Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation

Radha Kalluri
Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114 and Speech and Hearing Sciences Program, Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts 02139

Christopher A. Shera
Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114, Speech and Hearing Sciences Program, Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts 02139, and Department of Otology and Laryngology, Harvard Medical School, Boston, Massachusetts 02115

(Received 5 July 2000; revised 6 October 2000; accepted 27 October 2000)

This paper tests key predictions of the ''two-mechanism model'' for the generation of distortion-product otoacoustic emissions (DPOAEs). The two-mechanism model asserts that lower-sideband DPOAEs constitute a mixture of emissions arising not simply from two distinct cochlear locations (as is now well established) but, more importantly, by two fundamentally different mechanisms: nonlinear distortion induced by the traveling wave and linear coherent reflection off pre-existing micromechanical impedance perturbations. The model predicts that (1) DPOAEs evoked by frequency-scaled stimuli (e.g., at fixed f_2/f_1) can be unmixed into putative distortion- and reflection-source components with the frequency dependence of their phases consistent with the presumed mechanisms of generation; (2) The putative reflection-source component of the total DPOAE closely matches the reflection-source emission (e.g., low level stimulus-frequency emission) measured at the same frequency under similar conditions. These predictions were tested by unmixing DPOAEs into components using two completely different methods: (a) selective suppression of the putative reflection source using a third tone near the distortion-product frequency and (b) spectral smoothing (or, equivalently, time-domain windowing). Although the two methods unmix in very different ways, they yield similar DPOAE components. The properties of the two DPOAE components are consistent with the predictions of the two-mechanism model. [DOI: 10.1121/1.1334597]

PACS numbers: 43.64.Bt, 43.64.Jb, 43.64.Yp [BLM]

I. INTRODUCTION

Mammalian otoacoustic emissions (OAEs) have generally been regarded as originating through nonlinear electromechanical distortion (e.g., Kemp, 1978, 1997, 1998; Probst *et al.,* 1991; Allen and Neely, 1992; Allen and Lonsbury-Martin, 1993; Patuzzi, 1996). Shera and Guinan (1999), however, argue that OAEs arise by at least two fundamentally different mechanisms within the cochlea. These differences in mechanism, they suggest, can profitably be used to define an ''OAE family tree.'' The mechanism-based taxonomy groups emissions into two basic types: distortion-source emissions, which arise by nonlinear distortion induced by the traveling wave, and reflection-source emissions, which arise via linear reflection off pre-existing micromechanical impedance perturbations (Shera and Zweig, 1993b; Zweig and Shera, 1995). This distinction between distortion- and reflection-source emissions differs from the ''wave-'' and ''place-fixed'' dichotomy maintained by Kemp and Brown (1983) in that the latter was introduced and developed within an integrated framework that views all OAEs as manifestations of cochlear mechanical nonlinearity. The mechanism-based taxonomy, by contrast, emphasizes the fundamental differences between linear (reflection-source) and nonlinear (distortion-source) emission mechanisms.

The analysis underlying the taxonomy predicts that the two types of OAEs mix to form the evoked emissions measured in the ear canal. In any given measurement, the different emission types contribute in degrees dependent on species, stimulus parameters, and cochlear state. As an example of the process, Shera and Guinan suggest that the generation of lower-sideband distortion-product otoacoustic emissions (DPOAEs) can be understood in terms of the mixing of the two OAE types. Much of DPOAE fine structure apparently arises through the interference of emissions originating from two distinct cochlear locations (e.g., Kim, 1980; Kemp and Brown, 1983). Although the ''two-place model'' for DPOAEs now appears well established (e.g., Gaskill and Brown, 1990; Brown *et al.*, 1996; Engdahl and Kemp, 1996; Brown and Beveridge, 1997; Heitmann *et al.*, 1998; Fahey and Allen, 1997; Siegel *et al.*, 1998), the taxonomy identifies the two interfering emission components as arising not simply from two distinct locations, but, more importantly, via two different mechanisms.

The proposed generation process is illustrated in Fig. 1.

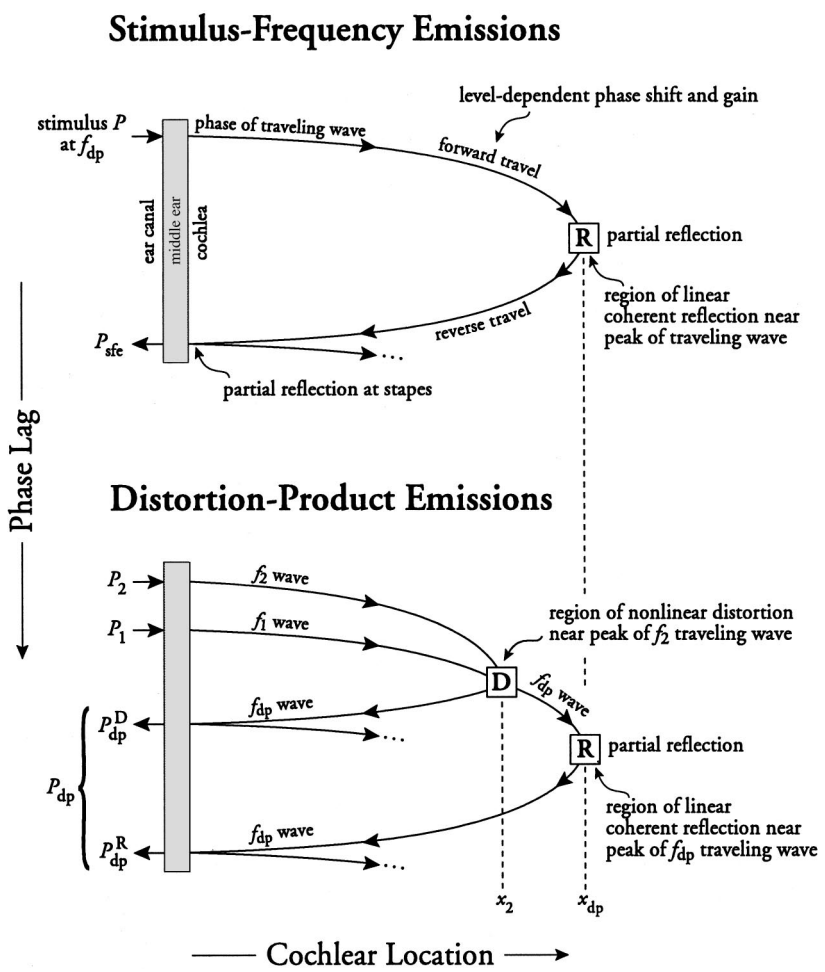

FIG. 1. Schematic diagram of the two-mechanism model. The figure illustrates the generation of SFOAEs (top) and DPOAEs (bottom) at low sound levels. Each panel shows phase lag relative to stimulus phase (lag increasing downward) of forward- and backward-traveling waves vs cochlear location. At low stimulus levels, SFOAEs (P_{sfe}) result from coherent reflection (**R**) in the region near the peak of the traveling-wave envelope. For DPOAEs, the primary traveling waves produce a region of nonlinear distortion (**D**), located near the peak of the f_2 wave (at x_2), where nonlinear distortion generates traveling waves at the frequency f_{dp} that travel in both directions (shown here for the case $f_{dp}=2f_1-f_2$, where f_{dp} equals the SFOAE frequency shown in the top panel). The backward-traveling wave propagates to the ear canal (where it appears as the distortion-source emission, P_{dp}^D). The forward-traveling wave propagates to its characteristic place (at x_{dp}), where it undergoes partial reflection (**R**) near the peak of its wave envelope, generating a second backward-traveling wave that propagates to the ear canal (the reflection-source emission, P_{dp}^R). The two types of emission combine to produce the DPOAE measured in the ear canal ($P_{dp}=P_{dp}^D+P_{dp}^R$). For simplicity, phase shifts due to propagation through the middle ear and reflection by the stapes are not shown. Adapted, with permission, from Shera and Guinan (1999).

The primary traveling waves, at frequencies f_1 and f_2 (with $f_2>f_1$), interact to produce a region of nonlinear distortion (**D**), located near the peak of the f_2 wave, which creates energy at distortion-product frequencies. In particular, traveling waves at the frequency $f_{dp}=2f_1-f_2$ are generated that travel in both directions. The backward-traveling wave propagates to the ear canal, where it appears as a distortion source emission. The forward-traveling wave propagates to its characteristic place, where it undergoes partial reflection (**R**) near the peak of its wave envelope, generating a second backward-traveling wave that propagates to the ear canal (a reflection-source emission). The two types of emission mix in the ear canal.[1]

The proposed model thus predicts that the two components originate not simply from two different regions of the cochlea but—more significantly—by two fundamentally different mechanisms. Similar predictions emerge from recent modeling studies (e.g., Talmadge *et al.*, 1998, 1999; Mauermann *et al.*, 1999a, b). Based on nonlinear cochlear models that meet the requirements detailed by the theory of coherent reflection filtering for the generation of realistic reflection emissions (Shera and Zweig, 1993b; Zweig and Shera, 1995),[2] these studies incorporate both classes of emission-generating mechanisms (i.e., nonlinear distortion and linear coherent reflection). The primary goal of the experiments reported here was to test this *two-mechanism model* for DPOAE generation.

According to the analysis underlying the taxonomy, distortion- and reflection-source emissions manifest very different frequency dependencies in their phase. In a nutshell, the argument runs roughly as follows:

Distortion-source OAEs: Nonlinear distortion depends upon the interaction between the two primary traveling waves. When produced using frequency-scaled stimuli, the combined excitation pattern of the primary traveling waves simply translates along the cochlear partition as the stimulus frequencies are varied. This approximate translation invariance (or ''shift similarity'') follows from local scaling symmetry and the logarithmic form of the cochlear frequency-position map. Approximate shift similarity ensures that the amplitudes and phases of the primary waves—and hence any nonlinear interactions between them—remain nearly invariant in a coordinate system that moves with the spatial envelope of the f_2 traveling wave as the frequencies are swept (Fig. 2, left). OAEs generated by frequency-scaled nonlinear distortion therefore manifest a nearly constant phase.

Reflection-source OAEs: According to the theory of coherent reflection filtering (Zweig and Shera, 1995), reflection-source emissions are generated when a forward-traveling wave reflects off ''random'' perturbations in the mechanics of the cochlea. The phase of each scattered wavelet depends on the phase of the forward-traveling wave at the location of scattering. Since the micromechanical impedance perturbations are fixed in space (unlike sources of nonlinear

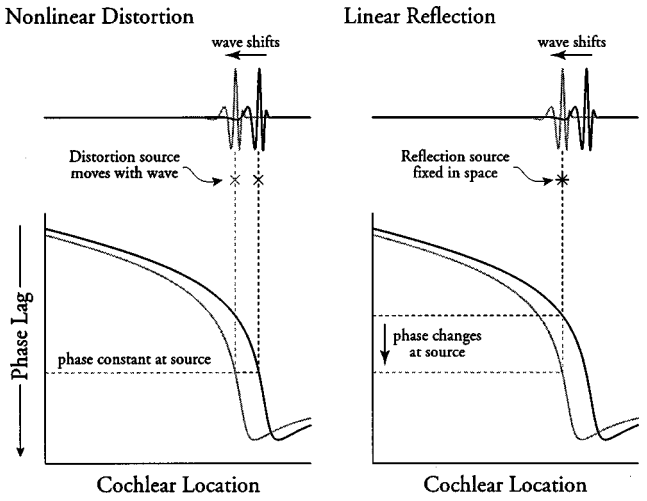

FIG. 2. Schematic diagram illustrating the consequences of scaling for the phase of distortion- and reflection-source emissions. The left-hand panel shows a snapshot of the f_2 traveling wave at two different frequencies (top) and the corresponding phase lags (bottom) vs cochlear location. The two frequencies are denoted by black and gray lines, respectively. For simplicity, the f_1 traveling waves are not shown. Distortion sources result from nonlinear interaction between the primary traveling waves. The sources illustrated here ($\times$) are idealized as points at the peak of the f_2 traveling wave. When the frequency ratio f_2/f_1 is held fixed during the measurement sweep, the primary traveling waves (and thus the resulting distortion sources) simply shift along the cochlear partition, maintaining a nearly constant relative phase relationship as the stimulus frequencies are varied. Note, for example, that the phases of the primary traveling waves at the distortion source remain constant as the frequency is increased and the wave pattern shifts ($\leftarrow$) along the partition. As a result, the phases of all resulting distortion products are essentially independent of frequency. The right-hand panel shows a similar diagram for a reflection source (e.g., a perturbation in the mechanics of the cochlea). Since the perturbation (*) is fixed in space the phase of the wave scattered by the perturbation changes considerably ($\downarrow$) as the stimulus frequency is varied. Consequently, the phases of OAEs generated by linear reflection vary rapidly with frequency.

distortion, which move with the excitation pattern as the frequency changes), the phase of the incident wave at each perturbation changes as the frequency of the forward-traveling wave is varied (Fig. 2, right). Consequently, OAEs generated by linear reflection manifest a phase that rotates rapidly with frequency.

In this paper we apply this reasoning to test the two principal predictions of the two-mechanism model, as suggested by the taxonomy and framework presented in Fig. 1. Specifically, we test the predictions that

(1) The total distortion-product emission, P_{dp}, represents the sum of distortion- and reflection-source components, P_{dp}^{D} and P_{dp}^{R}

$$P_{dp}=P_{dp}^{D}+P_{dp}^{R}, \tag{1}$$

where the components P_{dp}^{D} and P_{dp}^{R} manifest frequency dependencies in their phase consistent with their presumed mechanisms of generation. Specifically, the model predicts that when P_{dp} is evoked using frequency-scaled stimuli (e.g., with the ratio f_2/f_1 fixed), the phase of P_{dp}^{D} should be essentially independent of frequency, whereas the phase of P_{dp}^{R} should rotate rapidly.

(2) The putative reflection-source component, P_{dp}^{R}, of the total DPOAE closely matches the reflection emission measured at the same frequency under similar conditions. According to the taxonomy, stimulus-frequency emissions (SFOAEs) evoked at low stimulus levels are nearly pure reflection emissions (see Fig. 1). We thus test the prediction that

$$P_{dp}^{R}\approx P_{sfe}, \tag{2}$$

where P_{sfe} is the SFOAE at the same frequency. Once stimulus parameters have been adjusted to yield comparable overall emission levels, the predicted match includes the frequency dependence of both the amplitude (or spectral shape) and the phase.[3]

Testing these predictions requires a technique for unmixing the total DPOAE into putative distortion- and reflection-source components. Initially, we adopt an experimental approach based on selective suppression that exploits the spatial separation of the presumed distortion- and reflection-source regions within the cochlea. To explore the sensitivity of our results to variations in the methodology of unmixing, we compare our results obtained using suppression to an alternative unmixing procedure based on spectral smoothing or time-domain windowing. A preliminary account of this work has been presented elsewhere (Kalluri and Shera, 2000).

II. UNMIXING VIA SELECTIVE SUPPRESSION

Reference to Fig. 1 suggests that one can separate the two components, P_{dp}^{D} and P_{dp}^{R}, of the total DPOAE pressure if the reflection-source emission originating from the **R** region can be eliminated. The unmixing procedure would then be to (1) measure the total emission, P_{dp}, using frequency-scaled stimuli; (2) eliminate the **R** component and remeasure the DPOAE to obtain the pure distortion-source component, P_{dp}^{D}; and (3) compute the reflection-source component, P_{dp}^{R}, by subtraction, $P_{dp}^{R}=P_{dp}-P_{dp}^{D}$.

The spatial separation of the two source regions within the cochlea suggests trying to eliminate the **R** component by introducing a third, suppressor tone at a nearby frequency. The suppressor would act by reducing the amplitude of the wavelets incident upon and/or scattered back from the **R** region. Suppression techniques for separating OAE sources originating at different spatial locations in the cochlea were pioneered by Kemp and Brown (1983) and later refined by others (e.g., Heitmann *et al.*, 1998; Siegel *et al.*, 1998). The selective suppression strategy for unmixing yields the following estimates of P_{dp} and its components:

$$P_{dp}=P_{ec}(f_{dp}) \quad \text{(measured at fixed } f_2/f_1\text{)}; \tag{3}$$

$$P_{dp}^{D}\approx P_{ec}(f_{dp})\big|_{\text{with suppressor}} \quad \text{(with suppressor at } f_s\approx f_{dp}\text{)}; \tag{4}$$

$$P_{dp}^{R}\approx P_{dp}-P_{dp}^{D}. \tag{5}$$

In these expressions, $P_{ec}(f)$ denotes the complex ear-canal pressure at frequency f resulting from stimulation at primary frequencies f_1 and f_2.

A. Measurement methods

We measured emissions from one ear of each of four normal hearing humans. Treatment of human subjects was in accordance with protocols approved by the Human-Studies Committee at the Massachusetts Eye and Ear Infirmary. All measurements were performed with subjects reclining comfortably in a sound-proofed, vibration-isolated chamber (Ver *et al.*, 1975). Stimulus waveforms were generated and responses acquired and averaged digitally using a custom-built data-acquisition system. Acoustic signals were transduced using a Etymotic Research ER-10c DPOAE probe system supplemented with an ER-3A earphone whose sound-delivery tube was threaded through the ER-10c foam eartip. *In situ* earphone calibrations were performed at regular intervals throughout all measurement sessions. The calibrations were used to guarantee that the stimulus tones had constant level and zero starting phase in the ear canal at all frequencies. Real-time artifact rejection was implemented by comparing the time waveforms in successive data buffers before adding them to the averaged responses. In these and other respects, the methods and equipment used to obtain both SFOAEs and DPOAEs are generally similar to those described elsewhere (Shera and Guinan, 1999). We briefly summarize relevant differences here and provide detailed descriptions in the Appendix.

1. Measurement of DPOAEs

We measured distortion-product emissions at the frequency $2f_1-f_2$ using frequency-scaled stimuli (i.e., using frequency sweeps performed with the primary-frequency ratio, f_2/f_1, held constant). The measurements reported here were obtained using primary levels of $\{L_1,L_2\}=\{60,45\}$ dB SPL at the frequency ratio $f_2/f_1=1.2$. To ensure that our ability to maintain a constant f_2/f_1 ratio during the sweep was not systematically compromised by the frequency quantization imposed by digital stimulus generation, we modified our data-acquisition system to allow the sampling frequency to vary between measurement points. This flexibility enabled us to choose f_1 and f_2 so that the ratio f_2/f_1 varied by less than a thousandth of a percent between measurements (at our typical frequency spacing of about 15 Hz). The resulting sampling frequencies varied by less than 3% about the nominal value (59.94 kHz).

To allow any multiple reflections that might be present within the cochlea to settle into an approximately steady-state response, we measured $P_{\rm dp}$ and $P_{\rm dp}|_{\rm suppressed}$ over time intervals ($\approx$136 ms) much longer than the estimated round-trip travel time for cochlear waves ($\approx$10–15 ms). To guard against possible systematic variations in emission amplitude over time that might invalidate the unmixing procedure (e.g., due to efferent feedback or to changes in earphone calibration caused by subject movement or by temperature variations), we interleaved measurements of $P_{\rm dp}$ and $P_{\rm dp}|_{\rm suppressed}$ in time and averaged multiple repetitions (typically $n=64$). We set the suppressor frequency approximately 44 Hz below the distortion-product frequency (e.g., Siegel *et al.*, 1998; Dreisbach and Siegel, 1999). The suppressor level was adjusted (separately for each subject) to minimize the DPOAE fine structure while leaving the mean DPOAE amplitude (as averaged over several fine-structure periods) largely unchanged (cf. Heitmann *et al.*, 1998). The suppressor level chosen in this way was generally in the range 50–55 dB SPL.

2. Measurement of SFOAEs

We measured stimulus-frequency emissions using the suppression method (e.g., Guinan, 1990; Kemp *et al.*, 1990; Brass and Kemp, 1991, 1993; Souter, 1995; Shera and Guinan, 1999). In this method, the emission is obtained as the complex (or vector) difference between the ear-canal pressure at the probe frequency measured first with the probe tone alone and then in the presence of a stronger suppressor tone at a nearby frequency. Thus, the SFOAE pressure at the probe frequency, $P_{\rm sfe}(f_{\rm p})$, is defined as

$$P_{\rm sfe}(f_{\rm p}) \equiv P_{\rm ec}(f_{\rm p}) - P_{\rm ec}(f_{\rm p})|_{\text{with suppressor at } f_{\rm s}\approx f_{\rm p}}. \tag{6}$$

In the measurements reported here, the suppressor frequency was approximately 44 Hz below the probe ($f_{\rm s}\approx f_{\rm p}-44\,{\rm Hz}$). To prevent contamination from the considerable cross talk between output channels of the ER-10c, we generated the suppressor tone using a separate ER-3A earphone whose sound-delivery tube was threaded through the foam eartip. Unless otherwise noted, the probe and suppressor levels, $L_{\rm p}$ and $L_{\rm s}$, were 40 and 55 dB SPL, respectively (Shera and Guinan, 1999). Exploratory measurements at other nearby probe levels indicate that the spectral shape and phase of $P_{\rm sfe}$ are not strong functions of intensity at these levels. We found that probe levels of 40 dB SPL gave emission levels generally comparable to those of $P_{\rm dp}^{\rm R}$, especially after introduction of the primary-mimicking tone described below. As with the DPOAE measurements, we interleaved measurements of $P_{\rm ec}(f_{\rm p})$ and $P_{\rm ec}(f_{\rm p})|_{\rm suppressed}$ in time to minimize artifacts that might contaminate the difference. The measurement frequency resolution, approximately 15 Hz between points, was always sufficient to prevent ambiguities in phase unwrapping.

In some experiments, we measured SFOAEs in the presence of an additional continuous tone. The idea was to measure $P_{\rm sfe}$ under conditions matching as closely as possible those present during the measurement of $P_{\rm dp}^{\rm R}$. Thus, we introduced the additional tone at a frequency and level corresponding to the f_1 primary used in the measurement of DPOAEs. In terms of the probe frequency, the frequency of this additional tone (the ''f_1-primary mimicker'') was therefore given by $f_1=f_{\rm p}/(2-r)$, where r denotes the f_2/f_1 ratio we wished to mimic. We denote the SFOAE measured in the presence of the f_1-primary mimicking tone by $P_{\rm sfe}^1$:

$$P_{\rm sfe}^1 \equiv P_{\rm sfe}|_{\text{with } f_1 \text{ mimicker}}. \tag{7}$$

B. Results: Unmixing via selective suppression

1. P_{dp} and its components, P_{dp}^{D} and P_{dp}^{R}

Typical measurements of the total DPOAE and its components estimated using suppression are shown in Fig. 3. To illustrate the variation across subjects, we show results for three of our four subjects (those for whom the most data are available); similar results were obtained in the fourth subject.

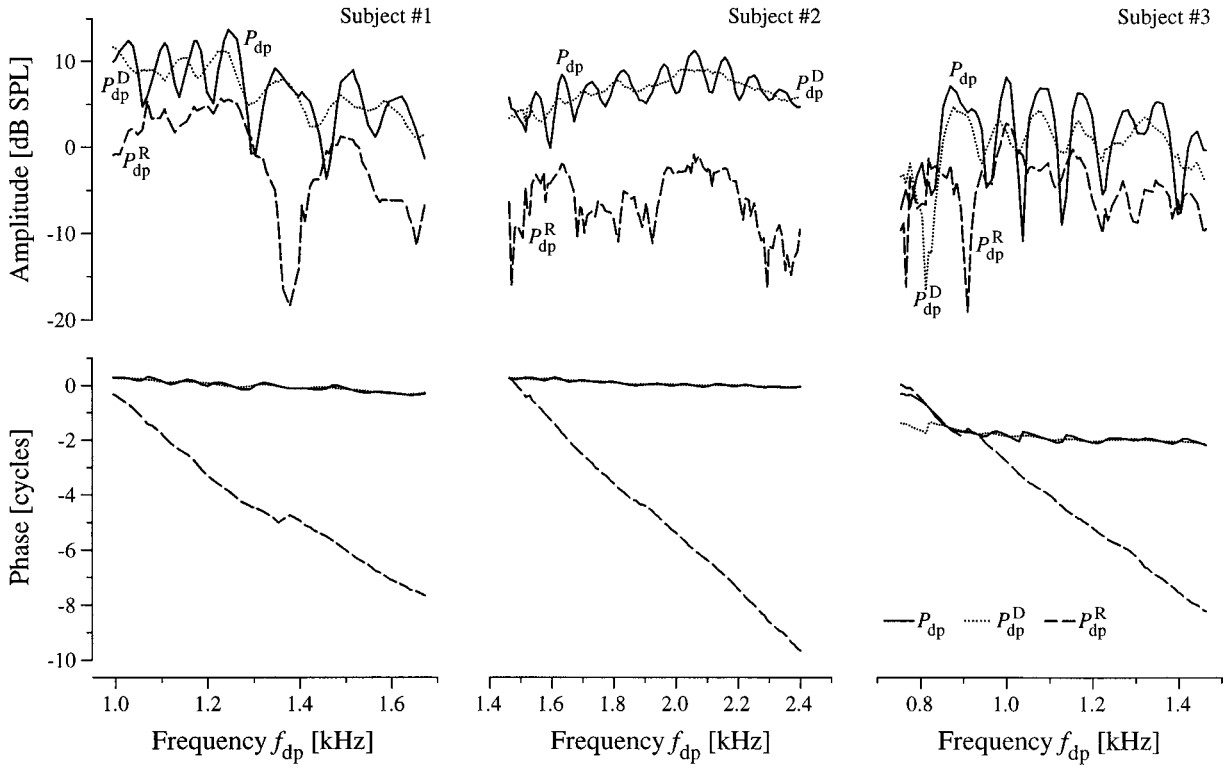

FIG. 3. The DPOAE P_{dp} and its estimated distortion- and reflection-source components, P_{dp}^{D} and P_{dp}^{R}, obtained using suppression. The figure shows typical measurements of the amplitude (top) and phase (bottom) of the $2f_1-f_2$ DPOAE and its components measured using a frequency-scaled stimulus (i.e., the primary frequencies f_1 and f_2 were swept with their ratio held constant at the value $f_2/f_1=1.2$). Left to right, the panels show data from subjects #1, #2, and #3, respectively; similar results were obtained in the fourth subject. In each case, the total DPOAE (solid line) was unmixed using a suppressor tone near the distortion-product frequency, f_{dp}. Although the phases of the estimated P_{dp}^{D} components (dotted lines) vary less than a period, the phases of the estimated P_{dp}^{R} components (dashed lines) fall through many cycles over the same frequency range, in agreement with Prediction #1. The measurement noise floor was approximately -25 dB SPL and the frequency resolution was always sufficient to prevent ambiguities in phase unwrapping. Stimulus levels for subject #1: $\{L_1,L_2,L_s\}=\{60,45,50\}$ dB SPL. Stimulus levels for subject #2: $\{L_1,L_2,L_s\}=\{60,45,50\}$ dB SPL. Stimulus levels for subject #3: $\{L_1,L_2,L_s\}=\{60,45,55\}$ dB SPL.

In each case, the putative distortion-source component, P_{dp}^{D}, is essentially a smoother version of the total DPOAE in which much of the quasiperiodic fine structure apparent in both the amplitude and phase of P_{dp} has been eliminated. In agreement with Prediction #1, the phase of P_{dp}^{D} is nearly constant, varying by less than a period. By contrast, the phase of the reflection-source component, P_{dp}^{R}, falls through many cycles (typically eight or more) over the same frequency range. These different frequency dependencies imply generation by fundamentally different mechanisms: The nearly constant phase of P_{dp}^{D} is consistent with generation by frequency-scaled nonlinear distortion and the rapidly rotating phase of P_{dp}^{R} with generation by linear coherent reflection (Shera and Guinan, 1999).

The fine-structure manifest in the total DPOAE arises because of alternating constructive and destructive interference between the two components, P_{dp}^{D} and P_{dp}^{R}, caused by the systematic rotation of their relative phase, a consequence of the very different slopes of their phase versus frequency functions. Thus, the components P_{dp}^{D} and P_{dp}^{R} ''beat'' against each other, producing an oscillatory interference pattern. In other words, DPOAE fine structure arises because DPOAEs are mixtures of emissions with distinctly different properties that reflect their different mechanisms of generation.

2. Comparison between P_{dp}^{R} and P_{sfe}

According to Prediction #2, the reflection-source component of the total DPOAE, P_{dp}^{R} (Fig. 1, lower panel), should closely match other reflection-source emissions measured under comparable conditions (e.g., SFOAEs at low stimulus levels as in Fig. 1, upper panel). Figure 4 tests this prediction by comparing P_{dp}^{R} and P_{sfe} measured in the same ear. In agreement with predictions, the phase slopes of P_{sfe} and P_{dp}^{R} are nearly identical. In addition, both P_{sfe} and P_{dp}^{R} have similar amplitude features (e.g., a deep notch near 1.4 kHz). These similarities support the idea that P_{sfe} and P_{dp}^{R} are generated by a similar mechanism. Note that deep spectral notches such as that apparent near 1.4 kHz are predicted by the theory of coherent reflection filtering (cf. Fig. 11 of Zweig and Shera, 1995). In the model, such notches arise from random spatial fluctuations in the irregularities that scatter the wave. At some frequencies, wavelets scattered from different locations within the scattering region combine nearly out of phase and cause near cancellation of the net reflected wave.

Although the overall match between P_{dp}^{R} and P_{sfe} is good—especially when one considers the substantial differences in the way that the two emissions are evoked and measured—details of the spectral shape (e.g., the precise lo-

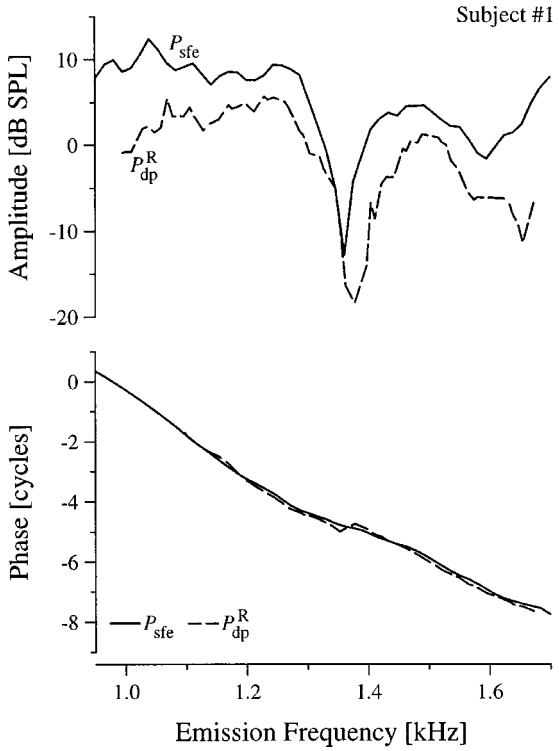

FIG. 4. Comparison between the estimated reflection-source component, P_{dp}^{R}, and the SFOAE, P_{sfe}. The figure shows the amplitude (top) and phase (bottom) of P_{dp}^{R} (dashed line), the reflection-source component of the total DPOAE obtained in Fig. 3 for subject #1. Shown for comparison is P_{sfe} (solid line), the SFOAE measured in the same subject at a probe level of 40 dB SPL. Note the considerable agreement in both amplitude and phase (e.g., P_{dp}^{R} and P_{sfe} have similar amplitude notches and phase slopes). SFOAE stimulus levels: $\{L_p, L_s\} = \{40,50\}$ dB SPL.

cation of the notch) do not match perfectly. Do these discrepancies suggest important differences between P_{dp}^{R} and P_{sfe} and their mechanisms of generation? Or do they reflect differences in measurement conditions that influence the magnitude and/or phase of the traveling-wave energy scattered back from **R**? For example, the primaries present during the DPOAE measurement may suppress the traveling wave near the f_{dp} place, thereby affecting the frequency dependence of P_{dp}^{R}.

3. Mimicking suppression by the primaries

To address these questions, we modified our P_{sfe}-measurement paradigm to better mimic the intracochlear conditions under which P_{dp}^{R} originated. Specifically, we measured P_{sfe} in the presence of an additional tone whose frequency and level were chosen to match those of the f_1 primary used during the measurement of P_{dp}^{R} (see Methods above). We mimic the f_1 primary because we expect it to have the greater effect; the f_1 primary is both closer in frequency to f_{dp} and higher in level than the f_2 primary. We define P_{sfe}^{1} as the value of P_{sfe} measured in the presence of the f_1-primary mimicker.

Measurements of P_{sfe}^{1} are shown and compared to those of P_{dp}^{R} in Fig. 5. The match between the two putative reflection-source emissions is now much closer. This result is consistent with the idea that the differences in Fig. 4 reflect differing intracochlear stimulus conditions; differences in the mechanisms of emission generation are not required. Thus, the similarity in both magnitude and phase between P_{dp}^{R} and P_{sfe}^{1} is in agreement with Prediction #2 and provides strong support for the model. Note that the changes in the overall amplitude and spectral shape of P_{sfe} caused by the f_1-primary mimicker suggest that the primaries have a significant effect on the reflection-source component of the DPOAE (presumably via suppression of the wave incident upon and/or reflected back from the **R** region).

III. UNMIXING VIA SPECTRAL SMOOTHING

A potential difficulty with suppression-based unmixing is that the suppressor tone, introduced with the intent of selectively suppressing the reflection-source component, may inadvertently modify the response in other ways. For example, the suppressor tone may also suppress the distortion-source component (either directly, or through its effects on the primaries) or ''catalyze'' the generation of additional distortion-sources at the frequency f_{dp} (Fahey *et al.*, 2000). As a test of these possibilities, and to investigate the sensitivity of our conclusions to the method of unmixing, we repeated our analysis using a completely different method. This method—spectral smoothing (or its equivalent, time windowing)—was suggested by the correspondence, in a linear system, between phase slope in the frequency domain and latency in the time domain (e.g., Papoulis, 1962). As unmixed by suppression, the two components P_{dp}^{D} and P_{dp}^{R} of P_{dp} have very different phase slopes, evidently reflecting fundamental differences in their mechanisms of generation. Consequently, if we apply Fourier analysis to our frequency-domain measurements of P_{dp}, we expect to see two components of very different latencies in the corresponding ''latency-domain response:''[4] namely, a short-latency component corresponding to P_{dp}^{D} and a long-latency component corresponding to P_{dp}^{R}. Thus, our suppression results suggest that the two components of P_{dp} should be separable using signal-processing strategies based on appropriate windowing in the latency domain. Techniques for analyzing OAEs in this way were introduced by Shera and Zweig (1993a; Zweig and Shera, 1995), who applied them to the study of SFOAEs; similar methods have since been applied to other emissions (e.g., Stover *et al.*, 1996; Brown *et al.*, 1996; Fahey and Allen, 1997; Knight and Kemp, 2000b; Ren *et al.*, 2000).

Multiplication by a window in the latency domain corresponds to convolution with a smoothing function in the frequency domain. Although the two approaches are entirely equivalent, we refer to the technique as ''spectral smoothing'' rather than ''time windowing'' because viewing the process in the frequency domain yields equations for the components that are more directly analogous to those of the suppression method [cf. Eqs. (3)–(5)]. The spectral smoothing strategy for unmixing thus yields the following estimates of P_{dp} and its components:

$$P_{dp} = P_{ec}(f_{dp}) \quad \text{(measured at fixed } f_2/f_1\text{)}; \tag{8}$$

$$P_{dp}^{D} \approx P_{dp}|_{\text{smoothed}} \quad \text{(convolved with smoothing filter)}; \tag{9}$$

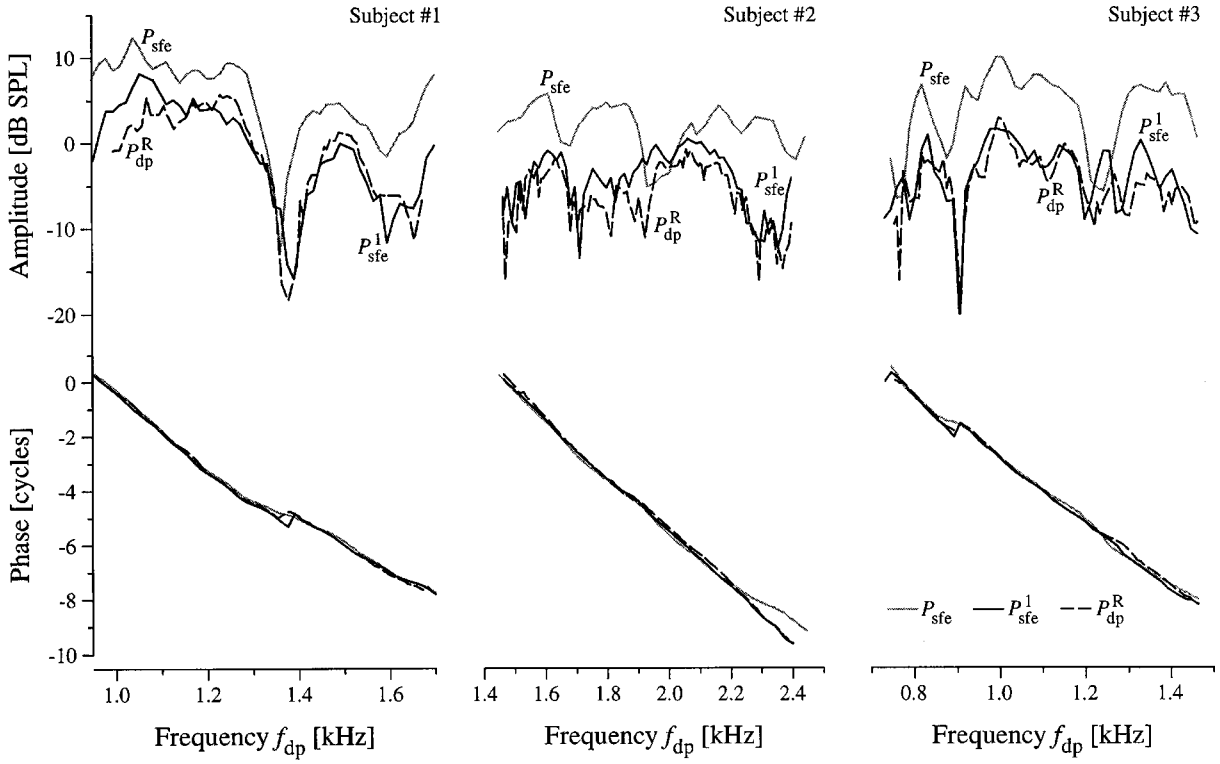

FIG. 5. Comparison between the estimated reflection-source component, P_{dp}^{R}, and the f_1-mimicked SFOAE, P_{sfe}^{1}. The figure compares the amplitude (top) and phase (bottom) of P_{dp}^{R} (dashed line, from Fig. 3) and P_{sfe}^{1} (black solid line), the value of P_{sfe} measured in the presence of an additional tone at the frequency and level of the f_1 primary present during the measurement of P_{dp}^{R}. Left to right, the three panels show data for subjects #1, #2, and #3, respectively. Shown for comparison is P_{sfe} (gray line). The match between the amplitude and phase of P_{dp}^{R} and P_{sfe}^{1} is generally excellent, in agreement with Prediction #2. The differences between P_{sfe} and P_{sfe}^{1} caused by the mimicker suggest that the primaries have a significant effect on the reflection-source component of the DPOAE. SFOAE stimulus levels: $\{L_p, L_s, L_1\} = \{40,55,60\}$ dB SPL.

$$P_{dp}^{R} \approx P_{dp} - P_{dp}^{D}. \tag{10}$$

A. Analysis methods

Measurements of transient-evoked and stimulus-frequency emissions indicate that reflection-emission latency varies with frequency (e.g., Kemp, 1978; Wilson, 1980; Norton and Neely, 1987; Neely *et al.*, 1988; Shera and Guinan, 2000a). This frequency dispersion tends to smear out the reflection-source component in time, making it more difficult to separate by windowing. To help compensate for this dispersion, it proves helpful to work in the log-frequency domain. Consequently, we perform Fourier transforms with respect to the dimensionless frequency variable[5]

$$\nu \equiv -\log(f/f_{ref}), \tag{11}$$

where f_{ref} is a reference frequency taken, for convenience, as the maximum frequency of hearing. Fourier transformation with respect to a log-frequency variable, suggested by the approximate local scaling symmetry of cochlear mechanics, results in sharper, more well-defined peaks in the Fourier-conjugate latency domain (Zweig and Shera, 1995; Knight and Kemp, 2000b).[6] The conjugate dimensionless latency variable, here denoted τ, represents emission latency expressed in periods of the emission frequency (Zweig and Shera, 1995).[7]

Unmixing by smoothing involves convolving P_{dp} with a smoothing function, S, of finite bandwidth (e.g., a Gaussian)[8]:

$$P_{dp}|_{smoothed} \equiv S * P_{dp}, \tag{12}$$

where $*$ denotes the operation of convolution. The convolution is equivalent to a multiplication (or windowing) in the τ domain. Thus,

$$P_{dp}|_{smoothed} = F^{-1}\{\hat{S} \times F\{P_{dp}\}\}, \tag{13}$$

where $F\{\cdot\}$ represents the operation of Fourier transformation (with respect to ν),[9] $F^{-1}\{\cdot\}$ the inverse transformation (with respect to τ), and the window, $\hat{S}(\tau)$, is the Fourier transform of S:

$$\hat{S} \equiv F\{S\}. \tag{14}$$

Separation of P_{dp} into meaningful components requires choosing the smoothing function (or, equivalently, the shape and duration of the latency window) appropriately. Ideally, the window $\hat{S}(\tau)$ should have a sharp cutoff in the τ domain—to cleanly separate emission components of different latencies—but avoid extensive spreading (or ''ringing'') in the frequency response (smoothing function). To approximate these desired characteristics we employ one of a class of ''recursive exponential filters'' (Shera and Zweig, 1993a).[10] The recursive-exponential filters are entire functions and have no poles, discontinuities, or other undesirable features in the complex plane to contribute large oscillations to the smoothing function.

In practice, measurements are only available over a finite frequency range, and the smoothing operation is compli-

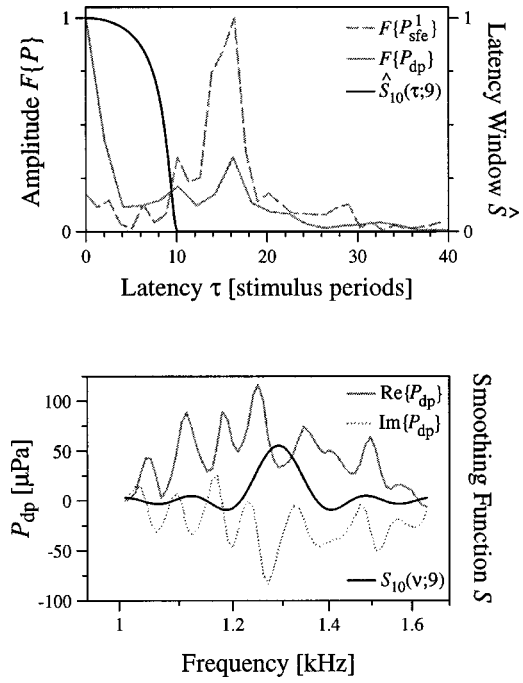

FIG. 6. The smoothing function and corresponding latency window. The figure shows both latency-domain (top) and corresponding frequency-domain representations (bottom) of P_{dp} and the matched smoothing function, S. The top panel shows the amplitudes of the Fourier transforms $F\{P_{dp}\}$ (solid gray line) and $F\{P^1_{sfe}\}$ (dashed gray line) vs τ, the emission latency expressed in stimulus periods. The τ-domain representations of both P_{dp} and P^1_{sfe} show a strong peak centered at a latency of about 15 cycles. Shown for comparison (solid black line) is the tenth-order recursive exponential filter, $\hat{S}_{10}(\tau;\tau_{cut})$, with a cutoff time of $\tau_{cut}=9$ used in subsequent analysis to separate short- and long-latency components of P_{dp}. The bottom panel shows frequency-domain representations of P_{dp} and S. Note that the frequency axis is logarithmic (linear in ν). The real and imaginary parts of P_{dp} are shown with the solid and dotted gray lines, respectively. A linear ramp has been subtracted from P_{dp} to render the function periodic on a cylinder. The black line shows the smoothing function $S_{10}(\nu;9)$ which, when convolved with P_{dp}, yields our estimate of P^D_{dp}. Note that the vertical scale for S, dependent on the number of points in our numerical Fourier transform, is not especially illuminating and has been left unspecified. The measurements of P_{dp} and P^1_{sfe} are from Figs. 3 and 5, respectively (subject #1).

cated by end effects. Throughout this paper, the analyzed frequency range was chosen to include an approximately integral number of spectral cycles, and smoothing was performed using periodic boundary conditions (the data were effectively wrapped around a cylinder). When necessary, a linear ramp was subtracted, and subsequently restored after smoothing, to remove any discontinuity at the ''seam.'' The estimate of P^D_{dp} so obtained was then discarded at each end over a frequency interval equal to the approximate bandwidth of the smoothing function.[11]

1. Determining the window duration

Unmixing via windowing (spectral smoothing) requires specification of the duration of the time window (bandwidth of the smoothing function) used to separate components with different latencies. The suppression studies reported above indicate that the long-latency component, P^R_{dp}, closely matches the characteristics of reflection emissions measured under comparable conditions (e.g., P^1_{sfe}). Consequently, an estimate of the appropriate window duration can be obtained from measurements of SFOAEs evoked at low stimulus levels. Analysis of such measurements indicates that in the 1–2-kHz range, reflection emissions are delayed by an average of about 15 periods of the stimulus frequency with a spread of roughly ±35% (Zweig and Shera, 1995; Shera and Guinan, 2000a). Multiplication by a window of duration $\tau_{cut}=8-9$ periods might therefore be expected to cleanly remove reflection-source components in this frequency range.

Figure 6 corroborates this analysis using our measurements of P_{dp} and P^1_{sfe}. Both short- and long-latency components are clearly apparent in the Fourier transform $F\{P_{dp}\}$, the latency-domain representation of P_{dp}. [By contrast, the long-latency component is almost entirely absent in the Fourier transform of P^D_{dp} obtained by suppression (not shown).] As expected, the long-latency component in $F\{P_{dp}\}$, centered at a latency of about 15 cycles, coincides with the peak in $F\{P^1_{sfe}\}$. The tenth-order recursive exponential filter, $\hat{S}_{10}(\tau;\tau_{cut})$, with a cutoff latency of $\tau_{cut}=9$ periods, is shown for comparison. In subsequent analysis, we use $\hat{S}_{10}(\tau;\tau_{cut}=9)$ to separate the short- and long-latency components of P_{dp}.

B. Results: Unmixing via spectral smoothing

1. P_{dp} and its components, P^D_{dp} and P^R_{dp}, revisited

Typical measurements of P_{dp} and its components unmixed by spectral smoothing are shown with the components obtained by suppression in Fig. 7. Qualitatively, the two methods unmix P_{dp} into similar components. For example, the estimates of P^R_{dp} obtained by the two methods have nearly identical phases and manifest similar frequency dependence in their amplitude curves. There are, of course, differences in the details. For example, the distortion-source components, P^D_{dp}, obtained by suppression unmixing have larger fine structure than the same components obtained by smoothing. We examine this issue further in the next section. Despite differences in detail, the qualitative agreement between the estimated components indicates that our tests of the two-mechanism model are not especially sensitive to the method of unmixing.

IV. ERRORS DUE TO INCOMPLETE UNMIXING

We explore in Fig. 8 the effects of varying key parameters in each of our two unmixing paradigms. For unmixing by suppression, the top panels show how estimates of P^D_{dp} and P^R_{dp} depend on suppressor level, L_s; for unmixing by smoothing (time windowing), the bottom panels show the dependence on the duration of the latency window, τ_{cut} (or, equivalently, the bandwidth of the smoothing function). Note how the fine-structure oscillations in P^D_{dp} (left) increase toward the bottom of each plot (i.e., at lower values of L_s or longer τ_{cut}). By contrast, the fine-structure oscillations in P^R_{dp} (right) increase towards the top (i.e., at higher values of L_s or shorter τ_{cut}).

These systematic trends can be understood using a simple model of the unmixing process. Let the model pressure P_{dp} be the sum of two components, D and R, with very different phase slopes. As a consequence of this difference, D and R beat against each other, producing an oscillatory

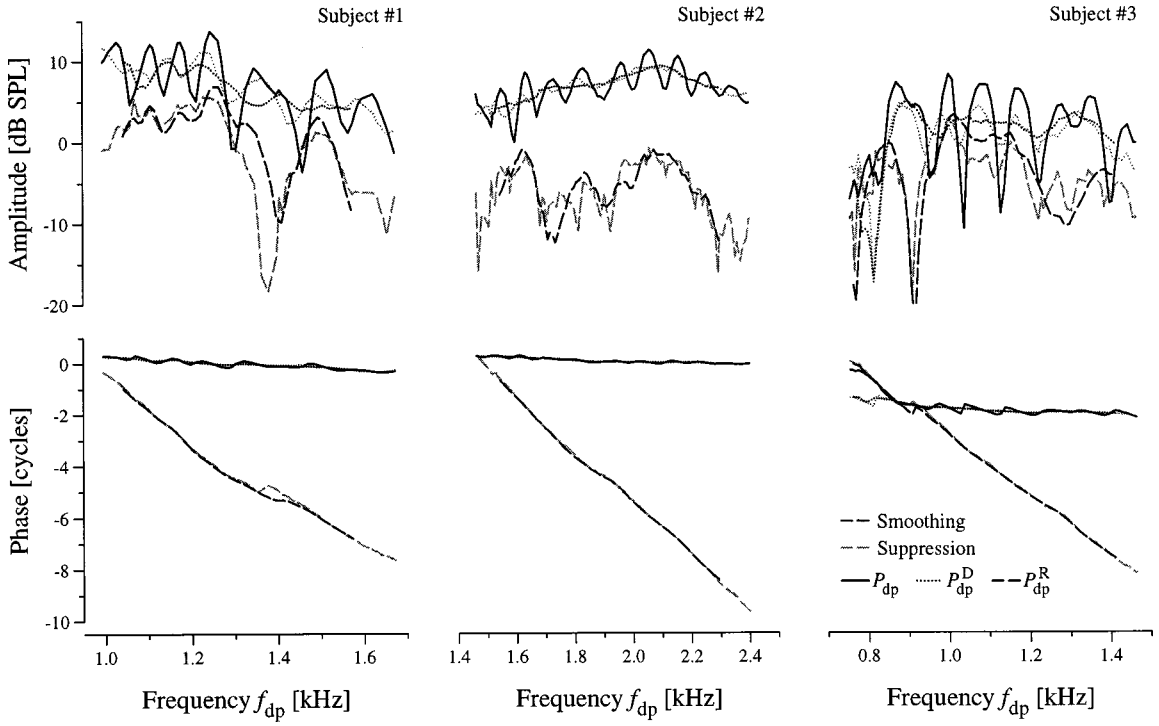

FIG. 7. The DPOAE P_{dp} and its estimated distortion- and reflection-source components, P^D_{dp} and P^R_{dp}, obtained using spectral smoothing (time windowing). The figure shows the amplitude (top) and phase (bottom) of P_{dp} from Fig. 3. Left to right, the three panels show data from subjects #1, #2, and #3, respectively. In each case, the total DPOAE (solid black line) was unmixed as described in the text using the tenth-order recursive exponential filter, $\hat{S}_{10}(\tau;9)$. The two components, P^D_{dp} (dotted black line) and P^R_{dp} (dashed black line), are qualitatively similar to those obtained using suppression (gray lines).

interference pattern in the amplitude and phase of P_{dp}. Imagine now that we attempt to unmix the components experimentally; let our estimates of the two components be denoted P^D_{dp} and P^R_{dp}, respectively. Perfect unmixing would yield values $P^D_{dp}=D$ and $P^R_{dp}=R$. In general, however, unmixing is incomplete, and the estimates contain contributions from both D and R:

$$\begin{pmatrix} P^D_{dp} \\ P^R_{dp} \end{pmatrix} = \begin{pmatrix} 1-\delta & \rho \\ \delta & 1-\rho \end{pmatrix} \begin{pmatrix} D \\ R \end{pmatrix}, \tag{15}$$

where the complex, frequency-dependent coefficients δ and ρ quantify the unmixing errors. Note that the coefficients satisfy the constraint $P^D_{dp}+P^R_{dp}=D+R$. Although perfect unmixing requires $\delta=\rho=0$, acceptable results occur with $|\delta| \ll 1$ and $|\rho| \ll 1$.

The unmixing errors δ and ρ depend on unmixing parameters such as the level of the suppressor and the duration of the latency window. To explicate the trends in Fig. 8, we consider three special cases of incomplete unmixing:

(1) Case $\delta=0$ and $\rho\neq 0$ so that

$$P^D_{dp}=D+\rho R; \tag{16}$$

$$P^R_{dp}=(1-\rho)R. \tag{17}$$

For suppression-based unmixing, this case results from using a weak suppressor that leaves D unchanged but only incompletely removes the R component from the mix (i.e., undersuppression); in the smoothing case, it corresponds to under-smoothing (i.e., to using too narrow a smoothing function or too long a latency window). Since the resulting estimate of P^D_{dp} appears contaminated by the R component, the magnitude $|P^D_{dp}|$ should oscillate with frequency. These features are found in the figure: at smaller values of L_s and longer values of τ_{cut}, the estimates L^D_{dp} manifest considerable fine structure.

(2) Case $\delta\neq 0$ and $\rho=0$ so that

$$P^D_{dp}=(1-\delta)D; \tag{18}$$

$$P^R_{dp}=R+\delta D. \tag{19}$$

Here, the suppressor is strong enough to completely remove the R component, but in so doing it modifies the D; for smoothing, this case results from oversmoothing (i.e., using an overly broad smoothing function or too short a latency window). In this case, the estimate P^R_{dp} is contaminated with part of the D component, and its magnitude should therefore oscillate. These features occur in Fig. 8: at the largest values of L_s and shortest values of τ_{cut}, the estimates L^R_{dp} show evidence of fine structure.

(3) Case $\delta\neq 0$ and $\rho\neq 0$ so that

$$P^D_{dp}=(1-\delta)D+\rho R; \tag{20}$$

$$P^R_{dp}=(1-\rho)R+\delta D. \tag{21}$$

In this more general case, the suppressor is neither strong enough to eliminate the R component nor weak enough not to affect the D component. For smoothing, this case results from a temporal overlap between the D and R components in the latency domain. In this situation, both

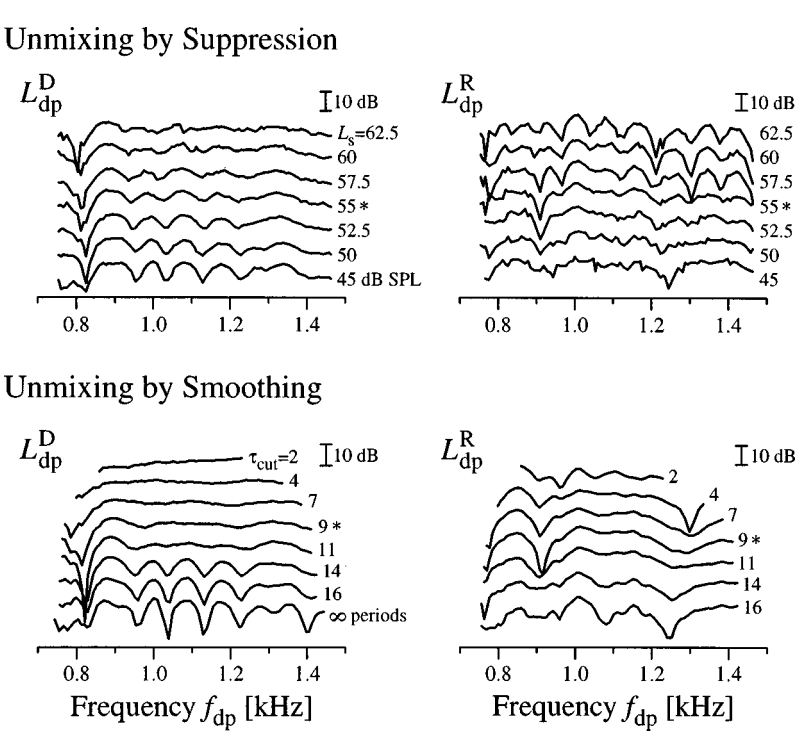

FIG. 8. Changes in estimates of $|P_{dp}^{D}|$ and $|P_{dp}^{R}|$ with variations in the parameters of the unmixing process. The figure shows the levels L_{dp}^{D} (left) and L_{dp}^{R} (right) obtained when the parameters for the suppression- and smoothing-based unmixing are varied systematically. The top panels show the results obtained by varying the level of the suppressor tone, L_s. The bottom panels show the results of varying the duration of the latency window, τ_{cut}. The original, unsmoothed measurement of L_{dp} is shown for comparison in the bottom left ($\tau_{cut}=\infty$). Note that because the estimates of P_{dp}^{D} and P_{dp}^{R} were discarded at each end over a frequency interval equal to the bandwidth of the smoothing function (see Sec. IV A), the estimates cover a more limited frequency range at smaller values of τ_{cut}. In all panels, the different curves have been offset from one another for clarity. Unmixing parameters used earlier in the paper ($L_s=55$ dB SPL and $\tau_{cut}=9$ periods) are marked with an asterisk. Data from subject #3 with stimulus levels of $\{L_1, L_2\}=\{60,45\}$ dB SPL.

P_{dp}^{D} and P_{dp}^{R} will show fine structure oscillations, as seen in Fig. 8 at certain intermediate values of L_s and τ_{cut}.

Our results with suppression unmixing suggests that there is no ''ideal'' suppressor level valid over a range of frequencies that simultaneously eliminates the reflection-source component while leaving the distortion-source essentially unaffected. Figure 8, for example, shows some residual fine structure in both P_{dp}^{R} and P_{dp}^{D} at this subject's ''optimal'' suppressor level of $L_s=55$ dB SPL. The choice of suppressor level involves a trade-off between minimizing $|\delta|$ and minimizing $|\rho|$, with their sum inevitably finite. With proper choice of the windowing function, the prospects for near ideal unmixing by spectral smoothing (time windowing) appear brighter. Judging by the almost negligible amplitude of the fine structure obtained at intermediate values of τ_{cut}, unmixing by smoothing appears able to effect a cleaner separation between the two components than is possible using suppression.

A. Estimating δ and ρ

We illustrate the trade-off between δ and ρ and give a feel for the suppression-based unmixing errors in Fig. 9, which shows estimates of $|\delta|$ and $|\rho|$ for three different suppressor levels. Computation of δ and ρ requires knowledge of D and R; the estimates in Fig. 9 were computed by substituting for D and R the components obtained by spectral smoothing (with $\tau_{cut}=9$). Since the two equations represented in matrix form in Eq. (15) are not independent (but are related by $P_{dp}^{D}+P_{dp}^{R}=D+R$), an additional constraint is necessary to determine δ and ρ uniquely. Since two parameters are available, a natural choice is to match both P_{dp}^{D} and its frequency derivative, $P_{dp}^{D'}$. We thus obtain values of δ and ρ by solving the pair of simultaneous equations

$$P_{dp}^{D}=(1-\delta)D+\rho R; \tag{22}$$

$$P_{dp}^{D'}=(1-\delta)D'+\rho R'. \tag{23}$$

The values of δ and ρ obtained in this way vary with frequency; at every point, the coefficients δ and ρ are chosen to match to the curve P_{dp}^{D}, both its value and its derivative, as closely as possible.[12] Because of the constraint $P_{dp}^{D}+P_{dp}^{R}=D+R$, these same coefficients also provide a match to P_{dp}^{R} and its derivative.

Since the true components D and R are not known, the unmixing errors δ and ρ shown in Fig. 9 were computed relative to the components obtained by spectral smoothing; they therefore provide only a rough guide to the actual errors. The results are, however, generally consistent with expectations based on the three special cases of Eq. (15) considered above. Note, for example, that $|\delta|$ and $|\rho|$ vary in opposite directions with changes in suppressor level. At the largest suppressor level, $|\delta|$ is relatively large and $|\rho|$ relatively small (corresponding to the expectations for strong suppression outlined in case #2 above). Similarly, at the smallest suppressor level, the relative magnitudes of δ and ρ are reversed (weak suppression, as in case #1). At the ''optimal'' suppressor level, the errors $|\delta|$ and $|\rho|$ are intermediate between these extremes. Not surprisingly, $|\delta|$ and $|\rho|$ can become large in frequency regions where the total DPOAE is itself poorly determined (e.g., near 0.8 kHz where $|P_{dp}|$ is relatively close to the noise floor) and/or where the estimated components P_{dp}^{D} and P_{dp}^{R} change rapidly (e.g., near notches of P_{dp}^{R}). Overall, however, the unmixing errors are fairly small for intermediate suppressor levels (typically $|\delta|\sim 0.1$ and $|\rho|\sim 0.2-0.3$). These findings corroborate the qualitative visual impression that the two methods, selective suppression and spectral smoothing, unmix into generally similar components.

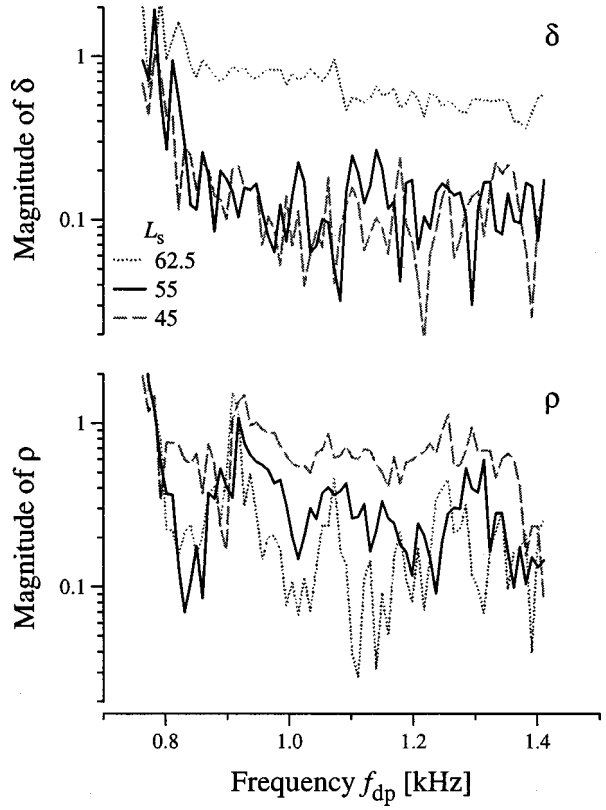

FIG. 9. Estimates of the unmixing errors δ and ρ at different suppressor levels. The figure shows magnitudes of the unmixing errors δ and ρ computed as the solution to Eqs. (22) and (23). The components P_{dp}^{D} and P_{dp}^{R} obtained by spectral smoothing (with $\tau_{cut}=9$) were used as estimates of D and R. Results for three different suppressor levels were computed using the data whose magnitudes are shown in the top panels of Fig. 8.

V. DISCUSSION

In this paper we tested the two key predictions of the two-mechanism model of DPOAE generation by successfully unmixing DPOAEs into components, P_{dp}^{D} and P_{dp}^{R}, with characteristics indicative of fundamentally different mechanisms of generation (i.e., nonlinear distortion vs linear reflection). In agreement with Prediction #1, the phase of the putative distortion-source component (P_{dp}^{D}) is nearly constant, whereas the phase of the reflection-source component (P_{dp}^{R}) varies rapidly with frequency. These differing phase slopes imply fundamental differences in the respective mechanisms of emission generation. In particular, the two slopes are consistent with generation by nonlinear distortion (P_{dp}^{D}) and linear coherent reflection (P_{dp}^{R}), respectively (Shera and Guinan, 1999). Furthermore, in agreement with Prediction #2, the spectral shape and phase of the reflection-source component closely match those of the SFOAE measured at the same frequency under comparable conditions (i.e., with the addition of an f_1-primary mimicker). Changes in the SFOAE caused by the mimicker suggest that the primaries have a significant influence on the reflection-source component of the DPOAE, presumably via suppression. To investigate the robustness of our conclusions, we unmixed DPOAE sources using two completely different methods: (a) selective suppression of the reflection source using a third tone near the distortion-product frequency, and (b) spectral smoothing (or, equivalently, time-domain windowing). Although the two methods unmix in very different ways, explicit analysis of the unmixing errors demonstrates that they yield similar DPOAE components, indicating that our results are not especially sensitive to the method of unmixing.

A. Source mechanism versus source location

The quasiperiodic fine structure often evident in DPOAE spectra is now generally regarded as resulting from the alternating constructive and destructive interference between backward-traveling waves originating in two separate regions of the cochlea (Kim, 1980). The physics underlying the interference pattern has generally been understood as follows (e.g., Brown *et al.*, 1996; Stover *et al.*, 1996; Fahey and Allen, 1997): Because the two sources are spatially separated, backward-traveling waves originating at the more apical location must travel further to reach the ear canal than waves originating at the basal location. Consequently, waves from the apical source are delayed relative to the basal source; in the frequency domain, this delay corresponds to a frequency-dependent phase shift. Thus, the relative phase of the two waves rotates with frequency, alternately passing through plus and minus 1. This rotation of relative phase creates the interference pattern—known as DPOAE fine structure—when the two waves are combined in the ear canal. Kim (1980) originally referred to the two DPOAE sources as the ''primary-place source'' and the ''characteristic-place source,'' and considerable evidence now suggests that the two backward-traveling waves do indeed originate at these locations (e.g., Furst *et al.*, 1988; Gaskill and Brown, 1990; Brown *et al.*, 1996; Engdahl and Kemp, 1996; Brown and Beveridge, 1997; Talmadge *et al.*, 1998, 1999; Heitmann *et al.*, 1998; Fahey and Allen, 1997; Siegel *et al.*, 1998; Mauermann *et al.*, 1999a, 1999b).

We demonstrate here, however, that this place-based nomenclature—and the conceptual model that underlies it (e.g., Brown *et al.*, 1996; Stover *et al.*, 1996; Fahey and Allen, 1997)—although apparently accurate in its specification of the locations of wave origin, fails to capture the critical distinction between the two sources. As suggested by Shera and Guinan (1999), the *fundamental distinction between the two sources is evidently not source* **location**, *but source* **mechanism**. Indeed, only by incorporating both classes of emission-generating mechanisms (i.e., nonlinear distortion and linear coherent reflection) have models been able to account for the known phenomenology of DPOAE fine structure (e.g., Talmadge *et al.*, 1998, 1999; Mauermann *et al.*, 1999a). Accordingly, our terminology distinguishes the two components not by their place of origin, but by their mechanism of generation (i.e., distortion- versus reflection-source components).

Our results support the two-mechanism model of DPOAE generation. To illustrate, consider how our experimental results would have differed if both sources in Fig. 1 had been distortion sources like **D**. When probed with the frequency-scaled stimuli used here, both sources would then have generated backward-traveling waves with phases essentially independent of frequency. Consequently, the relative phase of the waves from the two sources would have been

nearly constant, and no oscillatory fine structure would have appeared in the ear-canal pressure spectrum. *Note that this constancy of relative phase would have occurred despite the fact that the two waves originate at different spatial locations within the cochlea*. In other words, although the reflection-source region at **R** is further from the stapes than the distortion-source region at **D**, the difference in phase slope characterizing emissions from these two sources is not due to the differing locations of the **D** and **R** regions. Rather, contrary to standard assumption, phase slopes are ultimately determined by mechanisms of emission generation. For example, the theory of coherent reflection filtering (Shera and Zweig, 1993b; Zweig and Shera, 1995) implies that reflection-emission latency is determined not by the distance a wave travels to reach its characteristic place but by the characteristics of cochlear tuning it finds when it gets there (Shera and Guinan, 2000a, 2000b).

B. Comparison with other work

The experiments reported here were designed specifically to test Predictions #1 and #2 and therefore differ from most other studies of DPOAE components (e.g., Brown *et al.*, 1996; Stover *et al.*, 1996; Siegel *et al.*, 1998) in their use of frequency-scaled stimuli (i.e., fixed f_2/f_1). According to the analysis underlying the model (Shera and Guinan, 1999), distortion and reflection mechanisms yield qualitatively different phase behavior (i.e., nearly constant phase vs rapidly rotating phase) when emissions are evoked with frequency-scaled stimuli. Similar qualitative differences in phase are not found using other measurement paradigms, and the underlying differences in mechanism can therefore be considerably less transparent. For example, much more rapid phase rotation occurs when distortion emissions are measured using stimulus paradigms (e.g., fixed f_1, fixed f_2, or fixed f_{dp}) for which the cochlear wave pattern is not simply translated along the cochlear partition (e.g., Kimberley *et al.*, 1993; O'Mahoney and Kemp, 1995; Shera *et al.*, 2000). Unmixing DPOAEs measured using constant f_2/f_1-ratio sweeps, rather than one of the more common alternative paradigms, thus greatly facilitates recognition of the two emission mechanisms. By increasing the difference in phase slope between the distortion- and reflection-source components, our use of frequency-scaled stimuli also facilitates unmixing of the two components using spectral smoothing (time windowing) by maximizing the separation between the two components when the data are transformed into the ''latency domain'' using Fourier analysis.

Our tests of Prediction #2 contrast sharply with the findings of Brown *et al.* (1996), who performed a DPOAE unmixing analysis using a smoothing technique and compared the resulting ''DP residual'' (their analog of P_{dp}^{R}) to measurements of SFOAEs. Although they noted similarities in the phase slopes, they found ''little correspondence in the magnitude across frequency of the DP residual and SF[OA]E.'' Their reported discrepancy between emission components conflicts with earlier work (Kemp and Brown, 1983), which found at least qualitative agreement between the SFOAE and the DPOAE component believed to originate at the distortion-product place (as obtained, in this case, using a suppression paradigm). In contrast with these results, we find excellent agreement, both between DPOAE components unmixed via different paradigms and between P_{dp}^{R} and corresponding SFOAEs. Unfortunately, Brown *et al.* (1996) do not specify their smoothing algorithm in the detail necessary to enable a direct comparison with our method.[13] We note, however, that in our experiments the addition of the f_1-primary mimicker often improved the agreement between the magnitudes of P_{dp}^{R} and P_{sfe} considerably (cf. Fig. 5). This result indicates that suppressive and other effects of the primaries on P_{dp}^{R} must be taken into account in any such comparison.

C. Region of validity of the two-mechanism model

The tests of Predictions #1 and #2 reported here, together with more limited data at other (low to moderate) primary levels and at frequency ratios f_2/f_1 in the range 1.1–1.3, establish the validity of the two-mechanism model in humans for the DPOAE measurement parameters in common use [i.e., low to moderate sound-pressure levels with $L_1 \geqslant L_2$ and primary frequency ratios $f_2/f_1 \approx (f_2/f_1)_{optimal}$]. Knight and Kemp (2000a) provide a test of Prediction #1 over a broad range of frequency ratios ($1.01 \leqslant f_2/f_1 \leqslant 1.5$) in an unmixing analysis of their stunning $\{f_1, f_2\}$-area map (Knight and Kemp, 2000b). Their results, based on time windowing of DPOAEs measured using primary levels $L_1 = L_2 = 70$ dB SPL, are consistent with the two-mechanism model and indicate that the relative amplitudes of the components P_{dp}^{D} and P_{dp}^{R} vary systematically with f_2/f_1. Whether Prediction #2 also applies over a similarly broad range of parameter values remains an important open question.

Described and tested here in the frequency domain, Prediction #1 of the two-mechanism model evidently also applies in the time domain. Combining phase-rotation averaging (Whitehead *et al.*, 1996) with an elegant pulsed-primary technique, Talmadge *et al.* (1999) provide strong support for model predictions that amount, in effect, to time-domain analogs of Prediction #1. Since the responses involved arise in a nonlinear system, this conclusion is nontrivial. Time-domain tests of Prediction #2 await further experiment.

The validity of the model at high intensities also remains to be investigated. For example, at higher levels of intracochlear distortion, the emission evoked by the forward-traveling distortion component may contain, in addition to contributions from coherent reflection, significant energy from distortion-source waves created by nonlinear distortion (e.g., Withnell and Yates, 1998). Furthermore, the two emission sources may also begin to mix in ways more complicated than simple linear summation.[14] For example, the strength of the micromechanical impedance perturbations that scatter the traveling wave may depend on the local amplitude of basilar-membrane vibration.

D. Methods of unmixing

Our success at unmixing using two completely different methods (suppression and smoothing) demonstrates the robustness of our conclusions to the method of unmixing. The

two methods unmix in very different ways, and the systematic errors each introduces are presumably quite different. Whereas the suppression method separates components based on their differential modification by an external tone, the spectral-smoothing (or time-domain windowing) method separates components based on latency in the ''time-domain response'' obtained using Fourier analysis.[15] Despite these differences, the two methods unmix the total emission into rather similar components (at least for $f_2/f_1=1.2$ and low to moderate primary levels).[16] Whether the two methods yield similar results at other f_2/f_1 ratios and/or at higher stimulus levels remains an important open question. Differences between the methods would not be surprising at f_2/f_1 ratios close to 1—although the spectral-smoothing method does not depend on spatial separation of source regions in the cochlea, the ability of the suppression method to selectively eliminate one of the sources presumably deteriorates as the two sources draw closer to one another as f_2/f_1 approaches 1.

An advantage of the spectral-smoothing method is that it requires measurement of only a single quantity (namely, P_{dp}, whereas the suppression method requires both P_{dp} and $P_{\mathrm{dp}}|_{\mathrm{suppressed}}$). Unlike the suppression method, the smoothing method therefore allows each measurement of P_{dp} to serve as its own control against possible systematic changes (e.g., variations in overall emission level due to efferent effects) that may occur during the course of the measurement. In the suppression studies reported here, we sought to minimize these potential problems by interleaving measurements of P_{dp} and $P_{\mathrm{dp}}|_{\mathrm{suppressed}}$ in time. Although the spectral smoothing method depends only on P_{dp}, it requires knowledge of P_{dp} at multiple frequencies. Indeed, the method works best if applied to measurements that span a relatively wide frequency range (i.e., many periods of the microstructure) with good frequency resolution (i.e., many points per period). In addition, because of uncertainties introduced near the end points due to incomplete knowledge of P_{dp} outside the measured interval, the smoothing method requires measurements over an interval slightly larger than the desired frequency range. The suppression method, by contrast, imposes no such constraints; suppression unmixing requires measurement of P_{dp} and $P_{\mathrm{dp}}|_{\mathrm{suppressed}}$ only at the actual frequency (or frequencies) of interest.

E. Implications of unmixing DPOAEs

Uncontrolled mixing may be a substantial source of subject-dependent variability in DPOAE measurements. Indeed, our results imply that the interpretation of DPOAE responses appears doubly confounded. First, DPOAEs are mixtures of emissions originating from at least two different regions in the cochlea. This ''spatial blurring,'' now widely recognized, compromises the frequency selectivity of DPOAE measurements (e.g., Heitmann *et al.*, 1998). Second, DPOAEs are mixtures of emissions arising by fundamentally different mechanisms. This ''mechanistic blurring,'' established here, compromises the etiological specificity of DPOAE measurements. For although both distortion- and reflection-source emissions share a common dependence on propagation pathways from the cochlea to the ear canal, and are therefore both sensitive to modifications of that pathway (e.g., to middle-ear pathology or to reductions in cochlear amplification caused by damage to outer hair cells), their respective mechanisms of generation—and hence their dependence on underlying parameters of cochlear mechanics—remain fundamentally distinct. For example, whereas distortion-source emissions presumably depend on the form and magnitude of cochlear nonlinearities (e.g., on the effective ''operating point'' along hair-cell displacement-voltage transduction functions), reflection-source emissions depend strongly on the size and spatial arrangement of micromechanical impedance perturbations (e.g., on variations in hair-cell number and geometry). Distortion-product unmixing, using techniques such as those employed here, should therefore improve the power and specificity of DPOAEs as noninvasive probes of cochlear function.

ACKNOWLEDGMENTS

We thank our experimental subjects for their considerable patience and gratefully acknowledge the efforts of Paul Fahey, John Guinan, William Peake, and two anonymous reviewers, who provided valuable comments on the manuscript. This work was supported by Grants Nos. R01 DC03687 and T32 DC00038 from the NIDCD, National Institutes of Health.

APPENDIX MEASUREMENT METHODS

This Appendix describes in more detail the methods used to obtain the emission measurements reported here.

1. Measurement of DPOAEs

Distortion-product otoacoustic emissions at the frequency $2f_1-f_2$ were measured using frequency-scaled stimuli (i.e., with the ratio f_2/f_1 held constant). At each measurement frequency the acoustic stimulus had the form

$$\mathrm{stimulus}=\underbrace{\langle XX\cdots X\rangle}_{\#\geqslant M}, \tag{A1}$$

where X represents a periodic (5×4096)-sample ($\approx$342 ms[17]) segment consisting of three components:

$$X=\begin{cases}\pi_1^1\Pi_2^1\Pi_3^1\pi_4^1\pi_5^1\pi_6^1\Pi_7^1\pi_8^1 & (\text{primary earphone \#1})\\ \pi_1^2\Pi_2^2\pi_3^2\pi_4^2\pi_5^2\Pi_6^2\Pi_7^2\pi_8^2 & (\text{primary earphone \#2}).\\ o_1O_2O_3\langle{}_4\sigma_5\Sigma_6\Sigma_7\rangle_8 & (\text{suppressor earphone})\end{cases} \tag{A2}$$

Each component consisted of four long intervals (uppercase) and four short intervals (lowercase and angled brackets). The long intervals were each 4096 samples ($\approx$68 ms) in duration. The primary segments, Π_i^1 and Π_i^2, contained an integral number of periods of the primary frequencies, f_1 and f_2, respectively. The suppressor segments, Σ_i, contained an integral number of cycles of the suppressor frequency, f_s. The zero segments, O_i, were identically zero throughout. Waveform phases were adjusted, using information from the calibration procedure, so that each stimulus had zero (cosine) phase in the ear canal at the beginning of segment Π_2. The short intervals were one fourth the duration of the long intervals (i.e., 1024 samples or $\approx$17 ms) and did not, in gen-

eral, contain an integral number of periods of the corresponding waveform. The short intervals π_i, σ_i, and o_1 allowed for response settling time and contained segments of the primary, suppressor, or zero waveforms, respectively. The short intervals $\{\langle_4,\rangle_8\}$ were used to ramp the suppressor tone {on, off} using the {first, second} half of the Blackman window. The three components of X were synchronized and presented simultaneously through three separate earphones. Note that whereas the primary tones played continuously during the measurement, the suppressor tone cycled on and off repeatedly due to alternation of the zero and suppressor waveforms. Interleaving the measurements of P_{dp} and $P_{\mathrm{dp}}|_{\mathrm{suppressed}}$ in this way helps to minimize possible artifacts due to systematic variations over time (e.g., due to subject movement, drifts in earphone calibration, efferent feedback, etc.). Unless otherwise noted, the primary levels $\{L_1,L_2\}$ were {60, 45} dB SPL, respectively. Primary levels were chosen in approximate accordance with the formula $L_1 \approx 0.4L_2+39\,\mathrm{dB\,SPL}$, which tracks the ''ridge'' in the L_1L_2 plane that maximizes the $2f_1-f_2$ emission for $f_2/f_1\approx 1.2$ (Kummer *et al.*, 1998).

Measurements were made versus probe frequency by sweeping the primaries and suppressor from high frequencies to low, with $f_{\mathrm{s}}=f_{\mathrm{dp}}+\Delta f_{\mathrm{s}}$ and $\Delta f_{\mathrm{s}}=-44\,\mathrm{Hz}$. The periodic segments X were played repeatedly until M artifact-free responses were collected. In these measurements, M was typically 64 so that at each frequency the total stimulus duration was therefore $\gtrsim 64\times 342\,\mathrm{ms}\approx 22\,\mathrm{s}$. To reduce unwanted transients the probe waveform was ramped on and off by pre- and postpending two additional segments [indicated by the angled brackets $<$ and $>$ in Eq. (A1)] with envelopes of half Blackman windows with 2.5-ms rise and fall times. After digitizing the resulting ear-canal pressure, responses to all primary-alone segments (i.e., all segments Π_2 and Π_3) were averaged to form Y_{p}; similarly, the responses to all probe+suppressor segments (i.e., all segments Π_6 and Π_7) were averaged to form $Y_{\mathrm{p+s}}$. From these averaged response waveforms, the complex amplitudes of the f_{dp} components of the ear-canal pressure, denoted $P_{\mathrm{dp}}=P_{\mathrm{ec}}(f_{\mathrm{dp}})$ and $P_{\mathrm{dp}}^{\mathrm{D}}=P_{\mathrm{ec}}(f_{\mathrm{dp}})e^{-2\pi i\Delta N\Delta T f_{\mathrm{dp}}}|_{\mathrm{suppressed}}$, were extracted using Fourier analysis. The complex exponential compensates for the phase shift in the probe due to the time interval, $\Delta N\Delta T$, between the primary-alone and primary+suppressor segments. Here, ΔT is the sampling interval (reciprocal of the sampling rate), and ΔN represents the total number of these intervals that separate the two segments:

$$\Delta N=\#\mathrm{samples}(\Pi_2\Pi_3\pi_4\pi_5)=2\tfrac{1}{2}\times 4096=10\,240. \tag{A3}$$

Note that when the two segments are separated by an integral number of periods of the f_{dp} waveform, the phase shift modulo 2π is zero. The complex quantity $P_{\mathrm{dp}}^{\mathrm{R}}(f_{\mathrm{dp}})$ was then obtained as

$$P_{\mathrm{dp}}^{\mathrm{R}}=P_{\mathrm{dp}}-P_{\mathrm{dp}}^{\mathrm{D}}. \tag{A4}$$

2. Measurement of SFOAEs

Stimulus-frequency emissions were measured using the suppression method detailed elsewhere (Shera and Guinan, 1999). In some experiments, we measured SFOAEs in the presence of an additional continuous tone (the ''f_1-primary mimicker'') at a frequency and level corresponding to the f_1 primary in the measurement of DPOAEs detailed above.

At each measurement frequency the acoustic stimulus had the form given by Eq. (A1), with X representing a periodic (5×4096)-sample ($\approx$342 ms) segment consisting of three components:

$$X=\begin{cases}\pi_1\Pi_2\Pi_3\pi_4\pi_5\Pi_6\Pi_7\pi_8 & (\text{probe earhone})\\ o_1O_2O_3\langle_4\sigma_5\Sigma_6\Sigma_7\rangle_8 & (\text{suppressor earphone})\\ \mu_1M_2M_3\mu_4\mu_5M_6M_7\mu_8 & (\text{primary-mimicker earphone})\end{cases}. \tag{A5}$$

Each component consisted of four long (uppercase) and four short (lowercase and angled brackets) intervals. The long intervals were each 4096 samples ($\approx$68 ms) in duration and contained an integral number of periods of the probe (Π_i), suppressor (Σ_i), zero (O_i), or primary mimicker (M_i) waveforms, respectively. The phase of the probe waveform was adjusted, using information from the calibration procedure, so that the stimulus had zero (cosine) phase in the ear canal at the beginning of segment Π_2. The short intervals were one fourth the duration of the long intervals (i.e., 1024 samples or $\approx$17 ms) and did not, in general, contain an integral number of periods of the corresponding waveform. The short intervals π_i, σ_i, o_1, and μ_i allowed for response settling time and contained segments of the probe, suppressor, zero, and mimicker waveforms, respectively. The short intervals $\{\langle_4,\rangle_8\}$ were used to ramp the suppressor tone {on, off} using the {first, second} half of the Blackman window. The three components of X were synchronized and presented simultaneously through three separate earphones. Note that whereas the probe and primary mimicker tones played continuously during the measurement, the suppressor tone cycled on and off repeatedly due to alternation of the zero and suppressor waveforms. The probe and suppressor levels $\{L_{\mathrm{p}},L_{\mathrm{s}}\}$ were generally {40, 55} dB SPL. The primary mimicker was presented at a frequency and level corresponding to the f_1 primary in the measurement of DPOAEs (i.e., at a frequency equal to $f_1=f_{\mathrm{p}}/(2-r)$, where r denotes the f_2/f_1 ratio we wished to mimic, and a typical level of 60 dB SPL).

Other features of the stimulus paradigm and the subsequent data analysis used to compute P_{sfe}^1 are analogous to the measurement of DPOAEs detailed above and have been described elsewhere (Shera and Guinan, 1999).

[1]Note that for brevity this simple synopsis neglects contributions to the total reflection-source emission arising from multiple internal reflection within the cochlea (i.e., from multiple cycles of partial reflection at the stapes and linear coherent reflection within the **R** region).

[2]In a nutshell, the theory says that given ''almost any'' arrangement of micromechanical impedance perturbations (i.e., an arrangement with the appropriate spatial-frequency content, such as perturbations that are randomly and densely distributed), a model will produce realistic reflection emissions whenever the peak region of the traveling wave has a slowly varying wavelength and an envelope that is simultaneously both tall and broad.

[3]Some differences (e.g., in phase) between $P_{\mathrm{dp}}^{\mathrm{R}}$ and P_{sfe} are, of course, expected because the initial sources of forward-traveling cochlear waves at the emission frequency are at different spatial locations in the two cases (i.e., at the distortion-source region, **D**, for $P_{\mathrm{dp}}^{\mathrm{R}}$, and at the stapes for P_{sfe}).

[4]We put ''latency-domain response'' in quotes because the signal we obtain

by Fourier transforming the frequency response does not correspond with the time-domain impulse response of the system.

[5]The minus sign in Eq. (11) has the effect of converting a forward Fourier transform (with respect to ν) into an inverse transform (with respect to $\log f/f_{\text{ref}}$). We work with forward Fourier transforms (with respect to ν) for conceptual and numerical convenience.

[6]Measurements of tone-burst-evoked OAE and ABR latency (Neely *et al.*, 1988), as well as measurements of SFOAE group delay (Shera and Guinan, 2000a), all indicate a gradual breaking of scaling symmetry in the basal turns of the mammalian cochlea. For near-optimal compensation for traveling-wave dispersion, the measurements suggest working with the variable $-\sqrt{f/f_{\text{ref}}}$ (see also Shera *et al.*, 2000).

[7]Our τ scale differs from the time scale employed by Knight and Kemp (2000b), who apply a log-frequency transformation and plot the resulting Fourier transforms against an axis they call ''normalized time.'' Their normalized time has units of milliseconds and was computed based on the mean frequency step size in the log-frequency scale. The two scales differ by a multiplicative factor inversely proportional to the geometric mean frequency of the analyzed data.

[8]Note that unlike the more familiar case of time-domain filtering, the oscillatory function to be removed occurs here in the frequency response. In its reversal of the roles usually played by time and frequency, the technique used here is similar to cepstral analysis (Bogert *et al.*, 1963), although we work with a log-frequency variable, ν, and analyze P_{dp} rather than $\log(P_{\text{dp}})$. [In cepstral analysis, one takes the logarithm of the frequency response in order to decompose a presumed product of spectra into a sum. In our application, the pressure P_{dp} is represented directly as a sum of components (Prediction #1); taking the logarithm is therefore both unnecessary and undesirable.]

[9]To perform our transforms numerically, we resampled our measurements of P_{dp} at equal intervals in log frequency using cubic spline interpolation. Because our sampling rate was variable, our measurements of P_{dp} were not equally spaced in linear frequency.

[10]The nth order recursive-exponential filtering window is defined by (Shera and Zweig, 1993a)

$$\hat{S}_n(\tau;\tau_{\text{cut}}) \equiv 1/\Gamma_n(\lambda_n \tau/\tau_{\text{cut}}),$$

where the parameter τ_{cut} is the cutoff time (length of the window) and the function $\Gamma_n(\tau)$ is defined recursively:

$$\Gamma_{n+1}(\tau) = e^{\Gamma_n(\tau)-1}, \text{ with } \Gamma_1(\tau) = e^{\tau^2}.$$

The window $\hat{S}_n(\tau;\tau_{\text{cut}})$ has a maximum value of 1 at $\tau=0$; the scale factor λ_n is chosen so that the window falls to the value $1/e$ at $\tau=\tau_{\text{cut}}$:

$$\lambda_n = \sqrt{\gamma_n}, \text{ where } \gamma_{n+1} = \ln(\gamma_n+1) \text{ with } \gamma_1 = 1.$$

Note that the first-order filtering window is a simple Gaussian; in the limit $n\to\infty$, $\hat{S}_n$ approaches a rectangular (or boxcar) window. For intermediate n (e.g., the value $n=10$ used here), $\hat{S}_n$ has a much sharper cutoff than standard windowing functions (e.g., the Hamming, Blackman, etc.) and considerably less ''ringing'' in the smoothing function than the simple boxcar.

[11]The smoothing function has approximate width (Shera and Zweig, 1993a)

$$\Delta\nu = \Delta f/f \geqslant 1/\pi\tau_{\text{cut}}.$$

[12]The quantities δ and ρ are thus analogous to the osculating parameters used in the theory of linear differential equations (e.g., Mathews and Walker, 1964).

[13]To smooth their frequency-domain measurements, Brown *et al.* (1996) used a 101-point moving average (evidently tailored to a frequency spacing between points of approximately 1.2 Hz) but fail to specify the shape of their smoothing function. If all points in the moving average were weighted equally (i.e., if the smoothing function were rectangular), the corresponding time-domain window, a sinc function, would have been nonmonotonic, oscillating about zero with a period of roughly 8.25 ms.

[14]We are reminded here of the dialectic described by Levins and Lewontin (1985): ''A necessary step in theoretical work is to make distinctions. But whenever we divide something into mutually exclusive and jointly all-encompassing categories, it turns out on further examination that these opposites interpenetrate.''

[15]To unmix P_{dp} into two components we used a window with a ''low-pass'' characteristic in the time domain. The technique is easily generalized to the unmixing of multiple components with different latencies (e.g., by using multiple ''bandpass'' windows centered at different latencies or a succession of ''low-pass'' windows with different cutoffs).

[16]Working with SFOAEs at low stimulus levels, Shera and Zweig (1993a) established a similar equivalence between SFOAEs extracted using the vector-subtraction method (Kemp and Chum, 1980) and the method of spectral smoothing. The vector-subtraction method exploits the nonlinear saturation of the SFOAE—or ''self-suppression'' of the traveling wave (e.g., Kanis and de Boer, 1993)—at higher stimulus levels.

[17]Because we varied our sampling rate between measurement points, corresponding stimulus durations varied by up to $\pm 3\%$.

Allen, J. B., and Lonsbury-Martin, B. L. (**1993**). ''Otoacoustic emissions,'' J. Acoust. Soc. Am. **93**, 568–569.

Allen, J. B., and Neely, S. T. (**1992**). ''Micromechanical models of the cochlea,'' Phys. Today **45**, 40–47.

Bogert, B. P., Healy, M. J. R., and Tukey, J. W. (**1963**). ''The frequency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking,'' *Proceedings of the Symposium on Time Series Analysis*, edited by M. Rosenblatt (Wiley, New York), pp. 209–243.

Brass, D., and Kemp, D. T. (**1991**). ''Time-domain observation of otoacoustic emissions during constant tone stimulation,'' J. Acoust. Soc. Am. **90**, 2415–2427.

Brass, D., and Kemp, D. T. (**1993**). ''Suppression of stimulus frequency otoacoustic emissions,'' J. Acoust. Soc. Am. **93**, 920–939.

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.

Dreisbach, L. E., and Siegel, J. H. (**1999**). ''Level and phase relationships of distortion-product otoacoustic emission sources with varied primary frequency ratios in humans,'' Assoc. Res. Otolaryngol. Abs. **22**, 392.

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.

Fahey, P. F., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (**2000**). ''Nonlinear interactions that could explain distortion product interference response areas,'' J. Acoust. Soc. Am. **108**, 1786–1802.

Furst, M., Rabinowitz, W. M., and Zurek, P. M. (**1988**). ''Ear canal acoustic distortion at $2f_1-f_2$ from human ears: Relation to other emissions and perceived combinations tones,'' J. Acoust. Soc. Am. **84**, 215–221.

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.

Guinan, J. J. (**1990**). ''Changes in stimulus frequency otoacoustic emissions produced by two-tone suppression and efferent stimulation in cats,'' in *Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer, New York), pp. 170–177.

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. (**2000**). ''Are DPOAEs a mixture of emissions generated by different mechanisms?'' Assoc. Res. Otolaryngol. Abs. **23**, 480.

Kanis, L. J., and de Boer, E. (**1993**). ''Self-suppression in a locally active nonlinear model of the cochlea,'' J. Acoust. Soc. Am. **94**, 3199–3206.

Kemp, D. T. (**1978**). ''Stimulated acoustic emissions from within the human auditory system,'' J. Acoust. Soc. Am. **64**, 1386–1391.

Kemp, D. T. (**1997**). ''Otoacoustic emissions in perspective,'' in *Otoacoustic Emissions: Clinical Applications*, edited by M. S. Robinette and T. J. Glattke (Thieme, New York), pp. 1–21.

Kemp, D. T. (**1998**). ''Otoacoustic emissions: Distorted echoes of the cochlea's travelling wave,'' in *Otoacoustic Emissions: Basic Science and Clinical Applications*, edited by C. I. Berlin (Singular Group, San Diego).

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.

Kemp, D. T., and Chum, R. A. (**1980**). ''Observations on the generator mechanism of stimulus frequency acoustic emissions—Two tone suppres-

sion,'' in *Psychophysical Physiological and Behavioural Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft University Press, Delft), pp. 34–42.
Kemp, D. T., Brass, D., and Souter, M. (**1990**). ''Observations on simultaneous SFOAE and DPOAE generation and suppression,'' in *Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer, New York), pp. 202–209.
Kim, D. O. (**1980**). ''Cochlear mechanics: Implications of electrophysiological and acoustical observations,'' Hear. Res. **2**, 297–317.
Kimberley, B. P., Brown, D. K., and Eggermont, J. J. (**1993**). ''Measuring human cochlear traveling wave delay using distortion product emission phase responses,'' J. Acoust. Soc. Am. **94**, 1343–1350.
Knight, R., and Kemp, D. T. (**2000a**). ''Separation of 'wave' and 'place' fixed $2f_1-f_2$ DPOAE,'' Assoc. Res. Otolaryngol. Abs. **23**, 987.
Knight, R. D., and Kemp, D. T. (**2000b**). ''Indications of different distortion product otoacoustic emission mechanisms from a detailed f_1,f_2 area study,'' J. Acoust. Soc. Am. **107**, 457–473.
Kummer, P., Janssen, T., and Arnold, W. (**1998**). ''The level and growth behavior of the $2f_1-f_2$ distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss,'' J. Acoust. Soc. Am. **103**, 3431–3444.
Levins, R., and Lewontin, R. C. (**1985**). *The Dialectical Biologist* (Harvard University Press, Cambridge).
Mathews, J., and Walker, R. (**1964**). *Mathematical Methods of Physics* (Benjamin, New York).
Mauermann, M., Uppenkamp, S., van Hengel, P. W. J., and Kollmeier, B. (**1999a**). ''Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio f_2/f_1,'' J. Acoust. Soc. Am. **106**, 3473–3483.
Mauermann, M., Uppenkamp, S., van Hengel, P. W. J., and Kollmeier, B. (**1999b**). ''Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss,'' J. Acoust. Soc. Am. **106**, 3484–3491.
Neely, S. T., Norton, S. J., Gorga, M. P., and Jesteadt, W. (**1988**). ''Latency of auditory brain-stem responses and otoacoustic emissions using tone-burst stimuli,'' J. Acoust. Soc. Am. **83**, 652–656.
Norton, S. J., and Neely, S. T. (**1987**). ''Tone-burst-evoked otoacoustic emissions from normal-hearing subjects,'' J. Acoust. Soc. Am. **81**, 1860–1872.
O'Mahoney, C. F., and Kemp, D. T. (**1995**). ''Distortion product otoacoustic emission delay measurement in human ears,'' J. Acoust. Soc. Am. **97**, 3721–3735.
Papoulis, A. (**1962**). *The Fourier Integral and its Applications* (McGraw-Hill, New York).
Patuzzi, R. (**1996**). ''Cochlear micromechanics and macromechanics,'' in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer, New York), pp. 186–257.
Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (**1991**). ''A review of otoacoustic emissions,'' J. Acoust. Soc. Am. **89**, 2027–2067.
Ren, T., Nuttall, A. L., and Parthasarathi, A. A. (**2000**). ''Quantitative measure of multicomponents of otoacoustic emissions,'' J. Neurosci. Methods **96**, 97–104.
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. (**2000a**). ''Frequency dependence of stimulus-frequency-emission phase: Implications for cochlear mechanics,'' in *Recent Developments in Auditory Mechanics*, edited by H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike (World Scientific, Singapore), pp. 381–387.
Shera, C. A., and Guinan, J. J. (**2000b**). ''Reflection-emission phase: A test of coherent reflection filtering and a window on cochlear tuning,'' Assoc. Res. Otolaryngol. Abs. **23**, 545.
Shera, C. A., and Zweig, G. (**1993a**). ''Noninvasive measurement of the cochlear traveling-wave ratio,'' J. Acoust. Soc. Am. **93**, 3333–3352.
Shera, C. A., and Zweig, G. (**1993b**). ''Order from chaos: Resolving the paradox of periodicity in evoked otoacoustic emission,'' in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 54–63.
Shera, C. A., Talmadge, C. L., and Tubis, A. (**2000**). ''Interrelations among distortion-product phase-gradient delays: Their connection to scaling symmetry and its breaking,'' J. Acoust. Soc. Am. **108**, 2933–2948.
Siegel, J. H., Dreisbach, L. E., Neely, S. T., and Spear, W. H. (**1998**). ''Vector decomposition of distortion-product otoacoustic emission sources in humans,'' Assoc. Res. Otolaryngol. Abs. **21**, 347.
Souter, M. (**1995**). ''Stimulus frequency otoacoustic emissions from guinea pig and human subjects,'' Hear. Res. **90**, 1–11.
Stover, L. J., Neely, S. T., and Gorga, M. P. (**1996**). ''Latency and multiple sources of distortion product otoacoustic emissions,'' J. Acoust. Soc. Am. **99**, 1016–1024.
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., 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.
Ver, I. L., Brown, R. M., and Kiang, N. Y. S. (**1975**). ''Low-noise chambers for auditory research,'' J. Acoust. Soc. Am. **58**, 392–398.
Whitehead, M. L., Stagner, B. B., Martin, G. K., and Lonsbury-Martin, B. L. (**1996**). ''Visualization of the onset of distortion-product otoacoustic emissions and measurement of their latency,'' J. Acoust. Soc. Am. **100**, 1663–1679.
Wilson, J. P. (**1980**). ''Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus,'' Hear. Res. **2**, 233–252.
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.
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.