Hearing Research 296 (2013) 67–82

Contents lists available at SciVerse ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Two-source interference as the major reason for auditory-threshold estimation error based on DPOAE input–output functions in normal-hearing subjects

Ernst Dalhoff [1], Diana Turcanu [1], Aleš Vetešník [2], Anthony W. Gummer*

Eberhard-Karls-University Tübingen, Department of Otolaryngology, Section of Physiological Acoustics and Communication, Elfriede-Aulhorn-Straße 5, 72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 16 October 2012
Received in revised form
23 November 2012
Accepted 7 December 2012
Available online 23 December 2012

ABSTRACT

Fine structure in the frequency response of distortion product otoacoustic emissions (DPOAEs) can severely limit the usefulness of DPOAEs in estimating auditory thresholds. Here, fine structure is removed by extracting the primary-source DPOAE component using the onset-decomposition technique (Vetešník et al., 2009) and auditory threshold estimates are compared to those obtained from DPOAEs in response to conventional, continuous two-tone stimulation. Auditory thresholds are predicted using the estimated distortion product thresholds (EDPTs), obtained from linear regression of input–output (I/O) functions of DPOAE pressure amplitude versus second-tone stimulus level (Boege and Janssen, 2002). The accuracy of the auditory-threshold predictions is derived by comparison with measured auditory thresholds. The parameters of the two primary stimulus tones of frequency f_1 and f_2 and levels of L_1 and L_2 are chosen as: $f_2/f_1 = 1.2$ with $1.5 \leq f_2 \leq 2.5$ kHz, and $L_1 = 0.4L_2 + 39$ dB SPL, with $25 \leq L_2 \leq 65$ dB SPL. Data are from 12 normal-hearing subjects with profound DPOAE fine structure. 255 DPOAE I/O functions were measured for each of the two DPOAE paradigms. An EDPT value was accepted as reliable if: 1) the squared correlation coefficient, $r^2 \geq 0.8$, 2) the regression slope, $s_{I/O} \geq 0.2$ μPa/dB, and 3) the standard deviation of the EDPT, $\sigma_{EDPT} \leq 10$ dB. The proportion of rejected I/O functions was 8% for onset-decomposition DPOAEs, and 25% for continuous-tone DPOAEs. Removal of data points from the saturation region of the DPOAE I/O function by an automated algorithm reduced the rejection rate, to zero for onset-decomposition DPOAEs, but to only 13% for continuous-tone DPOAEs. In the absence of saturated DPOAE responses, auditory thresholds were predicted with standard deviation of only 4 dB for onset-decomposition DPOAEs, but 12 dB for continuous-tone DPOAEs. In summary, by extracting the primary-source component of the DPOAE by the method of onset-decomposition it is possible to predict human auditory threshold with hitherto unattainable accuracy.

1. Introduction

The potential of otoacoustic emissions (OAEs) for objective diagnosis of the state of the cochlea has been recognized ever since their discovery (Kemp, 1978). Within the decade to follow, the concept of an active, vulnerable cochlear amplifier emerged, and its relation to the OAEs was established [for review: Probst et al. (1991)]. Today, OAEs are viewed as a by-product of the active cochlear amplifier and, as such, as a tool for diagnosing the mechanical state of the cochlea. Clinically, OAEs are used for newborn screening and are also part of the audiometric test battery for routine diagnosis. However, their use is typically restricted to a dichotomous decision (normal/abnormal) about the state of the cochlear amplifier (functional/dysfunctional) [for review: Shera (2004)].

Abbreviations: BTH, Békésy threshold; 120-Hz BTH, 120-Hz bandwidth averaged Békésy threshold; CAP, compound action potential; DPOAE, distortion product otoacoustic emission; EDPT, estimated distortion product threshold; I/O, input/output; IHC, inner hair cell; OHC, outer hair cell; PT, pure tone; PT-BTH, pure-tone Békésy threshold; SD, standard deviation; SNR, signal-to-noise ratio.

* Corresponding author. Tel.: +49 7071 2988191; fax: +49 7071 294174.
E-mail addresses: ernst.dalhoff@uni-tuebingen.de (E. Dalhoff), diana.turcanu@uni-tuebingen.de (D. Turcanu), ales.vetesnik@fjfi.cvut.cz (A. Vetešník), anthony.gummer@uni-tuebingen.de (A.W. Gummer).
URL: http://www.uni-tuebingen.de/cochlea

[1] ED and DT contributed equally to this work.

[2] Present address: Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Nuclear Chemistry, Břehová 7, 115 19 Prague, Czech Republic.

0378-5955/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.heares.2012.12.003

Among the different types of OAE, transiently evoked OAE (TEOAE) and distortion product OAE (DPOAE) have been employed most extensively for diagnostic purposes. The commonly used DPOAE is the distortion product at the (so-called) low-side, cubic difference frequency, $2f_1 - f_2$, where f_1 and f_2 are the frequencies of the primary stimulus tones. DPOAEs in contrast to TEOAEs, have been said to provide: 1) frequency-specific information (Kimberley and Nelson, 1989), at least partially free from interfering contributions from frequencies other than that being analysed (Avan et al., 1997; Yates and Withnell, 1999), and 2) better signal-to-noise ratio (SNR), making it possible to measure DPOAEs up to a hearing loss of 50 dB HL (Gorga et al., 1993). In addition, DPOAEs differ in a unique manner from that of other OAEs with respect to their generation process. While all OAEs (up to moderate SPLs) rely on nonlinear cochlear amplification, this being the property promising diagnostic utility, all OAEs except DPOAEs also require a mechanical irregularity, which may be described as an irregularity in the impedance of the basilar membrane (Zweig and Shera, 1995). The degree of the irregularity can only be estimated by comparison between computed and measured OAEs, and to the best of our knowledge no quantitative information is available regarding its intersubject variability or its dependence on pathological mechanisms. In the absence of such information the responses resulting from these irregularities will likely introduce additional sources of error. Therefore, DPOAEs are attractive diagnostically because they appear to be more directly linked to the nonlinear amplification process than other OAEs. However, direct comparison of the diagnostic utility for hearing loss has not revealed marked differences between TEOAEs and DPOAEs (Gorga et al., 1993), showing only a modest advantage for DPOAEs at high frequency (4 kHz), and for TEOAE at low frequencies (500 Hz–1 kHz). A more recent investigation of DPOAEs by Burke et al. (2010) suggests that also at the lower frequency of 2 kHz an optimized stimulus paradigm (L_2 = 50 dB SPL, L_1 = 59 dB SPL, instead of $L_2 = 50$ dB SPL, $L_1 = 65$ dB SPL) can lead to DPOAE test performance superior to the TEOAE test performance reported by Gorga et al. (1993). On the other hand, in a large newborn screening study by Norton et al. (2000), DPOAE (but without using the optimized stimulus paradigm) did, if anything, slightly underperform at all frequencies compared with TEOAEs. Therefore, based on information from these larger screening studies, one might conclude that the DPOAE has not proven to be diagnostically superior to the TEOAE, in spite of its presumed superiority with respect to frequency specificity.

A confounding aspect for the interpretation of DPOAE measurements is the presence of a so-called fine-structure (Kemp and Brown, 1983; Gaskill and Brown, 1990; He and Schmiedt, 1993); that is, a periodicity in the frequency response of the DPOAE amplitude in the order of 1/10 octave and up to 20-dB peak-to-peak amplitude variation, which does not correlate with auditory threshold. Therefore, fine structure must contribute to the error in threshold estimation based on DPOAEs.

Today, there is good evidence that the DPOAE is generated mainly by two sources in the organ of Corti and that the DPOAE fine structure (at least up to moderate stimulus levels of, say, 55 dB SPL) is due to interference of the waves from the two sources (Shera and Guinan, 1999; Talmadge et al., 1999; Shera, 2004). The primary source is located at the place of maximum overlap of the two travelling-wave envelopes produced by the two stimulus tones, namely at approximately the f_2-place. Both stimulus components are processed simultaneously by the nonlinear mechanoelectrical transducer in the OHC stereocilia, leading to distortion products in the electromechanical force generated by the OHC soma that are then coupled into the cochlear fluids, organ of Corti and tectorial membrane. The distortion products propagate both retrograde to the stapes footplate and anterograde to their respective characteristic places along the cochlea. The anterograde waves are coherently reflected at these places from mechanical irregularities, giving rise to retrograde waves that can also propagate back to the stapes. The reflection process generating these secondary retrograde waves is called the secondary source. The retrograde waves emanating from the primary and secondary sources can interfere, the amount of constructive and destructive interference – the basis of the fine structure – depending on the relative amplitudes and phases of the waves from the two sources. Restricting discussion to the most prominent DPOAE in human – the DPOAE at $2f_1 - f_2$ evoked with a primary frequency ratio of $f_2/f_1 = 1.2$ –, the characteristic place of the secondary source lies only 0.6 octave apical to the characteristic place of the primary source (located near the f_2 place). Therefore, the secondary-source components can be comparable in strength to the primary-source components and, thus, lead to strong interference effects. If not accounted for, these effects can introduce errors in the interpretation of the DPOAE responses.

To date, four methods have been developed to remove the interfering effects of the secondary DPOAE source (Whitehead et al., 1996; Heitmann et al., 1998; Kalluri and Shera, 2001; Long et al., 2008) and of these, two have been investigated extensively. In the first extensively studied method, following Kemp and Brown (1983), Heitmann et al. (1998) used a third tone close to the DPOAE frequency (25 Hz above) to suppress the secondary-source component. In their group of five subjects, the stimulus conditions ($L_1 = 65$ dB SPL, $L_2 = 55$ dB SPL, $L_{\text{suppressor}} = 50$–$60$ dB SPL) led to substantial suppression of the DPOAE fine structure. The suppression technique is appealing for clinical applications, because it requires no additional stimulus time. However, later investigations with a larger number of subjects and lower primary levels showed considerable variability in the optimum suppressor level (Dhar and Shaffer, 2004; Johnson et al., 2006). Ultimately, in two large studies, no improvement in dichotomous test performance was detected (Johnson et al., 2007; Kirby et al., 2011). The second extensively studied method, called “spectral smoothing” or “time-windowing” (Kalluri and Shera, 2001), is based on filtering the frequency response of the DPOAE over a frequency band comprising a sufficiently large number of frequencies to resolve the fine structure on the one hand and at the same time using a sufficiently long stimulus duration to ensure that end effects are insignificant (Kalluri and Shera, 2001; Mauermann et al., 2004; Vetešník et al., 2009). Typically, measurements are performed at 40 stimulus frequencies (Kalluri and Shera, 2001). Whereas the suppression technique is fast, time windowing is inherently slow, presenting a major obstacle for clinical applications. Shaffer and Dhar (2006) compared the time-windowing technique with the suppression technique at moderate stimulus levels (45–65 dB SPL) in 10 subjects and found that time-windowing performed better in suppressing the fine structure. However, individual threshold estimation accuracy was not assessed. A third method of separating the components of the two DPOAE sources performs the analysis exclusively in the time domain. To distinguish the DPOAE from the primaries, several stimulus blocks are presented with different phase offsets of the two primary tones (Whitehead et al., 1996). The phase offsets and the averaging procedure are chosen such that, assuming time invariance, the two primary tones cancel in the resultant signal, whereas the DPOAE is constructively averaged. Thus, the DPOAE can be recorded without processing in the frequency domain and the time course of the DPOAE can be analysed after switching on one (or both) of the primary tones (Talmadge et al., 1999; Whitehead et al., 1996). Specifically, due to their different time delays, the transient responses of the two DPOAE sources are separated in the time domain. Recently, we