Level dependence of distortion product otoacoustic emission phase is attributed to component mixing

Carolina Abdala^{a)}

Division of Communication and Auditory Neuroscience, House Research Institute, 2100 West Third Street, Los Angeles, California 90057

Sumitrajit Dhar

Hugh Knowles Center, Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208

Radha Kalluri

Division of Communication and Auditory Neuroscience, House Research Institute, 2100 West Third Street, Los Angeles, California 90057

(Received 23 November 2010; revised 11 March 2011; accepted 14 March 2011)

Distortion product otoacoustic emissions (DPOAEs) measured in the ear canal represent the vector sum of components produced at two regions of the basilar membrane by distinct cochlear mechanisms. In this study, the effect of stimulus level on the $2f_1 - f_2$ DPOAE phase was evaluated in 22 adult subjects across a three-octave range. Level effects were examined for the mixed DPOAE signal measured in the ear canal and after unmixing components to assess level effects individually on the distortion (generated at the f_1, f_2 overlap) and reflection (at f_{dp}) sources. Results show that ear canal DPOAE phase slope becomes steeper with decreasing level; however, component analysis further explicates this result, indicating that *interference* between DPOAE components (rather than a shift in mechanics related to distortion generation) drives the level dependence of DPOAE phase measured in the ear canal. The relative contribution from the reflection source increased with decreasing level, producing more component interference and, at times, a reflection-dominated response at the lowest stimulus levels. These results have implications for the use of DPOAE phase to study cochlear mechanics and for the potential application of DPOAE phase for clinical purposes.

© 2011 Acoustical Society of America. [DOI: 10.1121/1.3573992]

PACS number(s): 43.64.Bt, 43.64.Kc, 43.64.Jb [BLM]

Pages: 3123–3133

I. INTRODUCTION

Distortion product otoacoustic emissions (DPOAEs) are considered dual-source and dual-mechanism responses (Kim, 1980; Talmadge *et al.*, 1998; Shera and Guinan, 1999; Knight and Kemp, 1999, 2000, 2001). This fundamental feature of the DPOAE can engender complications when they are applied to characterize cochlear mechanics or used for diagnostic purposes. The two mechanisms that produce and shape the DPOAE measured in the ear canal have distinct magnitude and phase characteristics as a function of frequency. Using inverse fast Fourier transformation (IFFT) and time windowing, the phase features of the DPOAE components allow for their separation and classification as either distortion, generated at the overlap of f_1 and f_2 traveling waves, or a reflection source produced at $2f_1 - f_2$ (f_{dp}). Each component exhibits unique phase-frequency functions elucidating their distinct mechanistic origins: the distortion source produced by outer hair cell nonlinearities and the reflection source, by physical irregularities fixed in position along the length of the basilar membrane, causing impedance perturbations and backward scattering of wavelets (Shera and Guinan, 1999).

The frequency-dependent phase rotation of each DPOAE source is its most telling characteristic. When recorded using frequency-scaled stimuli, the distortion-source phase is independent of frequency for most of the frequency range, consistent with local cochlear scaling (Shera *et al.*, 2000). Thus, when this distortion source dominates the ear canal signal, the resulting phase-frequency function is invariant. In contrast, the reflection-source phase rotates rapidly as its traveling wave is reflected off of mechanical irregularities in cochlear architecture at fixed loci. When the two component vectors combine in the ear canal, the resulting phase relationship between them can be constructive or destructive, producing oscillations in DPOAE level, termed DPOAE fine structure. The spacing of the DPOAE fine structure roughly follows the phase slope of the reflection source, showing minima as this component cycles 180° out-of-phase with relatively invariant distortion-source phase.

In general, in most normal, healthy ears and exclusively for lower-side band DPOAEs, the distortion source is the dominant contributor; consequently ear canal DPOAE phase is grossly frequency independent; however, component dominance depends on stimulus parameters (Knight and Kemp, 1999, 2000, 2001; Konrad-Martin *et al.*, 2001; Dhar *et al.*, 2002; Stover *et al.*, 1996). For example, contribution from the reflection source increases at f_2/f_1 ratios of <1.1 so that at narrow stimulus frequency ratios, the ear canal DPOAE

^{a)}Author to whom correspondence should be addressed. Electronic mail: cabdala@hei.org

shows more rapid phase rotation, following the reflection-source phase pattern (Knight and Kemp, 1999).

Stimulus level also influences DPOAE component contribution thereby impacting phase behavior of the DPOAE measured in the ear canal. Mauermann and Kollmeier (2004) explored the effectiveness of component-specific input/output functions for predicting audiometric threshold. They reported that the distortion source contribution was typically dominant; however, as stimulus level decreased from 71 (L_1) to 51 dB sound pressure level (SPL), the contribution of the components became more equal. When reflection dominated the response (i.e., showed higher relative magnitude), it determined the phase gradient, producing more rapid phase rotation as a function of frequency. Konrad-Martin and Keefe (2005) measured DPOAEs and stimulus frequency OAEs (SFOAEs) to tone pips and gated tones at two discrete frequencies (2.7 and 4 kHz) for moderate- to high-stimulus levels in adult subjects. Whereas SFOAE latency decreased with increasing stimulus level, DPOAE latency was minimally influenced by level. They suggest that a shift in the relative amplitudes of the DPOAE components explained the noted level dependence.

Defining the influence of stimulus parameters on magnitude and phase characteristics of the DPOAE is relevant to various areas of study. For example, newborns have shown steeper DPOAE phase slope than adults in the apical half of the cochlea (Abdala and Dhar, 2010). Though tempting to interpret this as a cochlear immaturity, the newborn auditory system is also characterized by inefficient energy transmission through the immature middle ear, producing lower driving levels at the infant cochlea for equivalent stimulus levels presented in the ear canal (Keefe *et al.*, 1993; Abdala and Keefe, 2006). Clearly, understanding the effects of stimulus level on DPOAE phase could further elucidate the origin of DPOAE-based age differences which have been hypothesized to be cochlear in nature.

Although previous investigations (as reviewed above) have reported level effects on OAEs phase at isolated frequencies and within a restricted range of stimulus levels, there has not been a targeted study to define the effect over a three-octave frequency range. Here, we extend earlier work to quantify the effect of stimulus level on the $2f_1 - f_2$ ear canal DPOAE phase in normal-hearing adults and, perhaps more importantly, to verify the origin of this effect by examining component interaction.

II. METHODS

A. Subjects

Subjects were 22 young adults ranging in age from 19 to 23 yr with a mean age of 20.9 yr. Four were males and 18 were females. Ten right and 12 left ears were tested. All subjects had audiometric thresholds <20 dB hearing level (HL) from 0.25 to 8 kHz and a negative history of noise exposure or otologic pathology. At the time of test each had a normal tympanogram recorded with a 226 Hz probe tone and a normal otoscopic examination. A screening DP-gram was conducted to ensure a present DPOAE with adequate signal-to-noise ratio (SNR) and to select the optimal ear for testing

(i.e., generally the ear with highest DPOAE levels). Subjects were students at Northwestern University and were tested while seated comfortably on a recliner within a sound-attenuated chamber.

B. Signal processing and procedure

Signal generation and recording were controlled using custom software and run on an Apple Macintosh G4 computer via a MOTU (Mark of the Unicorn, Cambridge, MA) 828 Mk II input/output device (24 bits/44 100 Hz). Stimuli were presented to the subjects' ear canal via ER2 transducers (Etymotic Research, Elk Grove Village, IL). The output of the ER10B+ microphone was pre-amplified and then passed through an analog high-pass filter with 300-Hz cutoff frequency before being digitized by the MOTU and stored on disk. DPOAE recordings were made over $2f_1 - f_2$ frequencies spanning three-octaves between 0.5 and 4 kHz at four primary tone levels with fixed level separation: $L_1 - L_2 = 65-55$, $55-45$, $50-40$, and $45-35$ dB SPL and a fixed f_2/f_1 ratio of 1.22. Primary tones, f_1 and f_2 , were logarithmically swept at 8 s/octave for a total of 24 s per sweep. Six sweeps were collected and averaged per condition.

DPOAE level and phase estimates were obtained using a least-squares-fit algorithm (LSF) (Long *et al.*, 2008) and yielded estimates every 2 Hz in the 0.5–1 kHz frequency range and every 7–9 Hz around the 3–4 kHz range, for a total of ~500 individual data points. The noise floor was estimated after phase-inverting every alternate sweep window. In this implementation of the LSF technique, models for the stimulus tones and DPOAE of interest are created. Signal components are then fitted to these models to minimize the sum of squared errors between the model and the data. The phase was “unwrapped” by sequentially subtracting 360° to all points beyond identifiable discontinuities. The final estimate of DPOAE phase was computed by subtracting $2\phi_1 - \phi_2$, (where $\phi_{1,2}$ are phases of the lower and higher frequency stimulus tones) from $-\phi_{dp}$ (the extracted phase at $2f_1 - f_2$).

C. Calibration

Calibrated stimuli were delivered to each subject compensated for the depth of probe insertion (Siegel, 2009). The depth of probe insertion was determined using the half-wave resonance peak after the ear canal response to a slow chirp was normalized to that recorded in a 50 ft long copper plumbing tube with internal diameter of 7.9 mm (approximating that of the average adult ear canal). The spectrum recorded in the plumbing tube (essentially reflection free) arguably represents the frequency response of the sound source. The spectrum recorded in the ear canal, on the other hand, represents the frequency response of the sound source as well as that of the ear canal. Thus the normalization process reveals the frequency response of the ear canal for any given insertion and allows an accurate estimation of the first half-wave resonance and, hence, of insertion depth. The pressure response recorded for a similar depth of insertion in a standard ear simulator (IEC 60318-4; Brüel and Kjaer 4157; Naerum, Denmark) was used to compensate the

frequency response of the sound sources such that the desired SPL was approximated at the subject's eardrum. The compensation factor was automatically implemented by the signal generation software, from a series of responses previously recorded with the insertion depth of the probe varying in ~ 1 mm steps.

D. Analysis

A SNR of 6 dB was used to eliminate noisy data points. The SNR estimate was made based on three-point median values of DPOAE level and compared against the noise floor. DPOAE fine structure features were not of strong interest in the present study and only analyzed to provide confirmatory evidence of trends more clearly elucidated by component separation. Fine structure maxima and minima were identified based on the first and second derivatives of the DPOAE level function and the relationship between them. Data points where the first derivative was equal to zero were identified as extrema (maximum or minimum) and then further classified as a maximum or minimum based on the second derivative being negative or positive, respectively. Fine structure depth for each period was computed as $20 \times \log_{10}(P_{\max}/P_{\text{avg}(\min)})$, where $P_{\max}$ is the DPOAE amplitude at a maximum and $P_{\text{avg}(\min)}$ is the average DPOAE amplitude of the preceding and following minima. The frequency spacing of fine structure was computed as the ratio $f/\Delta f$, where f is the geometric mean between two adjacent minima and Δf is the frequency separation between them. Estimates of prevalence, spacing, and depth were averaged into 1/3-octave intervals for a total of nine center frequencies. Multivariate analyses of variance (ANOVA) (level $\times$ frequency) were conducted to examine how these features changed as a function of stimulus level.

1. DPOAE phase

The phase versus frequency functions were fit with locally linear loess trend lines to examine and describe the trajectory and slope of phase for the group and capture global phase trends (Cleveland, 1993). Loess is a form of locally weighted scatterplot smoothing that is a modern version of classical linear and nonlinear least squares regression. Simple local models of linear and nonlinear least squares regression are fitted to localized subset of the data and adjacent fits are joined to create the overall fit. The loess fit essentially gives the deterministic portion of the variation in a data set without having to make any presumptions about the global fitting model, *a priori*.

To reduce non-meaningful phase variance, phase data across subjects was normalized to the measured phase at 0.8 kHz by integer values of one cycle. The procedure works because at any one frequency, phase differences of greater than one cycle are not meaningful; in other words a phase of 1/4 cycles is the same as a phase of 5/4 cycles (i.e., $1 + 1/4$ cycles). Here, the important measure is the relative phase accumulation across frequency. Since all frequency points are shifted by the same integer value, the normalization procedure does not affect the frequency-dependent phase accumulation measured in any one subject, but simply reduces

FIG. 1. DPOAE level fine structure in one subject at four stimulus levels. In the group data, the prevalence of DPOAE fine structure increases significantly, spacing narrows and fine structure becomes significantly deeper as stimulus level decreases.

the scatter that comes from the first point. We could have chosen any frequency as the reference point but chose 0.8 kHz because all subjects had good DPOAE data at this frequency.

2. Component separation

A MATLAB-based analysis software (Talmadge *et al.*, 1999) was used to separate the DPOAE distortion- and reflection-source components based on their respective group delays. During IFFT, DPOAE complex pressure measured in the frequency domain was multiplied by a moving Hann window in overlapping 50 Hz steps. The length of the Hann window was adjusted on a logarithmic scale in close approximation to the cochlear frequency map and ranged from 400 Hz at the lowest DPOAE frequency to 930 Hz at the highest, spanning the three-octave test range. Impulse response (IPR) functions were determined for each window and rectangular time-domain filters applied to each IPR to extract the desired delay. A search range of -2 to 10 ms was applied to window the short-latency (distortion) component; and 8–15 ms to window the long-latency (reflection) component. The time-domain window was computed based on the maximum in each IPR and was 6.25 ms in length at the lowest frequencies and 3.23 ms at the highest frequencies. The filtered windows of data were then transformed back to the frequency domain by fast Fourier transform (FFT) and the level and phase of the distortion and reflection components reconstructed.

IFFT-derived DPOAE components were examined in multiple ways. (1) the group phase versus frequency functions for each component were fit with locally linear loess trend lines to capture global phase trends for each DPOAE

component separately. Data at the extreme frequencies were eliminated as they were compromised by edge effects inherent in the time-windowing process; (2) magnitude for both the distortion and reflection components was averaged into 1/3-octave bins and examined at each primary tone level with a three-way ANOVA (component $\times$ center frequency $\times$ level); and (3) a difference score was calculated by subtracting component pressures to provide an index of the relative contribution of sources at each primary tone level.

III. RESULTS

A. DPOAE fine structure features

Figure 1 displays DPOAE fine structure data from one subject at the four primary tone levels. Stimulus level produced changes in fine structure periodicity ($F=3.49$; $p=0.015$), as did frequency. Fine structure spacing became narrower as level decreased, most markedly in the mid-frequency range between 0.9 and 2.2 kHz. This change in spacing suggests a steepening in the slope of phase for the reflection component. Both fine structure prevalence and depth increased significantly with decreasing level ($F=41.58$; $p<0.0001$; $F=22.25$; $p<0.0001$, respectively) consistent with more component interference (and more equivalent contribution) at lower stimulus levels. In the mid- to high-frequencies (~ 2.8 kHz), the mean depth of the fine structure period increased from 3 to 9 dB and the number of fine structure periods increased from 11 to 12 at the highest stimulus level to between 17 and 20 at the lowest level. Thus, fine structure periods became more numerous, deeper, and more narrowly spaced with decreasing stimulus level.

B. DPOAE phase

Twenty subjects provided DPOAE phase data at all four stimulus levels. Two subjects lacked data at 45–35 dB SPL, for a total of 86 averaged DPOAE phase-frequency functions. As noted in the analysis section, a cleaning algorithm was implemented prior to analysis to eliminate points with <6 dB SNR. This data elimination technique left gaps in the phase versus frequency function where noisy data had been eliminated. Here, our objective was to accurately characterize the slope and trajectory of these functions. Large gaps in this curve could influence estimates of phase slope since, in the absence of data, accuracy of interpolation (a step necessary for our chosen analyses) would be compromised; therefore, data were further treated to avoid this bias, as described below.

Of the 22 subjects, 10 showed gaps spanning >25 points (~ 100 Hz at 1 kHz), only in the 50–40 or 45–35 dB SPL level conditions and focused in the frequency range from 0.6 to 0.8 kHz where the noise floor is typically elevated. In these cases, the initial data point for analysis was reset to a frequency after the gap. Twelve of the 86 sweeps had a modified starting point between 0.6 and 0.8 kHz (compared to standard 0.5 kHz initial frequency). The remaining subjects had only small cleaning-related gaps (mean = 6.9 points) in their phase versus frequency functions.

Figure 2 shows the individual phase-frequency functions for all subjects (gray lines) and the loess trend lines (thick black line) superimposed on the individual data. The trend lines drawn through the group data indicate that, as primary levels are decreased from 65–55 to 45–35 dB SPL, the DPOAE shows greater phase accumulation, leading to steeper slope of phase. Also notable is the increased variability of phase slope at low-stimulus levels.

Figure 3 displays the individual phase-frequency functions shown in the far right panel of Fig. 2 to illustrate the various patterns of DPOAE phase across frequency at 45–35 dB SPL. Examining these patterns provides strong evidence of component mixing and further explicates the origin of the noted level dependence of ear canal DPOAE phase depicted in Fig. 2. Three basic phase versus frequency functions were evident: (1) phase that was relatively invariant across frequency with the exception of local oscillations that do not greatly disrupt the overall trend. These are shown in red (an ideal exemplar is highlighted with a thick line for each of three patterns) and are consistent with distortion source dominated phase as they are frequency independent over much of the range; (2) more rapid phase accumulation due to abrupt discontinuities, typically of half or whole cycle intervals, as depicted by the cyan blue lines. The segments of the function between these precipitous drops show a relatively flat phase trajectory. These abrupt discontinuities were typically aligned in frequency with deep minima or troughs in the DPOAE level fine structure, suggesting they were the result of destructive interference between components; and (3) overall rapid rotation of phase (~ 10 cycles/octave) typically initiated beyond 1.5 kHz, suggesting the response is dominated by the reflection-source component which is determining the phase accumulation (black lines). Note that abrupt discontinuities may be present in these black functions as well. However, the segments between these

FIG. 2. Individual phase-frequency functions for all subjects (gray lines) at four stimulus levels. The thick black line is the loess trend line fit to the group data to illustrate the global trend with decreasing stimulus level.

FIG. 3. Individual DPOAE phase-frequency functions from all adult subjects for the lowest stimulus level: 45–35 dB SPL. To illustrate the variability in configuration among subjects, the functions are categorized into three patterns denoted by color. The red lines are functions that have a predominantly flat phase for most of the frequency range; the cyan blue lines have one or more abrupt discontinuity and segments of both invariant and steep phase; the black lines are functions with rapid phase rotation for frequency segments spanning at least 500 Hz. The thick lines highlight one phase-frequency function from each category that exemplifies the dominant pattern.

discontinuities show a significantly steeper slope of phase than those in the family of cyan blue lines.

Most of the functions reflect combinations of these three dominant patterns which can shift within one subject throughout the frequency range. These individual configurations suggest that component interference and the shifting contribution of components play a significant role in determining the level dependence of ear canal DPOAE phase. This likelihood can be further tested by examining the effect of stimulus level on individual distortion- and reflection-source component magnitude and phase.

C. DPOAE unmixing

1. Component magnitude

Figure 4 shows the mean magnitude of distortion- and reflection-source components as a function of primary tone level derived from the 22 subjects. Component level is averaged into 1/3-octave intervals with center frequency (of each averaged bin) parameterized. As noted from Fig. 4, the distortion component magnitude was dominant at the moderate levels (65–55 dB SPL) where the mean difference between components approximated 20 dB. In contrast, mean component magnitude appeared relatively equal at the lowest level, 45–35 dB SPL. The distortion component magnitude decreased steadily as stimulus level decreased while reflection-source magnitude remained essentially invariant. This produces an enhanced relative contribution from the reflection source to the DPOAE measured in the ear canal. These mean data necessarily dilute the inter-subject variability; however, individual magnitude data indicate that, for many

FIG. 4. Mean component magnitude for distortion- (solid line) and reflection-source (dashed line) components at the four stimulus levels. The parameter is frequency (denoted by symbol); the data were averaged into 1/3-octave frequency bins and labeled by the center frequency of the bin. SEM is shown with error bars.

subjects, the reflection-source level exceeded that of the distortion component for portions of the frequency range at 45–35 dB SPL. This is consistent with expectations based on the DPOAE phase data shown in Fig. 3.

Figure 5 presents IFFT component magnitude output from two subjects, subject #16 and subject #31, at the four stimulus levels. Note that the data from subject #16 is used in the chimera simulation that follows. In these two examples, which are representative of the data overall, the magnitude of the distortion component is clearly dominant at high levels. As stimulus level progressively decreases, the reflection magnitude becomes equal to or surpasses distortion magnitude for segments of the frequency range. Overall the component contribution shifts systematically with level, however, it shifts differentially across frequency and uniquely for each of these two subjects (and, indeed, for all 22 subjects). The distinct and idiosyncratic pattern of shifting component contribution is clearly manifest in the individual phase data shown in Fig. 3.

A three-way (component $\times$ primary tone level $\times$ frequency) repeated measures ANOVA showed a significant interaction between component and primary tone level ($F = 90$; $p < 0.0001$) confirming that the distortion and the reflection components were differentially impacted by stimulus level. If a simple difference score between component pressures is calculated, this differences score changes significantly as a function of stimulus level ($F = 90.05$; $p < 0.0001$) again, confirming the changing relationship between components with changes in stimulus level.

2. Component phase

Figure 6 shows DPOAE phase data from all subjects as separate distortion- and reflection-source components. The

FIG. 5. IFFT output displaying component magnitude for two individual subjects (#16 and #31) at the four stimulus levels. The ear canal DPOAE level is shown in red, the distortion source component in cyan blue, and the reflection source in black. Note: Component magnitude data from subject #16 are used in the DPOAE chimera simulation described in Fig. 8.

individual data are shown in gray and the loess trend lines are displayed in thick red and cyan blue lines for each component at the four stimulus levels. The standard error of the mean (SEM) is also displayed by thin colored lines bracketing the trend lines. The SEM was computed by first fitting loess lines using identical parameters on all individual phase curves. Differences in phase accumulation between each component are evident as expected. The distortion source has an overall phase accumulation of less than 2–3 cycles,

whereas the reflection component has a phase accumulation >25 cycles over the same frequency range. There was no significant effect of stimulus level on the distortion-source phase. However, when the two extreme stimulus-level conditions were compared there was a significant effect of level on reflection-source phase. Figure 7 shows phase of the individual distortion- (cyan blue) and reflection-source (red) components for the two extreme stimulus levels: 65–55 and 45–35 dB SPL. It is noteworthy that the change in phase

FIG. 6. Phase-frequency functions for individual distortion- and reflection-source components from all subjects at four levels. The gray lines display individual phase data while the thick cyan blue (distortion) and red (reflection) lines display the loess fits elucidating the global trends. The SEM is shown with thin color lines. In cases of reduced variability, these thin SEM lines are not visible as they are merged with the loess line (e.g., 65–55 and 55–45 dB SPL for distortion emission).

characteristics of the reflection component with level is not equivalent to the level-induced changes observed in the ear canal phase (Fig. 2). Therefore, it does not appear that level dependence of the reflection-source phase alone could produce the level dependence of the ear canal DPOAE phase.

We created a simple simulation to test the above suggestion. The simulation constructs two artificial versions of the total ear canal DPOAE by swapping the phase curves measured at high-stimulus levels (65–55 dB SPL) with those measured at low-stimulus levels (45–35 dB SPL). In other words, the simulation constructs two DPOAE magnitude-phase chimeras. The DPOAE chimeras (P_{HL} and P_{LH}) are given by the following equations:

$$P_{HL} = D_H(f) e^{i\phi_L(f)} + R_H(f) e^{i\phi_L(f)},$$

$$P_{LH} = D_L(f) e^{i\phi_H(f)} + R_L(f) e^{i\phi_H(f)}.$$

The pressure of the distortion- and reflection-source emissions are complex numbers defined by their magnitudes (D and R , respectively) and phases (ϕ and φ , respectively) and denoted by either the subscript H (for 65–55 dB SPL) or L (for 45–35 dB SPL). Data used in the simulation are from subject #16 shown in Fig. 5.

FIG. 7. Mean loess trend lines and SEM for distortion- (cyan blue) and reflection-source (red) phase at the two extreme stimulus levels: 65–55 and 45–35 dB SPL.

FIG. 8. Simulated DPOAE chimeras test the impact of level-dependent variations in phase slope versus relative component magnitudes, on the phase of the total ear canal DPOAE. The P_{HL} chimera in the left column combines the distortion- and reflection-component magnitude measured at the highest stimulus levels [65–55 dB SPL (A) D_H and R_H , respectively] with the phase measured at the lowest stimulus level [45–35 dB SPL (C) ϕ_L and ϕ_L]. The P_{LH} chimera in the right column is constructed with the magnitude measured at low-levels [(B) D_L and R_L ; dashed and solid lines] and the phases measured at high-levels [(D) ϕ_H and ϕ_H ; dashed and solid lines]. Panels (E) and (F) present the total (mixed) DPOAE chimera phase with thick lines. (Actual DPOAE ear canal phase, thin dashed and solid lines, is included for comparison.) As noted in Panel (E), the phase of the chimera resembles the frequency invariant phase of the distortion source even though low-level DPOAE phase was used. In panel (F), even though high-level phase was used, there is component interaction which produces steepened phase slope; at frequencies where $R > D$ (non-shaded regions), the phase is dominated by the distortion source and does not vary markedly with frequency; at frequencies where $R < D$, the phase of the total emission is sloping (reflection-like). The arrow indicates a region of abrupt discontinuity, where the components are equal in magnitude.

Figures 8(A) and 8(B) show the component magnitudes of the chimeras while Figs. 8(C) and 8(D) show the component phase. Figures 8(E) and 8(F) show the phase of the total ear canal DPOAE for each chimera created (thick black line). The phase of the actual high- and low-level (thin dashed and solid lines, respectively) ear canal DPOAE is also included in Figs. 8(E) and 8(F).

The reader will note that the phase slope of the chimera using the low-level phase is similar to the shallow phase of the high-level DPOAE. Conversely, the phase slope of the chimera using the high-level reflection component phase is similar to the steep phase of the low-level DPOAE. There is a slight difference between phases of the chimera in Fig. 8(F) and the actual low-level DPOAE phase. The phase of this chimera is slightly shallower than phase of the actual low-level ear canal DPOAE because the reflection component used to construct the chimera has a shallower phase than the low-level reflection component. Note, however, this slight difference in phase accumulation (~ 2 cycles at 3.5 kHz) cannot account for the ~ 6 –7 cycles of phase difference between the actual low-level and high-level ear canal DPOAE shown in Fig. 2.

The simulation provides strong evidence that the relative magnitude of the two components is the more influential factor in determining phase slope. This is further elucidated in the magnitude data shown in Figs. 8(A) and 8(B). In Fig. 8(A), $D > R$ at all frequencies and consequently, the phase of

the chimera and the actual DPOAE, are relatively invariant across frequency. In Fig. 8(B), the pattern of relative magnitude across frequency is more complicated. In the non-shaded areas (best illustrated < 1.5 kHz), $D > R$ and, again, the phase in these segments is relatively frequency invariant. In the shaded areas, where $R > D$, the phase of both the chimera and the actual ear canal DPOAE accumulates more rapidly. The number and length of segments where $R > D$ plays a prominent role in determining the total phase accumulation. The idiosyncratic pattern of relative component magnitude across subjects (and frequencies) accounts for the inter-subject variability in phase accumulation illustrated in Fig. 3. Finally, where the components are nearly equal in magnitude (see arrow), an abrupt discontinuity is evident in chimera phase. These simulations prove that the stimulus-level-dependent steepening of the total ear canal DPOAE phase originates primarily from level-dependent changes in the relative magnitude of DPOAE components and less so from the level-dependent changes in the phase slope of the individual components.

In summary, DPOAE fine structure becomes more prevalent, manifesting deeper troughs and narrower spacing as stimulus level decreases. The deepening of fine structure is consistent with more equivalent component contribution and the narrowing of fine structure is consistent with the steepening slope of reflection-source phase at lower levels. Although the ear canal DPOAE phase gradient steepens with

decreased primary tone levels, unmixing the DPOAE further explicates this trend, suggesting that it is driven by component interference and shifting component contribution. In contrast to the marked change in phase accumulation of the ear canal DPOAE with stimulus level, the phase of the distortion source appears to be essentially level independent and that of the reflection source shows only a modest level dependence. This modest effect cannot account for the significant effects of stimulus level on the phase of the ear canal DPOAE. Simulations support this conclusion.

IV. DISCUSSION

A. Component mixing

The results of this investigation indicate that stimulus level impacts ear canal DPOAE phase, producing a steeper overall phase as level decreases; however, this steepening can be explained by two factors: (1) level-dependent component interference that produces abrupt, discontinuities in phase (and contributes to rapid phase accumulation) and (2) level-dependent shifts in the relative contribution of the reflection source to the ear canal DPOAE phase. Both factors bias the phase gradient toward steepness. Some steepening of the reflection component phase gradient was observed with decreasing level whereas the slope of phase attributed to the distortion component showed little level dependence. Others have, likewise, reported that the distortion-source phase is essentially independent of stimulus level (Mauermann and Kollmeier, 2004; Konrad-Martin *et al.*, 2001; Konrad-Martin and Keefe, 2005).

If ear canal DPOAE phase is to be used as a metric to study underlying cochlear function, the obvious impact of component mixing and shifting component contribution, in particular, at low stimulus levels, cannot be ignored. At low stimulus levels, when components are of similar magnitude, their destructive interference produces sharp, abrupt discontinuities that are aligned in frequency with deep minima in the level fine structure (shown in Fig. 9, top panel). These minima are a consequence of near-complete cancellation between the two DPOAE components. A corresponding null in DPOAE phase leads to the steep discontinuities noted in the thick cyan blue line of Fig. 3. These discontinuities produce an artifactual phase accumulation unrelated to underlying cochlear mechanics. The other clear trend is that the reflection source becomes the dominant source in the ear canal signal at low-levels where it is elicited most effectively. As noted in the Introduction, reflection phase cycles rapidly due to its initiation at fixed loci along the basilar membrane. This shift in component dominance also leads to an overall steepening of phase (Fig. 3, black lines).

The increased variability in DPOAE phase patterns at low stimulus levels is due to the fact that component interference and shifting dominance occurs idiosyncratically from subject to subject. This may have to do with slight differences in ear canal stimulus level due to variations in probe fit or middle ear transfer functions among subjects. If DPOAE phase is to be linked to aspects of cochlear mechanics or hearing status, stimulus level and its influence on component mixing needs to be carefully considered. If not, the

shifting relative contribution from the sources will potentially obscure the association between DPOAE phase and cochlear mechanics. Clearly, if ear canal DPOAE phase is to be used in the study of cochlear function or assessment of cochlear dysfunction, some form of component separation and analysis is desirable to avoid these confounds.

B. Reflection-source phase

Once separated, the phase of the component at f_{dp} , the reflection component, shows a modest dependence on level, becoming steeper at lower stimulus levels. Consistent with this observation, fine structure spacing became narrower as stimulus level decreased. A level dependence in the phase of reflection emissions is predicted by the theory of linear coherent reflection (Talmadge *et al.*, 1998; Shera and Guinan, 1999). Briefly, the theory associates the phase slope of reflection emissions with the phase slope of the traveling wave transfer function near the peak of the traveling wave where the reflection emissions are thought to arise (Shera and Zweig, 1993; Zweig and Shera, 1995). The theory predicts that delays of the basilar membrane mechanical transfer function are in turn related to underlying sharpness of cochlear tuning, with steeper slope of phase (longer delays) indicating sharper frequency tuning (Shera and Guinan, 2003; Shera *et al.*, 2002, 2010). Considering the reciprocal relationship predicted between reflection emission delay and tuning, both measures should respond in a predictable manner to changes in stimulus level. SFOAE and click-evoked OAEs have, in fact, shown prolonged delays (i.e., steeper slope of phase) and sharper tuning with decreasing stimulus level (Schairer *et al.*, 2006; Sisto and Moleti, 2007).

Alternative interpretations have been posited to explain the noted changes in reflection-source phase with level. Emission generators more basal to the traveling wave peak may be recruited with increasing level (Siegel *et al.*, 2005; Martin *et al.*, 2009), thus impacting the total phase accumulation. Basal DPOAE generators have been hypothesized in humans (Martin *et al.*, 2009) and rabbits (Martin *et al.*, 2010) though they are typically associated with high-level stimulation (≥ 75 dB SPL). It is also possible that reflection-emissions evoked by higher level primary tones contain energy from distortion sources (Talmadge *et al.*, 2000). The consequence of such mixing between distortion- and reflection-type emissions may also result in a level-dependent change in phase slope. A shallowing of phase slope has been hypothesized as the origin of the level dependence of SFOAE phase in guinea pigs (Goodman *et al.*, 2003). The impact of a possible interaction between nonlinearly generated and linearly generated “reflection” emissions would be effectively the same as the mixing between the distortion- and reflection-emission DPOAEs.

C. Maturation of cochlear function

The implications of the present work also extend to studies of cochlear development. The putative break from scaling symmetry, represented by the rapid accumulation of DPOAE phase below ~ 1.5 kHz, has been observed in human adults (Kemp and Brown, 1983; Shera *et al.*, 2000) as well

as newborns (Abdala and Dhar, 2010) and is thought to represent a transition from apical to basal cochlear function. As reviewed in the Introduction, distortion source phase is relatively invariant over most of the frequency range, presumably due to cochlear scale invariance. At the simplest level, scaling symmetry implies that a signal will accumulate the same phase at its characteristic frequency (CF) regardless of where it peaks along the cochlear partition. The phase of the DPOAE is linked to the *relationship* between the two traveling waves, f_1 and f_2 , and thus is determined essentially by the phases of the stimulus tones. When the frequency ratio between the stimulus tones is fixed as they are swept in frequency, the relative phases of the two primary tones do not vary significantly at their CF regions, hence nor does the phase of the ensuing DPOAE. A break from cochlear scaling would manifest as a frequency-dependent change in the phase of the emission.

A recent study (Abdala and Dhar, 2010) has reported that the DPOAE phase gradient (recorded at 65–55 dB SPL) is steeper in newborns than adults at low-frequencies, below the apical-basal transition frequency. This more rapid accumulation of phase and the break from scaling it elucidates could indicate a cochlear immaturity in the apex of newborns. However, an alternative explanation involves immaturity of the middle ear transfer function in newborns. This age-related inefficiency has been estimated to produce approximately 15 dB of attenuation in forward transmission of high-frequency signals through the newborn middle ear (Abdala and Keefe, 2006). Comparable estimates for low-frequency signals are not available.

Clearly, if reduced stimulus levels driving the infant cochlea produce a steeper slope of phase, then this conductive inefficiency might explain the non-adult-like phase slope in the apical half of the newborn cochlea. This developmental question can be directly addressed with the adult data from the present study since the range of stimulus levels presented here encompasses the range of attenuation estimated from middle ear immaturities. The level effect on the low-frequency segment of the phase versus frequency functions was assessed in 14 of 22 adult subjects. Subjects were included if they did not exhibit (a) substantive segments missing in their low-frequency phase-frequency function (due to inadequate SNR) or (b) abrupt phase discontinuities in the frequency range from 0.5 to 1.4 kHz. The upper panel of Fig. 9 illustrates the kind of DPOAE phase discontinuity used as criteria to eliminate subjects. The associated level fine structure is also presented in this panel to show the close correspondence of these phase discontinuities with deep minima, indicating destructive component interference. Only half or whole cycle discontinuities of this nature were eliminated. The middle panel in Fig. 9 illustrates the low-frequency segment of the phase-frequency function for one subject at four stimulus levels.

A linear regression equation was fit to each of these low-frequency segments at the four primary tone levels in 14 subjects and the resulting slope estimates were tested for level effects. The bottom panel of Fig. 9 shows the mean slope and SEM for the four level conditions. A one-way ANOVA found no effect of stimulus level on phase slope

FIG. 9. The top panel presents an example of the type of phase-frequency function eliminated from the targeted analysis of low-frequency phase slope: abrupt, approximately 0.5 or 1 cycle discontinuities corresponding to deep fine structure minima. The middle panel displays representative low-frequency phase data at 4 stimulus levels from 1 of 14 subjects included in this analysis; The bottom panel displays the mean slope of phase for the low-frequency segment between 0.5 and 1.4 kHz and the SEM.

measured from 0.5 to 1.4 Hz ($F = 0.957$; $p = 0.427$). This finding does not support the hypothesis that stimulus levels attenuated (by up to 20 dB) due to an immature middle ear easily explain the noted immaturities in newborn low-frequency phase slope; however, this question certainly merits further scrutiny.

V. CONCLUSIONS

DPOAE phase, while holding promise as an informative index of cochlear mechanics both in the adult and developing peripheral auditory system and a potential clinical index of auditory status, must be carefully considered. Clearly, scrutinizing the impact of parameters such as stimulus level and their influence on component contribution to the mixed, ear canal DPOAE is critical to interpretation. At moderate levels, the distortion-source component tends to dominate the ear canal signal and as such, the phase of the ear canal DPOAE does not change with level; however, at low-stimulus levels, though there is wide variability among subjects, phase is more likely to be strongly influenced by the reflection source, which produces rapid phase accumulation. Here, we quantified these changes across a three-octave frequency range and 20 dB stimulus-level range and verified the origin of the level effect, both by scrutinizing component contribution and by generating DPOAE chimeras. The noted level dependence of DPOAE ear canal phase can be explained only by component interaction.

A particular caution should be heeded if the DPOAE (phase or magnitude) is applied to assess auditory deficits

since each of these components is theorized to elucidate two distinct cochlear properties. It is not clear how the dual mechanisms of DPOAE generation uniquely reflect and/or detect auditory dysfunction. It would be unwise to assume that DPOAEs dominated by a reflection emission (recorded using low-level stimulus tones, for example) would behave like a distortion-source emission recorded with moderate-to high-level tones, in the presence of a given hearing deficit. Until this interesting and relevant clinical question is empirically studied, it will be incumbent upon diagnosticians to be aware of the complex interaction between the chosen clinical protocol and DPOAE component mixing or to select another OAE type less confounded by mixed sources.

ACKNOWLEDGMENTS

This research was supported by Grant No. DC003552 from the National Institutes of Health and by the House Ear Institute. The software for DPOAE data acquisition was developed by Carrick Talmadge and the level calibration routine by Jonathan Siegel. The analysis program for fitting loess trend lines was written by Christopher Shera who also provided valuable comments about this manuscript. Authors would like to acknowledge Abigail Rogers for data collection, Ping Luo for technical support, and Shannon Luckovich for data management.

- Abdala, C., and Dhar, S. (2010). "Distortion product otoacoustic emission phase and component analysis in human newborns," *J. Acoust. Soc. Am.* **127**, 316–325.
- Abdala, C., and Keefe, D. H. (2006). "Effects of middle-ear immaturity on distortion product otoacoustic emission suppression tuning in infant ears," *J. Acoust. Soc. Am.* **120**, 3832–3842.
- Cleveland, W. S. (1993). *Visualizing Data* (Hobart Press, Summit, NJ), pp. 88–101.
- Dhar, S., Talmadge, C. L., Long, G. R., and Tubis, A. (2002). "Multiple internal reflections in the cochlea and their effect on DPOAE fine structure," *J. Acoust. Soc. Am.* **112**, 2882–2897.
- Goodman, S. S., Withnell, R. H., and Shera, C. A. (2003). "The origin of SFOAE microstructure in the guinea pig," *Hear. Res.* **183**, 7–17.
- Keefe, D. H., Bulen, J. C., Arehart, K. H., and Burns, E. M. (1993). "Ear-canal impedance and reflection coefficient in human infants and adults," *J. Acoust. Soc. Am.* **94**, 2617–2638.
- Kemp, D. T., and Brown, A. M. (1983). "An integrated view of mechanical nonlinearities observable from the ear canal," in *Mechanics of Hearing*, edited by D. de Boer and M. A. Viergever (Martinus Nijhoff, The Hague, The Netherlands), pp. 75–82.
- Kim, D. O. (1980). "Cochlea mechanics: Implications of electrophysiological and acoustical observations," *Hear. Res.* **2**, 297–317.
- Knight, R. D., and Kemp, D. T. (1999). "Relationships between DPOAE and TEOAE amplitude and phase characteristics," *J. Acoust. Soc. Am.* **106**(3), 1420–1435.
- Knight, R. D., and Kemp, D. T. (2000). "Indications of different distortion product otoacoustic emission mechanisms from a detailed f_1 , f_2 area study," *J. Acoust. Soc. Am.* **107**(1), 457–473.
- Knight, R. D., and Kemp, D. T. (2001). "Wave and place fixed DPOAE maps of the human ear," *J. Acoust. Soc. Am.* **109**(4), 1513–1525.
- Konrad-Martin, D., and Keefe, D. H. (2005). "Transient-evoked stimulus-frequency and distortion-product otoacoustic emissions in normal and impaired ears," *J. Acoust. Soc. Am.* **117**(6), 3799–3815.
- Konrad-Martin, D., Neely, S. T., Keefe, D. H., Dorn, P. A., and Gorga, M. P. (2001). "Sources of distortion product otoacoustic emissions revealed by suppression experiments and inverse fast Fourier transforms in normal ears," *J. Acoust. Soc. Am.* **109**, 2862–2879.
- Long, G. R., Talmadge, C. L., and Lee, J. (2008). "Measuring distortion product otoacoustic emissions using continuously sweeping primaries," *J. Acoust. Soc. Am.* **124**, 1613–1626.
- Martin, G. K., Stagner, B. B., Fahey, P. F., and Lonsbury-Martin, B. L. (2009). "Steep and shallow phase gradient distortion product otoacoustic emissions arising basal to the primary tones," *J. Acoust. Soc. Am.* **125**, EL85–EL92.
- Martin, G. K., Stagner, B. B., and Lonsbury-Martin, B. L. (2010). "Evidence for basal distortion-product otoacoustic emission components," *J. Acoust. Soc. Am.* **127**, 2955–2972.
- Mauermann, M., and Kollmeier, B. (2004). "Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source," *J. Acoust. Soc. Am.* **116**(4), 2199–2212.
- Schairer, K. S., Ellison, J. C., Fitzpatrick, D., and Keefe, D. H. (2006). "Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears," *J. Acoust. Soc. Am.* **120**, 901–914.
- Shera, C. A., and Guinan, J. J., Jr. (1999). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," *J. Acoust. Soc. Am.* **105**, 782–798.
- Shera, C. A., and Guinan, J. J., Jr. (2003). "Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning," *J. Acoust. Soc. Am.* **113**, 2762–2772.
- Shera, C. A., Guinan, J. J., Jr., and Oxenham, A. J. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3318–3323.
- Shera, C. A., Guinan, J. J., Jr., and Oxenham, A. J. (2010). "Otoacoustic estimation of cochlear tuning: Validation in the chinchilla," *J. Assoc. Res. Otolaryngol.* **11**, 343–365.
- 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.
- Shera, C. A., and Zweig, G. (1993). "Noninvasive measurement of the cochlear traveling-wave ratio," *J. Acoust. Soc. Am.* **93**, 3333–3352.
- Siegel, J. (2009). "Estimating wideband eardrum sound levels in human," in *Abstracts of the Thirty-second Annual Midwinter Research Meeting, ARO Abstracts 32*, Baltimore, Maryland, pp. 11–12.
- Siegel, J. H., Cerka, A. J., Recio-Spinosa, A., Temchin, A. N., van Dijk, P., and Ruggero, M. A. (2005). "Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering," *J. Acoust. Soc. Am.* **118**, 2434–2443.
- Sisto, R., and Moleti, A. (2007). "Transient evoked otoacoustic emission latency and cochlear tuning at different stimulus levels," *J. Acoust. Soc. Am.* **122**, 2183–2190.
- 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., Long, G. R., Tubis, A., and Dhar, S. (1999). "Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions," *J. Acoust. Soc. Am.* **105**, 275–292.
- Talmadge, C. L., Tubis, A., Long, G. R., and Piskorski, P. (1998). "Modeling otoacoustic emission and hearing threshold fine structures," *J. Acoust. Soc. Am.* **104**, 1517–1543.
- Talmadge, C. L., Tubis, A., Long, G. R., and Tong, C. (2000). "Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure," *J. Acoust. Soc. Am.* **108**, 2911–2932.
- 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.