al.*'s model predictions.

Our findings also contradict the notion that CEOAEs arise via nonlinear interactions among the frequency components of the stimulus. Based on measurements in guinea pig in which they evoked CEOAEs using high-pass filtered clicks and identified significant OAE energy outside the stimulus passband, Yates and Withnell (1999b) proposed that CEOAEs result primarily from intermodulation distortion within the cochlea. CEOAEs, they suggest, are "predominantly composed of intermodulation distortion energy; each component frequency of a click stimulus presumably interacts with every other component frequency to produce a range of intermodulation distortion products" (Withnell *et al.*, 2000). Our finding that CEOAE and SFOAE transfer functions are almost identical argues against this interpretation, at least in humans.

Although the contribution of nonlinear intermodulation distortion mechanisms to human CEOAEs appears small at low and moderate levels, our use of the windowing technique to measure CEOAEs may have eliminated short-latency distortion components present in the response (e.g., Knight and Kemp, 1999; Withnell and McKinley, 2005). Because of a stimulus artifact due to earphone nonlinearities we were unable to quantify accurately the size of any short-latency physiological component using the nonlinear residual method. Nevertheless we can report that any such short-latency component is small enough to be indistinguishable from the distortion measured in a test cavity of similar impedance. Any short-latency nonlinear component in human ears is therefore small relative to the long-latency linear response. Similar conclusions apply also to human SFOAEs (Shera and Zweig, 1993).

Although the observed equivalence between CEOAEs and SFOAEs contradicts these inherently nonlinear models of CEOAE generation, the equivalence is entirely consistent with predictions of the coherent-reflection model (e.g., Zweig and Shera, 1995; Talmadge *et al.*, 1998; Shera and Guinan, 2007). In this model, OAEs are generated by a process equivalent to wave scattering by preexisting (place-fixed) micromechanical perturbations in the organ of Corti. Not only does the coherent-reflection model predict the empirical equivalence between $T_{\mathrm{CE}}(f;A)$ and $T_{\mathrm{SF}}(f;A)$, the model also predicts the observed spectral characteristics of the transfer functions across frequency (e.g., their slowly varying amplitudes punctuated by sharp notches and their rapidly rotating phases).

Because different stimuli are used to evoke them, CEOAEs and SFOAEs are conventionally classified as different OAE types. Our results establish, however, that at low and moderate stimulus intensities these two OAE "types" are really the same emission evoked in different ways—CEOAEs and SFOAEs are evidently better understood as members of the same emission family. Our findings thus support the mechanism-based classification scheme proposed elsewhere (Shera and Guinan, 1999; Shera, 2004).

ACKNOWLEDGMENTS

We thank Christopher Bergevin, John Guinan, and the anonymous reviewers for helpful comments on the manuscript. This work was supported by Grant No. R01 DC03687 from the NIDCD, National Institutes of Health.

[1]In subjects without SSOAE, the transfer function's standard deviation is computed either from the noise floor or, when possible, by finding the deviation of multiple runs. In subjects with SSOAE, we always made multiple measurements, particularly at low stimulus levels where the SSOAE were likely to have the most influence. Because SOAEs typically have constant magnitude but random phase, complex averaging of multiple runs allows one to partially eliminate the SSOAE. The standard deviation represents the uncertainty of the mean value.

[2]The peak-equivalent pressure of the click is the rms pressure of the pure tone with the same peak pressure.

[3]The equivalent rectangular bandwidths of the click stimuli used to collect the data in Fig. 7 were $\{3.2, 1.8, 1, 0.67\}$ kHz.

[4]At frequencies near a SSOAE, however, the across-day variability can be considerably larger (see Fig. 5), perhaps because the strength or synchronizability of the SSOAE varies from day to day.

Brown, A. M., and Beveridge, H. A. (**1997**). "Two components of acoustic distortion: Differential effects of contralateral sound and aspirin," in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. L. Hecht-Poinar (World Scientific, Singapore), pp. 219–225.

Brown, A. M., Harris, F. P., and Beveridge, H. A. (**1996**). "Two sources of acoustic distortion products from the human cochlea," J. Acoust. Soc. Am. **100**, 3260–3267.

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

de Boer, E., and Nuttall, A. L. (**2002**). "The mechanical waveform of the basilar membrane. IV. Tone and noise stimuli," J. Acoust. Soc. Am. **111**, 979–985.

Engdahl, B., and Kemp, D. T. (**1996**). "The effect of noise exposure on the details of distortion product otoacoustic emissions in humans," J. Acoust. Soc. Am. **99**, 1573–1587.

Fahey, P. F., and Allen, J. B. (**1997**). "Measurement of distortion product phase in the ear canal of the cat," J. Acoust. Soc. Am. **102**, 2880–2891.

Gaskill, S. A., and Brown, A. M. (**1990**). "The behavior of the acoustic distortion product, $2f_1-f_2$, from the human ear and its relation to auditory sensitivity," J. Acoust. Soc. Am. **88**, 821–839.

Gelb, A., and Vander Velde, W. E. (**1968**). *Multiple-Input Describing Functions and Nonlinear System Design* (McGraw Hill, New York).

Guinan, J. J., Backus, B. C., Lilaonitkul, W., and Aharonson, V. (**2003**). "Medial olivo-cochlear efferent reflex in humans: Otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs," J. Assoc. Res. Otolaryngol. **4**, 521–540.

Heitmann, J., Waldman, B., Schnitzler, H. U., Plinkert, P. K., and Zenner, H.-P. (**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., 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.

Kapadia, S., Lutman, M. E., and Palmer, A. R. (**2005**). "Transducer hysteresis contributes to "stimulus artifact" in the measurement of click-evoked otoacoustic emissions," J. Acoust. Soc. Am. **118**, 620–622.

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

Kemp, D. T., and Brown, A. M. (**1983**). "An integrated view of cochlear mechanical nonlinearities observable from the ear canal," in *Mechanics of*

Hearing, edited by E. de Boer and M. A. Viergever (Martinus Nijhoff, The Hague), pp. 75–82.
Kim, D. O. (**1980**). "Cochlear mechanics: Implications of electrophysiological and acoustical observations," Hear. Res. **2**, 297–317.
Knight, R. D., and Kemp, D. T. (**1999**). "Relationships between DPOAE and TEOAE amplitude and phase characteristics," J. Acoust. Soc. Am. **106**, 1420–1435.
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.
Krylov, N. M., and Bogolyubov, N. N. (**1947**). *Introduction to Nonlinear Mechanics* (Princeton U. P., Princeton).
Lapsley-Miller, J. A., Boege, P., Marshall, L., and Jeng, P. S. (**2004a**). "Transient-evoked otoacoustic emissions: Preliminary results for validity of TEOAEs implemented on Mimosa Acoustics T2K Measurement System v3.1.3," Technical Report No. 1232, Naval Submarine Medical Research Lab, Groton, CT.
Lapsley-Miller, J. A., Boege, P., Marshall, L., Shera, C. A., and Jeng, P. S. (**2004b**). "Stimulus-frequency otoacoustic emissions: Validity and reliability of SFOAEs implemented on Mimosa Acoustics SFOAE Measurement System v2.1.18," Technical Report No. 1231, Naval Submarine Medical Research Lab, Groton, CT.
Maat, B., Wit, H., and van Dijk, P. (**2000**). "Noise-evoked otoacoustic emissions in humans," J. Acoust. Soc. Am. **108**, 2272–2280.
Nobili, R. (**2000**). "Otoacoustic emissions simulated by a realistic cochlear model," in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 402–408.
Nobili, R., Vetešník, A., Turicchia, L., and Mammano, F. (**2003a**). "Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model," J. Assoc. Res. Otolaryngol. **4**, 478–494.
Nobili, R., Vetešník, A., Turicchia, L., and Mammano, F. (**2003b**). "Otoacoustic emissions simulated in the time domain by a hydrodynamic model of the human cochlea," in *Biophysics of the Cochlea: From Molecules to Models*, edited by A. W. Gummer (World Scientific, Singapore), pp. 524–530.
Oxenham, A. J., and Shera, C. A. (**2003**). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," J. Assoc. Res. Otolaryngol. **4**, 541–554.
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. (**1986**). "Spontaneous, click-, and toneburst-evoked otoacoustic emissions from normal ears," Hear. Res. **21**, 261–275.
Recio, A., and Rhode, W. S. (**2000**). "Basilar membrane responses to broadband stimuli," J. Acoust. Soc. Am. **5**, 108.
Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.
Shera, C. A. (**2004**). "Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions," Ear Hear. **25**, 86–97.
Shera, C. A., and 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.
Shera, C. A., and 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., and Guinan, J. J. (**2007**). "Mechanisms of mammalian otoacoustic emission," in *Active Processes and Otoacoustic Emissions*, edited by G. Manley, B. L. Lonsbury-Martin, A. N. Popper, and R. R. Fay, in press (Springer-Verlag, New York).
Shera, C. A., and Zweig, G. (**1993**). "Noninvasive measurement of the cochlear traveling-wave ratio," J. Acoust. Soc. Am. **93**, 3333–3352.
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.
Siegel, J. H., Cerka, A. J., Recio-Spinoso, A., Temchin, A. N., van Dijk, P., and Rugero, M. A. (**2005**). "Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering," J. Acoust. Soc. Am. **118**, 2434–2443.
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.
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., 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.
Withnell, R. H., Yates, G. K., and Kirk, D. L. (**2000**). "Changes to low-frequency components of the TEOAE following acoustic trauma to the base of the cochlea," Hear. Res. **139**, 1–12.
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. (**1999a**). "Reply to "Comment on 'Enhancement of the transient-evoked otoacoustic emission produced by the addition of a pure tone in the guinea pig'" [J. Acoust. Soc. Am. 105, 919–921 (1999)]," J. Acoust. Soc. Am. **105**, 922–924.
Yates, G. K., and Withnell, R. H. (**1999b**). "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.
Zwicker, E. (**1983**). "Delayed evoked oto-acoustic emissions and their suppression by Gaussian-shaped pressure impulses," Hear. Res. **11**, 359–371.
Zwicker, E., and Schloth, E. (**1984**). "Interrelation of different oto-acoustic emissions," J. Acoust. Soc. Am. **75**, 1148–1154.