J Am Acad Audiol 17:279–292 (2006)

DPOAE Component Estimates and Their Relationship to Hearing Thresholds

Lauren A. Shaffer*
Sumitrajit Dhar†

Abstract

Distortion product otoacoustic emissions (DPOAEs) recorded in the ear canal are a composite or vector sum of two underlying components. The relationship between hearing thresholds and DPOAE-component level, rather than composite level, has been of recent interest. Two different signal-processing methods, inverse fast Fourier transform (IFFT) with time-windowing and low-pass filtering, were used to obtain estimates of the levels of the two components. Component estimates were then correlated to behavioral thresholds. Improvement in the strength of the correlation was not significant over that of the composite. While the signal processing methods were found to yield similar estimates of the generation component, application of the IFFT with time-windowing method was more complex due to the overlap of the components in the time domain. This time domain variability was observed both within and between subjects. These results highlight the complexities of DPOAE generation and the related difficulties of component separation.

Key Words: Distortion product otoacoustic emissions, hearing threshold

Abbreviations: CF_{dp} = characteristic frequency place of the distortion product; DPOAE = distortion product otoacoustic emission; FFT = fast Fourier transform, IFFT = inverse fast Fourier transform; SFOAE = stimulus frequency otoacoustic emission; sgDPOAE = single generator distortion product otoacoustic emission

Sumario

Las emisiones otoacústicas por productos de distorsión (DPOAE) registradas en el canal auditivo son la resultante o la suma de vectores de dos componentes subyacentes. La relación entre el umbral auditivo y el nivel del componente de las DPOAE, más que el nivel de la resultante, ha recibido reciente interés. Se utilizaron dos métodos diferentes de procesamiento de la señal – transformaciones rápidas inversas de Fourier (IFFT) con una ventana temporal, y filtros de pasa-bajo – para estimar los niveles de los dos componentes. Los estimados de los componentes fueron luego comparados con los umbrales conductuales. La mejoría en la fortaleza de la correlación no fue significativa sobre aquella de la resultante. Mientras que los métodos de procesamiento de la señal mostraron estimados similares del componente generado, la aplicación de la IFFT con un método de ventana temporal fue más complejo debido al traslape de los componentes en el dominio temporal. Esta variabilidad en el dominio temporal se observó tanto entre sujetos como en relación a cada

*Department of Speech Pathology and Audiology, Ball State University; †Department of Communication Sciences and Disorders, Northwestern University

Sumit Dhar, Hugh Knowles Center, Department of Communication Sciences and Disorders, 2240 North Campus Drive, Northwestern University, Evanston, IL 60208; Phone: 847-491-2470; E-mail: s-dhar@northwestern.edu

This research was supported by an NIH/NIDCD grant (R03 DC005692-01A1 to SD) and a grant from the Deafness Research Foundation (SD and LAS).

Part of this work was presented at the annual meeting of the American Auditory Society in Scottsdale, AZ, March 2005.

sujeto particular. Estos resultados destacan la complejidad de la generación de DPOAE y las dificultades relacionadas con la separación de componentes.

Palabras Clave: Emisiones otoacústicas por productos de distorsión, umbral auditivo

Abreviaturas: CF_{dp} = lugar característicos de frecuencia del producto de distorsión; DPOAE = emisión otoacústica por producto de distorsión; FFT = transformación rápida de Fourier; IFFT = transformación rápida inversa de Fourier; SFOAE = emisión otoacústica de la frecuencia del estímulo; sgDPOAE = emisión otoacústica por producto de distorsión de un generador único

The idea of a "single generator" distortion product otoacoustic emission (sgDPOAE) was first proposed by Plinkert et al (1997) when they found that the amplitude of the $2f_1$-f_2 distortion product otoacoustic emission (DPOAE) could be altered by introducing a suppressor tone 25 Hz above the frequency of $2f_1$-f_2. According to the authors, moderate suppressor levels of 50–60 dB SPL "always clamped the DP-amplitude and extinguished the DP-gram fine structure" (Plinkert et al, 1997, p. 909).

The concept of an sgDPOAE arising from suppression of activity at the characteristic frequency place of $2f_1$-f_2 (CF_{dp}) follows logically from a two-source model of DPOAE generation (Talmadge et al, 1998; Mauermann et al, 1999). Energy of the DPOAE arises due to nonlinearities at the region of overlap of the basilar membrane excitation patterns associated with the stimulus tones, f_1 and f_2. The distortion product energy then travels both basally toward the ear canal and apically toward the CF_{dp} where the basilar membrane vibrates maximally to the frequency $2f_1$-f_2. At the CF_{dp} location, the distortion product energy is reflected by random inhomogeneities in the mechanics of the basilar membrane (Zweig and Shera, 1995; Talmadge et al, 1998). The $2f_1$-f_2 DPOAE measured in the ear canal is the vector sum of this reflected energy (or "reflection component") and the basally-traveling energy from the overlap region (the "generation component"). The fine structure pattern arises due to the inherently different phase properties of these two components, such that their interference creates a pattern of amplitude variation (peaks and valleys) that can be seen only when the frequency-resolution of the DPOAE recording is sufficiently high (e.g., 10–50 Hz). For a more thorough discussion of fine structure and models of OAE generation, the reader is referred to Talmadge et al (1998), Mauermann et al (1999), or Shaffer et al (2003). For the purposes of this paper, we use the terms "generator" and "reflection" components primarily to be consistent with our previous work. Our intent is to distinguish the DPOAE components arising from the overlap and CF_{dp} regions, exact mechanisms operational at these regions notwithstanding.

In light of the model described above, it is obvious that complete suppression of the CF_{dp} location will yield a DPOAE resulting from only one component, the generation component. This is what Plinkert et al (1997) refer to as the sgDPOAE. In the case of complete suppression of the CF_{dp}, the fine structure pattern is removed. If, however, suppression of the CF_{dp} is not complete, then residual fine structure remains. Recently, we published a paper (Dhar and Shaffer, 2004) disputing the claim by Plinkert et al (1997) that moderate suppressor levels of 50–60 dB extinguish the fine structure pattern. Rather, we found that there is no "universal" suppressor level that will completely suppress the fine structure pattern at all frequencies.

Talmadge et al (1999) showed that a suppressor tone yields complex effects that depend on the relative amplitudes of the generation and reflection components. These effects include both increased and reduced depth of fine structure.[1] In normal-hearing subjects the relative amplitudes of the generation and reflection components vary across frequency such that in some frequency regions the generation component may have greater amplitude while in other frequency regions the reflection component may dominate (Shaffer et al, 2003). The relative amplitudes of the components will also vary depending on choice of stimulus parameters. When the generation component amplitude is dominant, application of a suppressor tone will reduce the depth of fine structure. When the reflection component amplitude is larger, however, fine structure may initially deepen because partial suppression of the reflection component causes the relative amplitudes of the two components to be closer to equal. Further increasing the level of the suppressor tone will ultimately reduce the contribution of the reflection component, thereby reducing fine-structure depth. It should also be noted that a suppressor tone that is too high in level may have an additional and unwanted effect of suppressing the generation component, due to spread of suppression into the overlap region of the stimulus traveling waves.

Choosing a suppressor level that produces complete suppression of the reflection component is, therefore, not a simple task. We showed that among 20 normal-hearing ears, suppressor tones of 45, 55, and 65 dB SPL could cause fine structure depth to decrease, increase, or remain unchanged (Dhar and Shaffer, 2004). This pattern of effects was variable both across frequency and between subjects, making the clinical application of suppression paradigms so complex as to limit the viability of this approach.

The original clinical appeal of an sgDPOAE was that by removing the amplitude variability associated with fine structure, perhaps DPOAE amplitude would be more robustly correlated with behavioral thresholds, thereby improving the prediction of hearing thresholds from the DP-gram. Unfortunately, this hope has not been realized by our findings. Correlations of behavioral thresholds to the sgDPOAE were not improved when compared to correlations observed in the unsuppressed condition (Dhar and Shaffer, 2004). The results, however, were somewhat ambiguous because no single suppressor level was able to readily yield an sgDPOAE. The question as to whether an sgDPOAE would yield better correlation with hearing thresholds, therefore, remains open.

Suppression of the reflection component is not the only means by which to separate the two components. Inverse fast Fourier transform (IFFT) has also been used successfully to isolate the components (Stover et al, 1996; Kalluri and Shera, 2001; Konrad-Martin et al, 2001). Briefly, IFFT converts a high resolution DP-gram into its time-domain equivalent. Time-windowing is then used to isolate the two components, which due to their different phase properties, peak at different points in time. Errors in the estimation of the two components have been previously described (Kalluri and Shera, 2001) and depend on how completely the components separate in time and how well time-windowing can be applied to resolve the time differences.

Additionally, a low-pass filtering method can be used to isolate the components. Brown et al (1996) applied such a method to first estimate the generator component and then to extract the DPOAE residual (reflection component) from the composite DPOAE using vector subtraction. While they were primarily interested in comparing the magnitude and phase properties of stimulus frequency otoacoustic emissions (SFOAEs) to the DPOAE residual, it is possible to use this method to estimate both the generation and reflection components.

In this paper, we apply the IFFT/time-windowing method, as well as the low-pass filtering method to the unsuppressed data from Dhar and Shaffer (2004). In pursuing the goal of determining whether DPOAE components correlate more strongly to behavioral thresholds, we also ask how easily separable are the two components in normal-hearing ears, and do the IFFT/time-windowing and low-pass filtering methods yield similar estimates of the DPOAE components.

METHODS

Results from this data set have been published previously (Dhar and Shaffer, 2004). With an alternate goal, the data are analyzed with different techniques, to answer questions pertinent to this paper.

Participants

With the approval of and following the guidelines of the Indiana University Human Subjects Committee, pure-tone thresholds and DPOAE recordings were obtained from 20 ears of 10 normal-hearing individuals between 18 and 27 years of age. Participants qualified only if they presented with hearing thresholds better than 20 dB HL (re: ANSI,

2006) between 250 and 8000 Hz, negative history of otologic disease and noise exposure, and normal middle ear function indicated by a type A tympanogram, defined by static compliance between 0.4 and 1.5 cc and peak pressure between ±150 daPa.

Measurements

DPOAE and hearing threshold measurements were made in a double-walled, sound-treated IAC booth, with the subject comfortably seated in a recliner. Hearing thresholds were measured in 2 dB steps using a modified Hughson-Westlake procedure at 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz with signals being presented through Etymotic Research ER3 earphones from a Grason-Stadler Instruments GSI 61 audiometer. Limitations related to DPOAE data acquisition and analysis allowed us to use hearing thresholds information between 1000 and 6000 Hz only, for reasons discussed in the results section.

DPOAE fine structure was recorded using levels of 65 and 55 dB SPL for the low- (L_1) and high-frequency (L_2) stimulus tones with a constant stimulus frequency ratio of 1.22 between f_2 frequencies of 1000 and 8000 Hz. Stimulus pairs were chosen to yield DPOAE data points separated by 0.025 mm on the basilar membrane according to the Greenwood map (Greenwood, 1990). The corresponding frequency resolution was approximately 4 and 18 Hz at DPOAE frequencies of 1000 and 6000 Hz, respectively. For a subset of the experiments in Dhar and Shaffer (2004), a suppressor tone, 25 Hz below the DPOAE frequency, was added at 45, 55, and 65 dB SPL. While the suppression data are not directly presented here, some of the data are used for comparison.

Custom-developed software was used to control signal generation and recording on an Apple Macintosh computer running OS X. Signals were delivered to the subjects' ears from a MOTU 828 eight-channel input/output device after being amplified by TDT HB6 headphone amplifiers via Etymotic Research ER2 tube phones. The output of the tube phones was delivered via the probe assembly of a custom-built, three-port Etymotic Research ER10B microphone. The ear-canal signal recorded by the ER10 microphone assembly was passed through its preamplifier (20 dB gain) and a battery-operated Stanford Research SR560 low-noise voltage amplifier and filter. The gain provided at this stage varied across subjects and was compensated for in the analysis. The signal was also band-pass filtered between 300 and 10,000 Hz at this stage. The output of the SR560 was digitized by the MOTU 828 at 44100 Hz and stored on disk for analysis offline. System distortion measured in a hard-walled cavity, approximately 2 cc in size, was approximately -35 dB SPL or below between the frequencies of interest for the stimulus conditions used in these experiments. Instrumentation used for both hearing threshold and DPOAE measurements were calibrated in a Zwislocki coupler. Voltage calibration for DPOAE recording systems does lead to variations in signal levels at the tympanic membrane (Siegel, 2002). Such variations should not affect the inferences drawn from these data.

Data Analyses

DPOAE level, phase, and noise floor information was extracted from the four-second recordings of the ear-canal sound pressure using a least-squares-fit algorithm (Talmadge et al, 1999). The known relationship between the frequencies of the stimulus tones and the DPOAE frequency allows construction of a sinusoidal model of the DPOAE for every stimulus pair. The data are then fitted to this model to extract the level and phase information. Using this method also allows estimation of the noise floor at the frequency of interest as opposed to using adjacent bins in the more traditional FFT method. The initial estimates of DPOAE level and phase were "cleaned" by eliminating data points where the noise floor was higher than a preset level, which was determined individually for each ear. The rejection threshold was typically around -10 dB SPL for the noisiest subjects. A standard cut-off was not used, in order to preserve the sharp minima in the DPOAE-level fine structure. The "cleaned" DPOAE data were interpolated to give a frequency spacing of 2 Hz, to comply with the demand of essentially continuous data for the inverse FFT analysis used later. This close spacing of data points also allowed the exact matching of DPOAE and hearing-threshold frequencies when comparisons were made between them.

Component Separation

Two methods, an IFFT algorithm combined with time-windowing and moving-window averaging, were used to isolate the DPOAE components from the composite ear-canal signal. The IFFT method is commonly used for this purpose (Stover et al, 1996; Kalluri and Shera, 2001; Dhar et al, 2002; Shaffer et al, 2003), and the procedural details are available elsewhere (Shaffer et al, 2003). Briefly, the frequency domain DPOAE level and phase data are processed through an IFFT algorithm to attain a (pseudo) time domain representation of the DPOAE, allowing separation of the early and late components (generation and first reflection peaks, respectively) through the application of an appropriately sized time window. The time-domain representation of these components is then passed through a regular FFT transform to revert back to frequency-domain representations. The analysis is done in narrow frequency bands in an attempt to maintain a flat phase for the generation component within any given analysis window. Analyses were performed in overlapping windows of 400 Hz with 100 Hz steps in between adjacent windows. Selection of the time cut-off for separation of the components is critical. We chose to determine these time windows on an individual basis for each frequency window within each ear based on visual inspection of the time-domain data. We also examined the effect of applying a standard cut-off across the entire frequency range. Results for such manipulation are presented, and issues related to the choice of an optimum cut-off are discussed.

The second method used to estimate the level of the generation component was adapted after Brown et al (1996). Based on the simple assumption that the level of the generation component varies very slowly as a function of frequency, this method involves obtaining an average of the ear canal DPOAE level over a narrow frequency range. Our implementation of this method involved an overlapping 50 Hz averaging window with 25 Hz overlap between windows. Given the final frequency spacing of 2 Hz, this would be approximately equivalent to the 101-point window used by Brown et al (1996). Using overlapping averaging windows in this fashion essentially low-pass filters DPOAE data presumably resulting in an estimate of the generation component. A comparison of estimates of the generation component from the IFFT analysis and low-pass filtering for one ear from our data pool are presented in Figure 1.

Fine-Structure Depth

Fine-structure depth was averaged for

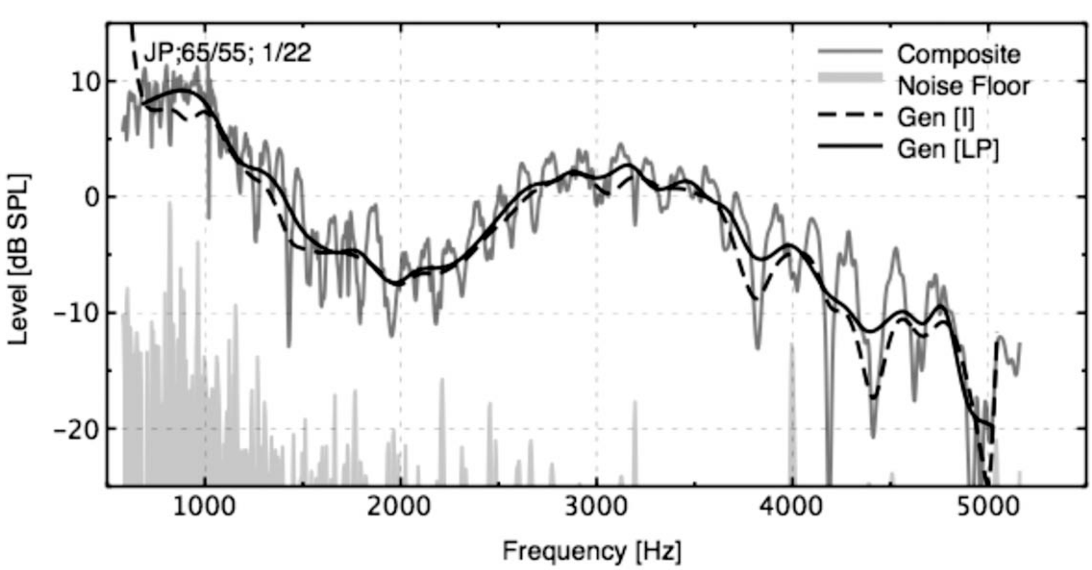

Figure 1. Comparison of estimates of the generator component from the IFFT and moving-window averaging paradigms for subject JP. Note the edge effects at either limit of the frequency range for the IFFT results. In order to avoid contamination from such effects, data points (from the IFFT analysis) over 100 Hz ranges at either extreme of the frequency range were omitted from any subsequent analysis.

three fine-structure periods centered about each hearing-threshold frequency, with depth for each period computed as

$$FS_{depth} = 20 * log_{10}(P_{max}/P_{avg(min)}),$$

where P_{max} is the DPOAE amplitude at a maximum and $P_{avg(min)}$ is the average DPOAE amplitude of the preceding and following minima. This initial computation was done for the composite DPOAE data and repeated for the estimates of the generation component from the IFFT and low-pass methods. Boundaries for fine-structure periods used in the initial analysis were used for the computations made on the estimates of the generation component.

Statistical Analyses

Available estimates of the generation and reflection components from various analysis techniques allowed us to make several comparisons between component levels and hearing thresholds. Pearson's product moment correlation tests were performed between hearing thresholds and the (a) DPOAE level recorded in the ear canal *(composite)*; (b) level of the generation component as estimated using the IFFT analysis *($gen_{[I]}$)*; (c) level of the reflection component (f_2 = hearing threshold frequency) as estimated using the IFFT analysis *($refl_{[I]}$)*; (d) level of the reflection component, when the $2f_1$-f_2 frequency is matched to the hearing-threshold frequency *($reflm_{[i]}$)*; (e) level of the generation component, as estimated using the moving-window averaging (low-pass filtering) *($gen_{[lp]}$)*; and (f) estimate of the generation component using a suppressor at 65 dB SPL *($gen_{[s]}$)* from Dhar and Shaffer (2004). The SAS statistical package was used to perform the analyses. In computing correlation between hearing thresholds and composite or component DPOAE levels in individual ears, we make assumptions of independence of data obtained from two ears within one individual, and normal distribution of these variables about each other.

RESULTS

The results are presented in three general sections. Average data for hearing thresholds, DPOAE levels, and DPOAE-component levels estimated using different techniques are presented first. Fine-structure depth computed for the composite DPOAE-level data and the data processed through IFFT and low-pass filtering is presented next along with results of the correlation analyses. General results of the IFFT analyses and the effect of variation of the time window on estimates of the generation-component level are presented last.

Mean hearing thresholds, DPOAE composite and component levels, along with error bars representing ±1 standard deviation, are presented in Figure 2. Average level-estimates of the reflection component are consistently lower than those of the composite DPOAE or the generation

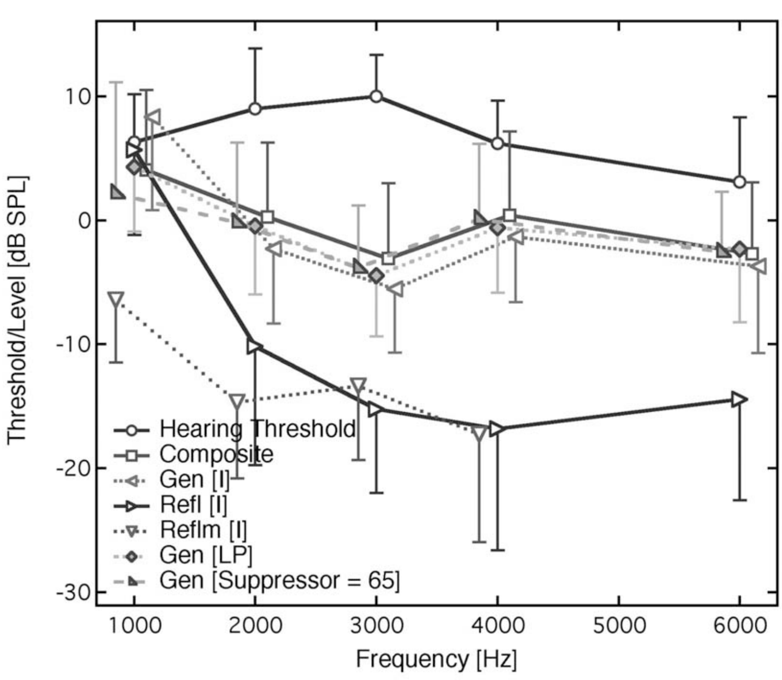

Figure 2. Average hearing thresholds, DPOAE levels, and DP-component levels estimated using various techniques. Error bars represent ±1 standard deviation. Data for the suppression paradigm are from Dhar and Shaffer, 2004.

component. The only exception is observed at 1000 Hz where the average level of the reflection component (when f_2 = hearing-threshold frequency) is approximately equivalent to average level-estimates of the generation component. An estimate of the reflection component, where the DPOAE and hearing threshold frequencies were matched, is not available at 6000 Hz, as the required f_2 was out of the test range. The average level-estimates of the generation component are equivalent for the three (IFFT, low-pass, suppression) analysis paradigms. Data for the suppression paradigm are used from Dhar and Shaffer (2004). The average level of the generation component is also equivalent to that of the composite DPOAE level measured in the ear canal. The variance of the reflection component is larger than that of the generation component.

Average fine-structure depth measured from the composite DPOAE level function and from estimates of the generation component level obtained using the three analysis methods are presented in Figure 3. Error bars represent ±1 standard deviation, and the data for the suppression paradigm are from Dhar and Shaffer (2004). Mean fine-structure depth and variance are greatest for the composite DPOAE level with some reduction observed for the suppression paradigm.

Average fine-structure depth as well as variance is notably reduced for estimates of the generation component obtained using the IFFT and low-pass paradigms. In fact, average depth is roughly equivalent to zero signifying absence of fine structure for both of these analyses types. The reader is reminded that fine-structure boundaries from the composite DPOAE data were used to calculate fine-structure depth from the estimates of the generation component.

Pearson's product moment correlation values for hearing thresholds and various estimates of DPOAE composite and component levels are presented in Table 1. Note that a correlation of -1 is most desired assuming that ears with higher (poorer) hearing thresholds would lead to lower DPOAE component or composite levels. The only significant condition ($p < 0.01$) is observed for correlation between hearing thresholds and composite DPOAE level for 1000 Hz. In general, no consistent pattern of correlation between hearing threshold and composite or component DPOAE level emerges.

As an intermediate step to obtaining estimates of DPOAE component levels, the magnitude and phase of the DPOAE signal

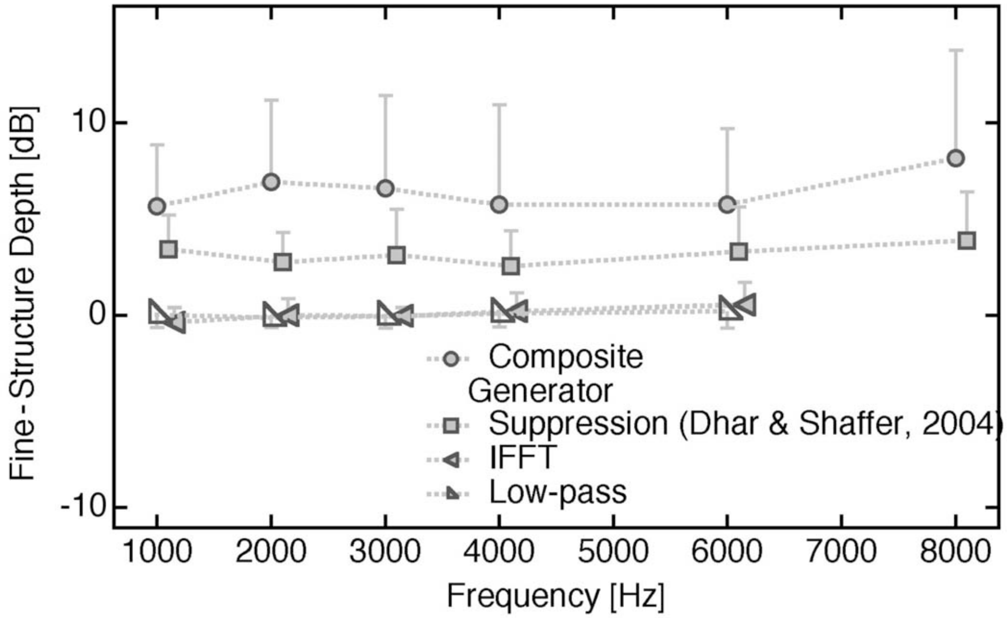

Figure 3. Average fine-structure depth computed from raw DPOAE data and from component estimates using various separation techniques. Error bars represent ±1 standard deviation.

Table 1. Pearson's Product Moment Correlation Computed between Hearing Thresholds and Different Measures of DPOAE Composite or Component Levels

	1000	2000	3000	4000	6000
Composite	-0.5936*	-0.39941	-0.48726	0.31969	-0.03405
$Gen_{[I]}$	-0.44806	-0.45479	-0.37094	0.05192	0.32716
$Gen_{[LP]}$	-0.48254	-0.42032	-0.37761	0.15601	0.3822
$Gen_{[s]}$	-0.31182	0.33084	-0.5632	0.19372	-0.33753
$Refl_{[I]}$	-0.45951	-0.00498	0.02267	0.48242	0.29403
$Reflm_{[I]}$	-0.55634	-0.30913	-0.01742	0.14186	

* $p < 0.01$

Note: "Composite" represents the DPOAE recorded in the ear canal. "$Gen_{[I,LP,S]}$" represents estimates of the generator component using the IFFT, low-pass, and suppression paradigms. The suppression data are from Dhar and Shaffer (2004). "$Refl_{[I]}$" and "$Reflm_{[I]}$" represent estimates of the reflection component with f_2 and $2f_1$-f_2 matched to the threshold frequency, respectively.

recorded in the ear canal is processed through an IFFT algorithm. The output of this stage of the analysis process, from every ear in the data pool, for one frequency window (2600-3000 Hz) is displayed in Figure 4. Subject number and the contributing ear are displayed in each panel. The frequency window was chosen randomly without any *a priori* knowledge of the variety of responses observed. Results in panels 4/R and 8/R could be considered typical of a two-source model, where the two main "lobes" of energy are representative of the generation (early) and reflection (late) components. In panels 4/R and

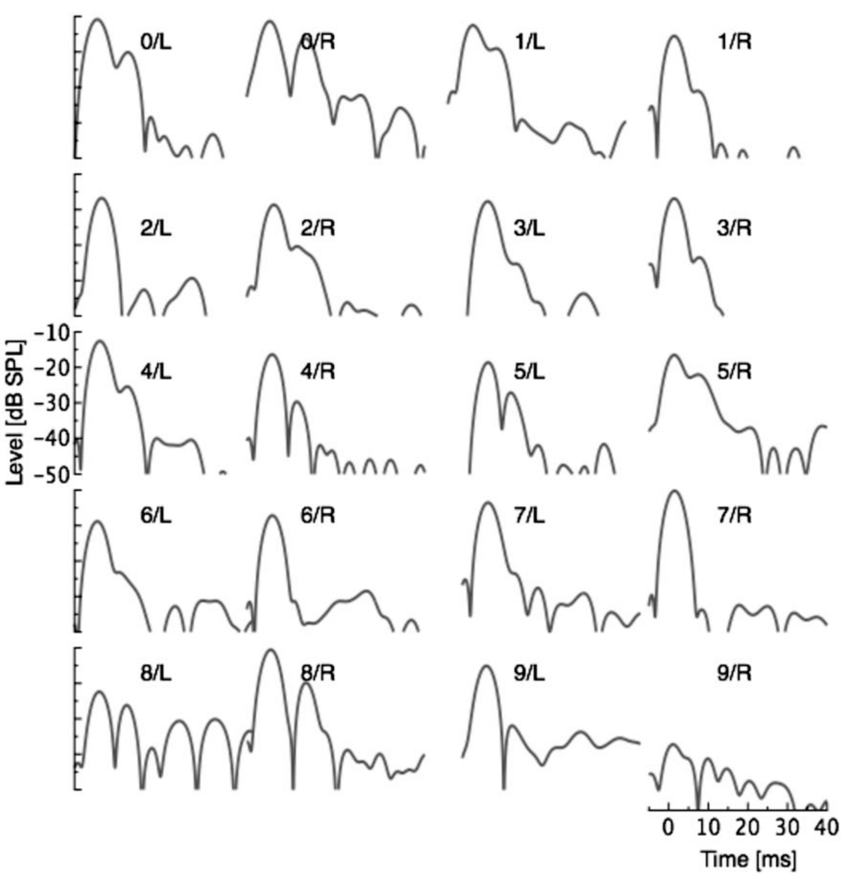

Figure 4. Results of IFFT analyses on data from individual ears for frequencies between 2600 and 3000 Hz for 20 ears. The subject number (0 to 9) and the ear (L/R) are marked in each panel. Each trace is plotted on a time axis (abscissa) between -5 and 40 msec, and a level axis (ordinate) between -50 and -10 dB SPL.

8/R, the clear separation between the two components allows for ready selection of the upper limit of the time window to separate the generation component. However, great variety is observed in the morphology of the time-domain response across individuals and between ears within individuals. The variability and ambiguity in the "cut-off" between the two components is noticeable across ears. This variability, even within one frequency window, motivated us to custom-select the cut-off for each ear within individual frequency windows. The peak-level of the generation component shows less variability across ears as compared to the peak-level of the reflection component. In the extreme, the reflection component is very small or absent in some cases (e.g., 0/L, 6/R, 7/R). The results of using a standard cut-off are presented below with the consequences examined in the "Discussion" section.

The effects of variations in the upper limit of the time window used to isolate the generation component in two ears are presented in Figures 5 and 6. The lighter trace in each panel represents the composite DP-level function with the darker trace representing an estimate of the generation component with a certain cut-off. Estimates of the generation component with increasingly shorter time windows are displayed for 8, 6, 4, and 2 msec windows in the panels from the top to the bottom of the figures. The bottom-most panel displays an additional trace in a dashed line. This is the estimate of the generation component obtained with custom-selected windows. These data were used in our estimation of fine-structure depth and correlation analysis. Elimination of fine structure spreads toward the higher frequencies with the shortening of

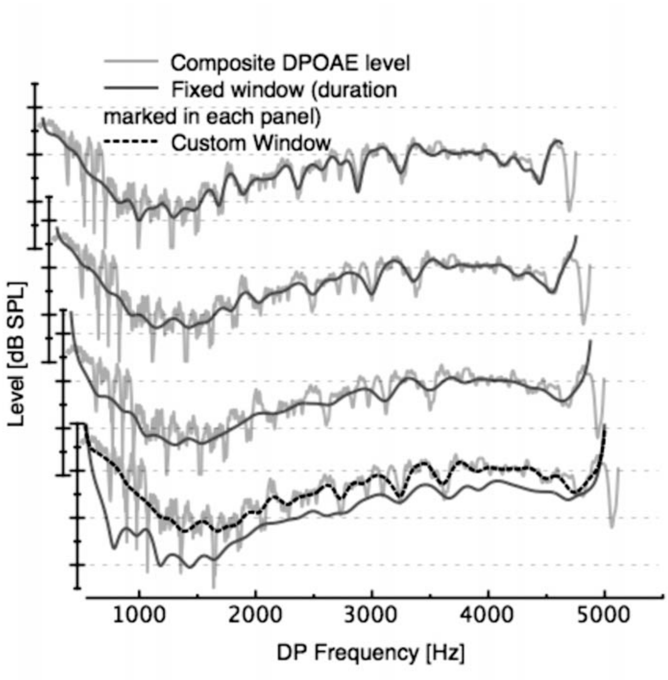

Figure 5. Effects of time-window variation on estimates of the generator component from one subject (KG). Each panel, except the bottom-most, displays two traces. The light trace, used as reference, represents the composite DPOAE level. The darker trace in each panel represents an estimate of the generator component computed using a specific-duration time window. Decreasing window-lengths of 8, 6, 4, and 2 msec were used to generate the data in the panels from the top to the bottom of the figure. The bottom-most panel has an additional trace (dashed line) representing the estimate of the generator component from custom-selected time windows. In this process, we examined the time-domain output of the IFFT analysis in each frequency window and individually selected the upper limit of the time window. The range of the ordinate is constant across panels, but the minimum and maximum vary.

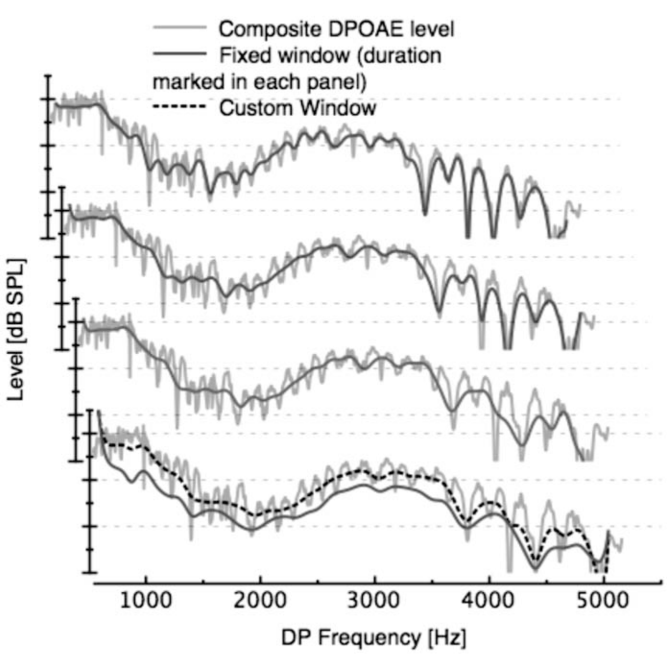

Figure 6. Effects of time-window variation on estimates of the generator component from another subject (JP). The format of this figure is similar to Figure 5.

the time window in both Figures 5 and 6. Fine-structure spacing appears to broaden before disappearance as the time window is shortened as well. A reduction in the overall level of the DPOAE is observed when a 2 msec time window is used. Some "structure," albeit broadly spaced, remains even for the shortest time window or the custom-selected time window. At the very lowest frequencies, some structure reappears for the 2 msec time window.

DISCUSSION

Fine structure has been identified as a potential limitation to the diagnostic utility of DPOAEs (Shaffer et al, 2003; Dhar and Shaffer, 2004; Shera, 2004). The argument that the natural variation in DPOAE level, as a result of the interaction between two DPOAE components, limits the correlation between DPOAE level and hearing thresholds is intuitively appealing. The problem is perhaps made worse by the presence of threshold microstructure, which adds another source of variability. A possible solution would be to isolate the DPOAE components from the overlap and CF_{dp} regions and use the level of the appropriate component as a predictor of hearing status in a certain region of the cochlea. It is with this backdrop and interest in the complexities of isolating DPOAE components from normal-hearing human ears that we present our current work.

In the recent past we have attempted to isolate the DPOAE components using a universal suppressor tone close to the DPOAE frequency (Dhar and Shaffer, 2004). However, our results questioned the viability of such a protocol as we found great variation in the effect of a (fixed-level) suppressor tone across ears and across frequency in a given ear. The possibility exists that the exploration of a greater stimulus-suppressor-parameter matrix could lead to the identification of a stimulus-suppressor combination resulting in better isolation of the generation component. The ideal solution would be the identification of a stimulus condition that biases the ear canal DPOAE signal heavily toward the generation component. A low-level suppressor tone could then be used to eliminate the component from the CF_{dp} region without adversely affecting the generation component. The stimulus space for DPOAE recordings is rather complex, as variables include stimulus-frequency ratio as well as relative stimulus levels. The stimulus-frequency ratio of 1.22, routinely used in clinical protocols, appears to serve the purpose of biasing the ear canal DPOAE signal toward the generation component. A greater bias could perhaps be achieved by using a ratio of 1.26 but at the cost of lowered overall level, thereby adversely affecting signal-to-noise ratio (Dhar et al, 2005). Published data, however limited, also indicate that the composite DPOAE signal in the ear canal has the greatest bias toward the generation component at higher stimulus levels (e.g., Konrad-Martin et al, 2001; Dhar, et al, 2005). Unfortunately, there also is evidence that the sensitivity of DPOAEs to physiological changes in the hearing mechanism diminishes with increasing stimulus levels (e.g., Whitehead et al, 1992). Thus, finding the *ideal* suppressor appears to be a complex problem.

Another method, popularly used in isolation of DPOAE components, is the IFFT and time-window combination (Stover et al, 1999; Kalluri and Shera, 2001; Knight and Kemp, 2001; Konrad-Martin et al, 2001; Dhar et al, 2005). In taking this approach, we asked two questions, (1) How separable are the two DPOAE components in normal-hearing human ears? and (2) Is the single-component DPOAE level a better predictor of hearing status?

Correlation

In our data, correlation between hearing thresholds and DPOAE level does not show any improvement when DPOAE component rather than composite level is used in the analysis. This remains true irrespective of the DPOAE component used in the comparison, namely the generation component estimated using the IFFT method or the low-pass filtering method, or the reflection component, whether or not the DP-frequency is matched to the threshold-frequency. We recognize that the small *N* of this study, as well as the exclusive inclusion of normal-hearing ears, may have limited the extent of the improvement in the magnitude of correlation.

Component Separation

Separation of DPOAE components by IFFT and time-windowing is a three-step procedure. In the first, the frequency-domain DPOAE level and phase data are passed through an inverse FFT algorithm, which results in a time-domain representation of the DPOAE signal. An idealized schematic of the results of this step is presented in Figure 7(a). The two components are represented as independent "lobes" of energy clearly separated in time, with the early and late lobes associated with the generation and reflection components, respectively. Success in separating the two components is integrally tied to the temporal separation between the components at this stage. The rectangular representation of the time window captures the entire generation component without including the reflection component.

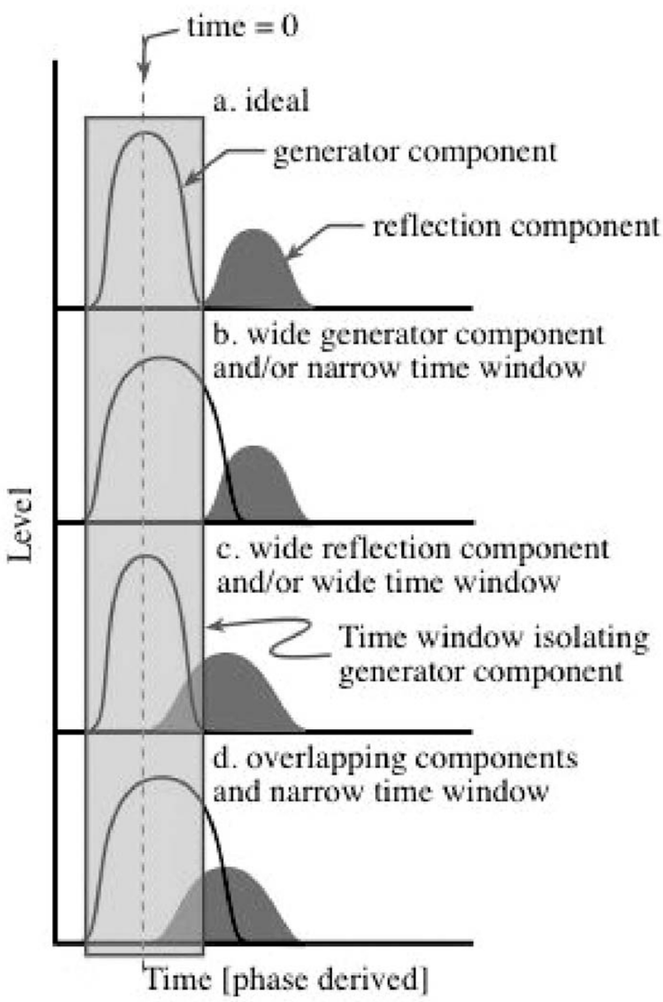

Figure 7. Schematic of possible results of IFFT procedure and the consequence of time-windowing. The abscissa represents a measure of time as derived from the phase slope of frequency-domain DPOAE data. The early and late "lobes" of energy represent the generator and reflection components, respectively. The reflection component is filled while the generator component is not. The rectangular box overlaid across the panels represents the time window used to isolate the generator component.

Examination of Figure 4 shows that such clear separation between the two components is rare at best, even when data from the same frequency range are considered from different ears. The variable that largely controls the results of the IFFT analysis and causes deviation from the idealized result shown in Figure 7(a) is the slope of the phase of the DPOAE components. The time axis in the results of any IFFT analysis represents a measure of the group delay of each DPOAE component. Group delay is defined as the negative of the slope of the phase versus frequency function. Thus, the generation component can be centered on the zero of the time axis if the phase of that component is constant as a function of frequency. Theoretically, constant phase results only if DPOAEs are recorded with a fixed stimulus frequency ratio (f_2/f_1), and if an assumption of cochlear scaling symmetry can be made. Practically, these conditions are satisfied relatively easily by using a fixed-ratio paradigm and by performing the IFFT in small frequency-ranges over which scaling symmetry is a reasonable assumption.

The slope of the phase of the reflection component also determines its position along the abscissa, with a linear phase-frequency relationship resulting in a well-defined lobe of energy. The reflection component has been shown to be equivalent to stimulus frequency OAEs (SFOAEs) recorded under comparable stimulus conditions (Kalluri and Shera, 2001). SFOAE group delays have been demonstrated to be inversely proportional to frequency. More importantly, for our purposes, considerable variability has been demonstrated in SFOAE group delays (Shera and Guinan, 2003) and other reflection source (place fixed) emissions (Goodman et al, 2003). This variability in group delay has the potential of causing overlap between the two components (Figure 7b-d), making them harder to separate.

Choosing a Time Window

The length of the time window used to isolate the generation component determines the accuracy of the separation process. A time window of appropriate duration is expected to "capture" the entire generation component with minimal contamination from the reflection component. This, in turn, should significantly reduce or eliminate fine structure without causing a decrease in the overall DPOAE level. These conditions are perhaps best approximated by the 4 msec window in Figures 5 and 6. Too short a time window (Figure 7b) captures only part of the generation component, thereby resulting in a reduction in the overall DPOAE level. On the other hand, too long a time window causes the estimate of the generation component to be contaminated by part of the reflection component (Figure 7c).

Kalluri and Shera (2001) treat the issue of the appropriate time window in some detail. Their approach is to use a window duration tied to the estimated group delay of an SFOAE at a given frequency. Based on a measured 35% variability about an estimated SFOAE group delay of 15 periods of the stimulus frequency, Kalluri and Shera (2001) use a nine-period long window to isolate the generation component. The use of a period-based window results in a frequency-dependent window duration. The justification for a window duration that varies with frequency is evident in Figures 5 and 6, where fine structure is eliminated at the lowest frequencies for the longest time windows. The "corner frequency" of fine-structure elimination moves toward higher frequencies as the duration of the time window is reduced. This observation is consistent with the inverse proportionality of SFOAE group delay with frequency. Curiously, some fine structure returns at the very lowest frequencies for the 2 msec window.

In our application we customize the selection of window duration further by visually examining the IFFT results for each frequency range in each ear and choosing the best possible window duration. Kalluri and Shera's (2001) application of the nine-period window is efficient in avoiding contamination of the generation component by the reflection component. However, this is achieved at the potential cost of not capturing the entire generation component. This would result in an undesirable reduction in the overall DPOAE level due to a failure to capture the entire generation component (Figure 7b).

The duration of the time window emerges as the most important variable in this process of separating DPOAE components. Unfortunately, the choice of the optimum window duration also emerges as a complex problem to solve. While our protocol of individually choosing the window duration in small frequency ranges in each ear attempts to overcome the problems associated with the variability in group delays of either component, it does little to resolve the other source of complexity—overlap between the two DPOAE components. This appears to be a common occurrence in Figure 4.

Let us assume, for this discussion, that we are able to discern the two components separately even when there is overlap between them—much like in Figure 7b–d. One of three choices in window duration can be made under such circumstances. First, we could choose a window narrow enough to avoid the reflection component (Figure 7b). This would successfully eliminate fine structure but underestimate the generation component. Second, a window wide enough to encompass the entire generation component could be chosen (Figure 7c). Such a window would allow the earliest parts of the reflection component to be included in the estimate. In turn we would see residual fine structure in the estimate of the generation component. The periodicity of the fine structure would be broader (wider frequency spacing of adjacent maxima or minima) as the reflection component causing the interference pattern would have a short delay. This indeed is observed in Figures 5 and 6, with fine structure depth increasing and frequency-spacing decreasing with increasing window duration. Finally, we could choose a time window that is too narrow to encompass the entire generation component but wide enough to include parts of the reflection component (Figure 7d). We would expect residual fine structure with a reduction in the overall DPOAE level, in this case. While Kalluri and Shera (2001) have demonstrated the first two types of errors, they deemed the third possibility hypothetical. We demonstrate this type of error in both Figures 5 and 6, again highlighting the preponderance of cases where there is significant overlap between the two components.

In the second scenario above, we argued that some fine structure, albeit broadly spaced, appears in the generator component due to the inclusion of the "earliest" parts of the reflection component in its estimate. Recall that we computed fine-structure depth for the generator component using fine-structure boundaries from the composite DPOAE data. Thus, such broad fine structure was not accounted for, and our estimate of fine-structure depth is biased from that point of view. While it would be easy to adjust the frequency limits of each fine-structure period when evaluating the level function of the generator component, doing so would neglect the possibility that some variation in the level of the generator component could be arising from sources other than interference with the reflection component—changes in cochlear nonlinearity, as a function of frequency, could be one such source.

As a parallel process, we analyzed the composite DPOAE level data using a moving-window average, in effect low-pass filtering the composite data to arrive at an estimate of the generation component. In the subsequent correlation analysis, these data performed no worse than the data processed through the significantly more complex and involved IFFT algorithm. These preliminary results along with our observation of the prevalent overlap between DPOAE components force us to question the potency of the IFFT process, even when applied to a comparatively small population.

We did not estimate the reflection component by vector subtracting the generator component (estimated using low-pass filtering) from the composite DPOAE data. Such an estimate of the reflection component would include multiple reflections from the CF_{dp} region following internal reflections of both the generator and initial reflection components from the middle ear boundary (Dhar et al, 2002). Because we were interested in estimating the energy specifically from the overlap and CF_{dp} regions, the inclusion of multiple reflections would contaminate the estimate of the reflection component. To deal with this issue, only the first reflection "lobe" was isolated in the IFFT/time-windowing analysis; therefore the reflection component estimates that we report correspond to the first reflection from the CF_{dp} region.

CONCLUSION

In closing, we return to the clinical motivation for separating DPOAE components. In light of the complexities in component isolation, either through the suppression method or the IFFT method, is this a goal worth pursuing? The motivation to isolate the components comes from the theory that fine-structure level variation is partly responsible for the overlap in DPOAE-level distributions of normal-hearing and hearing-impaired populations (Gorga et al, 1997). Any method that could reduce fine structure might aid in separating the two populations. Objectively, DPOAEs have been repeatedly demonstrated in animal models to be extremely sensitive to physiological changes in the cochlea (e.g., Davis et al, 2004; Davis et al, 2005). In this area attempts have been made to relate DPOAE level with permanent threshold shifts (PTS—as measured through evoked potentials) and cochleograms (to measure hair cell loss). The conclusion often is that DPOAEs are more sensitive than PTS measures as they are reflective of morphological changes in outer hair cell populations shy of complete hair cell loss. Experiments in these animal models are "cleaner" as (relatively) controlled lesions can be introduced and their effects examined through direct (e.g., cochleograms) and indirect (e.g., DPOAEs) means. Thus, isolating DPOAE components would appear to hold clinical promise.

As applied to human ears with hearing loss, we are at a disadvantage due to unclear etiologies and overlapping pathologies that are commonplace. A behavioral measure of hearing sensitivity in humans may very well encompass a physiological domain greater than that involved in DPOAE generation. This questions the validity of attempting to correlate behavioral hearing thresholds with any measure of DPOAEs. Indeed there is evidence that individuals classified as "normal hearing" according to their behavioral audiogram may demonstrate lowered DPOAE levels (Zhao and Stephens, 2000). Interestingly, these individuals often demonstrate deficit in other areas of auditory performance, such as speech discrimination in noise. This finding once again suggests that whether or not DPOAE measures can ever be used to predict hearing thresholds, DPOAEs still hold promise as a measure that may be more sensitive than the audiogram for certain cochlear pathologies.

Acknowledgments. We thank Ms. Angela Thomsen for her help with data analyses. Thanks also to two anonymous reviewers for constructive and expeditious reviews.

NOTE

1. Depth is calculated as the ratio of the amplitude of a fine-structure maximum (peak) to the amplitude of the adjacent minima (valleys). The greatest depth of fine structure results when the two components are approximately equal in amplitude such that when the components are in phase they sum creating an amplitude maximum, and when they are 180 degrees out of phase, they cancel, creating an amplitude minimum.

REFERENCES

American National Standards Institute. (2006) *American National Standard Reference Quantities for Acoustical Levels* S1.8-1989 (R2006). New York: American National Standards Institute.

Brown AM, Harris FP, Beveridge HA. (1996) Two sources of acoustic distortion products from the human cochlea. *J Acoust Soc Am* 100:3260–3267.

Davis B, Qiu W, Hamernnik RP. (2004) The use of distortion product emissions in the estimation of hearing and sensory cell loss in noise damaged cochleas. *Hear Res* 41:12–24.

Davis B, Qiu W, Hamernnik RP. (2005) Sensitivity of distortion product otoacoustic emissions in noise-exposed chinchillas. *J Am Acad Audiol* 16:69–78.

Dhar S, Long GR, Talmadge CL, Tubis A. (2005) The effect of stimulus-frequency ratio on distortion product otoacoustic emissions components. *J Acoust Soc Am* 117:3766–3776.

Dhar S, Shaffer LA. (2004) Effects of a suppressor tone on distortion product otoacoustic emissions fine structure: why a universal suppressor level is not a practical solution for obtaining single-generator dp-grams. *Ear Hear* 25:573–585.

Dhar S, Talmadge CL, Long GR, Tubis A. (2002) Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. *J Acoust Soc Am* 112:2882–2897.

Goodman SS, Withnell, RH, deBoer E, Lilly DJ, Nuttall AL. (2003) Cochlear delays measured with amplitude-modulated tone-burst-evoked OAEs. *Hear Res* 188:57–69.

Gorga MP, Neely ST, Ohlrich B, Hoover B, Redner J, Peters J. (1997) From the laboratory to clinic: a large scale study of distortion product otoacoustic emissions in ears with normal hearing and ears with hearing loss. *Ear Hear* 18:440–455.

Greenwood DD. (1990) A cochlear frequency position function for several species—29 years later. *J Acoust Soc Am* 87:2592–2605.

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

Knight RD, Kemp DT. (2001). Wave and place fixed DPOAE maps of the human ear. *J Acoust Soc Am* 109:1513–1525.

Konrad-Martin D, Neely ST, Keefe DH, Dorn PA, Gorga MP. (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.

Mauermann M, Uppenkamp S, van Hengel PWJ, Kollmeier B. (1999) 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.

Plinkert P, Heitmann J, Waldmann B. (1997) Single generator distortion products (sgDPOAE). Precise measurements of distortion product otoacoustic emissions by three tones stimulations. *HNO* 45:909–914.

Shaffer LA, Withnell RH, Dhar S, Lilly DJ, Goodman SS, Harmon KM. (2003) Sources and mechanisms of DPOAE generation: implications for the prediction of auditory sensitivity. *Ear Hear* 24:367–379.

Shera CA. (2004) Mechanisms of mammalian otoacoustic emissions and their implications for the clinical utility of otoacoustic emissions. *Ear Hear* 25:86–97.

Shera CA, Guinan JJ. (2003) Stimulus-frequency-emission group delay: a test of coherent reflection filtering and window on cochlear tuning. *J Acoust Soc Am* 113:2762–2772.

Siegel J. (2002) Calibrating otoacoustic emissions probes. In: Robinette MS, Glattke TJ, eds. *Otoacoustic Emissions: Clinical Applications.* New York: Thieme Medical Publishers, 416–441.

Stover LJ, Neely ST, Gorga MP. (1996) Latency and multiple sources of distortion product emissions. *J Acoust Soc Am* 99:1016–1024.

Stover LJ, Neely ST, Gorga MP. (1999) Cochlear generation of intermodulation distortion revealed by DPOAE frequency functions in normal and impaired ears. *J Acoust Soc Am* 106:2669–2678.

Talmadge CL, Long GR, Tubis A, Dhar S. (1999) Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. *J Acoust Soc Am* 105:275–292.

Talmadge CL, Tubis A, Long GR, Piskorski P. (1998) Modeling otoacoustic and hearing threshold fine structure. *J Acoust Soc Am* 104:1517–1543.

Whitehead ML, Lonsbury-Martin BL, Martin GK. (1992) Evidence for two discrete sources of 2f1-f2 distortion-product otoacoustic emissions in rabbits. II: Differential physiological vulnerability. *J Acoust Soc Am* 92:2662–2682.

Zhao F, Stephens D. (2000) Subcategories of patients with King-Kopetzky syndrome. *Br J Audiol* 34:241–256.

Zweig G, Shera C. (1995) The origins of periodicity in the spectrum of evoked otoacoustic emissions. *J Acoust Soc Am* 98:2018–2047.