Sources and Mechanisms of DPOAE Generation: Implications for the Prediction of Auditory Sensitivity

Lauren A. Shaffer, Robert H. Withnell, Sumit Dhar, David J. Lilly, Shawn S. Goodman, and Kelley M. Harmon

Otoacoustic emissions (OAEs) have become a commonly used clinical tool for assessing cochlear health status, in particular, the integrity of the cochlear amplifier or motor component of cochlear function. Predicting hearing thresholds from OAEs, however, remains a research challenge. Models and experimental data suggest that there are two mechanisms involved in the generation of OAEs. For distortion product, transient, and high-level stimulus frequency emissions, the interaction of multiple sources of emissions in the cochlea leads to amplitude variation in the composite ear canal signal. Multiple sources of emissions complicate simple correlations between audiometric test frequencies and otoacoustic emission frequencies. Current research offers new methods for estimating the individual components of OAE generation. Input-output functions and DP-grams of the nonlinear component of the $2f_2$-f_2 DPOAE may ultimately show better correlations with hearing thresholds. This paper reviews models of OAE generation and methods for estimating the contribution of source components to the composite emission that is recorded in the ear canal. The clinical implications of multiple source components are discussed.

(Ear & Hearing 2003;24;367–379)

Despite research spanning more than 20 yr since the discovery that the inner ear produces sound (Kemp, 1978), as a clinical tool, otoacoustic emissions (OAEs) have not evolved beyond a test of "normal" or "impaired" auditory status (Gorga, Neely, Ohlrich, Hoover, Redner, & Peters, 1997). Why is it that we have not been able to develop an OAE test that can predict auditory thresholds? The correlation of OAE amplitude or OAE thresholds with behavioral or electrophysiological thresholds in nonhuman mammalian species is high, suggesting that the threshold of hearing in mammals should be predictable from OAE amplitude (Brown & Gaskill, 1990; Kossl, 1994). Similar research in humans, however, has been thwarted by the tremendous variability observed in the OAE amplitude of normally hearing humans (Gaskill & Brown, 1990; Kim, Paparello, Jung, Smurzynski, & Sun, 1996; Kimberley, Hernadi, Lee, & Brown, 1994; Nelson & Kimberley, 1992).

The source of OAE amplitude variability in humans is due, in part, to the complexity of their generation. In this paper, the models and experimental findings that underlie our current understanding of the mechanisms and sources of OAE generation are reviewed. The limitations of current clinical testing protocols are explored, and current research methods that may enhance the clinical utility of OAE testing are discussed with emphasis on distortion product otoacoustic emissions (DPOAEs).

Mechanisms of OAE Generation

Evoked otoacoustic emissions can be classified by their underlying mechanisms of generation. Two distinct mechanisms are currently recognized: 1) nonlinear distortion and 2) linear coherent reflection (Shera & Guinan, 1999; Talmadge, Tubis, Long, & Piskorski, 1998; Zweig & Shera, 1995). Nonlinear distortion arises from the action of the cochlear amplifier producing a wave-related mechanical interaction on the basilar membrane and depends on inherent physiological nonlinearities of the cochlear amplifier. Because this mechanism is associated with the traveling wave, it has historically been called a "wave-fixed" phenomenon (Kemp, 1986; Kemp & Brown, 1983).

In contrast, the reflection mechanism, often denoted as a "place-fixed" phenomenon, involves reflection of energy from "inhomogeneities" that are distributed randomly, but fixed in position, along the cochlear partition (Zweig et al., 1995). Although the nature of these inhomogeneities is not known, they are conceptualized as impedance irregularities or spatial corrugations in the cochlear mechanics or anatomy (Zweig et al., 1995). Variation in the number and spacing of outer hair cells in primates has been suggested as an example of a micromechanical irregularity that could lead to reflection of energy

Department of Speech and Hearing Sciences, Indiana University (L.A.S., R.H.W., S.D., S.S.G., K.M.H.), Bloomington, Indiana; and VA National Center for Rehabilitative Auditory Research, Veterans Affairs Medical Center (D.J.L.), Portland, Oregon.

DOI: 10.1097/01.AUD.0000090439.16438.9F

0196/0202/03/2405-0367/0 • Ear & Hearing • Copyright © 2003 by Lippincott Williams & Wilkins • Printed in the U.S.A.

(Lonsbury-Martin, Martin, Probst, & Coats, 1988; Wright, 1984). Alternately, variation in the gain of the OHC active feedback process has also been suggested (Strube, 1989). Such irregularities create reflections from multiple sites that sum with different phases. Only those reflections that sum constructively and arise from the tip of the basilar membrane excitation pattern in an active cochlea will have sufficient amplitude to be recorded in the ear canal as an emission (Zweig et al., 1995).

Both nonlinear distortion and linear coherent reflection are believed to contribute to all evoked otoacoustic emissions, but which mechanism is dominant is both stimulus level and site of origin dependent. At low levels, stimulus frequency otoacoustic emissions (SFOAEs; evoked by a single pure-tone stimulus) and transient evoked otoacoustic emissions (TEOAEs; evoked by a click stimulus or toneburst) are thought to arise predominantly from linear coherent reflection (Shera et al., 1999; Talmadge et al., 1998; Zweig et al., 1995). At higher levels, however, SFOAEs and TEOAEs are believed to arise from a combination of reflection and nonlinear distortion (Long, Shaffer, Dhar, & Talmadge, 2001; Shera et al., 1999, Talmadge, Long, Tubis, & Tong, 2000; Yates & Withnell, 1999). DPOAEs (evoked by two pure tone stimuli) are also a combination of both generation mechanisms (Talmadge, Long, Tubis, & Dhar, 1999).

Sources of OAE Generation

The distinction between "sources" and "mechanisms" is particularly important in attempting to relate measurements of OAEs and auditory sensitivity. In this usage, "source" refers to a cochlear location or region from which energy arises by either mechanism. Integral to the problem of attempting to use OAEs to predict auditory sensitivity is determining whether there is a direct relation between behavioral test frequencies and OAE frequencies, that is, do they represent the same cochlear locations?

Avan, Bonfils, Loth, and Wit (1993) showed that changes in human hearing threshold at 6 to 8 kHz caused a 6 to 20 dB change in the TEOAE amplitude at 1 kHz, despite there being no change in the 1 kHz behavioral threshold. In a later study, Avan, Bonfils, Loth, Elbez, and Erminy (1995) extended these findings to guinea pigs by using acoustic overstimulation to elevate high-frequency compound-action-potential (CAP) thresholds. Elevation of high-frequency CAP thresholds affected TEOAE components at lower frequencies where thresholds remained unchanged. Although these studies do not rule out the possibility of damage to lower frequency regions that were not evident from CAP thresholds, it is also possible that high-frequency regions may contribute to the amplitude of lower frequency TEOAE components. This was further supported by studies showing that TEOAEs in the guinea pig are composed of significant amounts of intermodulation distortion and that TEOAE energy at low frequencies is produced in part by nonlinear stimulus-related interactions at more basal regions of the cochlea (Withnell, Yates, & Kirk, 2000; Yates et al., 1999). The findings of these studies suggest that there is not a direct correspondence between the frequencies of TEOAEs and behavioral threshold test frequencies.

Therefore, TEOAEs, although widely used in the clinic, are arguably the most complex emission with respect to sources and mechanisms of generation. They arise from stimulation of a broad region of the cochlear partition. Any given TEOAE frequency measured in the ear canal may, in fact, represent energy from multiple cochlear locations (Avan et al., 1993, 1995; Avan, Bonfils, Loth, Narcy, & Trotoux, 1991; Withnell et al., 2000; Yates et al., 1999). Although this does not detract from the clinical utility of TEOAEs for determining whether cochlear function is impaired or normal, it may explain why TEOAE amplitudes or thresholds have not held much predictive power for determining auditory sensitivity (Avan et al., 1991, 1993; Wagner & Plinkert, 1999).

In contrast to TEOAEs, SFOAEs arise predominantly from a single cochlear source, the characteristic frequency region of the stimulus tone. At low stimulus levels SFOAEs are believed to be dominated by a single generation mechanism (reflection) (Shera et al., 1999). Stimulus frequency otoacoustic emissions may hold promise as a frequency-specific measure of auditory sensitivity; however, SFOAE testing is not presently available in most clinical instruments.

Distortion product otoacoustic emissions are generated by a two-tone complex that results in the production of distortion products arising from specific regions of the cochlea. Models and experimental findings of DPOAEs have converged suggesting that the cubic distortion product, $2 f_2$-f_2, arises from two distinct cochlear sources (Brown, Harris, & Beveridge, 1996; Gaskill & Brown, 1996; Kemp et al., 1983; Kim, 1980; Kummer, Janssen, & Arnold 1995; Popelka, Osterhammel, Nielsen, & Rasmussen, 1993; Talmadge et al., 1998, 1999). One source is located at the region of overlap* of the stimulus tone

*The term "overlap region" refers to a region of interaction of the two traveling waves that result from the two stimulus tones. The extent of the region from which distortion is generated will depend on the stimulus levels and frequency ratio. Under many stimulus conditions, the region can be approximated as the "region of f2," and is often referred to as such in the OAE literature. We use the term overlap for simplicity.

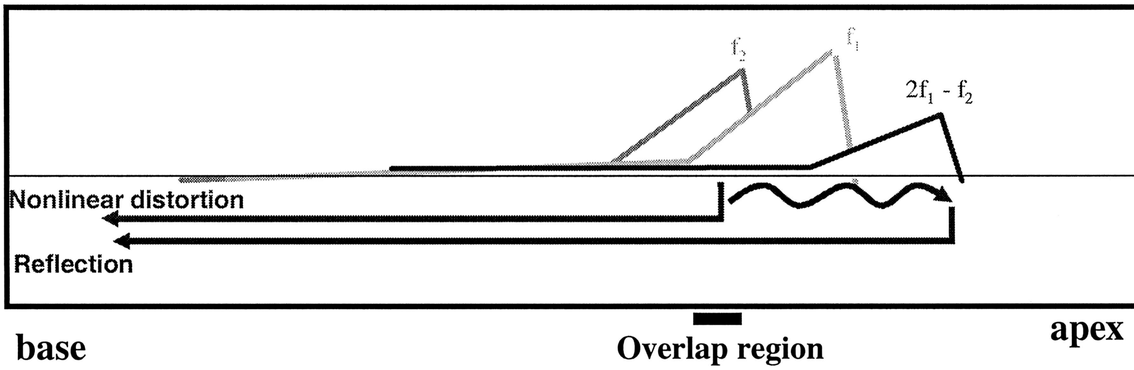


Figure 1. schematic shows the origin of nonlinear distortion at the overlap region* of the stimulus traveling waves, f_1 and f_2. The energy of 2 f_2-f_2 travels basally toward the ear canal and apically where it is reflected at its own characteristic frequency place (CF_{dp}) on the basilar membrane. Reflections from the characteristic frequency of 2 f_2-f_2 interact with 2 f_2-f_2 energy from the overlap region resulting in destructive and constructive amplitude interference. The level of 2 f_2-f_2 recorded in the ear canal is a vector sum of the amplitude and phase interactions of these two sources, which arise from distinct mechanisms: nonlinear distortion (overlap) and reflection (CF_{dp}).

traveling waves and the other is located at the characteristic frequency place of 2 f_2-f_2 (CF_{dp}) (Fig. 1). The remainder of this paper will explore the clinical implications of these two sources, and the possibility that new analysis methods for separating the contributions of the mechanisms of OAE generation may improve DPOAEs as a clinical tool for estimating auditory sensitivity.

Measurement of DPOAES in the Ear Canal

Energy from the two sources interacts to yield the DPOAE amplitude and phase recorded in the ear canal. A different generation mechanism is dominant at each source. Nonlinear distortion predominates at the overlap region where the emission first arises, whereas the CF_{dp} region is dominated by reflection of energy that has traveled apically from the overlap region. On route to the ear canal, energy from the CF_{dp} region sums with energy from the overlap region creating the composite wave that is recorded in the ear canal.

Therefore, the DPOAE measured in the ear canal is a vector sum of two underlying components, both having the frequency of 2 f_2-f_2, but coming from different cochlear locations via different mechanisms of generation. The relative amplitudes and phases of these two components are determined in part by stimulus conditions and also by subject-related factors. When there is little reflected energy, the nonlinear distortion energy from the overlap region is largely responsible for the level of the emission recorded in the ear canal. In this case, DPOAE amplitude plotted as a function of f_2, which approximates the frequency of the overlap region, may reflect the auditory sensitivity of the f_2 location. However, in the case of a significant reflection component from the CF_{dp} region, the 2 f_2-f_2 DPOAE measured in the ear canal will represent auditory sensitivity from two disparate cochlear locations. This may partly explain why attempts to correlate 2 f_2-f_2 amplitude with behavioral thresholds in humans have yielded both high and low correlation coefficients (Gaskill & Brown, 1990, 1993; Kim et al., 1996; Kimberley et al., 1994; Nelson et al., 1992).

DPOAE Amplitude Variation

When fine frequency resolution DPOAE measurements are obtained, a pattern of somewhat regularly spaced amplitude minima and maxima is observed This quasi-periodic amplitude variation, which can vary by as much as 20 dB peak-to-peak, is known as "fine structure" (Fig. 2). Fine structure has been observed in all types of evoked otoacoustic emissions and in fine resolution measurements of behavioral thresholds (Talmadge et al., 1998). Kemp (1979) was the first to recognize that there was similarity in the frequency spacing of the amplitude minima and maxima of OAE fine structure and behavioral threshold fine structure, often called threshold microstructure. This same periodicity was evident in the minimum frequency spacing between adjacent spontaneous otoacoustic emission peaks (Talmadge, Long, Murphy, & Tubis, 1993; Zwicker & Peisl, 1990). The consistency of this spacing suggests that OAE fine structures have a common cochlear origin, presumably from the underlying mechanisms of generation (Talmadge et al., 1998).

DPOAE fine structure arises from the constructive and destructive interference of the components from the overlap and CF_{dp} regions. As the components sum with differing amplitudes and phases, variation is observed in both the amplitude and

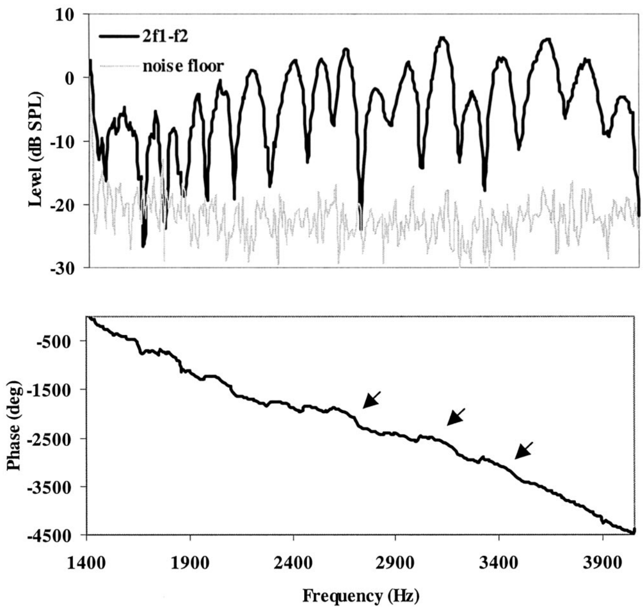


Figure 2. (Top) A sample of quasi-periodic amplitude variation known as DPOAE fine structure from a normal-hearing subject (gg). (Bottom) DPOAE phase for the same subject. The variation in the slope of the phase is indicative of multiple source components. Arrows indicate a "sawtooth" pattern, characteristic of dominance of the component from the overlap region (Talmadge et al., 1999). Frequency ratio was 1.2; stimulus levels were L1 = 60, L2 = 45. For a detailed description of the methods used to collect and analyze DPOAEs, see Talmadge et al. (1999).†

phase of the composite signal recorded in the ear canal. Figure 2 (bottom) shows the phase of 2 f_2-f_2 after subtraction of the phases of the stimulus tones. The variation in the slope of the phase with frequency is indicative of multiple underlying components. The phase pattern can be indicative of shifts in the relative amplitude of the two source components (Talmadge et al., 1999). Using a vector model for the contribution of the two sources, Talmadge et al. (1999) showed that the phase of the 2 f_2-f_2 DPOAE displays a ramp-like pattern when energy from the CF_{dp} region is dominant, and a sawtooth pattern when energy from the overlap region is dominant.

Experimental findings support the theory that interference from multiple sources is responsible for the observed fine structure. Use of a suppressor tone near the frequency of the 2 f_2-f_2 distortion product removes fine structure (Heitmann, Waldmann, Schnitzler, Plinkert, & Zenner, 1998) but does not abolish the emission, suggesting that suppression of the CF_{dp} removes only one of the two sources. In an interesting clinical study, Mauermann, Uppenkamp, van Hengel, and Kollmeier (1999) showed that when a region of mild to moderate hearing impairment encompassed the CF_{dp} region but not the overlap region, fine structure was absent from the measured DPOAE. However, when there was impairment in the overlap region and not the CF_{dp} region, fine structure was present for all cases where there was a measurable DPOAE. This study supports the theory that the emission energy is generated in the overlap region, and that interference of reflected energy from the CF_{dp} results in the fine structure pattern.

DP-Gram and Two Sources of DPOAE

Early research focused on optimizing stimulus parameters to produce a robust 2 $f_{2\text{-}f2}$ DPOAE (Abdala, 1996; Dhar, Long, & Culpepper, 1998; Gaskill et al., 1990; Harris, Lonsbury-Martin, Stagner, Coats, & Martin, 1989; Whitehead, McCoy, Lonsbury-Martin, & Martin, 1995; Whitehead, Stagner, McCoy, Lonsbury-Martin, & Martin, 1995). The clinical applications of these findings include the use of a fixed ratio paradigm where the frequencies of the stimulus tones are held at a ratio (f_2/f_1, where $f_2 > f_1$) of 1.2 and swept across frequency. The optimal stimulus level difference (L_1-$L_{2,}$ in which L_1 is the level of f_1 and L_2 is the level of f_2) is 10 to 15 dB for a moderate level of L_1 (60 dB). The results of the fixed ratio sweep are plotted with 2 f_2-f_2 level as a function of frequency (typically stimulus frequency, e.g., f_2) yielding the clinical "DP-gram."‡

Typical resolution of a clinical DP-gram is ⅓ octave. This resolution is expedient for testing, but is too coarse to observe the characteristic fine structure (Fig. 2). Moreover, fine structure amplitude minima can fall below the range of the normative data used to evaluate clinical test results. Figure 3 shows data for a normal-hearing subject with robust fine structure and large peak-to-peak amplitude variation. Normative data from Gorga et al. (1997) are given for comparison. Two minima fall below the normative range. If these minima were observed in the course of a clinical test, the clinician might conclude that these points were indicative of impaired auditory function.

† All data used to illustrate issues raised in this tutorial were obtained with the approval of the Indiana University Bloomington Campus Committee for the Protection of Human Subjects.

‡The term "DP-gram" is commonly used to refer to the $2f_2$-f_2 DPOAE level versus f_2 frequency function obtained with a fixed stimulus frequency ratio (1.22) during clinical testing. For rapid clinical testing, the resolution of a clinical DP-gram is usually coarse (⅓ octave). Throughout this paper, the term DP-gram is used generically to represent $2f_2$-f_2 DPOAE level plotted as a function of either f_2 or the distortion product frequency. This broad usage includes fixed ratio DPOAE testing at either coarse or fine frequency resolution.

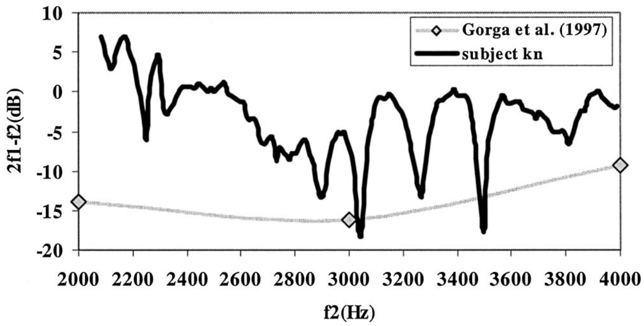


Figure 3. Fine resolution (10 Hz) DP-gram for f_2 frequencies from 2 to 4 kHz for a normal-hearing subject (kn). When compared to normative data (Gorga et al., 1997), it is evident that fine structure amplitude minima (dips) can fall below the level of the normative data. Stimulus levels were moderate (L1 = 65, L2 = 55 dB SPL) and frequency ratio was 1.2. Background noise level was −25 dB SPL.

DPOAEs as Predictors of Auditory Thresholds

Attempts to correlate DPOAE amplitude and behavioral thresholds in normal-hearing humans have yielded varying results with correlation coefficients across studies ranging from −0.94 to 0.83 (Gaskill & Brown, 1990, 1993; Kim et al., 1996; Kimberley et al., 1994; Nelson et al., 1992). The range of observed correlations suggests that, in some individuals amplitude and threshold are highly correlated, whereas in others there is only a weak relationship. Studies in guinea pigs indicate a more robust relationship between DPOAE amplitude and threshold (Brown et al., 1990).

The varied relation between DPOAE amplitude and behavioral thresholds in humans has resulted in research focused on finding criteria that separate ears with normal hearing from those with hearing loss (Kim et al., 1996; Kimberley et al., 1994; Gorga et al., 1997; Gorga, Stover, Neely, & Montoya, 1996; Nelson et al., 1992; Stover, Neely, & Gorga, 1996). Clinical decision theory has been used in the analysis of DPOAEs and audiometric data, resulting in improved criteria for separating normal from impaired ears (for a review see Gorga, Neely, & Dorn, 2002). Still, Gorga et al., (1997) suggest that there is no set of criteria that will unambiguously separate normal and impaired ears due to an overlap in the DPOAE amplitude distributions between ears with normal hearing and those with cochlear hearing impairment.

The presence of fine structure in human DPOAEs has been suggested as a partial explanation for the overlap in response property distributions and the variability of correlations between DPOAE amplitude and behavioral thresholds (Heitmann, Waldmann, & Plinkert, 1996; Talmadge et al., 1999). Interestingly, the fine structure of rodents has reduced peak-to-peak amplitude variation and much larger frequency spacing between amplitude minima (Long et al., 1999; Shaffer, Withnell, & Kirk, 2002), which may explain why statistically significant correlation coefficients have been observed. Of course, fine structure can only be suggested as an explanation for individuals with normal-hearing thresholds or individuals with outer hair cell impairment. In individuals with inner hair cell or auditory nerve dysfunction and normal outer hair cell function (as in auditory neuropathy) in which DPOAE characteristics are unaltered or in individuals with mixed outer and inner hair cell/auditory nerve dysfunction, correlations of OAE amplitude to hearing threshold would be highly variable (Martin, Ohlms, Harris, Franklin, & Lonsbury-Martin, 1990). The prevalence of this type of loss in the general hearing-impaired population is speculative, and therefore the extent to which it contributes to the variability of correlations in large-scale studies of the hearing impaired is not known.

Potential Clinical Applications of Current Research

The DP-Gram • Reducing normal amplitude variation and obtaining separate estimates for the amplitude of the DPOAE generation components could simplify interpretation of the DP-gram. The two components of 2 f_2-f_2 have been separated from each other using two different techniques: 1) a third tone near the 2 f_2-f_2 frequency to suppress the CF_{dp} (Heitmann et al., 1998) and 2) Time-domain windowing of the magnitude of the inverse fast Fourier Transform (IFFT), which has been used to evaluate the time-domain properties of the DP-gram (Kalluri & Shera, 2001) (see the Appendix).

Use of a suppressor tone, close in frequency to the 2 f_2-f_2 DPOAE, suppresses the component from the CF_{dp} region and removes the amplitude variation (Heitmann et al., 1998). Figure 4 shows the amplitude and phase of 2 f_2-f_2 recorded with and without a suppressor tone. Much of the amplitude and phase variation are removed with the addition of the suppressor tone indicating that there is only one source in the suppressed condition, the component from the overlap region.

The addition of a suppressor tone to test protocols could have some clinical benefit, but it is difficult to know a-priori what level of suppressor tone will fully suppress the component from the CF_{dp} region in any given individual. Figures 4 and 5 compare suppression DP-grams, for two different subjects. For both subjects a 55 dB SPL suppressor at 50 Hz below the frequency of 2 f_2-f_2, produced different amounts of suppression. The variation in the amplitude and

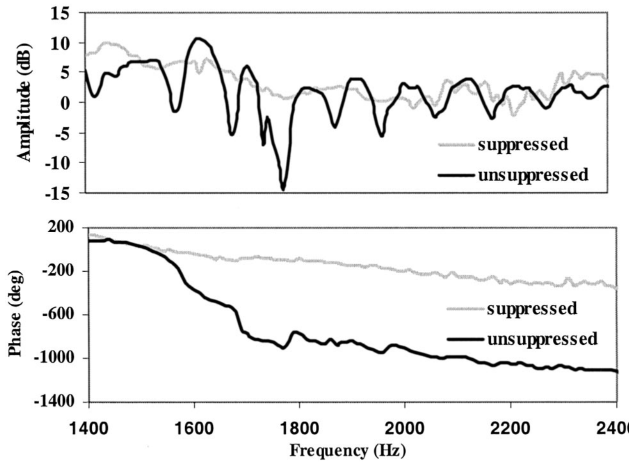


Figure 4. (Top panel) In the unsuppressed condition, fine structure is robust for this normal-hearing subject (ag). When a suppressor tone is added 50 Hz below the frequency of $2f_2$-f_2 the amplitude variation is significantly reduced. Changes are also seen in the phase of the DPOAE (bottom panel). In the unsuppressed condition, variation in the slope of the phase suggests multiple sources; with the addition of the suppressor tone the slope is relatively linear across the frequency range, consistent with a single source with constant group delay. Stimulus ratio was 1.2 and stimulus levels were $L_1 = 60$ and $L_2 = 45$ dB SPL.

phase data in Figure 5 indicates incomplete suppression. In contrast, the lack of amplitude or phase variation in Figure 4 suggests effective suppression of the CF_{dp} region.

The two-mechanism model of DPOAE generation (Shera et al., 1999; Talmadge et al., 1998) suggests that the phases of the distortion product produced by each mechanism will be distinct. In a fixed stimulus ratio paradigm, the stimuli are swept in frequency. Regardless of frequency, the traveling wave that results from a pure tone stimulus has an approximately constant number of cycles to the characteristic frequency place of a stimulus tone. Because of this scaling symmetry, the phase of the traveling wave does not vary as stimulus frequency is swept. Hence, in response to a stimulus complex consisting of two pure tones, the resulting nonlinear distortion, which depends on stimulus traveling wave interactions, also exhibits an approximately constant phase with frequency as its site of origin moves with the stimulus traveling waves.

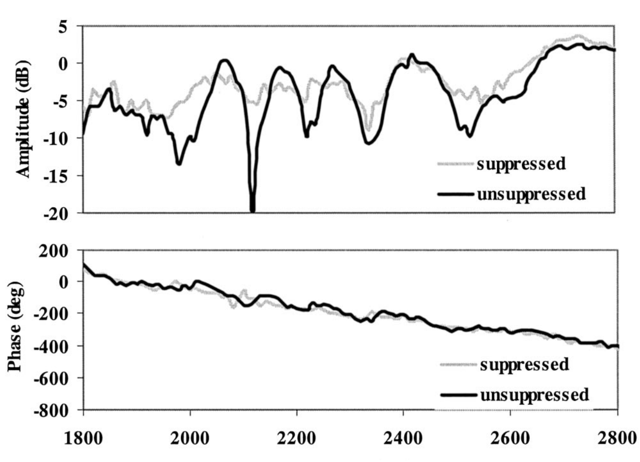


Figure 5. Suppressed and unsuppressed amplitude (top) and phase (bottom) from a normal-hearing subject (ls). For this subject, a 55-dB suppressor does not completely suppress amplitude or phase variation, indicating that there are still multiple source interactions for the suppressed condition. Stimulus levels, frequency ratio, suppressor frequencies and suppressor levels are the same as in Figure 4.

Part of the distortion product energy generated by the nonlinear mechanism in the overlap region travels apically toward the CF_{dp} region. Reflections of this apically traveling wave from place-fixed irregularities in the cochlear partition show rapid phase accumulation. Therefore, in contrast to nonlinear distortion, the phase of the component that arises from the reflection mechanism rotates rapidly with frequency. In fact, the spectral periodicity of the fine structure is a construct of this rapid variation in phase of the CF_{dp} component with frequency (Zweig et al., 1995).

Because the two mechanisms show the different phase properties of their generation, the underlying components of 2 f_2-f_2 can be separated in the time domain. Stover et al. (1996) were the first to use inverse fast Fourier transform analysis to examine the time domain properties of the DP-gram. They suggested that multiple long latency peaks in the time domain were indicative of reflections. Subsequent research has shown that suppression of the CF_{dp} removes long latency peaks from the time domain, and provides additional evidence that long latency peaks represent energy of reflection (Knight & Kemp, 2000; Konrad-Martin, Neely, Keefe, Dorn, & Gorga, 2001). A single study using the IFFT method to look at DPOAEs from normal-hearing and hearing-impaired subjects showed that there are differences in the relative contributions of the two mechanisms between the two groups (Konrad-Martin, Neely, Keefe, Dorn, Cyr, & Gorga, 2002). Attempts to predict differences based on audiogram configurations, however, were only successful in a small subset of subjects.

An example of a DP-gram converted to the time-domain using IFFT is given schematically in Figure 6, and actual data are shown in Figure 7 (see the Appendix for a complete description of the analysis method). The IFFT technique converts the DP-gram into a time-domain response. The multiple peaks seen in the time-domain are indicative of the generation mechanisms. The initial peak near time zero

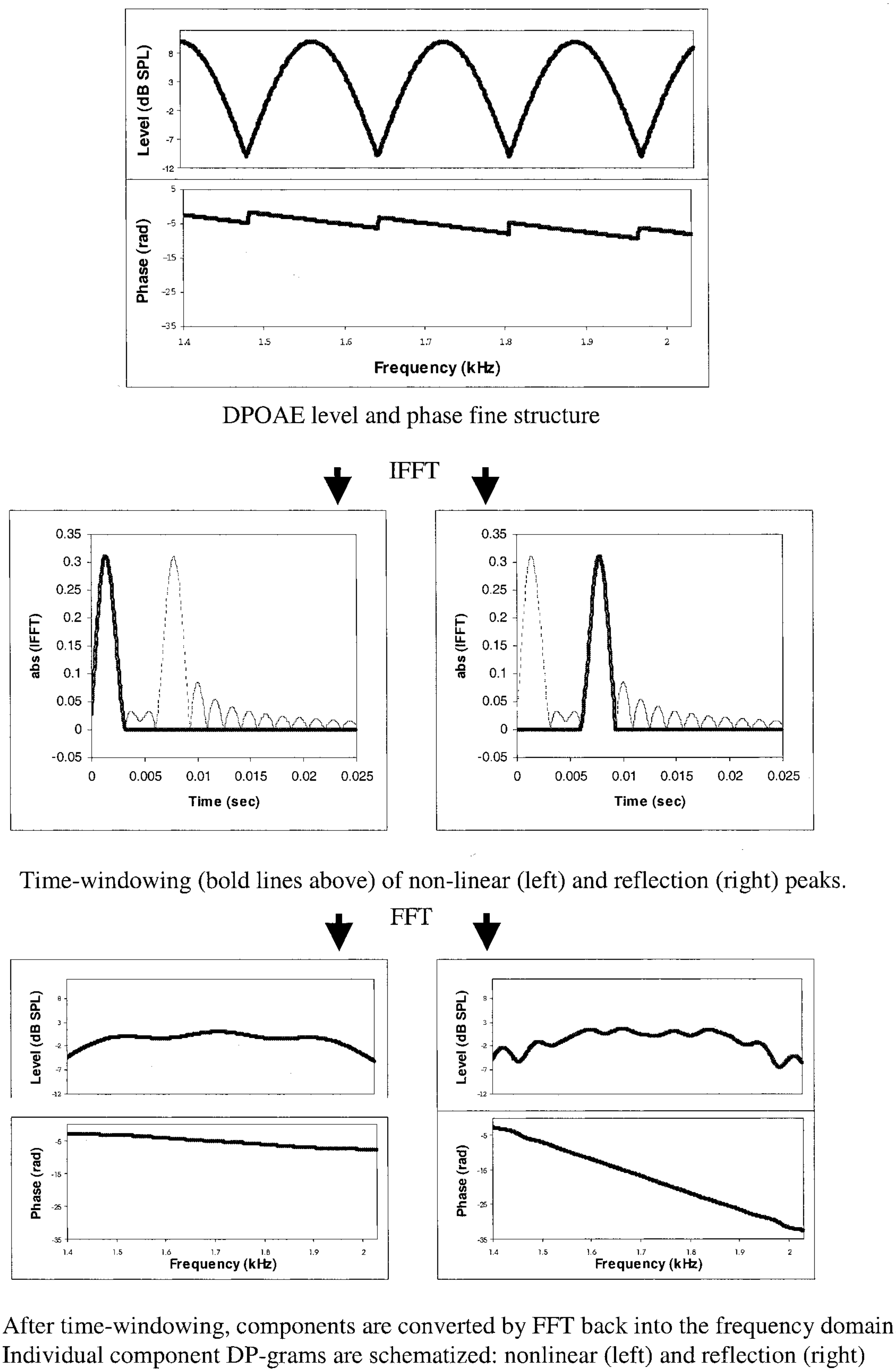


Figure 6. Schematic of the IFFT and time-windowing analysis. Inverse FFT is applied to the level and phase fine structure of the top panel resulting in the time domain representations of the middle panels. Time-windowing of the dominant peaks is schematized in the middle panel (nonlinear distortion component left; reflection component right). The time-windowed data are then converted back into the frequency domain by FFT (bottom panel). For more complete details of the method see the appendix.

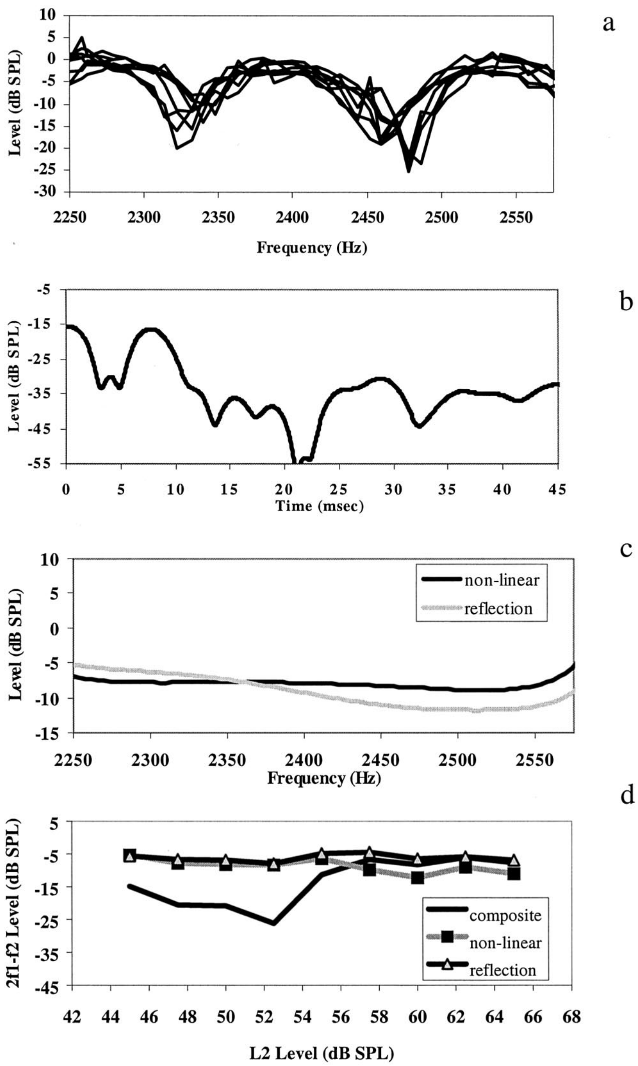


Figure 7. (a) Several DP-grams (L_1 = 55 and L_2 from 47.5 to 65 dB SPL) from a normal-hearing subject (mb) with robust fine structure at a stimulus frequency ratio of 1.3. The time-domain representation of the fine structure data from L_1 = 55 and L_2 = 47.5 is given in the panel (b) as an example. The first peak near time zero is the nonlinear distortion component. The second peak at approximately 7 msec is the reflection component. (c) After time-windowing of the individual peaks, FFT is used to convert the data back into the frequency domain yielding separate DP-grams for each component. When the IFFT and time-windowing analysis is performed on data at all the different stimulus levels an input-output function can be created for each component. Composite, nonlinear and reflection component input-output functions for the level data (a) are given in the bottom panel (d) for the frequency, 2320 Hz.

represents the energy of the response that is a product of nonlinear generation from the overlap region. The second peak represents energy of reflection from the CF_{dp} region. Multiple reflections can occur and would yield additional peaks above the level of the noise floor.

The amplitude and phase of the generation components can be separated by time-domain windowing (filtering) of the individual peaks. Time-windowing is schematized in Figure 6. After the nonlinear and reflection peaks are isolated by filtering, FFT is performed to convert the data back into the frequency domain, effectively giving DP-grams for each component. Figure 7 shows human data to which the IFFT and time-windowing analysis were applied. The top panel (Fig. 7a) shows amplitude fine structure for multiple stimulus levels. When IFFT was applied to one of the functions (L_1 = 55, L_2 = 47.5) from Figure 7a, two dominant peaks result in the time domain (Fig. 7b). Time-windowing was used to isolate each of the peaks and FFT was then performed on this time-windowed data to yield the amplitude spectrum of each component (Fig. 7c).

Kalluri et al. (2001) showed that IFFT analysis and time-windowing of the nonlinear component gives results equivalent to using a suppression paradigm. The IFFT analysis has the advantages of avoiding potential suppression effects on the stimulus tones that can occur with the use of a suppressor tone and ensuring that the f_2 component can be isolated from the reflection component (a suppressor tone may not completely suppress the CF_{dp} emission, as was seen in Fig. 5). This analysis has yet to be applied clinically, and it remains to be seen whether DP-grams of only the nonlinear distortion component will be better predictors of auditory sensitivity than the traditional DP-gram. However, a preliminary study suggests that when a suppression paradigm is used, a statistically significant correlation of the nonlinear component to behavioral threshold is obtained (unpublished observation).

Input/Output Functions • DPOAEs are a by-product of nonlinear mechanical amplification processes within the cochlea. The $2f_2$-f_2 DPOAE shows the same compression with increasing stimulus level as the basilar membrane response (Withnell & Yates, 1998). For the $2f_2$-f_2 DPOAE, when suppression effects are small, fixing the level of the f_1 stimulus and varying the level of the f_2 stimulus provides a DPOAE input-output (I/O) function analogous to the basilar membrane input-output function (Withnell et al., 1998). The slope in the region of compressive growth provides an estimate of cochlear sensitivity and can be indicative of cochlear impairment. Kummer, Janssen, and Arnold (1998) showed that for hearing losses greater than 30 dB SPL, the slope of the I/O function approached unity, indicating a linearization of the basilar membrane response, that is a loss of normal compression.

There have been attempts to correlate the slope or threshold of the I/O function to behavioral thresholds (Boege & Janssen, 2002; Kummer, Janssen, &

Arnold, 1998). Recently, Boege et al. (2002) reported that a DPOAE threshold estimate could be extrapolated by linear regression from semi-log plots of DPOAE pressure as a function of the level of f_2 (using the optimized level paradigm: $L_1 = 0.4L_2 + 39$). When criteria for the correlation coefficient of the linear regression, the slope of the I/O function and the standard error of the DPOAE threshold estimate were applied, 70% of I/O functions met all criteria. Among the data set meeting the criteria for inclusion, correlations of behavioral threshold and DPOAE threshold levels were statistically significant ($r = 0.65$).

One of the problems in correlating I/O function slope to behavioral threshold has been variability in the slope with frequency. The I/O functions of rodents are somewhat less variable with frequency than human I/O functions, but exhibit a characteristic notch at moderate stimulus levels that is suggestive of multiple source interference. Lukashkin and Russell (1999, 2001) and Lukashkin, Lukashkin, and Russell (2002) have shown that a saturating nonlinear model for the mechanoelectrical transduction (MET) of the hair cell bundle can explain 1) notches observed in the rodent I/O function 2) the decrease in DPOAE amplitude seen for low to moderate level stimulus tones at narrow ratios 3) the differential behavior of DPOAEs generated from low and high stimulus levels in response to insults to the efficiency of cochlear amplification, such as the injection of furosemide. These findings argue against multiple source interference theories in rodents. Indeed, the fine structure of rodents has smaller peak-to-peak amplitude, indicating a smaller reflection component (Withnell, Shaffer, Talmadge, 2003). The small amount of amplitude variation coupled with the increased spectral period in rodents, suggests that there should be less variation in the slope of I/O functions with frequency, and that a single saturating nonlinearity may well explain notches in the rodent I/O function.

Unlike rodents, notches are an idiosyncratic and inconsistent feature of human DPOAE functions (Whitehead, 1998). The work of Lukashkin et al. (1999) suggests that the operating point of the nonlinear function that describes MET is situated, in humans, near the point of inflection or maximum sensitivity. At this position, notches in the I/O function are less likely to be observed. Conversely, the operating point in rodents may be positioned further from the point of inflection. This theory is supported by studies showing that there are differential changes in the levels of the quadratic (e.g., f_2-f_1) and cubic distortion products (e.g., $2f_2$-f_2) in the gerbil when the operating point is shifted by acoustical or electrical biasing of the cochlear partition, or by injection of salicylate

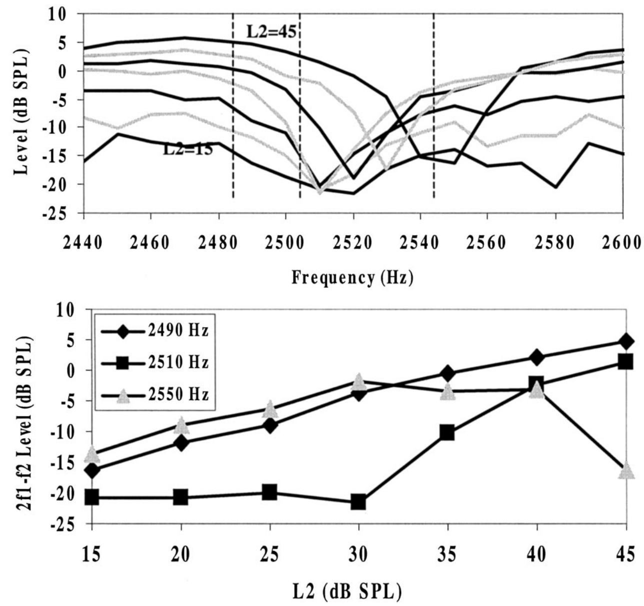


Figure 8. Shift in fine structure with changing L_2 level (15 - 45 in 5 dB steps); L_1 was fixed at 60 dB SPL (top). The frequency of the amplitude minima has shifted for the highest three L_2 levels (35 to 45). In the bottom panel, I/O functions are plotted for three different frequencies taken from the fine structure above (dashed lines). The shape of the I/O function varies depending on its frequency position in the fine structure, illustrating that the shape of the I/O function is impacted by multiple source interactions (Goodman, Shaffer, Dhar, Lilly, & Withnell, 2002).

(Frank & Kossl, 1996, 1997). In fact, these studies raise the interesting question of whether the quadratic distortion product, f_2-f_1, might be a more sensitive indicator of operating point shifts that could occur with cochlear pathology.

Although notches in the I/O function could be partly attributable to cochlear nonlinearities, the contribution of multiple generation components and the resulting fine structure also complicates the measurement of the I/O function. Due to fine structure and stimulus level dependent shifts in fine structure, the slopes of I/O functions can vary dramatically within a small frequency range (He & Schmiedt, 1993). The data in Figure 8 show how I/O functions vary when measurements are taken from different frequencies along a subject's fine structure. I/O functions are plotted for three different frequencies. Notice that the 2.51 kHz frequency is an amplitude minimum (i.e., a dip in the fine structure) for all levels, except the highest level where the minimum has shifted to 2.55 kHz. The I/O functions that result show different shapes, all of which are 'normal' for this subject. Therefore, having insufficient resolution to identify fine structure-related variations challenges the clinical interpretation of I/O functions.

The IFFT analysis method can be used to derive I/O functions for each of the generation components (Fig. 7; Dhar, Talmadge, Harmon, & Tubis, 2002). Figure 7d illustrates the relationship between the composite I/O function recorded in the ear canal and the underlying component I/O functions derived by IFFT analysis. The striking feature of this figure is the difference in level between the overall and component I/O functions. A notch in the overall level of $2f_2$-f_2 for an L_2 level of 55 dB is approximately 20 dB lower than the levels of the component I/Os, indicating the effect of destructive interference between the sources.

SUMMARY

The interactions of multiple sources of OAEs arising by distinct mechanisms complicate the simple interpretation that OAE frequencies and hearing test frequencies assess the integrity of the same cochlear locations. Correlations between hearing threshold and DPOAE level or threshold are likely to be variable not only due to the "mismatch" between the cochlear locations of test frequencies, but also because interactions of multiple sources cause interference that alters the amplitude of the emission recorded in the ear canal. Normative DPOAE amplitude data will tend to "smooth" the amplitude variation as large data sets are averaged. It is, therefore, possible that an individual's fine structure minima may fall below the level of normative data or below the level of the noise floor, and yet still be "normal". Only high-resolution measurements in the frequency region around points that fall below norms will reveal whether such points indicate normal fine structure or cochlear impairment.

Research methods that have employed suppressor tones and IFFT analysis hold the promise of allowing a separate assessment of the "sources" of DPOAE generation. A further advantage of using either an IFFT analysis or a suppression paradigm to derive I/O functions for the nonlinear component is that the slope of the I/O function may then reflect the underlying basilar membrane mechanics, without the complications of multiple component interactions. It remains to be seen whether the slope of the nonlinear distortion component measured in the compressive region can be correlated reliably with hearing threshold, or whether nonlinear distortion component DP-grams are better indicators of hearing sensitivity.

ACKNOWLEDGMENTS:

This work was supported in part by a NIH-NIDCD T32 DC00012 Training Grant (L.A.S. and S.S.G.), by a NIH-NIDCD RO3 DC 04921-01A1 Grant (R.H.W.), by a grant to the National Center for Rehabilitative Auditory Research (RCTR S97-0160) (D.J.L.), by a Merit-Review Award from the Department of Veterans Affairs Rehabilitation Research and Development Service (C2225R) (D.J.L.), by a VA Merit-Review Awards from the VA Rehabilitation Research and Development Service (BR2568R) (D.J.L.) and from the VA Medical Research Service (55034) (D.J.L.)., and by a NSF Science Undergraduate Research Fellowship (K.M.H.). Portions of this paper were presented at the Association for Research in Otolaryngology Midwinter Meeting, St. Petersburg, Florida, February 2002. The authors thank Dr. Paul Avan, Dr. Michael Gorga, Dr. Judith Widen, and one anonymous reviewer for their thoughtful comments on this manuscript. They also thank Hiren Patel for his assistance in the preparation of this document.

Address for correspondence: Lauren A. Shaffer, Ph.D., Department of Speech Pathology and Audiology, Ball State University, Arts and Communications AC 215A, Muncie, IN 47306. E-mail: lashaffer@bsu.edu.

Received August 22, 2002; accepted April 17, 2003

REFERENCES

Abdala, C. (1996). Distortion product otoacoustic emission ($2f_2$-f_2) amplitude as a function of f2/f1 frequency ratio and primary tone level separation in human adults and neonates. *Journal of the Acoustical Society of America*, *100*, 3726–3740.

Avan, P., Bonfils, P., Loth, D., Elbez, M., & Erminy, M. (1995). Transient-evoked otoacoustic emissions and high-frequency acoustic trauma in the guinea pig. *Journal of the Acoustical Society of America*, *97*, 3012–3020.

Avan, P., Bonfils, P., Loth, D., Narcy, P., & Trotoux, J. (1991). Quantitative assessment of human cochlear function by evoked otoacoustic emissions. *Hearing Research*, *52*, 99–112.

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

Boege, P., & Janssen, T. (2002). Pure-tone threshold estimation from extrapolated distortion product emission I/O functions in normal and cochlear hearing loss ears. *Journal of the Acoustical Society of America*, *111*, 1810–1818.

Brown, A. M., & Gaskill, S. A. (1990). Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses. *Journal of the Acoustical Society of America*, *88*, 840–849.

Brown, A. M., Harris, F. P., & Beveridge, H. A. (1996). Two sources of acoustic distortion products from the human cochlea. *Journal of the Acoustical Society of America*, *100*, 3260–3267.

Dhar, S., Long, G. R., & Culpepper, N. B. (1998). The dependence of distortion product $2f_2$-f_2 on primary levels in non-impaired human ears. *Journal of Speech, Language, Hearing Research*, *41*, 1307–1318.

Dhar, S., Talmadge, C. L., Harmon, K. M., & Tubis, A. (2002). A detailed analysis of the effects of primary levels on DPOAE components. *Association for Research in Otolaryngology: Abstracts of the Twenty-fifth Midwinter Meeting*, *767*, 202.

Frank, G., & Kossl, M. (1996). The acoustic two-tone distortions $2f_2$-f_2 and f_2-f_1 and their possible relation to changes in the operating point of the cochlear amplifier. *Hearing Research*, *98*, 104–115.

Frank, G., & Kossl, M. (1997). Acoustical and electrical biasing of the cochlea partition. Effects on the acoustic two-tone distortions f_2-f_1 and $2f_2$-f_2. *Hearing Research*, *113*, 57–68.

Gaskill, S. A., & Brown, A. M. (1990). The behavior of the acoustic distortion product, $2f_2$-f_2, from the human ear and its relation to auditory sensitivity. *Journal of the Acoustical Society of America*, *88*, 821–839.

Gaskill, S. A., & Brown, A. M. (1993). Comparing the level of the acoustic distortion product $2f_2$-f_2 with behavioural threshold audiograms from normal-hearing and hearing-impaired ears. *British Journal of Audiology*, *6*, 397–407.

Gaskill, S. A., & Brown, A. M. (1996). Suppression of human acoustic distortion product: Dual origin of $2f_2$-f_2. *Journal of the Acoustical Society of America*, *100*, 3268–3274.

Goodman, S., Shaffer, L. A., Dhar, S., Lilly, D. J., & Withnell, R. H. (2002). Mechanisms of OAE production in humans. *Association for Research in Otolaryngology: Abstracts of the Twenty-fifth Midwinter Meeting*, *768*, 202.

Gorga, M. P., Neely, S. T., & Dorn, P. A. (2002). Distortion product otoacoustic emissions in relation to hearing loss. In M. S. Robinette & T. J. Glattke (Eds.), *Otoacoustic Emissions: Clinical Applications* (pp. 243–272). New York: Thieme Medical Publishers.

Gorga, M. P., Stover, L., Neely, S. T., & Montoya, D. (1996). The use of cumulative distributions to determine critical values and levels of confidence for clinical distortion product otoacoustic emission measurements. *Journal of the Acoustical Society of America*, *100*, 968–977.

Gorga, M. P., Neely, S. T., 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 & Hearing*, *18*, 440–455.

Harris, F. P., Lonsbury-Martin, B. L., Stagner, B. B., Coats, A. C., & Martin, G. K. (1989). Acoustic distortion products in humans: Systematic changes in amplitudes as a function of f2/f1 ratio. *Journal of the Acoustical Society of America*, *89*, 280–286.

He, N. H., & Schmiedt, R. A. (1993). Fine structure of the 2f1-f2 acoustic distortion product: Changes with primary level. *Journal of the Acoustical Society of America*, *94*, 2659–2669.

Heitmann, J., Waldmann, B., & Plinkert, P. K. (1996). Limitations in the use of distortion product otoacoustic emissions in objective audiometry as the result of fine structure. *European Archives of Otorhinolaryngology*, *253*, 167–171.

Heitmann, J., Waldmann, B., Schnitzler, H. P., Plinkert, P. K., & Zenner, H. P. (1998). Suppression of distortion product otoacoustic emissions (DPOAE) near 2f1-f2 removes DP-gram fine structure: Evidence for a second generator. *Journal of the Acoustical Society of America*, *103*, 1527–1531.

Kalluri, R., & Shera, C. A. (2001). Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation. *Journal of the Acoustical Society of America*, *109*, 622–637.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America*, *64*, 1386–1391.

Kemp, D. T. (1979). The evoked cochlear mechanical response and auditory microstructure: Evidence for a new element in cochlear mechanics. *Scandinavian Audiology Supplement*, *9*, 35–47.

Kemp, D. T. (1986). Otoacoustic emissions, traveling waves and cochlear mechanisms. *Hearing Research*, *22*, 95–104.

Kemp, D. T., & Brown, A. M. (1983). An integrated view of the cochlear mechanical nonlinearities observable in the ear canal. In E. de Boer & M. A. Viergever (Eds.), *Mechanics of Hearing* (pp. 75–82). The Hague, The Netherlands: Martinus Nijhoff.

Kim, D. O. (1980). Cochlear mechanics: implications of electrophysiological and acoustical observations. *Hearing Research, 2*, 297–317.

Kim, D. O., Paparello, J., Jung, M. D., Smurzynski, J., & Sun, X. (1996). Distortion product otoacoustic emission test of sensorineural hearing loss: Performance regarding sensitivity, specificity and receiver operating characteristics. *Acta Otolaryngologica*, *116*, 3–11.

Kimberley, B. P., Hernadi, I., Lee, A. M., & Brown, D. K. (1994). Predicting pure tone thresholds in normal and hearing-impaired ears with distortion product emissions and age. *Ear and Hearing*, *15*, 199–209.

Knight, R. D., & Kemp, D. T. (2000). Indications of different distortion product otoacoustic emissions mechanisms from a detailed *f*1 *f*2 area study. *Journal of the Acoustical Society of America*, *107*, 457–473.

Konrad-Martin, D., Neely, S., Keefe, D., Dorn, P., Cyr, E., & Gorga, M. (2002). Sources of distortion product otoacoustic emissions revealed by suppression experiments, inverse fast Fourier transforms and SFOAEs in impaired ears. *Journal of the Acoustical Society of America*, *111*, 1800–1809.

Konrad-Martin, D., Neely, S., Keefe, D., Dorn, P., & Gorga, M. (2001). Sources of distortion product otoacoustic emissions revealed by suppression experiments and inverse fast Fourier transforms in normal ears. *Journal of the Acoustical Society of America*, *109*, 2862–2879.

Kossl, M. (1994). Otoacoustic emissions from the cochlea of the 'constant frequency' bats. *Pteronotus parnellii* and *Rhinolophus rouxi*. *Hearing Research*, *72*, 59–72.

Kummer, P., Janssen, T., & Arnold, W. (1995). Suppression tuning characteristics of the $2f_2$-f_2 distortion product otoacoustic emission in humans. *Journal of the Acoustical Society of America*, *98*, 197–210.

Kummer, P., Janssen, T., & Arnold, W. (1998). The level and growth behavior of the $2f_2$-f_2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal and cochlear hearing loss. *Journal of the Acoustical Society of America*, *103*, 3431–3444.

Long, G. R., Shaffer, L. A., Dhar, S., & Talmadge, C. L. (1999). Cross species comparison of otoacoustic fine-structure. In H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, & T. Koike (Eds.), *Recent Developments in Auditory Mechanics* (pp. 367–373). Singapore: World Scientific Press.

Long, G. R., Talmadge, C. L., & Thorpe, C. A. (2001). Experimental measurement of level dependence of stimulus frequency otoacoustic emissions fine structure. *Association for Research in Otolaryngology: Abstracts of the Twenty-fourth Midwinter Meeting*, *46*, 13.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R., & Coats, A. C. (1988). Spontaneous otoacoustic emissions in the nonhuman primate. II. Cochlear anatomy. *Hearing Research*, *33*, 69–94.

Lukashkin, A. N., & Russell, I. J. (1999). Analysis of the f_2-f_1 and $2f_2$-f_2 distortion components generated by the hair cell mechanoelectrical transducer: Dependence in the amplitudes of the primaries and feedback gain. *Journal of the Acoustical Society of America*, *106*, 2661–2668.

Lukashkin, A. N., & Russel, I. J. (2001). Origin of the bell-like dependence of the DPOAE amplitude on primary frequency ratio. *Journal of the Acoustical Society of America*, *110*, 3097–3106.

Lukashkin, A. N., Lukashkin, V. A., & Russell, I. J. (2002). One source for distortion product otoacoustic emissions generated by low- and high-level primaries. *Journal of the Acoustical Society of America*, *111*, 2740–2748.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Cambridge: Academic Press.

Martin, B. K., Ohlms, L. A., Harris, F. P., Franklin, D. J., & Lonsbury-Martin, B. L. (1990). Distortion product emissions in humans III. Influence of sensorineural hearing loss. *Annals of Otology, Rhinology, & Laryngology*, *99*, 30–42.

Mauermann, M., Uppenkamp, S., van Hengel, P. W. J., & 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. *Journal of the Acoustical Society of America*, *106*, 3484–3491.

Nelson, D. A., & Kimberley, B. P. (1992). Distortion-product emissions and auditory sensitivity in human ears with normal hearing and cochlear hearing loss. *Journal of Speech and Hearing Research*, *35*, 1141–1159.

Popelka, G., Osterhammel, P., Nielsen, L., & Rasmussen, A. (1993). Growth of distortion product otoacoustic emissions with primary tone level in humans. *Hearing Research*, *71*, 12–22.

Shaffer, L. A., Withnell, R. H., & Kirk, D. L. (2002). Reflection emissions in the guinea pig. *Association for Research in Otolaryngology: Abstracts of the Twenty-fifth Midwinter Meeting*, *323*, 85.

Shera, C. A., & Guinan, J. J. Jr (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. *Journal of the Acoustical Society of America*, *105*, 782–798.

Stover, L. J., Neely, S. T., & Gorga, M. P. (1996). Latency and multiple sources of distortion product emissions. *Journal of the Acoustical Society of America*, *99*, 1016–1024.

Strube, H. W. (1989). Evoked otoacoustic emissions as cochlear Bragg reflections. *Hearing Research*, *38*, 35–45.

Talmadge, C. L., Long, G. R., Tubis, A., & Dhar, S. (1999). Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. *Journal of the Acoustical Society of America*, *105*, 275–292.

Talmadge, C. L., Long, G. R., Murphy, W. J., & Tubis, A. (1993). New off-line method for detecting spontaneous otoacoustic emissions in human subjects. *Hearing Research*, *71*, 170–182.

Talmadge, C. L., Long, G. R., Tubis, A., & Tong, C. (2000). Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure. *Journal of the Acoustical Society of America*, *108*, 2911–2932.

Talmadge, C. L., Tubis, A., Long, G. R., & Piskorski, P. (1998). Modeling otoacoustic and hearing threshold fine structure. *Journal of the Acoustical Society of America*, *104*, 1517–1543.

Wagner, W., & Plinkert, P. K. (1999). The relationship between auditory threshold and evoked otoacoustic emissions. *European Archives of Otorhinolaryngology*, *256*, 177–188.

Whitehead, M. L. (1998). Species differences of distortion-product otoacoustic emissions. *Journal of the Acoustical Society of America*, *103*, 2740–2742.

Whitehead, M. L., McCoy, M. J., Lonsbury-Martin, B. L., & Martin, G. K. (1995). Dependence of distortion product otoacoustic emissions on primary level in normal and impaired ears. I. Effects of decreasing L_2 below L_1. *Journal of the Acoustical Society of America*, *97*, 2346–2358.

Whitehead, M. L., Stagner, B. B., McCoy, M. J., Lonsbury-Martin, B. L., & Martin, G. K. (1995). Dependence of distortion product otoacoustic emissions on primary levels in normal and impaired ears. II. Asymmetry in L_1, L_2 space. *Journal of the Acoustical Society of America*, *97*, 2359–2377.

Withnell, R. H., Shaffer, L. A., & Talmadge, C. L. (2003). Generation of DPOAEs in the guinea pig. *Hearing Research*, *178*, 106–117.

Withnell, R. H., & Yates, G. K. (1998). Onset of basilar membrane non-linearity reflected in cubic distortion tone input-output functions. *Hearing Research*, *123*, 87–96.

Withnell, R. H., Yates, G. K., & Kirk, D. L. (2000). Changes to low-frequency components of the TEOAE following acoustic trauma to the base of the cochlea. *Hearing Research*, *139*, 1–12.

Wright, A. A. (1984). Dimensions of the cochlear stereocilia in man and in guinea pig. *Hearing Research*, *13*, 89–98.

Yates, G. K., & Withnell, R. H. (1999). The role of intermodulation distortion in transient-evoked otoacoustic emissions. *Hearing Research*, *136*, 49–64.

Zweig, G., & Shera, C. (1995). The origins of periodicity in the spectrum of evoked otoacoustic emissions. *Journal of the Acoustical Society of America*, *98*, 2018–2047.

Zwicker, E., & Peisl, W. (1990). Cochlear processing in analog models, in digital models, and in human inner ear. *Hearing Research*, *44*, 206–216.

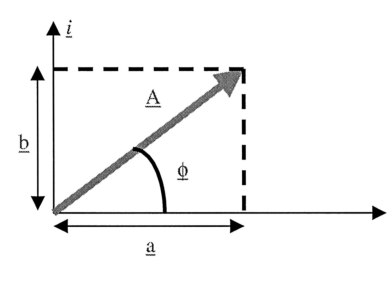
APPENDIX

To evoke DPOAEs, two pure tone stimuli of frequencies f1 and f2 are simultaneously presented to the ear. A microphone in the ear canal will measure the ear canal sound pressure over time, which is a composite of the stimulus tones and the OAEs. Fourier analysis§ of this ear canal signal in response to any combination of stimulus tones f1 and f2 will produce the complex amplitude versus frequency of this ear canal signal. Although any DPOAE could be reported (e.g., f_2-f_1, $3f_1$-$2f_2$), clinical DPOAE systems typically report only the amplitude of the stimulus tones and the $2f_1$-f_2 DPOAE. However, the $2f_1$-f_2 DPOAE has both amplitude and phase, where the complex amplitude a + ib is related to amplitude (A) and phase (ϕ) by

a = A.cosϕ

b = A.sinϕ

i.e., in which the x-axis is the real component and the y-axis is the imaginary component and $i = \sqrt{(-1)}$.



So, at the frequency $2f_1$-f_2, one will obtain a complex amplitude a + ib. If one then sweeps stimulus frequency while holding the stimulus frequency ratio and stimulus level constant, one obtains a data set of complex $2f_1$-f_2 amplitude values versus $2f_1$-f_2 frequencies.

§For a sampling rate of 40960 Hz and an averaging epoch of 100 msec, the complex amplitude spectrum will range from 0 to 20480 Hz (this value represents the anti-aliasing or Nyquist frequency, and is typically about 0.46 times the sampling rate) in 10Hz steps for an FFT with 4096 points.

IFFT Analysis

The complex amplitude data versus $2f_1$-f_2 frequency can then be converted to its analogous time domain representation by performing an inverse fast Fourier transform (IFFT) on the data set. To be able to perform an IFFT, the complex amplitude spectral data set must have constant frequency spacing between adjacent data points (either on a linear or logarithmic scale), and the IFFT must be performed over a sufficient number of data points to encompass the frequencies contained in the data set.‖ To facilitate Fourier analysis, the complex amplitude data set is zero-buffered i.e., where there is no data a zero is inserted. By not mirroring the complex amplitude data set, this zero-buffered data set represents the Fourier transform of an analytic signal (Mallat, 1998).

An example of an IFFT analysis is shown in Figure 7b. The component at time zero represents the emission arising from nonlinear distortion, the peak at 0.007 sec represents the emission arising from reflection. The peaks following the component at time zero can be removed from the response by windowing, leaving only the nonlinear component. Fourier analysis on the windowed component provides the complex amplitude in the frequency domain of the emission arising from nonlinear distortion, from which the amplitude versus frequency of this component can be obtained. Subtraction of the Fourier transform of the windowed component from the initial complex amplitude data set (the measured 2f1-f2 DPOAEs) gives the total reflection component. As an alternative to subtraction, the reflection peaks following the component near time zero can be isolated by windowing. Fourier analysis on the windowed peaks provides the complex amplitude in the frequency domain of the emission arising from the reflection.

Caveats exist in performing Fourier analysis on DPOAE complex amplitude data. Such caveats are not presented here but are discussed in Withnell et al. (2003). Additionally, the reader is referred to Kalluri et al. (2001) for a discussion of the effects of time window duration on the estimation of source components.

‖For example, if the complex amplitude data set consists of $2f_1$-f_2 complex amplitudes over a frequency range 1000 to 4000 Hz (with 10 Hz resolution), then the IFFT must be performed over a frequency range of at least 0 to 8000 Hz (see Kalluri et al., 2001, for an alternative method).