Separation Anxiety: DPOAE Components Refuse to be Apart

Sumitrajit Dhar
Fellow, Hugh Knowles Center
Roxelyn & Richard Pepper Department of Communication Sciences and Disorders
Northwestern University
Evanston, Illinois, USA

In two relatively recent papers, (Dhar & Shaffer, 2004; Shaffer & Dhar, 2006) have explored the possibility of improving the prediction of behavioral hearing thresholds from distortion product otoacoustic emission (DPOAE) level. DPOAEs are sounds generated in the cochlea (Kemp, 1979) that can be recorded in the ear canal using appropriate recording and analysis equipment. When a normally functioning cochlea is stimulated simultaneously with two pure tones, say at frequencies f_1 and f_2 ($f_1 < f_2$), the cochlea generates energy at various other frequencies related to those of the stimulus tones. The DPOAE produced at the frequency $2f_1-f_2$ is commonly used for clinical purposes as it is easily recorded from human ears under certain stimulus conditions. DPOAEs at frequencies lower than those of the stimulus tones, such as that at the frequency $2f_1-f_2$, can be called "apical" DPOAEs as the tonotopic organization of the basilar membrane dictates that their characteristic frequency place lies apical to that of the stimulus tones.

In typical clinical applications, DPOAEs are recorded with relatively coarse frequency spacing. For example, a clinical device may default to measuring DPOAEs at 3 frequencies per octave. The resultant DPOAE level versus frequency function is referred to as the DP gram. When DPOAEs are recorded with significantly greater frequency resolution, the resultant DPOAE level versus frequency function exhibits an alternating pattern of peaks and valleys. This pattern is referred to as *fine structure* in the literature and the peaks and valleys are referred to as maxima and minima, respectively (Kemp, 1979). An example of one such recording is displayed in Figure 1. Note the pseudo periodic variation in DPOAE level over the entire recording frequency range.
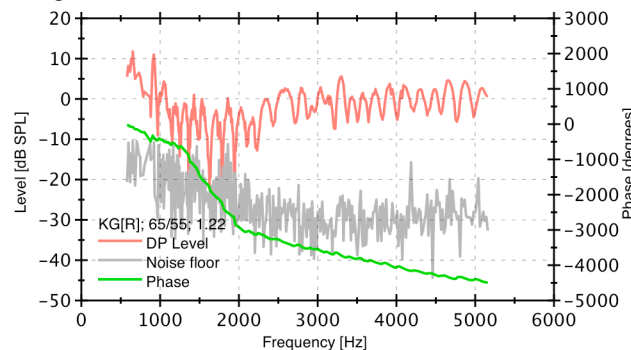

Figure 1. Example of DPOAE fine structure measured from a normal-hearing young human adult. DPOAE level, phase, and noise floor are represented using the orange, green, and gray traces, respectively.

Prevailing models generate DPOAE fine structure by constructive and destructive interference between two components of the DPOAE at $2f_1-f_2$ (e.g., Talmadge et al., 1998; Talmadge et al., 1999). One component is envisioned to be generated in the overlap region between the traveling wave patterns due to the stimulus tones. This component has been termed the overlap, generator, place-fixed, and distortion component by various authors. The second component is envisioned to be generated in the characteristic frequency (CF) region of the DPOAE. This component has been termed the DP CF, reflection, wave-fixed, and the reflection component. We will refer to the two components as the overlap and the DP CF components here.

The initial recognition that (apical) DPOAEs recorded in the ear canal may comprise of more than one contributing component was based on the observation of sharp notches in DPOAE level (Kim et al., 1980). The insight into the source of these notches came from the observation of activity in neural

populations associated with both the overlap region and the DP CF region. Mechanical activity on the basilar membrane at distortion frequencies were also reported at both these regions (Robles & Ruggero, 2001). Confirmation of the presence of (at least) two components in the ear canal signal has since come from several lines of experiments. A suppressor tone close in frequency to 2f1-f2 has been shown to eliminate the fine structure seen in Figure 1 without altering the overall DPOAE level (Heitmann et al., 1998). Similarly, fine structure is either reduced or eliminated when the DP frequency happens to fall in regions of hearing loss (Mauermann et al., 1999). The two components have also been observed in isolation at signal onset and offset when one of the stimulus tones has been pulsed on and off (Talmadge, Long et al., 1999). Based on this evidence, there is now general consensus about the presence of two distinct DPOAE components.

```
*******************************************************************************************
                                two-source.swf animation
*******************************************************************************************
```

Figure 2. Animation of two-source model of DPOAEs. See text for details.

The general idea of the two-source model of DPOAEs can be visualized in the animation in Figure 2. The activity related to the traveling wave patterns of the stimulus tones are depicted in red and blue. Energy at the DPOAE frequency is generated in the area of overlap between these activity patterns. In the case of apical DPOAEs, Part of the DPOAE energy travels towards the ear canal while another part travels towards the apex of the cochlea. This apically traveling energy reaches a peak in activity in the CF region of the DPOAE and is returned as a second component to the ear canal. The inward DPOAE energy is depicted in the purple traveling wave pattern. The two DPOAE components are then depicted with arrows traveling towards the outer ear. It is the constructive and destructive interference between these two components that leads to the fine structure observed.

For the two DPOAE components at the same frequency ($2f_1-f_2$) to give rise to the observed fine structure pattern, their phases have to change at different rates as a function of frequency. Indeed the phase of the component from the overlap region has been shown to be relatively invariant as a function of frequency. On the other hand, the phase of the component from the DP CF region rotates rapidly with frequency (Knight & Kemp, 2001). This difference in phase behavior and the resultant fine structure is depicted in the animation in Figure 3. In the left half of the animation, a blue vector represents the component from the overlap region. Neither the phase nor the magnitude of this vector is altered as a function of frequency in the animated model. The component from the DP CF region is depicted with a red vector. The magnitude of this vector is kept fixed while its phase is rotated as a function of frequency. The resultant green vector represents the DPOAE recorded in the ear canal. The magnitude of the green vector is traced in the right half of the animation by an orange ball that traverses along the blue trace representing the fine structure that is recorded in the ear canal. Thus, at least in this cartoon, the systematically fluctuating DPOAE level recorded in the ear canal is a result of the phase rotation of the component from the DP CF region.

```
*******************************************************************************************
                                   vector.swf animation
*******************************************************************************************
```

Figure 3. Animation of generation of fine structure due to interference between two DPOAE components. See text for details.

In introducing the two DPOAE components, we had listed a variety of terms. These terms can now be understood based on various differences between the two DPOAE components. First, the difference in the location of their generation leads to the nomenclature that we will use in this paper: *overlap* and *DP CF* components. The fact that the phase of one of the components does not change as a function of frequency (overlap) while that of the other does (DP CF) leads to the terms *wave-* and *place-fixed* components (Knight & Kemp, 2000). In this classification the names stem from the fact that the phase of a component would not change if its phase were tied to the phase of the stimulus tones due to presumed scaling symmetry in the cochlea. On the other hand, if the phase of a component were tied to a certain place on the basilar membrane, signals of different frequencies would register different phases. This idea

is extended further in the use of the terms *generator/distortion* and *reflection* components. In the use of this nomenclature, the two DPOAE components are thought to be generated by two distinct mechanisms: the component from the overlap region due to nonlinear distortion and that from the DP CF region due to coherent linear reflections (Shera & Guinan, 1999). As we delve deeper into techniques for separating DPOAE components, it will be evident that each technique exploits either the difference in location or phase behavior of the two. But first let us ask, "why bother"?

Through out the 1990s there were concerted efforts to evaluate the predictability of hearing thresholds using DPOAEs. This would certainly seem to be an attractive proposition given the non-invasive, objective, and frequency-specific nature of DPOAEs. However, several experiments yielded coefficients of correlation between 0.2 and 0.5 (Gaskill & Brown, 1993; Gorga et al., 1994; Kimberley et al., 1994). It was readily apparent that the great variability in DPOAE level, either in a normal-hearing population or in a population of ears with similar hearing loss was (at least) partly the problem. The argument was then forwarded that this variability at least in a normal-hearing population was due to fine structure. Although, fine structure patterns were stable in frequency in a given ear, they were not identical across ears. In other words, fine structure patterns from several ears would not be aligned if they were superimposed. Thus at any given test frequency, the variable phase relationship between the two DPOAE components across ears would lead to a variety of DPOAE levels, even though all these ears were functioning normally. Following this argument, (Plinkert et al., 1997) suggested using a recording paradigm that would result in a *single-generator* [*sg*]DPOAE thereby improving the predictability of hearing thresholds from DPOAE levels. The *sg*DPOAE is recorded by presenting a suppressor tone very close in frequency to the DPOAE along with the stimulus tones. We will discuss the details and complications of this and other techniques later but let us return to the initial notion of improving the prediction of behavioral hearing thresholds from DPOAE levels. It turns out that neither the *sg*DPOAE nor any other technique has been successfully used to improve this prediction to date. The problem turns out to be in the great variability in the relationship between the two DPOAE components across ears and across frequency in a given ear. The rest of this paper is dedicated to a discussion of various techniques for separating DPOAE components.

The use of a suppressor tone near the DPOAE frequency, as suggested in the *sg*DPOAE technique, is motivated by the disparate regions of generation of the DPOAE components on the basilar membrane. An appropriate suppressor near the DPOAE frequency is expected to selectively suppress the DP CF component. Under such conditions, the DPOAE recorded in the ear canal would comprise only of the overlap component. Various groups have demonstrated the feasibility of this technique (Heitmann, Waldmann et al., 1998; Talmadge, Long et al., 1999). However, using this technique did not result in any improvement in the correlation between DPOAE level and hearing thresholds (Dhar & Shaffer, 2004; Johnson et al., 2007). This lack of improvement was not due to a failure of the technique but rather due to the variable relationship between DPOAE components across ears and across frequency within ears. A compulsion of any clinical test is that it has to be "one size fits most" if not "one size fits all". A suppressor level that would eliminate the DP CF component without affecting the overlap component universally in all ears would be required. Dhar and Shaffer (2004) did not find a universal suppressor level that would fit this bill.

Yet another variable that adds to the complexity of DPOAE generation is stimulus level. The relative contribution of the overlap and DP CF components to the DPOAE level in the ear canal changes at different stimulus levels (Konrad-Martin et al., 2001). Thus, a suppressor of different strengths would be appropriate for each of these stimulus combinations. This complex interaction between the relative strengths of DPOAE components and the suppressor tone was speculated to be the breakdown in the idea of the *sg*DPOAE obtained using a suppressor (Dhar & Shaffer, 2004). However, this result could not be used to drive clinical practice due to the small sample size (10) of this study. The basic result has since been confirmed in a much larger sample (n=205, Johnson, Neely et al., 2007). The failure of the suppressor technique can be demonstrated using the simplistic model animation in Figure 4.

****

sup1.swf animation here
sup3.swf animation here

sup2.swf animation here

****

Figure 4. Animation of effects of suppressor tone on DPOAE components and fine structure. See text for details.

The ideal outcome of using a suppressor tone is displayed in the top panel of Figure 4. Green arrows mark CF regions for the stimulus tones. The magnitudes of the overlap (nonlinear) and DP CF (reflection) components are marked by the purple rectangles. The animated arrow in purple represents a suppressor tone of increasing intensity. As the intensity of the suppressor tone increases, it causes a reduction in the level of the DP CF component. The highest suppressor level eliminates the DP CF component without affecting the overlap component. The outcome of each suppressor condition is demonstrated in the right half of the animation. We start out with pronounced fine structure, which is eliminated for the highest suppressor level without any change in the overall level of the DPOAE. Note that in this case we start out with an overlap component that is greater in magnitude than the DP CF component.

The middle and bottom panels of Figure 4 depict situations where fine structure is either not eliminated or the suppressor alters the overlap component in addition to the DP CF component. In the middle panel of Figure 4 we start out with a DP CF component that is larger than the overlap component. The suppressor causes a reduction in the level of the DP CF component but is unable to completely eliminate it even at the highest suppressor level. An added complication is observed for the suppressor of intermediate strength. This suppressor causes a reduction in the DP CF component but makes it equal to the overlap component leading to an increase in fine structure depth. In the bottom panel of Figure 4 we start out with a DP CF component that is considerably smaller than the overlap component. The suppressor is able to eliminate the DP CF component, but at its highest strength the suppressor alters the magnitude of the overlap component as well. The result is the elimination of fine structure followed by an unwanted reduction in overall DPOAE magnitude. Of the three possibilities depicted in Figure 4, the clinical goal would be to achieve the outcome in the top panel in all ears.

While the suppressor technique differentiated between the two DPOAE components based on their generation location on the basilar membrane, other techniques exploit the differences in phase behavior between these components. As implied by the wave- versus place-fixed nomenclature, the phase of the overlap component changes little as a function of frequency while that of the DP CF component changes rapidly as a function of frequency. Models of DPOAE generation have attributed this difference to fundamental differences in mechanisms responsible for the generation of each of these components (Shera & Guinan, 1999). The approximately invariant phase of the overlap component implies that this component has a group delay (negative of the slope of the phase) close to 0. On the other hand, the rapidly changing phase of the DP CF component results in a much greater group delay. Thus, if the magnitude and phase of the ear-canal DPOAE is subjected to an inverse FFT, the two components appear separated along the x-axis, where the x-axis represents group delay. The output of one such operation on DPOAEs recorded from a human ear is displayed in Figure 5.
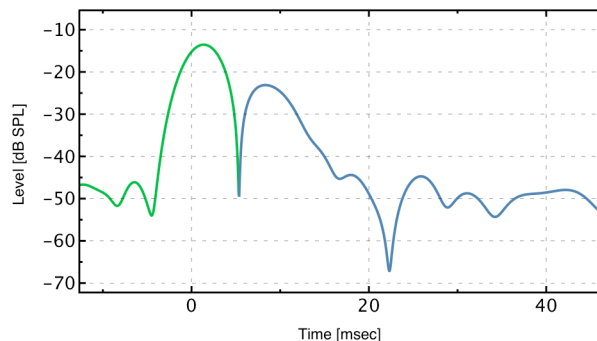

Figure 5. Output of inverse FFT operation on DPOAE level and phase recorded from a normal-hearing human ear. The two DPOAE components are separated by color.

Two peaks of energy are evident in Figure 5, the first one centered approximately around 0 ms and the second close to 10 ms. These two peaks are separated by color in the figure. In using this technique to separate DPOAE components, appropriate time windows would be applied to separate the two components in time. Each component would then be returned to the frequency domain through a regular FFT resulting in independent level and phase estimates of the two DPOAE components.

The iFFT technique requires essentially continuous data for the results to be valid. In terms of DPOAEs this translates to recordings that are very dense in frequency, thereby rendering the technique impractical for clinical applications. However, to answer the theoretical question as to whether eliminating fine structure improves the correlation between DPOAE level and hearing thresholds, (Shaffer & Dhar, 2006), applied the iFFT technique to their data set from Dhar and Shaffer (2004). Although the elimination of fine structure was more complete than when the suppressor technique was used, the correlation between hearing thresholds and DPOAE levels did not show improvement. It would not be fair to attempt to answer such a clinical question with this limited data set. However, the results of Shaffer and Dhar (2006) showed interesting results in the variability of the relationship between DPOAE components. Take for example the result of the iFFT technique applied on data between 2200 and 2600 Hz from 16 ears. The results are displayed in Figure 6.
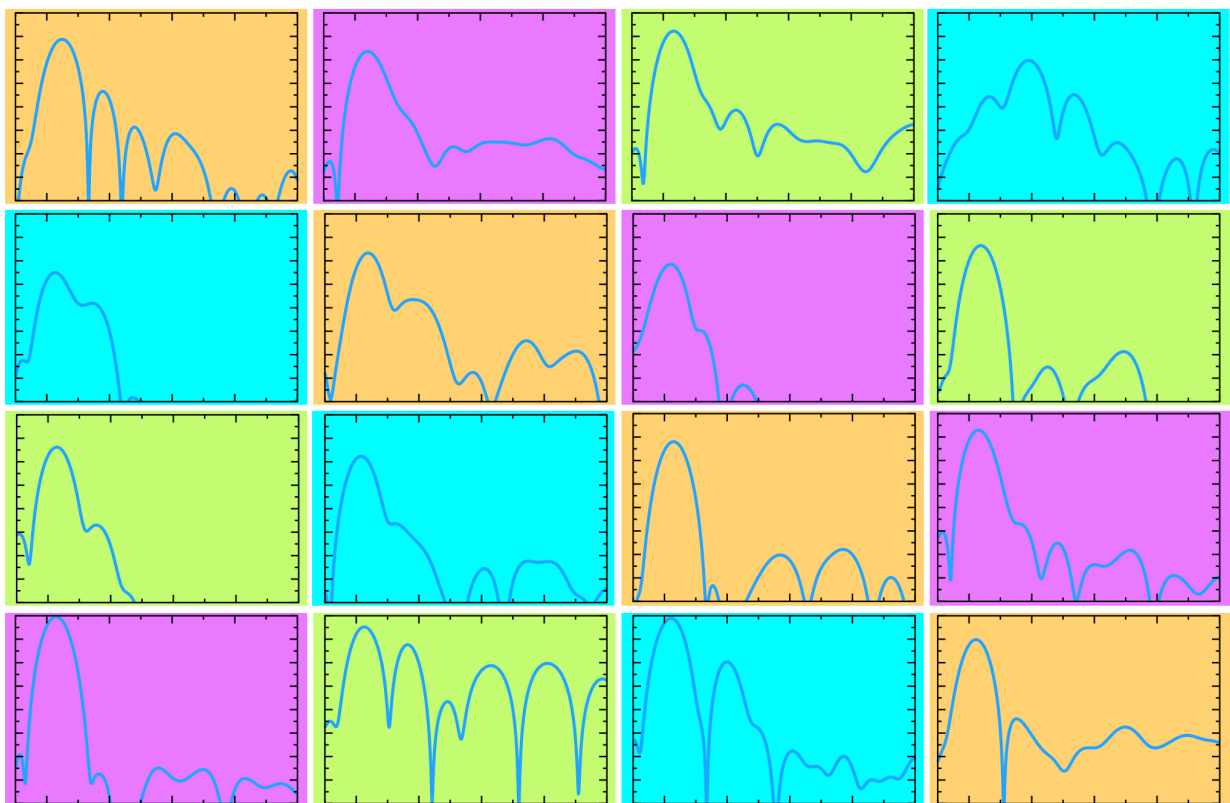

Figure 6. Differences in the results of iFFT analysis on data obtained between 2200 and 2600 Hz from 16 ears.

 While the two DPOAE components are clearly segregated in some panels of Figure 6, the second (DP CF) component is absent in some and a clear separation cannot be viewed in other panels. The non-existence of the DP CF component is not a problem, per se. However, the blurring of the two components in time without a distinct border makes it difficult, if not impossible, to separate them. What could cause the blurring of this boundary? The confinement of all the energy from the overlap component around a delay of 0 ms is determined by the shift symmetry of traveling wave patterns on the basilar membrane. That is, it is only when a cochlea is perfectly shift symmetric that the delay of the overlap component is exactly zero. Any deviation from shift symmetry would cause the overlap component to have a finite

delay. Similarly, the delay of the DP CF component is determined by the slope of the phase of that component. A linear and invariant slope as a function of frequency would result in a well-defined delay of the DP CF component. If the slope is large enough, this component would be clearly separated from the overlap component in the results of the iFFT. Whenever, both these conditions are not met, the result is a blurring of the border between the DPOAE components and subsequent difficulty in separating them. Even from the limited data set of Shaffer and Dhar (2006), it is evident that all conditions that lead to distinct and easily separable DPOAE components are not always present. The outcome, therefore, is the same as that seen with the suppressor technique – a single DPOAE component may well be more predictive of hearing thresholds, but it remains elusive.

Component separation is a worthy exercise; from these results we can learn about the mechanisms responsible for generating each DPOAE component. Studying the properties of each component in isolation is also likely to provide a deeper understanding of the overall process of DPOAE generation and their relationship to cochlear mechanics. However, separation of DPOAE components for the sake of predicting hearing thresholds appears to be a very difficult exercise that could turn out to be futile in the end.

References

Dhar S. & Shaffer L.A. 2004. Effects of a suppressor tone on distortion product otoacoustic emissions fine structure: why a universal suppressor level is not a practical solution to obtaining single-generator DP-grams. *Ear Hear*, 25, 573-585.
Gaskill S.A. & Brown A.M. 1993. Comparing the level of the acoustic distortion product 2f1-f2 with behavioural threshold audiograms from normal-hearing and hearing-impaired ears. *Br J Audiol*, 27, 397-407.
Gorga M.P., Neely S.T., Bergman B.M., Beauchaine K.L., Kaminski J.R., et al. 1994. Towards understanding the limits of distortion product otoacoustic emission measurements. *J Acoust Soc Am*, 96, 1494-1500.
Heitmann J., Waldmann B., Schnitzler H.U., Plinkert P.K. & Zenner H.P. 1998. Suppression of distortion product otoacoustic emissions (DPOAE) near f1-f2 removes DP-gram fine structure - Evidence for a secondary generator. *J Acoust Soc Am*, 103, 1527-1531.
Johnson T.A., Neely S.T., Kopun J.G., Dierking D.M., Tan H., et al. 2007. Distortion product otoacoustic emissions: cochlear-source contributions and clinical test performance. *J Acoust Soc Am*, 122, 3539-3553.
Kemp D.T. 1979. The evoked cochlear mechanical response and the auditory microstructure - evidence for a new element in cochlear mechanics. *Scand Audiol Suppl*, 35-47.
Kim D.O., Molnar C.E. & Matthews J.W. 1980. Cochlear mechanics: nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure. *J Acoust Soc Am*, 67, 1704-1721.
Kimberley B.P., Kimberley B.M. & Roth L. 1994. A neural network approach to the prediction of pure tone thresholds with distortion product emissions. *Ear Nose Throat J*, 73, 812-823.
Knight R.D. & Kemp D.T. 2000. Indications of different distortion product otoacoustic emission mechanisms from a detailed f1,f2 area study. *J Acoust Soc Am*, 107, 457-473.
Knight R.D. & Kemp D.T. 2001. Wave and place fixed DPOAE maps of the human ear. *J Acoust Soc Am*, 109, 1513-1525.
Konrad-Martin D., Neely S.T., Keefe D.H., Dorn P.A. & Gorga M.P. 2001. Sources of distortion product otoacoustic emissions revealed by suppression experiments and inverse fast Fourier transforms in normal ears. *J Acoust Soc Am*, 109, 2862-2879.
Mauermann M., Uppenkamp S., van Hengel P.W. & 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.K., Heitmann J. & Waldmann B. 1997. ["Single generator distortion products"(sgDPOAE).Precise measurements of distortion product otoacoustic emissions by three-tone stimulations]. *Hno*, 45, 909-914.

Robles L. & Ruggero M.A. 2001. Mechanics of the mammalian cochlea. *Physiological reviews*, 81, 1305-1352.

Shaffer L.A. & Dhar S. 2006. DPOAE component estimates and their relationship to hearing thresholds. *J Am Acad Audiol*, 17, 279-292.

Shera C.A. & Guinan J.J., Jr. 1999. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *J Acoust Soc Am*, 105, 782-798.

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. *J Acoust Soc Am*, 105, 275-292.

Talmadge C.L., Tubis A., Long G.R. & Piskorski P. 1998. Modeling otoacoustic emission and hearing threshold fine structures. *J Acoust Soc Am*, 104, 1517-1543.