Multi-Frequency Acquisition of DPOAE Input-Output Functions for Auditory-Threshold Estimation

D. Zelle, J. P. Thiericke, A. W. Gummer, E. Dalhoff
Dept. of Otolaryngology, Eberhard-Karls-University Tübingen, Tübingen, Germany
dennis.zelle@uni-tuebingen.de

Abstract

Distortion product otoacoustic emissions (DPOAEs) evolve as a byproduct of the nonlinear amplification process of two stimulus tones in the cochlea and comprise a nonlinear-generation and a coherent-reflection component. Wave interference between these components may limit the diagnostic validity of DPOAEs for assessing the function of the inner ear. Utilizing the extracted nonlinear-generation component from DPOAE signals acquired with pulsed f_2-tones increases the accuracy of auditory threshold estimates computed from semi-logarithmic input-output (I/O) functions. However, the acquisition of pulsed DPOAEs required up to now a considerably longer measurement time to attain sufficient signal-to-noise ratio. The measurement time may be decreased by shortening both the f_2-pulse and also the acquisition blocks for ensemble averaging. However, undersized blocks bear the risk that slowly decaying time responses interfere with the DPOAE signal in the subsequent block, resulting in reduction of the accuracy of the auditory-threshold estimate. Here, to further decrease the measurement time, we introduce a new method for presenting pulse-train stimuli for quasi-simultaneous acquisition of multiple I/O-functions.

DPOAE I/O-functions were acquired from 16 normal-hearing subjects with the new acquisition technique for a set of frequencies with f_2 = 1, 1.5, 2, 3 and 4 kHz. Estimated distortion product thresholds (EDPTs) derived from I/O-functions were compared with thresholds obtained by Békésy audiometry. EDPTs correlate with Békésy thresholds ($r^2 > 0.36$) with high significance ($P < 0.001$). The standard deviation of the auditory-threshold estimates was 5.2 dB SPL for the pooled data. The average measurement time was 13.6 ± 3.7 min. This algorithm enables auditory-threshold estimation for a set of frequencies with high accuracy and reasonable cost of measurement time; it represents a promising diagnostic tool for objective auditory-threshold assessment.

1 Introduction

Presenting two stimulus tones to the inner ear evokes distortion-product otoacoustic emissions (DPOAEs) as a byproduct of the nonlinear cochlear amplification process which may be used to assess the function of the outer hair cells objectively [1]. According to a widely accepted model, DPOAEs consist of two source components, a nonlinear-generation and a coherent-reflection component [2]. The presence of two source components can be visualized during the onset and offset of the DPOAE response when the second stimulus tone is pulsed [3]. Wave interference between both components limits the diagnostic validity of DPOAEs and is the major reason for the high standard deviation of auditory-threshold estimation [4] by means of semi-logarithmic DPOAE input-output (I/O) functions [5].

Because the generation site of the nonlinear-generation component is close to the tonotopic place of the frequency f_2 of the second stimulus tone, this component is the signal of interest when estimating auditory threshold on the basis of DPOAEs. Using a pulsed f_2-tone, the nonlinear generation component can be extracted without significant interference from the coherent-reflection component by a technique called onset-decomposition (OD), which exploits the different latencies of both source components and enables the quantification of the nonlinear-generation component in the time domain [6]. Extracting the nonlinear-generation component with f_2-pulses of 100 ms duration enhances the accuracy of auditory-threshold estimates derived from semi-logarithmic I/O-functions but at the cost of increased measurement time [4].

Recently, we introduced a refined measurement paradigm, called short-pulse DPOAE acquisition, reducing the length of the f_2-tone to 8 ms [7]. This method enables auditory-threshold estimation by means of OD while decreasing the measurement time considerably compared with f_2-tones of 100-ms length, due to shortening of the acquisition blocks used for ensemble averaging. However, undersized blocks bear the risk that slowly decaying time responses interfere with the DPOAE signal in the subsequent block resulting in an accompanying reduction of signal fidelity. In this paper, we present a new technique incorporating pulse-train stimuli for quasi-simultaneous acquisition of multiple I/O-functions in order to decrease the measurement time while retaining accuracy of the auditory-threshold estimates.

2 Methods

2.1 DPOAE Signal Acquisition

DPOAEs were recorded from 16 normal-hearing subjects (age: 30 ± 6.4 yr.) with auditory thresholds better than 20 dB HL, using six primary-tone level pairs according to the scissor paradigm $L_1 = 0.4L_2 + 39$ dB SPL to account for the different compression of the traveling waves evoked by the stimulus tones in the cochlea [8]. L_2 varied from 25 to 65 dB SPL. DPOAEs were recorded for a set of frequencies with f_2 = 1, 1.5, 2, 3, and 4 kHz and a constant ratio of $f_2/f_1 = 1.2$. DPOAE acquisition was performed unilaterally using an ER-10 C DPOAE probe system (Etymotic Research, Elk Grove Village, IL) connected to an IBM-compatible PC equipped with a 16-bit analog output card and a 24-bit signal acquisition card (NI PCI 6733 and NI PCI 4472, National Instruments, Austin, TX). The sampling frequency was 102.4 kHz.

Both stimulus generation and signal acquisition were controlled by measurement software implemented in LabVIEW (Ver. 12.0, National Instruments, Austin, TX). The f_1-tone comprised sequenced pulses each of 35 ms steady-state length and cosine-shaped rising and falling edges of 2.5 ms. The Hanning-shaped f_2-pulses started 10 ms after the onset of their corresponding f_1-pulse and had variable half-widths, $T_{HW} = c/f$, with $c = 13.07$ estimated from the results of Vetešník et al. [6], to account for the frequency-dependent latencies of the DPOAE source components. The total length of a single acquisition block was 200 ms. Data acquisition was terminated if a signal-to-noise ratio (SNR) of 10 dB was reached at f_2 = 1 kHz, or at a maximum acquisition of 800 blocks.

Consecutive phase shifts of the f_1- and f_2-tones by 90° and 180°, respectively, enabled cancellation of the primary tones in the ensemble-averaged signal [9]. Blocks with excessive noise were discarded from the averaging by a suitable artifact threshold. Post-processing was done in Matlab (Ver. 8.2, MathWorks, Natick, MA). Because a single recording contained multiple DPOAE signals associated with multiple stimulus frequencies, we call this measurement paradigm "multi-frequency acquisition".

2.2 Computation of I/O-functions

Consecutive stimulus pairs were arranged in such a way as to provide sufficient distance in the frequency domain to allow band-pass filtering and to avoid overlap between stimulus tones and DPOAEs (see Fig. 1). Zero-phase band-pass filtering with auto-regression extrapolation to reduce filter-edge effects enabled the extraction of the cubic distortion products at frequencies $f_{dp} = 2f_1$-f_2 from the multi-frequency recordings. Circular expansion resulted in DPOAE signals without discontinuities even in cases where the signal exceeds the end of the acquisition block.

DPOAE primary-source extraction was achieved by OD [6], i.e. sampling the envelope of the DPOAE signal at a time instant before wave interference occurs (see Fig. 2b). The envelope was computed as the absolute value of the Hilbert transform of the DPOAE signal $|H\{p(t)\}|$ and sampling instants were commensurate to T_{HW}. Semi-logarithmic I/O-functions (Fig. 2c) were derived from the extracted primary-source components to compute the estimated distortion product threshold (EDPT) by means of linear regression [5].

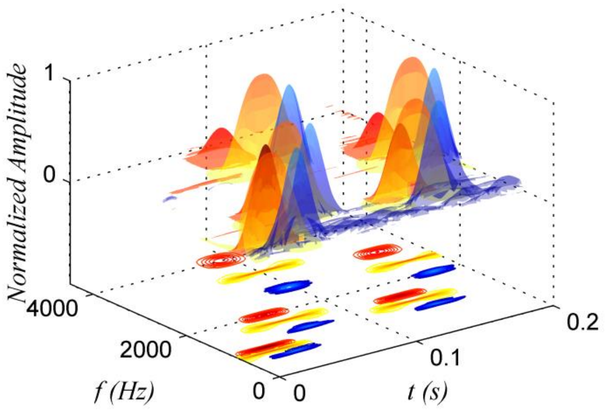

Figure 1: Short-time Fourier transforms (STFT, N = 1024) of the f_1-tones (yellow), f_2-tones (red), and the corresponding band-pass filtered DPOAE signals (blue) with normalized amplitude. For visualization purposes, values below a certain threshold are not displayed. The x-y plane shows the contour plots of the associated STFTs. (Subject S001; L_2 = 45 dB SPL).

Extracted DPOAE signals with a SNR ≥ 10 dB were accepted for the computation of I/O-functions. A minimum of three data points was necessary to extrapolate distortion product thresholds from the semi-logarithmic I/O-functions. DPOAEs showing saturation behavior at high stimulus levels (open red circle in Fig. 2c) were excluded from the computation of the regression lines by an iterative algorithm which maximizes the correlation coefficient *r* of the I/O-functions [4].

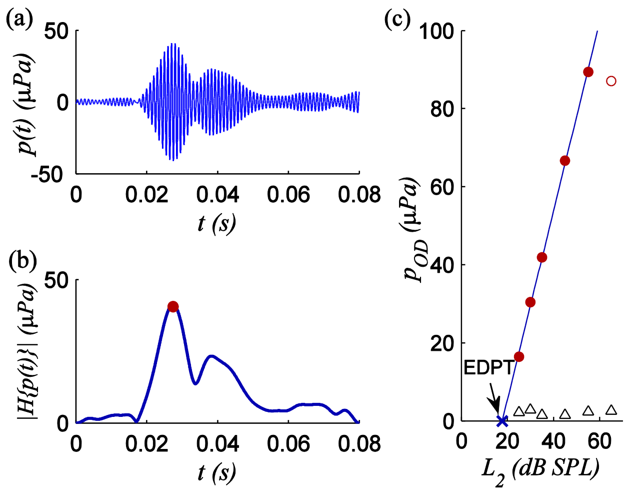

Figure 2: (a) DPOAE signal p(t) extracted from a multi-frequency recording by band-pass filtering (Subject S004; f_2 = 4 kHz, L_2 = 45 dB SPL). (b) $|H\{p(t)\}|$ of the DPOAE signal (blue) shown in (a) and the OD value (red dot) used to quantify the amplitude of the primary-source component P_{OD}. (c) Semi-logarithmic I/O-function obtained from the amplitudes of the extracted primary-source components. The intersection of the extrapolated regression line with the abscissa yields the EDPT. Triangles represent noise values.

Biomed Tech 2014; 59 (s1) © 2014 by Walter de Gruyter • Berlin • Boston. DOI 10.1515/bmt-2014-5011

The computed EDPT values were accepted for auditory-threshold estimation if the corresponding I/O-functions satisfied the objective evaluation criteria $r^2 \geq 0.8$ and $\sigma_{EDPT} \leq 10$ dB SPL [5]. Accepted EDPTs were compared to thresholds obtained by Békésy audiometry. Békésy thresholds (BTHs) were computed as the mean of three successively recorded measurements at frequencies f_2 and neighboring frequencies within the bandwidth ($B = 1/T_{HW}$) of the f_2-pulse and with frequency spacing of $\Delta f = 20$ Hz. Correction for outliers and averaging over frequency yielded BTHs which correspond to responses to stimulation of the basilar membrane with short f_2-pulses [4].

3 Results

The multi-frequency method yields I/O-functions from which reliable auditory-threshold estimates are obtainable. Highly significant correlations ($P < 0.001$) of the BTH with the EDPT are obtained for the pooled data and for $f_2 \geq 3$ kHz (Fig. 3a). Fig. 3b shows the standard deviation of the auditory-threshold estimate assuming a linear relationship between the BTH and the EDPT with either a constant slope, s, of 1.18 taken from a study by Boege and Janssen [5], or variable slope for $f_2 = 3$ kHz and 4 kHz because at these frequencies correlations with $r^2 > 0.42$ were obtained. The corresponding values are $s = 0.88$ ($t_{14} = 6.27$; $P = 10^{-5}$) and 0.70 ($t_{14} = 3.10$; $P = 0.004$). The value $s = 1.18$ can be assumed to be reliable due to the high threshold range in their study and the significant correlation ($r^2 = 0.42$). The standard deviation σ of the pooled data was 5.2 dB SPL. σ varies with stimulus frequency, exhibiting a minimum of 2.6 dB SPL at $f_2 = 3$ kHz. Fig. 3b also plots the standard deviation of the BTH measurements (gray bars).

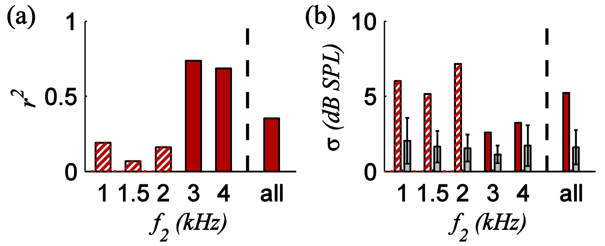

Figure 3: (a) Squared correlation coefficient r^2 of BTH with EDPT as function of f_2 and for the pooled data ("all"). (b) shows the corresponding standard deviation σ of the auditory-threshold estimate. Filled bars emphasize data exhibiting a highly significant correlation ($P < 0.001$). Gray bars in (b) depict the standard deviations of the BTH measurements.

Fig. 4 shows the BTH as a function of the sound pressure level of the EDPT for the pooled data (Fig. 4a), $f_2 = 3$ kHz (Fig. 4b), and 4 kHz (Fig. 4c). For comparison, the black circles in Fig. 4a show data acquired with continuous stimulus tones in the frequency range $1.5 \leq f_2 \leq 2.5$ kHz from a recent study [4]. The estimated thresholds obtained with the conventional measurement paradigm, i.e. continuous stimulation, exhibits pronounced scatter and high standard deviation of 10.4 dB SPL. The agreement of the red regression lines with the measurement data indicates a linear relationship between the BTHs and the EDPT values computed from the I/O-functions when the extracted non-linear generation component is used.

For some I/O-functions the *a priori* determined sampling instant for onset-decomposition was insufficient to extract the primary-source component, causing an exclusion rate of 3.8% (3/80) according to the objective evaluation criteria for I/O-functions, as well as outliers in the auditory-threshold estimate.

Figure 5 shows the median values of the slopes of the semi-logarithmic I/O-functions (red) (a) and the normalized slopes (b) as function of f_2. The gray area covers 68.3% of the data computed as the 15.85%- and 84.15%-quantiles. The magnitude of the slopes varies inter-individually but exhibits a similar frequency dependence for all subjects with a maximum slope at $f_2 = 1.5$ kHz, except for three subjects with peak slopes at 1 and 2 kHz. Above frequency of the maximum the slope decreases monotonically with 6dB/octave.

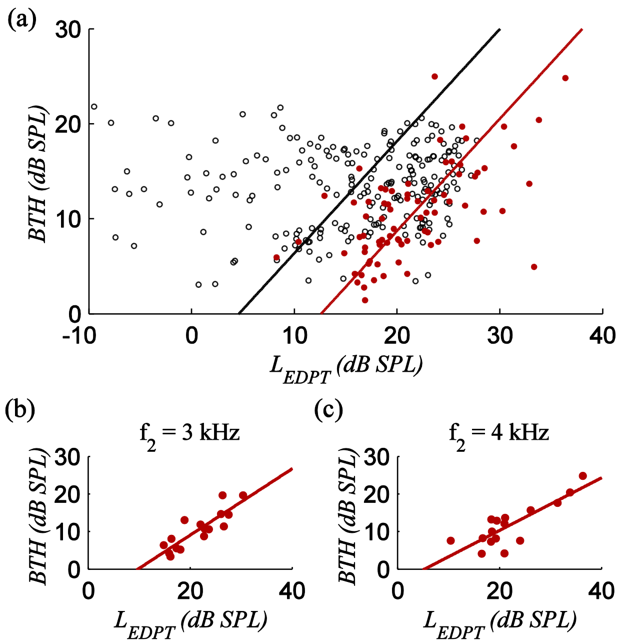

Figure 4: Auditory threshold acquired with Békésy audiometry as a function of the sound pressure level of the estimated threshold L_{EDPT} (red dots) gained from DPOAE I/O-functions for pooled data (a) and for frequencies $f_2 = 3$ kHz (b) and 4 kHz (c). Black circles in (a) represent data acquired with continuous stimulus tones from a recent study [4] exhibiting a standard deviation of 10.4 dB SPL. Black and red lines depict the corresponding regression lines with slope $s = 1.18$ (a), 0.88 (b), and 0.70 (c). Values for the squared correlation coefficient are $r^2 = 0.74$ ($f_2 = 3$ kHz) and $r^2 = 0.69$ ($f_2 = 4$ kHz).

The finding of maximum slope at 1.5 kHz is in good agreement with middle-ear transfer functions, which show on average an amplitude maximum at 1.2 kHz [10]. It is known that the slope of the I/O-function depends on the forward and reverse transmission properties of the middle ear [11]. The average measurement time to acquire all EDPTs per subject was 13.6 ± 3.7 min. Because the termination condition for data acquisition was applied to the DPOAE presumably exhibiting the smallest SNR, i.e. for f_2 = 1 kHz, the signal quality is considerably better for higher frequencies as a result of the large number of additional acquisition blocks for those frequencies.

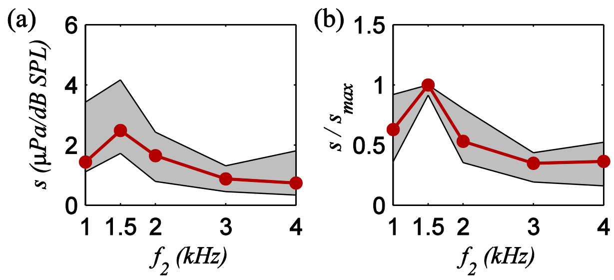

Figure 5: (a) Red line depicts the median values of the slope for the pooled data while the gray area shows the coverage of 68.3% of the data. (b) Same as (a), but with slope values normalized to the maximum of each subject.

4 Conclusion

Multi-frequency acquisition of short-pulse DPOAE I/O-functions enables quasi-simultaneous auditory-threshold estimation for a set of frequencies with high accuracy and reasonable cost of measurement time. The sequential arrangement of the primary tones allows shortening of the DPOAE stimulus to 40 ms while the analysis window increases to the combined length of all f_1-tones, i.e. 200 ms in this study. Thus, the I/O-functions exhibit a reduced susceptibility to long DPOAE decay times, for instance due to multiple internal reflections [12] or the synchronization of spontaneous otoacoustic emissions to the distortion product [13].

The presented method achieves an auditory-threshold estimate with low standard deviation covering a clinically relevant frequency range between f_2 = 1 and 4 kHz. For $f_2 \geq 3$ kHz the standard deviation of the auditory-threshold estimates is only 1.5 dB SPL higher than the corresponding standard deviation of the BTH measurements. Furthermore, the consistency of the normalized slope values over subjects (Fig. 5b) suggests that the slope of the DPOAE I/O-functions can be used as an additional diagnostic parameter to differentiate between conductive and sensorineural hearing loss [4].

The multi-frequency acquisition of DPOAE I/O-functions allows objective estimation of the auditory threshold with lower standard deviation than hitherto possible and with reasonable expenditure of time. Therefore, this method represents a promising diagnostic tool for objective auditory-threshold assessment, especially for the newborn hearing screening and in the field of pediatrics.

4 References

[1] Lonsbury-Martin B, Martin GK. Otoacoustic emissions. Curr Opin Otolaryngol Head Neck Surg 2003; 11: 361-366.

[2] Shera C. Mechanisms of mammalian otoacoustic emissions and their implications for the clinical utility of otoacoustic emissions. Ear Hear 2004; 25: 86-97.

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

[4] Dalhoff E, Turcanu D, Vetešník A, Gummer AW. Two-source interference as the major reason for auditory-threshold estimation error based on DPOAE input-output functions in normal-hearing subjects. Hear Res 2013; 296:67-82.

[5] Boege P, Janssen T. Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission I/O-functions in normal and cochlear hearing loss ears. J Acoust Soc Am 2002; 111: 1810-1818.

[6] Vetešník A, Turcanu D, Dalhoff E, Gummer AW. Extraction of sources of distortion product otoacoustic emissions by onset-decomposition. Hear Res 2009; 256: 21-38.

[7] Zelle D, Gummer AW, Dalhoff E. Extraction of otoacoustic distortion product sources using pulse basis functions. J Acoust Soc Am 2013; 134: EL64-69.

[8] Kummer P, Janssen P, Hulin P, Arnold W. Optimal L_1-L_2 primary tone level separation remains independent of test frequency in humans. Hear Res 2000; 146: 47-56.

[9] Whitehead ML, Stagner BB, Martin GK, Lonsbury-Martin BL. Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency. J Acoust Soc Am 1996; 100: 1663-1679.

[10] Aibara R, Welsh JT, Puria S, Goode RL. Human middle-ear sound transfer function and cochlear input impedance. Hear Res 2001; 152: 100-109.

[11] Kummer P, Schuster EM, Rosanowski F, et al. The influence of conductive hearing loss on DPOAE threshold. The effect of an individually optimized stimulation. HNO 2006; 6: 457-467.

[12] Dhar S, Talmadge CL, Long GR, Tubis A. Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. J Acoust Soc Am 2002; 112: 2882-2897.

[13] van Dijk P, Wit HP. Synchronization of spontaneous otoacoustic emissions to a $2f_1$-f_2 distortion product. J Acoust Soc 1990; 88: 850-856.