VNiVERSiDAD
D SALAMANCA

PHYSIOLOGICAL AND PSYCOACOUSTICAL ESTIMATION OF HUMAN COCHLEAR INPUT/OUTPUT CURVES

Ph.D. Thesis

Peter Tinggaard Johannesen

November 2013

Instituto de Neurociencias de Castilla y León

Universidad de Salamanca

Enrique Alejandro López Poveda, Profesor Contratado Doctor Permanente, adscrito al Departamento de Cirugía de la facultad de Medicina de la Universidad de Salamanca e investigador del Instituto de Neurociencias de Castilla y León,

CERTIFICA

que la tesis doctoral titulada *"Physiological and psycoacoustical estimation of human cochlear input/output curves"* describe el trabajo de investigación realizado por D. Peter Tinggaard Johannesen bajo mi dirección durante los últimos seis años.

La memoria de tesis describe un conjunto de estudios psicofísicos, fisiológicos y de simulación diseñados con el objetivo de aclarar cómo inferir mejor las curvas de entrada/salida de la cóclea humana. Los resultados de este trabajo indican que sería posible inferir las curvas de entrada/salida en la región basal mediante la técnica de otoemisiones acústicas, al menos en personas con audición normal. Los resultados constituyen un paso adelante para poder realizar un ajuste más individualizado de los audífonos, lo cual solo sería posible conociendo estas curvas. Por ello, este estudio no sólo tiene interés científico, sino también clínico.

Cada uno de los capítulos de esta tesis ha sido publicado, en su totalidad o en parte, en artículos científicos en prestigiosas revistas especializadas de investigación.

Considero, por tanto, que esta tesis doctoral reúne la calidad y el rigor científico necesarios para ser defendida publicamente en nuestra Universidad como requisito parcial para que D. Peter Tinggaard Johannesen opte al grado de Doctor.

En Salamanca, a 28 de Octubre de 2013

Enrique A. López Poveda
Director de la tesis

ABSTRACT

Human cochlear input/output (I/O) curves are not completely understood because they can be obtained using indirect methods only. The temporal masking curve (TMC) method is a favoured psychophysical method to infer human cochlear I/O curves, but is inconvenient for clinical applications. Distortion Product Otoacoustic Emissions (DPOAE) I/O curves share many of the characteristics of cochlear I/O curves and could be a useful alternative to the TMC method, but its generation mechanisms are not completely known, and it is uncertain how DPOAE I/O curves relates to the cochlear I/O curve. The aim of the present work is to test if the two methods can be used indistinctly to infer human cochlear I/O curves in normal hearing listeners. The approach is to compare individual estimates of cochlear I/O curves inferred with the two methods. If the results are consistent, this would provide support for the assumptions of both methods.

The results showed reasonably good correspondence between I/O curves of the two methods for frequencies above ~2 kHz but not for lower frequencies. At lower frequencies, the DPOAE I/O curves frequently presented plateaus and notches, which were not present in the I/O curves inferred from TMCs. The DPOAE I/Os were measured using the group average stimulation level parameters (primary level rule) of Kummer *et al.* (1998).

Simulations are presented that aimed at testing if individual differences from the DPOAE group average primary level rule could explain the plateaus and notches in the DPOAE I/O curves at lower frequencies. The results suggest that the primary level rule that maximizes the DPOAE level is also a good primary level rule to estimate the I/O curve of the underlying non-linearity and that even a small deviation from this rule may lead to notches in the DP I/O curve.

It is a common hypothesis that maximum DPOAE levels occur when the basilar membrane excitation by the two primary tones is equal at the cochlear site

tuned to the higher of the two primary frequencies. A novel TMC-based approach is also presented here designed to test this hypothesis. The results support this hypothesis as the levels required for equal excitation inferred from the TMC method coincide with the empirically found DPOAE primary levels that produce maximum response.

Cochlear I/O curves as inferred from TMCs were finally compared to estimates obtained from DPOAEs I/O using two additional primary level rules. The first rule consisted in individualized primary levels inferred from TMCs that obtained equal cochlear excitation and the second used empirically found primary levels optimized individually for maximum DPOAE level.

The results showed that the correspondence between I/O curves inferred from TMCs and DPOAEs remained high at frequencies above ~2 kHz but did not improve at lower frequencies, independently of the primary level rule used.

Reasons for the poorer correspondence between I/O curves inferred from TMCs and DPOAEs at lower frequencies and particularly for the plateaus and notches in the DPOAE I/Os are discussed. Directions of future research are suggested.

ACKNOWLEDGEMENTS

First of all, I would like to thank Enrique for his inspiration and enthusiasm for science that he generously spreads to anyone around him. Enrique has an incredible talent for producing new ideas and not just follow the “common thinking”, and this leads to novel science. In fact, many of the good ideas in this thesis should be attributed to Enrique. Although I had previous experience in scientific research, Enrique should clearly be considered my “scientific father”. He taught me many good lessons of science and helped me to reach a much higher level as a researcher.

I am in debt with the many volunteers who “lent me their ears” for many hours, and without whom the data could not have been obtained. The only compensation for many of them was that they were contributing to science and thereby improving the quality of life of future generations.

Many thanks also go to my laboratory fellows, especially Almudena and Patricia who “suffered” more hours of testing than anybody else.

Each chapter of this thesis is adapted, with permission, from published articles. I am grateful to the editors and reviewers for their contributions to improve these articles.

I am also very grateful to my wife Obdulia and my son Marco who supported me especially during the preparation of this thesis. I hope not to have been too absent from their lives during the last months.

This work was carried out over a number of years with support from several institutions:

- IMSERSO 131/06
- The Spanish Ministry of Science and Education (BFU2006-07536 and CIT-390000-2005-4)
- The Junta de Castilla y León (GR221)
- The William-Demant Oticon Foundation.

LIST OF ABBREVIATIONS AND ACRONYMS

ANSI	American National Standards Institute
BM	Basilar Membrane
CF	Characteristic Frequency
dB	Decibels
DP	Distortion Product
DPOAE	Distortion Product Otoacoustic Emissions
f_P	Probe Frequency
f_M	Masker Frequency
f_1, f_2	DPOAE primaries frequencies
FFT	Fast Fourier Transform
GOM	Growth Of Masking
HI	Hearing Impaired
HL	Hearing Level
I/O	Input/Output
L_M	Masker Level
L_1, L_2	Level of f_1 and f_2
LR	Linear Regression
NH	Normal Hearing
OHC	Outer Hair Cell
PT	Pure Tone
RMS	Root Mean Square
SPL	Sound Pressure Level
SD	Standard Deviation
SE	Standard Error
SL	Sensation Level
TMC	Temporal Masking Curve

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

1.1 Motivation and background

Humans can perceive sounds over a wide range of sound pressure levels (SPL) most likely thanks to the functioning of the outer hair cells (OHC) and their amplification effect on basilar membrane (BM) vibrations (e.g., Oxenham and Bacon 2003). Indeed, the healthy BM is sensitive to very low level sounds and the magnitude of its response grows compressively with increasing sound level (see Fig. 1.1 and 1.2), thus accommodating a range of ~120 dB SPL to a narrower range of physiological responses (e.g., Robles and Ruggero 2001). The damaged BM, by contrast, shows reduced sensitivity and linearized responses (see Fig. 1.1), which likely explains why hearing-impaired listeners show abnormally high thresholds and narrower dynamic ranges (e.g., Moore 2007). The BM response characteristics are therefore important for describing hearing and are often quantified as BM I/O curves, which are graphical representations of the relationship between stimulus sound level and mechanical BM response magnitude (Fig. 1.1). When measured directly in rodents using the characteristic frequency (CF) of the recording site, the BM I/O curves typically consist of a linear response growth (~1 dB/dB) at low input levels followed by a compressive segment (growth < 1 dB/dB) at medium input levels and finally sometimes it has a linear response growth again at very high input levels (Robles and Ruggero 2001).

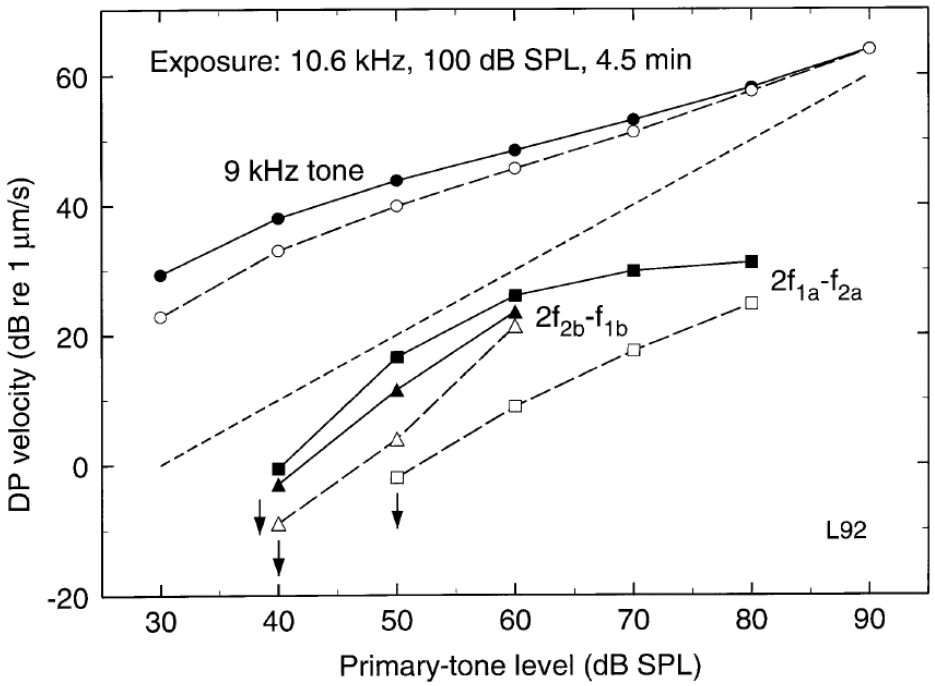

Figure 1.1. Example of BM measurement of I/O curves. Filled circles represent an I/O curve for a pure tone at the characteristic frequency (9 kHz) of the recording site. The filled squares and triangles show I/O curves for the $2f_1$-f_2 and $2f_2$-f_1 DP components, respectively. The dotted line indicates linear growth. Open symbols represent pure tone and DP I/O curves after noise exposure. Figure taken from Robles *et al.* (1997).

Despite its importance to hearing, the I/O characteristics of the human BM response are not completely understood. Human BM responses cannot be measured directly so indirect techniques are used to infer I/O curves. There exist various methods to infer BM I/O curves from perceptual data (e.g., Lopez-Poveda and Alves-Pinto 2008; Nelson *et al.,* 2001; Oxenham and Plack 1997; Plack and Oxenham 2000). These behavioral methods are reasonably well grounded (e.g., Bacon and Oxenham 2004; Oxenham and Bacon 2004). The different methods yield similar within-subject results (Lopez-Poveda and Alves-Pinto 2008; Nelson *et al.,* 2001; Rosengard *et al.,* 2005) and they are replicable across different measurement sessions.

The temporal masking curve (TMC) method of Nelson *et al.* (2001) (see also Lopez-Poveda *et al.*, 2003) is a favoured psychoacoustical method [reviewed in Bacon (2004)] because it minimizes off-frequency listening effects that might occur

when the probe level is varied. This method consists of measuring the level of a tonal forward masker required to just mask a *fixed* tonal probe as a function of the time interval between the masker and the probe. A TMC is a graphical representation of the resulting masker levels against the corresponding masker–probe intervals. Because the probe level is fixed, the masker level increases with increasing masker–probe time interval and hence TMCs have positive slopes. Nelson *et al.* (2001) argued that the slope of any given TMC depends simultaneously on the amount of BM compression affecting the masker at a cochlear place whose CF equals approximately the probe frequency and on the rate of decay (or recovery) of the internal (postcochlear or postcompression) masker effect. By assuming that the decay rate is identical across masker frequencies and levels, BM I/O functions may be estimated by plotting the masker levels of a linear reference TMC (i.e., the TMC for a masker that is processed linearly by the BM) against the levels for any other masker frequency, paired according to masker–probe delays. Nelson *et al.* (2001) provide a full justification of these assumptions [see also Lopez-Poveda and Alves-Pinto (2008)].

The TMC method has been widely used to infer BM I/O curves in normal-hearing and hearing-impaired listeners (e.g., Lopez-Poveda *et al.*, 2003, 2005; Lopez-Poveda and Alves-Pinto, 2008; Nelson *et al.*, 2001; Nelson and Schroder, 2004; Plack *et al.*, 2004; Rosengard *et al.*, 2005). It is arguably a reliable method but time consuming and requires the active participation of the listener. This would make it inconvenient for clinical purposes, particularly for testing non-cooperative subjects, like newborns or the elderly.

The human ear is not a high-fidelity system. It distorts acoustic signals within the cochlea (Ruggero 1993). The distortions can be perceived as audible sounds (Goldstein 1967) and are emitted from the cochlea back to the ear canal as otoacoustic emissions (Kemp 1978). Indeed, emitted distortions are a sign of a healthy ear: the weaker emission, the greater the cochlear damage (Dorn *et al*. 2001; Lonsbury-Martin and Martin 1990). The level of these emissions also depends on the parameters of the sounds used to evoke them. Typically, two pure tones (or

primaries) of slightly different frequencies (f_1 and f_2; f_2/f_1~1.2) are used and the level of the $2f_1-f_2$ emitted distortion in the ear canal is regarded as an indicator of the physiological status of the cochlea (Gorga *et al.* 1997). We will refer to this indicator as the distortion product otoacoustic emission (DPOAE). The graphical representation of the $2f_1-f_2$ distortion product (DP) emission as a function of the level, L_2, of the primary tone f_2 is referred to as a DPOAE I/O curve (Fig. 1.2).

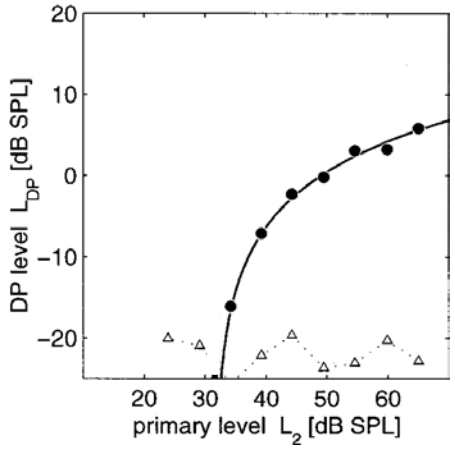

Figure 1.2. A typical DPOAE I/O curve for a normal hearing subject (filled circles and solid line). Open triangles represent the measurement noise floor. Figure taken from Boege and Janssen (2002).

DPOAE I/O curves share many characteristics with BM I/O functions (Cooper and Rhode, 1997; Dorn *et al.*, 2001; Neely *et al.*, 2003). Specifically, both of them are generally linear at low levels but become compressive above a certain compression threshold (Dorn *et al.*, 2001; Kummer *et al.*, 1998) and both of them are similarly labile to OHC dysfunction (see also Fig. 1.1) (Rhode, 2007). This suggests that DPOAE I/O functions could be used as a faster and universal way to infer individual BM I/O functions in clinical conditions (Müller and Janssen, 2004). Measuring DPOAEs does not require active participation from the listeners and so it has been suggested that DPOAE I/O curves could provide a useful alternative to behavioral methods to infer BM I/O curves (e.g., Jansen and Müller 2008; Müller and Janssen 2004). Unfortunately, it is still uncertain that DPOAE I/O curves constitute reasonable estimates of BM I/O curves on an *individual* basis (e.g., Williams and Bacon 2005).

The long term goal of the present research is to define DPOAE stimuli and conditions that would allow using DPOAE I/O curves as a reliable alternative to behavioral methods for inferring individualized, frequency-specific cochlear I/O curves. Although the long-term goal is to extend the study to hearing-impaired listeners, the focus here is on normal-hearing listeners.

1.2 Hypotheses

The overall hypothesis is that behaviorally inferred cochlear I/O curves and DPOAE I/O curves should be mutually consistent as both depend on the cochlear I/O characteristics.

The specific hypotheses are:

1. Behavioral TMC and physiological DPOAE methods can be used indistinctly to infer human cochlear I/O curves.
2. DPOAE I/O curves resemble closer the I/O curve of the underlying causal nonlinearity when measured with parameters optimized for maximum DP response.
3. DPOAE response is maximal when the primary tones evoke approximately the same excitation at the BM site tuned to f_2.
4. The correspondence between I/O curves inferred from TMCs and DPOAE I/O curves improves by using DPOAE parameters optimized individually for maximal response.

1.3 Objectives

The objectives of this thesis address one by one the hypotheses:

1. To compare I/O curves in NH listeners as measured by DPOAE and inferred psychophysically by the TMC method.

2. To simulate DP I/O curves for two simple nonlinearities and for two primary level rules where one is designed for maximal response.
3. To compare level rules inferred from TMCs designed to evoke equal BM excitation with empirically found DPOAE primary level rules that produce maximal response.
4. To compare I/O curves inferred from TMCs with DPOAE I/O curves measured with primaries individually optimized for maximal response.

1.4 Overview of this thesis

This thesis is organized so that each chapter addresses one of the proposed hypothesis/objectives.

Chapter 2 addresses the first hypothesis/objective by presenting a comparison between I/O curves inferred from TMC with DPOAE I/O curves obtained using the group average stimulus rule of Kummer *et al.* (1998).

Chapter 3 presents simulations aimed at testing the second hypothesis. DP I/O curves were simulated for two different nonlinearities and for two primary level rules and compared to the underlying causal nonlinearities. One primary rule had equal primary levels and the other was optimized in the sense of generating maximum distortion. The impact on the estimated I/O curve when the primaries deviate slightly from the desired primary level rule was also analyzed.

Chapter 4 presents a novel psychophysical approach to finding the individual optimal primary rule and it also tests the third hypothesis that maximum DPOAE generation occurs when the BM excitation of the two primaries is equal at the BM site tuned to f_2.

In Chapter 5 individually optimized primary level rules are used to measure DPOAE I/O curves to test if the correlation between DPOAE I/O curves and psychophysically inferred I/O curves improves (fourth hypothesis/objective).

Chapters 2, 4 and 5 are each based on papers already published and can be read independently. The original published articles are reproduced in Appendix B. Chapter 3 was originally published as an appendix to the published article on which Chapter 5 is based.

Chapter 6 contains a general discussion of issues common to all chapters and directions for future work are suggested.

Lastly, the conclusions from the research are summarized.

1.5 Original contributions

The main original contributions of this thesis are:

1. A demonstration that normal hearing human cochlear I/O curves as estimated with TMCs and DPOAEs are consistent for frequencies above ~2 kHz but not for lower frequencies.
2. Simulations that suggest that DP I/O curves resemble better the underlying non-linearity when the primaries are optimized for maximal DP response.
3. Novel psycoacoustical support to the common hypothesis that maximum DPOAE levels occur when the primaries evoke equal BM excitation at the BM site tuned to the higher of the DPOAE primaries (f_2).
4. Experimental evidence that the correspondence between TMC and DPOAE inferred I/O curves do not improve using primaries individually optimized for maximal response.

1.6 Units

All units follow the conventions of International System of Units (SI).

CHAPTER 2.

COCHLEAR NONLINEARITY IN NORMAL-HEARING SUBJECTS AS INFERRED PSYCHOPHYSICALLY AND FROM DISTORTION-PRODUCT OTOACOUSTIC EMISSIONS[1]

2.1 Introduction

This chapter describes a first effort to determine whether DPOAE I/O functions may be used to estimate the degree of *individual* compression in normal-hearing listeners. This hypothesis is tested by measuring the degree of correlation between BM I/O functions as inferred from the psychophysical TMCs and from DPOAEs in the same subject. Clearly, only if the results of the two methods correlate well, will it be possible to add support to the use of DPOAEs to infer *individual* cochlear response characteristics. If they disagree, however, it will be difficult to resolve which of the two methods (if any) is more appropriate to reveal the true nonlinear characteristics of the underlying BM responses.

Several earlier studies have addressed this or related questions. Müller and Janssen (2004) investigated the similarity of loudness and DPOAE I/O curves in the same subject sample [Neely *et al.* (2003) had done it previously using different subject samples and slightly different methods]. They found a high resemblance between the characteristics gain and compression of the two sets of *average* I/O

[1] This chapter is adapted from the published article: Johannesen, P. T., and Lopez-Poveda, E. A. **2008**. "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **124**, 2149–2163. Some portions of the article are reproduced here with permission from the Acoustical Society of America.

curves in normal-hearing and hearing-impaired listeners. Müller and Janssen (2004) acknowledged, however, that loudness may be affected by retrocochlear mechanisms (see also Heinz and Young, 2004) and it is also thought that loudness is affected by off-frequency effects e.g., different spreads of excitation at different levels, which make it difficult to establish a one-to-one relationship between loudness and underlying BM I/O curves (Moore, 2003). This undermines the conclusions of Müller and Janssen (2004). Furthermore, their conclusions applied to *average* I/O curves and frequencies of 2–4 kHz, and thus may not be valid individually or for other frequencies, particularly 0.5 and 1 kHz.

Gorga *et al.* (2007) measured the degree of cochlear compression in a very large sample (N=103) of normal hearing listeners as estimated from DPOAE I/O functions at 0.5 and 4 kHz. As a consequence, the I/O functions they reported likely provide a good description of *average* normal responses. Their results supported the conclusion of earlier psychophysical studies that the degree of compression is similar for apical and basal cochlear sites (Lopez-Poveda *et al.*, 2003; Plack and Drga, 2003). However, their study did not include within-subject psychophysical/physiological comparisons.

Williams and Bacon (2005) inferred cochlear I/O curves from TMCs and DPOAEs in four listeners and for frequencies of 1, 2, and 4 kHz. The results revealed that both methods yield similar *average* compression estimates. Like the above-mentioned studies, this study was not intended to investigate within-subject correlations between the results of both methods. Furthermore, their DPOAE I/O curves could have been influenced by the DP fine structure. Indeed, Gaskill and Brown (1990) showed rapid variations (known as “fine structure”) of the magnitude of the $2f_1 - f_2$ DPOAE with changing the frequencies of the primaries (f_1 and f_2, with f_2 / f_1 =1.21) only slightly. This fine structure is thought to be the result of constructive and destructive interference between DPs generated at two spatially distant sites (Kummer *et al.*, 1995; Stover *et al.*, 1996; Gaskill and Brown, 1996; Heitmann *et al.*, 1998; Talmadge *et al.*, 1998; 1999; Mauermann *et al.*, 1999; Mauermann and Kollmeier, 1999; Shera and Guinan, 1999). The principal

generation site of DPOAEs is the BM region of maximum overlap between the excitation caused by the two primaries, that is the BM site with CF= f_2 (e.g., Kummer *et al.*, 1995). This component propagates back toward the oval window but also to the cochlear site with CF=$2f_1-f_2$ where it excites a second generation source termed the reflection source. The DP generated at this second source propagates back toward the oval window and is summed vectorially with the response of the first source. The fine structure is thought to originate from vector summation of these two components, whose varying phases give rise to constructive or destructive interference and thereby to peaks and valleys in the DP-gram. The f_2 generator site is the dominant source at high stimulus levels [Fig. 3 of Mauermann and Kollmeier (2004)], which explains why the fine structure is more pronounced at low levels.

The fine structure has a large influence on DPOAE I/O curves, especially at low levels. He and Schmiedt (1993) mentioned that the DPOAE magnitude can change by as much as 20 dB for a change in f_2 of 1/32 octave. Mauermann and Kollmeier (2004) reported that the response varied by 10–15 dB when f_2 varied over the interval from 2250 to 2610 Hz. The influence of the fine structure is greater for individual than for average (across subjects) I/O curves, but can also affect average curves when the sample size is small. The sample size was small (N=4) in the study of Williams and Bacon (2005). Thus, fine-structure effects may have complicated the interpretation of their results or even led to wrong conclusions.

The present work extends these earlier studies in several respects. First, the focus here is on within-subject as opposed to average psychophysical/physiological correlations. Second, psychophysical BM I/O curves were inferred using what is arguably the most accurate method available to date for this purpose [see Nelson *et al.* (2001) for a full justification; but see also Sec. 2.4 of the present chapter]. Third, special care was exercised to reduce the influence of the fine structure on individual DPOAE I/O curves by averaging the magnitude of the $2f_1- f_2$ DPOAE for five f_2 frequencies near the frequency of interest. Fourth, the frequency range considered (0.5–4 kHz) included low and high frequencies. Fifth, the

physiological/psychophysical comparisons extended to parameters pertaining to cochlear nonlinearity other than compression magnitude; specifically, compression threshold and the level at which maximum compression occurs.

It will be shown that reasonable correlation exists between the characteristics of individual TMC-based and DPOAE I/O curves at 4 kHz but not at 0.5 and 1 kHz. Reasons for the observed discrepancies at low frequencies will be discussed here and also in Section 6.1-6.3.

2.2 Methods

2.2.1 Subjects

Ten normal-hearing subjects participated in the study. Their age ranged from 20 to 39 years. Their absolute thresholds were measured using a two-down, one-up adaptive procedure. Signal duration was 300 ms, including 5-ms cosine-squared onset and offset ramps. All subjects had thresholds within 20 dB HL at the frequencies considered in this study (0.5, 1, 2, and 4 kHz, see Table 2.1).

Table 2.1. Thresholds in dB SPL measured with Etymotic ER2 insert earphones for all subjects and for tone durations of 300 ms absolute threshold, 110 ms masker threshold, and 10 ms probe threshold, respectively. n.a. stands for not available.

Frequency (kHz)	*0.5*	*1*	*4*
Tone duration (ms)	*300 / 110 / 10*	*300 / 110 / 10*	*300 / 110 / 10*
S1	13 / 15 / 39	6 / 8 / 30	13 / 13 / 33
S2	14 / 21 / 43	10 / 15 / 30	21 / 21 / 40
S3	13 / 16 / 39	7 / 8 / 32	4 / 10 / 25
S4	17 / 20 / n.a.	10 / 16 / 34	21 / 24 / 44
S5	23 / 24 / n.a.	13 / 17 / 41	0 / 3 / 23
S6	9 / 10 / 36	10 / 11 / 31	4 / 4 / 24
S7	6 / 10 / n.a.	10 / 11 / n.a.	12 / 15 / 31
S8	11 / 18 / 34	5 / 7 / 38	11 / 10 / 30
S9	12 / n.a. / n.a.	9 / n.a. / n.a.	10 / 15 / 31
S10	5 / n.a. / n.a.	–2 / n.a. / n.a.	3 / 1 / 23

2.2.2 TMC stimuli

TMCs were measured for probe frequencies (f_P) of 0.5, 1, 2, and 4 kHz and for masker frequencies equal (on-frequency) to the f_P. Additional TMCs were measured for a probe frequency of 4 kHz and a masker frequency of 1.6 kHz (f_M=0.4f_p). The latter were selected as the linear references (Lopez-Poveda and Alves-Pinto, 2008) and used to infer BM I/O curves for all probe frequencies (Lopez-Poveda *et al.*, 2003). The masker–probe time intervals ranged from 5 to 100 ms in 5-ms steps with an additional interval of 2 ms. The duration of the masker was 110 ms, including 5-ms cosine-squared onset and offset ramps. The probe duration was 10 ms, including 5-ms cosine-squared onset/offset ramps and no steady state portion. The level of the probe was fixed at 9 dB sensation level (SL)

(i.e., 9 dB above the individual absolute threshold for the probe), except for subject S5 for whom it was 15 dB SL. Stimuli were generated with a Tucker Davies Technologies Psychoacoustics Workstation (System 3) operating at a sampling rate of 48.8 kHz and with analog to digital conversion resolution of 24 bits. If needed, signals were attenuated with a programmable attenuator (PA-5) before being output through the headphone buffer (HB-7). Stimuli were presented to the listeners through Etymotic ER-2 insert earphones. TMC sound pressure levels (SPL) were calibrated by coupling the earphones to a sound level meter through a Zwislocki DB-100 coupler. Calibration was performed at 1 kHz only and the obtained sensitivity was used at all other frequencies because the earphone manufacturer guarantees an approximately flat (±2 dB) frequency response between 200 Hz and 10 kHz.

2.2.3 TMC procedure

The procedure was identical to that of Lopez-Poveda and Alves-Pinto (2008). Masker levels at threshold were measured using a two-interval, two-alternative forced-choice paradigm. Two sound intervals were presented to the listener in each trial. One of them contained the masker only and the other contained the masker followed by the probe. The interval containing the probe was selected randomly. The subject was asked to indicate the interval containing the probe. The initial masker level was set sufficiently low that the subject always could hear both the masker and the probe. The masker level was then changed according to a two-up, one-down adaptive procedure to estimate the 71% point on the psychometric function (Levitt, 1971). An initial step size of 6 dB was applied, which was decreased to 2 dB after three reversals. A total of 15 reversals were measured. Threshold was calculated as the mean of the masker levels at the last 12 reversals. A measurement was discarded if the standard deviation (SD) of the last 12 reversals exceeded 6 dB. Three threshold estimates were obtained in this way and their mean

was taken as the threshold. If the SD of these three measurements exceeded 6 dB, a fourth threshold estimate was obtained and included in the mean.

The maximum SPL was set to 104 dB to prevent subject discomfort and/or temporary threshold shifts. A measurement run was stopped and discarded when the subject reached this limit on more than two consecutive trials over the last 12 reversals. Masker levels at threshold were measured for masker–probe time intervals in increasing order. This was done to minimize the possibility that the measurements would be affected by potential temporary thresholds shifts that might have occurred if intervals had been presented in random order and a long interval high masker level immediately preceded a short interval. An attempt was made to measure masker levels for all masker–probe time intervals. Missing data indicate that the maximum output level (104 dB SPL) was reached for the time interval in question or that it was impossible within six to ten attempts to obtain three threshold estimates with SD $\leq$ 6 dB.

The listeners' absolute threshold for the maskers and probes were measured using the same equipment and conditions used to measure the TMCs. At least three measurements were obtained and averaged. Results are shown in Table 2.1. Listeners were trained in the forward-masking task for several hours, at first with a higher probe level of 15 dB SL and later with a probe level of 9 dB SL, until performance became stable. Listeners sat in a double-wall sound attenuating chamber during all measurements.

2.2.4 Inferring BM I/O functions from TMCs

BM I/O functions were inferred from TMCs by plotting the levels for the linear reference TMC against the levels for any other masker frequency paired according to masker–probe time interval (Nelson *et al.*, 2001). The off-frequency TMC for f_P=4 kHz was used as the linear reference to infer I/O curves for all other frequencies, as suggested by Lopez-Poveda *et al.* (2003). It was sometimes necessary to extrapolate the linear references to longer masker–probe time intervals

to infer BM I/O functions over the wider possible range of levels. In similar situations, some authors have fitted the linear reference TMC with a straight line (e.g., Lopez-Poveda *et al.*, 2003; 2005; Nelson *et al.*, 2001; Plack *et al.*, 2004). There is strong evidence, however, that the decay of forward masking is better described with two time constants (e.g., Lopez-Poveda and Alves-Pinto, 2008; Meddis and O'Mard, 2005; Oxenham and Moore, 1994; Plack and Oxenham, 1998). Based on this, the individual linear-reference TMCs were fitted here using a least-squares procedure with a double exponential function of the form:

$$L_m = L_0 - 20log10\left[\alpha e^{(-t/\tau_a)} + (1-\alpha)e^{(-t/\tau_b)}\right], \quad (2.1)$$

where $L_m(t)$ is the masker level required to mask the probe at masker–probe time interval t; L_0 is masker level for a masker–probe interval of zero; τ_a and τ_b are time constants; and α determines the level at which the second exponential takes over from the first one. α, τ_a, τ_b, and L_0 were fitting parameters and were allowed to vary freely within certain boundaries: α was restricted to the interval [0, 1]; L_0 was restricted to vary within [50,120] dB; and τ_a and τ_b to the interval [1,200] ms. The lowest correlation between actual and predicted masker levels was r=0.94, which shows that the goodness of fit was excellent.

2.2.5 DPOAE stimuli

DPOAE I/O curves were obtained by plotting the magnitude (in dB SPL) of the $2f_1 - f_2$ DP emission as a function of the level, L_2, of the primary tone f_2. DPOAEs were measured only for f_2 frequencies equal to the probe frequencies for which TMCs had been previously measured (0.5, 1, 2, and/or 4 kHz). The f_2/f_1 ratio was fixed at 1.2. L_2 ranged from 20 to 75 dB SPL in 5 dB steps, except for f_2 =0.5 kHz for which it ranged from 45 to 75 dB SPL. L_1 and L_2 were related according to L_1=0.4L_2+39 dB, the rule proposed by Kummer *et al.* (1998) to obtain maximum-level DPOAEs for $L_2 \leq$65 dB SPL. In the current study, this rule was extrapolated to L_2>65 dB SPL.

In an attempt to reduce the potential variability of the I/O curves caused by the DP fine structure, DPOAE I/O curves were measured for five f_2 frequencies near the frequency of interest, and the resulting I/O curves were averaged. This procedure is supported by Kalluri and Shera (2001) and Mauermann and Kollmeier (2004), who showed that a "cleaned" DP-gram i.e., a DP-gram where the fine structure has been accounted for resembles very closely a moving average of the original DP-gram with fine structure. The five adjacent frequencies were selected to differ by as much as 2% of the frequency of interest based on a suggestion of Mauermann and Kollmeier (2004) that this frequency spacing is appropriate to reveal or to account for the influence of the fine structure. That meant measuring DPOAE I/O curves for frequencies of $0.98f_2$, $0.99f_2$, f_2, $1.01f_2$, and $1.02f_2$. For instance, the final DPOAE I/O curve at 4 kHz was the mean of five I/O functions for f_2 ={3920, 3960, 4000, 4040, 4080 Hz}. To further assess the potential influence of the fine structure around the frequency of interest, DPOAEs were measured for four additional adjacent f_2 frequencies on each side of the five main frequencies. In the case of 4 kHz these were: f_2 ={3760, 3800, 3840, 3880, . . . ,4120, 4160, 4200, 4240 Hz}. DPOAEs for these latter frequencies were measured only for L_2=70, 60, 50, 40 dB and the resulting I/O curves were not included in the final mean I/O curve.

The influence of standing waves must be taken into consideration when measuring DPOAE at high frequencies. Siegel (1994) and Whitehead *et al.* (1995) found that restricting DPOAE measurements to f_2<6 kHz avoids the majority of the idiosyncratic variations of the sound pressure at the eardrum due to standing waves. For this reason, it was decided to restrict DPOAE measurements to $f_2 \leq 4$ kHz. Measuring DPOAEs for f_2<1 kHz is also problematic due to increased physiological subject noise. The moderate-to-high level part of the DPOAE I/O curve for f_2=0.5 kHz could still be measured for most listeners, however, by considering L_2 above 45 dB SPL only.

2.2.6 DPOAE stimulus calibration

DPOAE stimuli were calibrated with a Zwislocki DB-100 coupler for each f_1, f_2 frequency. In some studies, calibrated levels are further adjusted with the probe *in situ* to account for the acoustic effects associated with ear-canal resonances. This *in situ* adjustment, however, was not applied here for two reasons. First, Siegel (1994) has shown that it does not always work for frequencies above 2–3 kHz because of the errors in predicting the level at the eardrum from measurements made at the plane of the probe where the standing waves interact. Second, as this is a comparison study, it was deemed important that the two methods applied the same stimulus level control. Given that the psychophysical equipment did not allow easy *in situ* level adjustment, this option was disabled in the OAE instrument.

2.2.7 DPOAE system artifacts

When measuring DPOAE I/O curves at high primary levels it is necessary to control for cubic distortion produced by the measurement instrument. This system artifact was assessed by measuring the magnitude of the $2f_1 - f_2$ DP in two different couplers: a DB-100 Zwislocki coupler and a plastic syringe having a volume of approximately 1.5 cc. The test was performed for L_2 from 50 to 80 dB SPL with the same equipment and in the same conditions used to measure DPOAEs. The measurement time was prolonged to maximize the chances of discovering any artifacts. The magnitude of the cubic DP would be $-\infty$ dB SPL for an ideal OAE system. The system artifact limit was set to the higher of the responses for the two couplers that was also 2SD above the mean level of ten adjacent frequency bins in its corresponding spectrum. This procedure was repeated for each of the f_2s considered when measuring DPOAEs.

It is common to accept DPOAE measurements when they are above the system artifact limit. This may be tolerable in a clinical context where it comes to making a “pass/refer” decision. The magnitude of the measured DPOAE is the *vector* sum of DP contributions from any nonlinearity along the signal path, be it

from the instrument or from the subject. If the clinical rule was applied, then the true physiological response would be any value within the range [−∞, +6] dB around the measured DPOAE in the worst possible case (i.e., when the measured DPOAE magnitude just exceeds the artifact limit and the physiological DP has opposite or equal phase to the system's DP, respectively). This uncertainty range, however, seems too broad for the present study where the slope of the I/O curve is of interest. Therefore, a more restrictive rule was applied. A DPOAE measurement was accepted as valid only when it exceeded the artifact limit by 6 dB or more. This guaranteed that the physiological DP contribution was within the range [−6, +3.5] dB in the worst possible case (i.e., when the DPOAE measurement just met the present criterion). Therefore measurements were rejected if they were less than 6 dB above the system artifact limit. This was the case for subject S4 at 4 kHz, for whom most data points were discarded based on this criterion, and also for subject S3 at 1 kHz at high stimulus levels (75 dB SPL).

2.2.8 DPOAE procedure

DPOAE measurements were obtained with an IHS Smart system (with SmartOAE software version 4.52) equipped with an Etymotic ER-10D probe. During the measurements, subjects sat comfortably in a double-wall sound attenuating chamber and were asked to remain as steady as possible.

The probe fit was checked before and after each measurement session. The probe remained in the subject's ear throughout the whole measurement session to avoid measurement variance from probe fit. DPOAEs were measured for a preset measurement time. For f_2=4 kHz, the measurement time ranged from 60 s at L_2=20 dB SPL to 8 s at high L_2; for f_2=0.5 kHz, it ranged between 60 and 30 s for L_2 between 45 and 75 dB SPL, respectively. A DPOAE measurement was considered valid when it was 2SD above the measurement noise floor (defined as the mean level over ten adjacent frequency bins in the spectrum). When a response did not meet this criterion, the measurement was repeated and the measurement time was

increased if necessary. The probe remained in the same position during these remeasurements. If the required criterion was not met after successive attempts, the measurement point was discarded.

Each recording session consisted of measuring DPOAE I/O curves for one frequency of interest (consisting of five adjacent frequencies) (see Sec. 2.2.5) and was allowed to take up to 1 h. Three DPOAE measurements were obtained per condition (i.e., per f_2 and L_2) and averaged, except for subject S2 for whom only one measurement was obtained. Therefore, each point in the final I/O curves was the mean of 15 (3 measurements x 5 adjacent frequencies) measurements for each L_2 level. Occasionally, it was not possible to obtain all 15 points, particularly for the lower L_2 levels. In those cases, the direct mean of the available points would have been biased toward the DPOAE values for the frequencies giving the stronger responses. To minimize this potential bias, a mean was calculated only when two (out of the three possible) measurements were available per frequency and when eight of the new ten possible measurements were available. Otherwise, the corresponding point was neglected for further analysis.

2.3. Results

2.3.1 TMCs

Figure 2.1 shows the TMCs for the ten subjects for probe frequencies of 0.5, 1, 2, and 4 kHz. Note that TMCs were not measured for all frequencies and for all subjects. Subject S5 was unable to perform the TMC task using a probe level of 9 dB SL, thus a probe level of 15 dB SL was used in that case. Even so, it was not possible to measure high masker levels at 4 kHz for subject S5 and the resulting TMC did not allow estimating the true degree of compression (open circles in panel S5 of Fig. 2.1). Therefore these data were discarded from further analysis.

The shapes of the present TMCs are generally consistent with those of previous studies (e.g., Nelson *et al.*, 2001; Lopez-Poveda *et al.*, 2003; Lopez-

Poveda and Alves-Pinto, 2008; Plack *et al.*, 2004; Rosengard *et al.*, 2005). The linear reference TMCs (open squares) can be described either by a straight line (e.g., S4, S5) or by a shallow and gradually saturating function (e.g., S3, S8). The latter justifies the decision to fit the linear-reference TMCs with a double exponential function. The continuous, thick lines illustrate these fits. Several on-frequency TMCs show a shallow segment for short masker–probe time intervals (or gaps) followed by a steeper segment for moderate intervals. Others, however, are better described by a segment steeper than the linear reference followed sometimes by a shallower section at high masker levels (e.g., S1, S2, S6, and S8). In any case, all of the on-frequency TMCs have segments that are much steeper than the linear reference TMC. Assuming that the off-frequency masker condition used to generate the linear reference is processed linearly by the BM and that the rate of decay of the internal masker effect is identical across frequencies (Lopez-Poveda *et al.*, 2003; Lopez-Poveda and Alves-Pinto, 2008), the steeper segments may be interpreted to indicate BM compression (Nelson *et al.*, 2001). The validity of these assumptions is discussed in Sec. 2.4.6.

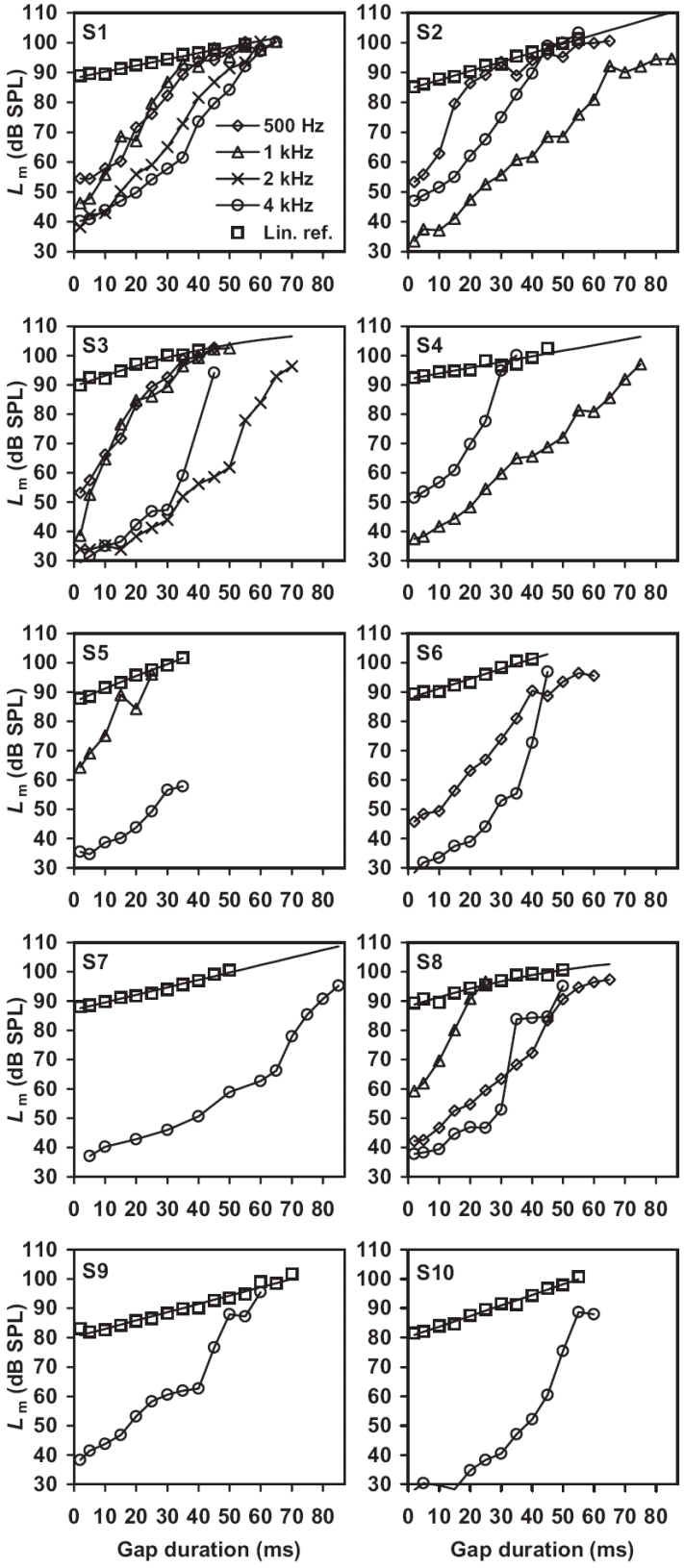

Figure 2.1. TMCs for all subjects and probe and masker frequencies. Each panel illustrates the data for one subject. Open squares illustrate the linear reference TMC; i.e., the TMC for a probe frequency of 4 kHz and a masker frequency ($0.4f_P$) of 1.6 kHz. The smoother continuous lines illustrate fits to the linear reference TMC with a double-exponential function. Other symbols illustrate on-frequency TMCs for different probe frequencies (as indicated by the inset in the top-left panel). The probe level was 9 dB SL except for subject S5 for whom it was 15 dB SL.

2.3.2 DPOAEs

Figure 2.2 shows a typical example of the amount of data measured to estimate one DPOAE I/O curve (2 kHz in this particular case). Each data point is the average of three measurements. Figure 2.2 also serves to illustrate the influence of the DP fine structure on the resulting I/O function. The inset indicates L_2 (in dB SPL). Note that DPOAEs were measured for a wider range of stimulus levels (from 20 to 75 dB SPL) for the five central adjacent frequencies only. A narrower range of levels (40, 50, 60, and 70 dB SPL) was considered for frequencies outside this frequency range. This was sufficient, however, to get an idea of the surrounding DPOAE fine structure and its potential influence on the DPOAE I/O curve at the frequency of interest.

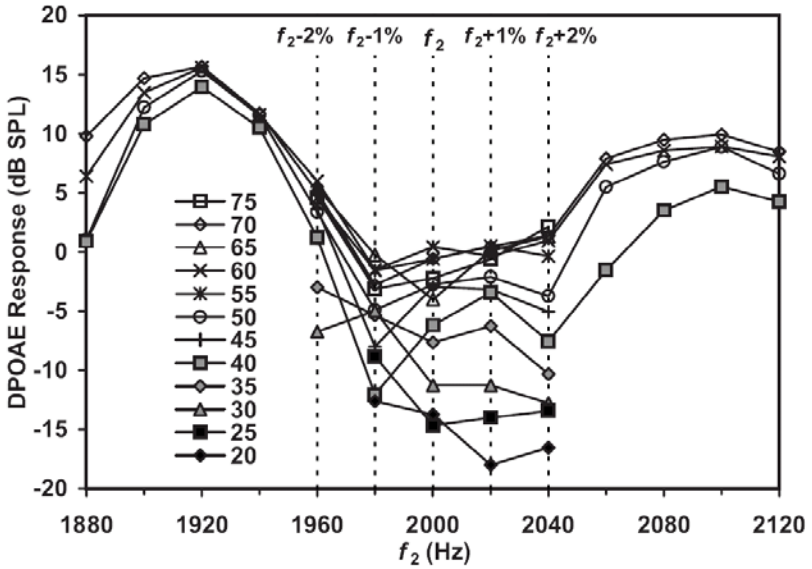

Figure 2.2. An example of the influence of the DPOAE fine structure at 2 kHz. An example data set recorded to estimate every I/O curve is also illustrated. The DPOAE magnitude is shown for 12 f_2 frequencies around the frequency of interest (2 kHz) and for different L_2 levels (in dB SPL) as indicated in the inset.

Obviously the DPOAE I/O curve would change considerably by changing f_2 within a narrow frequency range of 100 Hz, which emphasizes the need to take into account the effect of fine structure when estimating the actual I/O curve. The final

I/O curve, representing the BM I/O function at 2 kHz, would be the mean of the five I/O curves for the five central frequencies.

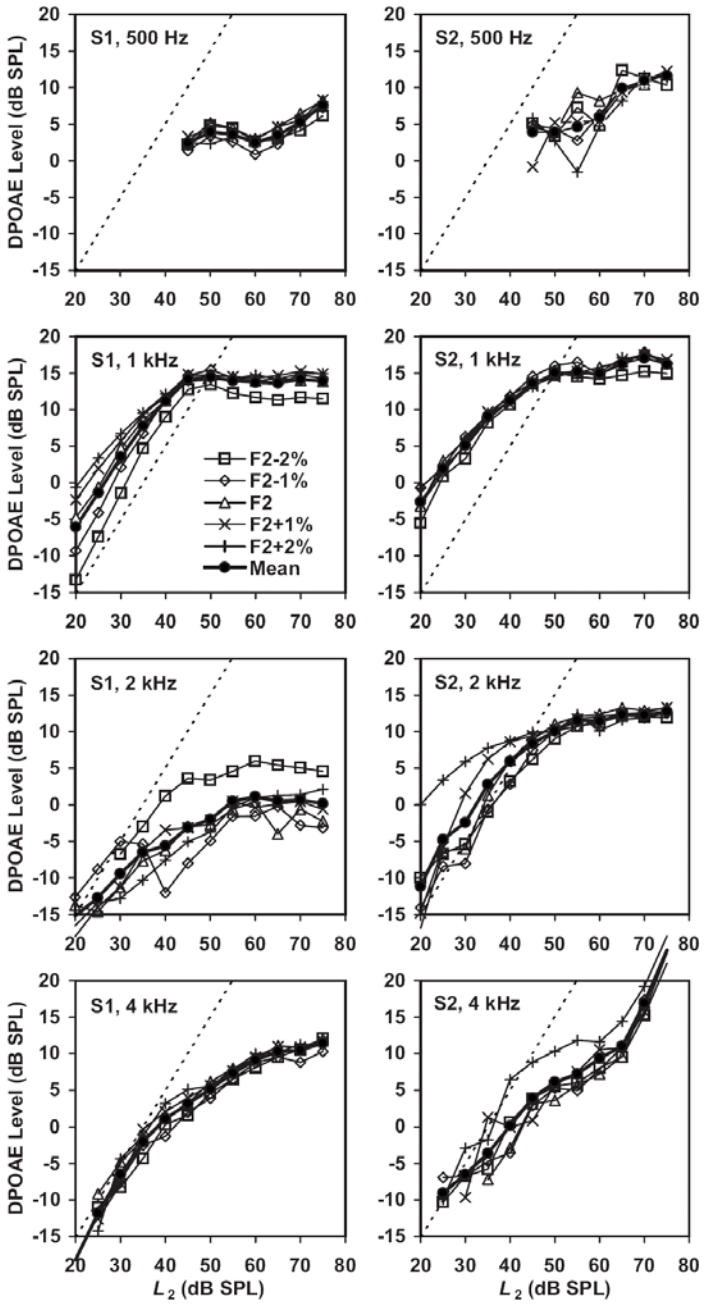

Figure 2.3. Example DPOAE I/O curves for subjects S1 (left panels) and S2 (right panels), for frequencies from 0.5 (top) to 4 kHz (bottom) in octave steps. Each final I/O curve (closed circles, thick line) was obtained as the average of five I/O curves for frequencies of $0.98f_2$, $0.99f_2$, f_2, $1.01f_2$, and $1.02f_2$, illustrated with different symbols according to the inset. The dashed line illustrates a linear response for comparison.

Figure 2.3 illustrates these five I/O curves for subjects S1 and S2 and for all frequencies considered in the present study. As can be seen, there is variability

across subjects and across frequencies. Nevertheless, variability appears to be greater for the lower stimulus levels. This seems reasonable given that the fine structure is pronounced at low stimulus levels (Mauermann and Kollmeier, 2004).

2.3.3 Comparison of BM I/O curves inferred from TMCs and DPOAEs

Figures 2.4–2.7 allow within-subject comparisons of DPOAE I/O curves (open squares) with BM I/O curves inferred from TMCs (closed circles) for corresponding cochlear sites with CFs of 0.5, 1, 2, and 4 kHz, respectively. Each panel illustrates the results for one subject. The associated solid lines are third-order polynomials fitted to the data. Error bars denote one standard error (SE) of the mean. For DPOAE I/O curves, the SE was based on up to 15 measurement points for each stimulus level, which explains why error bars are so short. The average error across all subjects and levels was 0.7 dB. The vertical and horizontal error bars of TMC-based I/O curves illustrate the SE of the linear reference or the TMC in question, respectively, based on at least three measurements. The average errors across all subjects and time gaps for the linear reference and the on-frequency TMCs were 0.9 and 2.3 dB, respectively.

The open triangles in the top-right panel of Figs. 2.4–2.7 and their associated error bars illustrate the mean DPOAE noise floor plus 2SD (based on three measurements for the five adjacent frequencies considered) for one example subject. The noise levels were similar for the other subjects. In general, the noise has the effect of increasing the DPOAE magnitude, particularly at low levels. However, the strict criteria used here should have avoided this influence.

In general, both DPOAE and TMC-based I/O curves are similar in that they are linear at low levels and become gradually compressive with increasing level. There is a tendency in both sets for I/O curves to become linear again at the highest levels tested (e.g., S6 at 0.5 kHz in Fig. 2.4; or S2 and S5 at 4 kHz in Fig. 2.7). The degree of similarity between the two sets of I/O curves is greatest at 4 kHz (Fig.

2.7). At this frequency, both sets of I/O curves indicate equally mild compression for subjects S1, S2, S5, and S7, and equally strong compression for S3 and S8. The strongest disagreement between the two sets of I/O curves at 4 kHz occurs for subject S9 (Fig. 2.7). S4 deserves a special mention because her DPOAEs could not be measured for levels outside the 40–50 dB SPL level range despite her having normal hearing at 4 kHz (her DPOAE readings for L_2<40 dB SPL were below the physiological noise level and for L_2>50 dB SPL they did not meet the instrument artifact criterion, see Sec. 2.2.7).

The degree of similarity between the shapes of the two sets of I/O curves is, however, much lower for 0.5 and 1 kHz. It is noteworthy that, for these frequencies, some DPOAE I/O curves show plateaus and notches that do not have a clear correlate with TMC-based I/O curves (e.g., S1, S2, and S8 at 0.5 kHz in Fig. 2.4; or S3, S4, and S5 at 1 kHz in Fig. 2.5). Possible explanations for the notches and plateaus are discussed in Sec. 2.4.2.

It was not possible to draw conclusive results regarding the degree of correlation between the two sets of I/O curves at 2 kHz because data at this frequency were collected for two subjects only (Fig. 2.6). Good correspondence (akin to what was observed at 4 kHz) was found for one of the two subjects S3. The DPOAE I/O curve for the other subject S1, however, exhibited a negative gradient at moderate-to-high levels that was not present in the TMC-based I/O curves. This negative gradient was typical of the I/O curves at frequencies of 0.5 and 1 kHz.

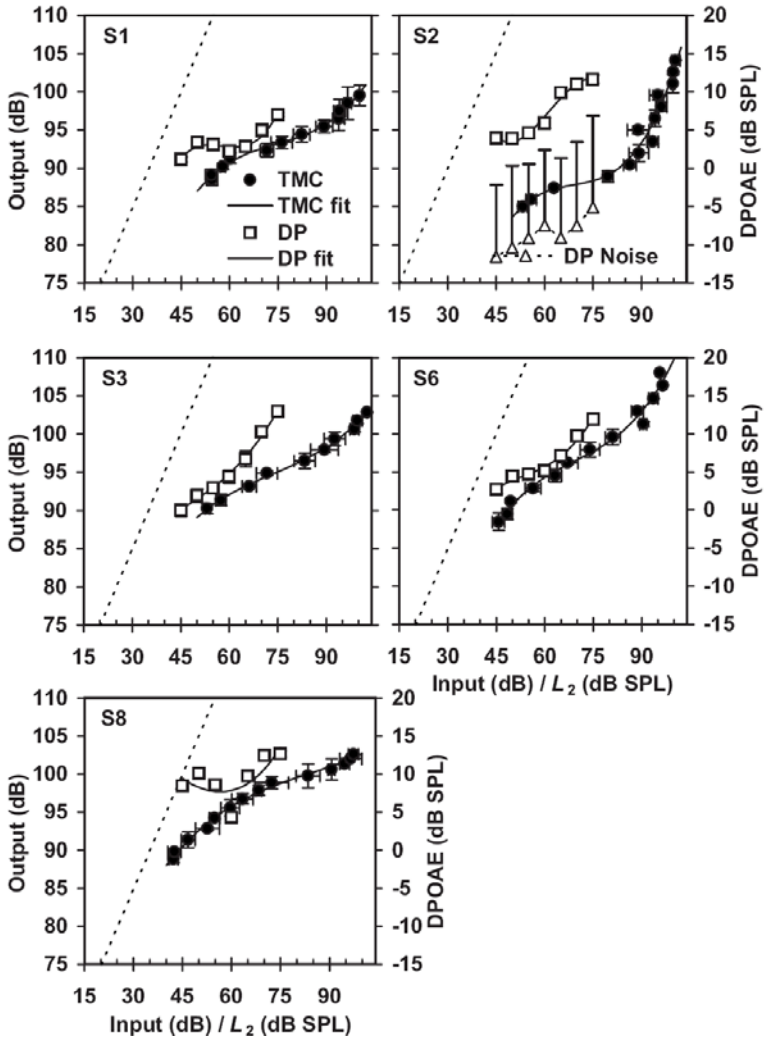

Figure 2.4. Experimental DPOAE (open squares) and TMC-based (closed circles) I/O curves at 0.5 kHz (=$fp = f_2$). Continuous lines illustrate third order polynomial fits to the experimental I/O curves. Error bars denote one SE of the mean. Horizontal bars (only for TMC-based curves) represent the standard error of the input level (i.e., the standard error for the on-frequency masker level). Each panel illustrates the result for a different subject. The panel for subject S2 also illustrates the mean DP noise floor and its corresponding 2SD. Thin dashed lines illustrate a linear response for comparison.

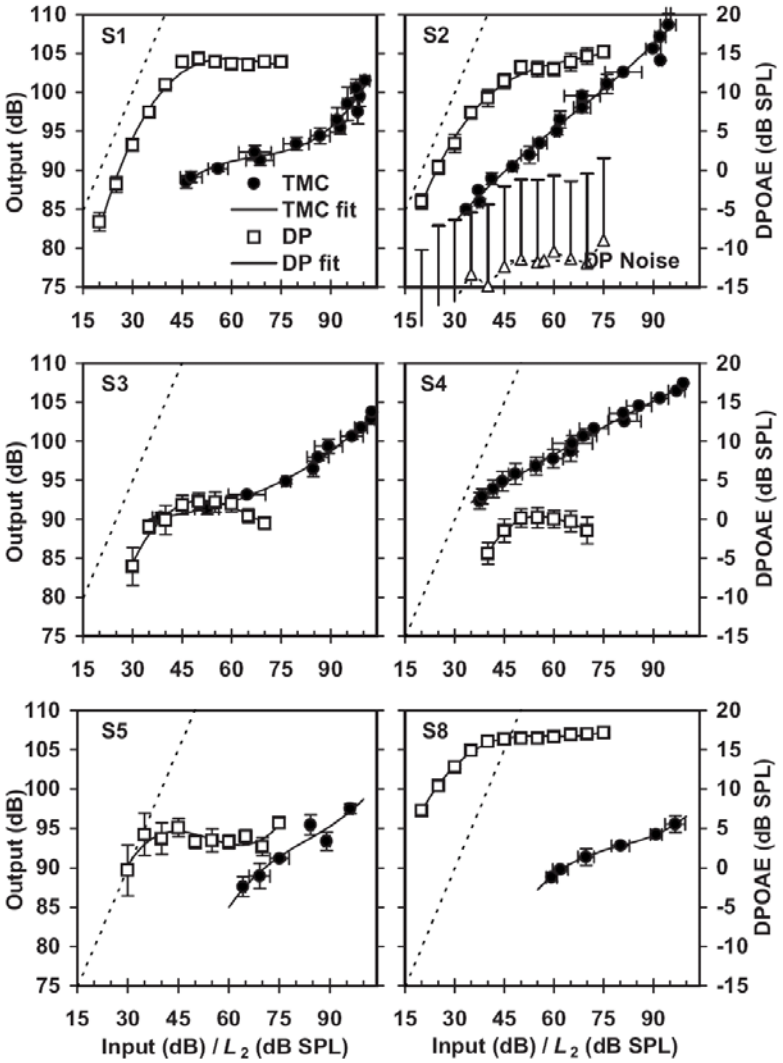

Figure 2.5. The same as Fig. 2.4, but for a frequency of 1 kHz.

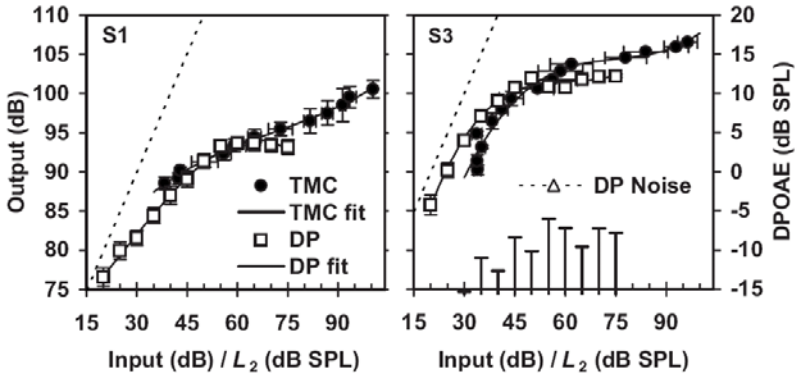

Figure 2.6. The same as Fig. 2.4, but for a frequency of 2 kHz.

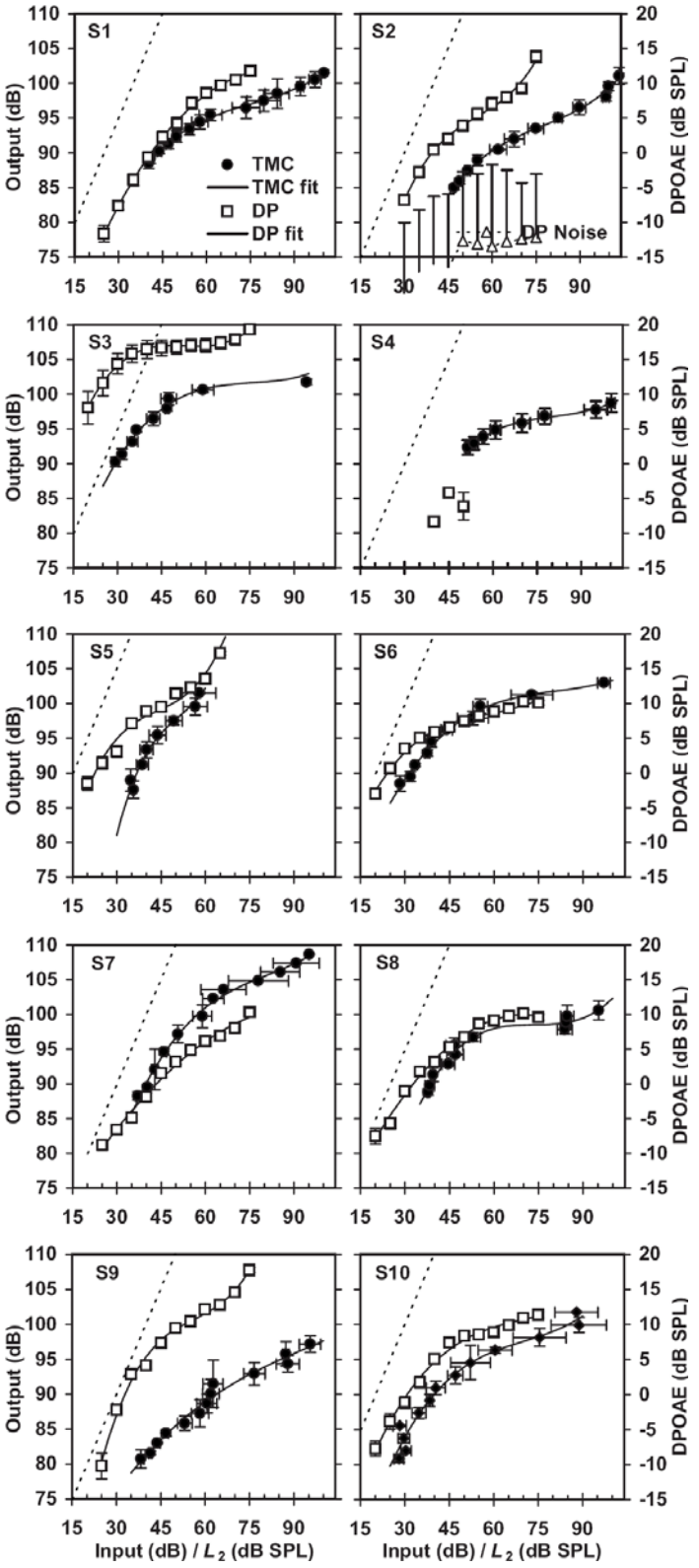

Figure 2.7. The same as Fig. 2.4, but for a frequency of 4 kHz.

2.3.4 Comparison of derived cochlear nonlinearity parameters

Third-order polynomials were fitted by least squares to the TMC-based and DPOAE I/O curves (continuous lines in Figs. 2.4–2.7) and used to derive the following parameters pertaining to cochlear nonlinearity: minimum compression exponent, the level at which it occurred, compression threshold, cochlear gain, and the level at which the I/O curves returned to linearity at high levels.

The minimum compression exponent was estimated as the minimum slope of the fitted I/O curves over the measured range of input levels. Figure 2.8(A) compares the minimum compression exponent as inferred from the I/O curves obtained with the two methods. Closed symbols show the results for 4 kHz separately. Open symbols illustrate the results for the other frequencies (0.5, 1, and 2 kHz), as indicated by the inset. The solid and dashed lines illustrate LR fits to the 4 kHz data only or to all data points (including also the 4 kHz points), respectively. Both regression lines were forced to pass through the origin of the graph. Pearson's correlation coefficients were (high) 0.92 for the 4 kHz subsample and (low) 0.19 when all data points (all frequencies) were considered. The mean compression exponents across frequencies were 0.13 and 0.04 for the TMC-based and the DPOAE data sets, respectively. They, however, were 0.10 and 0.11, respectively, considering only the 4 kHz data points. The mean difference was significant when all data points were considered ($p<0.005$ for a paired two tailed Student's t-test) but not for the 4 kHz subsample ($p=0.81$). This is clearly seen from Fig. 2.8A in that DPOAE based compression estimates are lower overall than TMC-based ones for 0.5 and 1 kHz.

The reason for the disagreement between the results at 0.5 and 1 kHz is, almost certainly, that many DPOAE I/O curves for the lower frequencies showed plateaus or notches, which resulted in unexpectedly low compression exponent slopes (slopes < 0 dB/dB). Table 2.2 summarizes the number of cases across frequencies when the fitted I/O curve exhibited plateaus or negative compression exponents.

Table 2.2. Incidence of plateaus and notches in DPOAE I/O curves as a function of f_2.

	Frequency (kHz)			
	0.5	1	2	4
Number of measured I/O curves	5	6	2	10
Number of I/O curves with plateaus or notches	3	5	1	2

These results contrast with those of Williams and Bacon (2005), who reported a high correlation between compression estimates inferred from TMC-based and DPOAE I/O curves also at low frequencies. The compression estimates of Williams and Bacon (2005), however, were calculated differently. They were calculated as the slope of linear segments fitted to the DPOAE and TMC-based I/O curves over the range of input levels from 40 to 65 dB approximately (they determined the actual level range by visual inspection). This fitting method was also employed here as it could prove less sensitive to notches. Figure 2.8(B) shows the correlation between compression exponent estimates for the two methods derived by LR fits to the I/O curves over the level range 40–65 dB. The 4 kHz data are shown by the closed symbols; open symbols illustrate the results for frequencies of 0.5, 1, and 2 kHz. The solid and dashed lines are LR fits to the 4 kHz data and to all data points (including also the 4 kHz data), respectively. Pearson's correlation coefficient considering all data points was 0.32 and Fig. 2.8(B) illustrates that DPOAE-based compression exponents were generally lower (most points are below the diagonal) than those inferred from TMCs. Group mean compression exponents were 0.31 and 0.18 for TMCs and DPOAEs, respectively. This difference was statistically significant ($p < 0.0005$). A moderately higher correlation was found when analyzing the 4 kHz data separately (r=0.53), but the associated mean difference (0.36 and 0.25 for TMCs and DPOAEs, respectively) was still statistically significant (p<0.01).

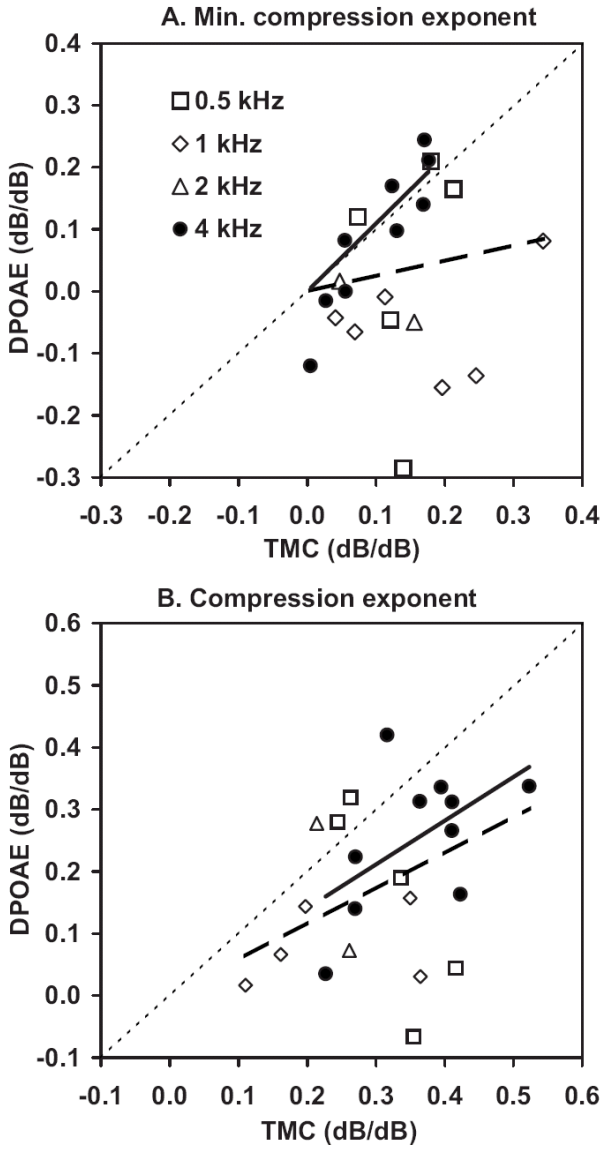

Figure 2.8. Correlation between compression-exponent estimates obtained from DPOAE- and TMC-based I/O curves (i.e., each point represents data for one subject). (A) Minimum compression exponent based on third-order polynomial fits to the I/O curves. (B) Compression exponent based on LR fits to the I/O curves over the input-level range 40–65 dB SPL. Closed and open symbols illustrate the 4 kHz data and the data at other frequencies (0.5, 1, and 2 kHz), respectively. Solid lines show LR fits to the 4 kHz data points constrained to cross through the origin. Thick dashed lines are LR fits considering all data points (including 4 kHz) constrained to cross through the origin. Diagonal thin dashed lines illustrate perfect correlation.

Another characteristic of cochlear nonlinearity is the level at which maximum compression (or, equivalently, minimum compression exponent) occurs. Correlations were sought between estimates of this parameter for the I/O curves obtained with the two methods. Figure 2.9(A) reveals that high correlation occurs in a few cases only (points close to or on the diagonal). Most points are below the diagonal, which means that the level at which minimum compression occurred was lower for DPOAE I/O curves than for I/O curves inferred from TMCs. The average

input level for minimum compression exponent was 61 and 76 dB for DPOAEs and TMC-based I/O curves, respectively. The difference was statistically significant ($p<0.001$). Results were similar when the data for each frequency were analyzed separately.

Compression threshold was defined as the input level at which the slope of the fitted polynomial decreased from a value close to one at low levels to 0.4 dB/dB at a higher level. This is an arbitrary definition, but seems reasonable for our purpose. When the slope of the I/O curve at the lowest level for which a data point existed was reasonably close to 0.4 dB/dB, the fitted polynomial was extrapolated up to 5 dB to identify the compression threshold and the value was noted for further analysis.

An extrapolated compression estimate was thus included in the analysis only if it was less than 5 dB from an existing data point. Figure 2.9(B) illustrates the correlation between the compression threshold estimates inferred from DPOAE and TMC-based I/O curves. Three compression estimates were considered outliers [symbols surrounded by an open circle in Fig. 2.9(B)] and thus excluded from the statistical analysis. The possible reasons for these outliers will be discussed in Sec. 2.4.4. The regression line was constrained to cross through the graph origin. The line indicated a high degree of correlation between the estimates of the two methods and, indeed, Pearson's correlation coefficient was reasonably high ($r=0.8$). The average compression threshold estimates were lower for the DPOAE than for the TMC-based I/O curves (40 and 47 dB SPL, respectively), which is also clearly seen in Fig. 2.9(B). This difference was statistically significant ($p<0.0001$).

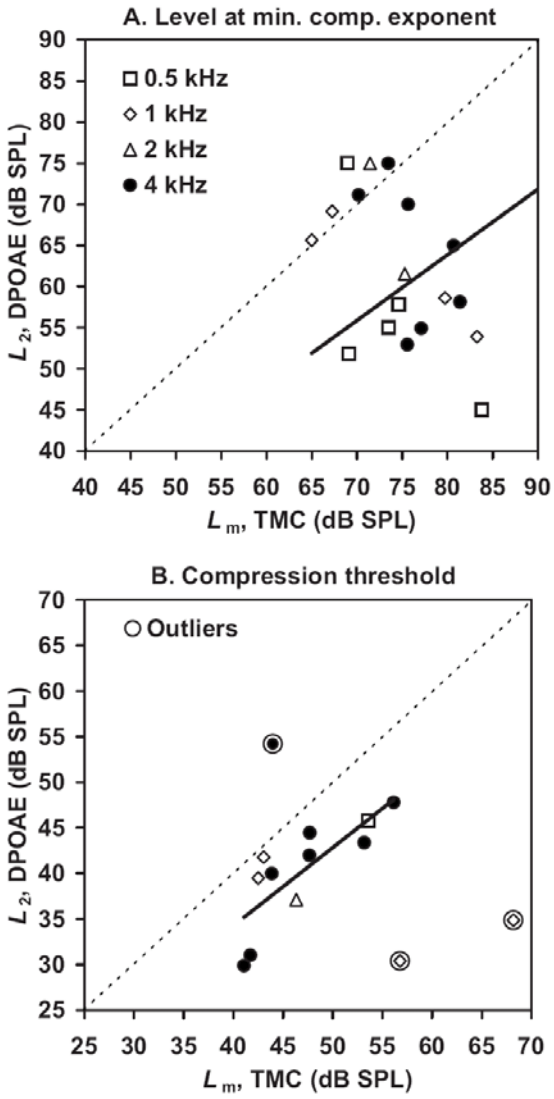

Figure 2.9. (A) Correlation between the level at which maximum compression occurs based on the I/O curves inferred from DPOAE and TMCs (i.e., each point represents data for one subject). The solid line shows a LR fit to the data constrained to cross through the origin. (B) Correlation between compression-threshold estimates obtained with the two methods. Threshold was defined as the input level at which the I/O curve slope decreased from linearity to 0.4 dB/dB. The solid line shows the LR constrained to cross through the origin. Symbols surrounded by open circles illustrate three data points regarded as outliers. Diagonal thin-dashed lines illustrate perfect correlation.

2.3.5 Cochlear nonlinearity dependency on characteristic frequency

This section addresses whether the parameters considered in the preceding section vary with frequency similarly when they are inferred from TMCs or DPOAEs. Two additional parameters were considered based on TMCs, namely gain and the threshold of return to linearity at high levels. I/O curves may be generally described as having a linear segment (slope of 1 dB/dB) at low levels, followed by a compressive segment (slope < 1 dB/dB) at moderate levels, followed by linear segment at high levels (e.g., Lopez-Poveda *et al.*, 2003). Perhaps, the clearest

examples of this pattern were the TMCs for S6 at 0.5 kHz and S2 at 4 kHz. Few TMC-based I/O curves showed a linear segment at high levels (e.g., S2 at 0.5 kHz or S1 at 1 kHz). The slope of the I/O curves, however, always increased with increasing level beyond the inflection point of the curve (defined as the level at which maximum compression occurs or, equivalently, the level at which the second derivative of the I/O curve equals zero). This suggests that the I/O curves might approach linearity at very high levels. Since the minimum slopes were always < 0.4 dB/dB [Fig. 2.8(A)], the return-to-linearity threshold was arbitrarily defined here as the level at which the slope of the fitted I/O curve equated to 0.4 dB/dB for levels above the inflection point. Gain was defined as the difference between the return-to-linearity and the compression thresholds in decibels.

Figure 2.10 shows the frequency dependency of the minimum compression exponent (top panels), the level at which it occurs (middle panels), and the compression threshold (lower panels), respectively. Left and right panels show the value of these parameters as inferred from TMCs and DPOAE I/O curves, respectively. Figure 2.10 illustrates that all these parameters remain approximately constant across frequencies. Indeed, no significant differences were found between the mean values of every parameter across frequencies. The same applied to parameters estimated with both methods.

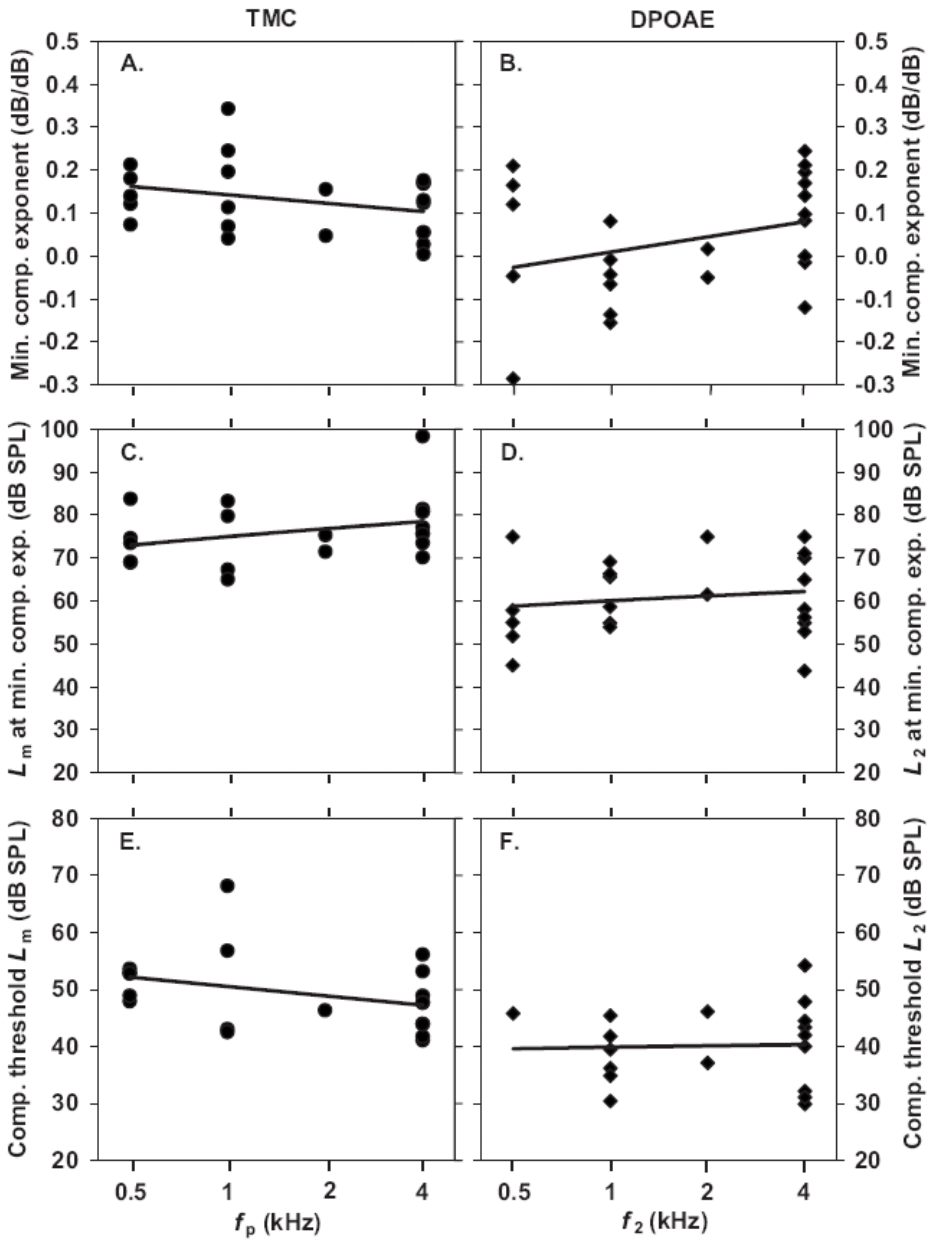

Figure 2.10. Frequency dependency of cochlear nonlinearity parameters as estimated from TMCs (left panels) and DPOAEs (right panels). (A), (B) Minimum compression exponent. (C), (D) The level at which maximum compression occurs. (E), (F) Compression threshold.

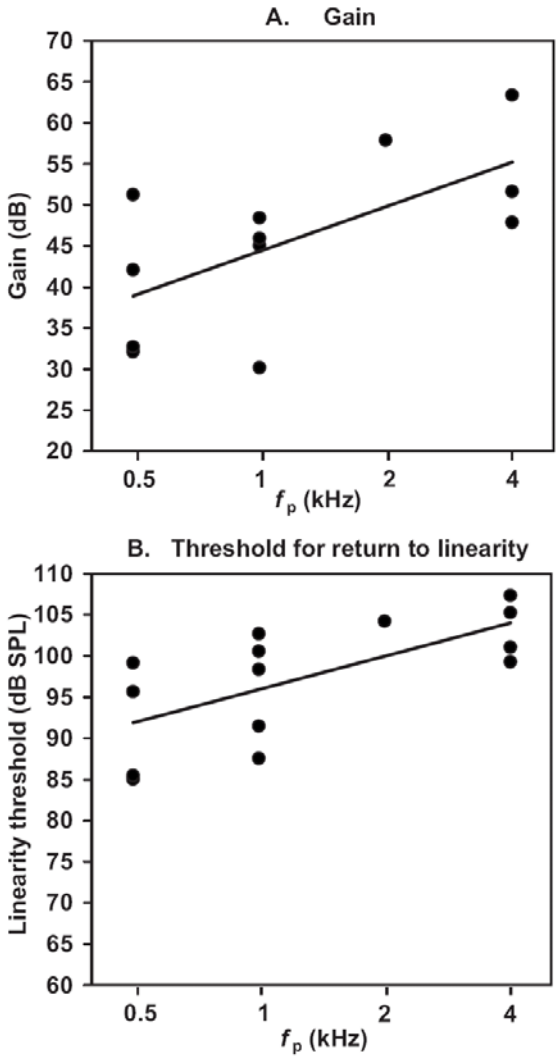

Figure 2.11. Frequency dependency of gain (A) and the threshold of return to linearity (B) derived from TMC I/O curves. The latter was defined as the input level at which the I/O curve reached a slope of 0.4 dB/dB from a lower value for increasing input level.

Figure 2.11 shows the gain and the return-to-linearity threshold as inferred from TMCs only. Both parameters tend to increase with increasing frequency. Since the compression threshold remained approximately constant across frequencies [Fig. 2.10(E)], the frequency dependency of the gain is fully attributed to the increase of the return-to-linearity threshold with increasing frequency. In any case, a one-to-one correspondence between these two parameters should not be expected because they are based on data for different subjects. Indeed, it was not possible to

estimate the return to linearity threshold in several cases (one different subject for each frequency of 0.5, 1, and 2 kHz, and six subjects at 4 kHz). The I/O curves for these subjects may show a return to linearity at input levels higher than those considered in the present study. That the majority of these cases occurred at 4 kHz supports the idea that the return to linearity at higher frequencies occurs at higher input levels than those considered here.

2.4 Discussion

The goal of this study was threefold. The first objective was to compare cochlear nonlinearity parameters inferred from TMCs and DPOAEs, and, if coinciding, add support to the notion that they are two equivalent manifestations of cochlear nonlinearity. A second aim was to evaluate the feasibility of using DPOAE I/O curves as a fast tool for estimating individual parameters of cochlear nonlinearity. A third objective was to investigate the frequency dependency of parameters describing cochlear nonlinearity, as inferred from DPOAEs and TMCs.

2.4.1 Equivalence between DPOAE and TMC-based I/O curves

The degree of correlation between compression estimates inferred with the two methods was high (0.92) at 4 kHz but much lower at 0.5 and 1 kHz. The assumption has been made that DPOAE I/O curves reflect the characteristics of the BM response to single tones at the f_2 place. One possible explanation for the lack of coincidence is that mutual suppression between the two primaries f_1 and f_2 could have influenced the shape of the DPOAE I/O curves especially at low CFs (Table 2.2). This is discussed further in Section 6.3.

The group mean compression-exponent estimates at 4 kHz obtained in the present study (0.10 and 0.11 dB/dB for TMC and DPOAE, respectively) are in agreement with those from Gorga *et al.* (2007), who reported a minimum slope value of ~0.12 dB/dB at 4 kHz based on DPOAEs (estimated from their Fig. 4 at

L_2=60 dB SL). As for compression estimated using LR over the midlevel range [Fig. 2.8(B)], the group mean exponent values obtained in the present study (0.36 and 0.25 for TMC and DPOAE, respectively) were moderately higher than those reported by Williams and Bacon (2005) (0.26 and 0.15 for TMC and DPOAE at 4 kHz, respectively). In any case, the present study shows that the estimated degree of compression differs considerably depending on the method used to infer it e.g., polynomial versus LR fits, which emphasizes the need to specify clearly the method used in every study.

The minimum compression exponent found here at 4 kHz (based on polynomial fits) is lower than previously reported values obtained with the same (or different) methods [e.g., 0.14, third-order polynomial, Nelson and Schroder (2004), 0.20, third-order polynomial, Plack and Drga (2003), 0.13, sum of linear and sigmoidal function, and 0.23, straight line, Rosengard *et al.* (2005), 0.20, straight line, Plack *et al.* (2004), 0.25, straight line, Lopez-Poveda *et al.* (2003)]. The difference in TMC-based estimates may relate to differences in the linear reference used by different studies. Here, the linear reference was the TMC for a masker frequency of $0.4f_P$, whereas the above-mentioned studies used the TMC for a masker frequency between $0.5f_P$ and $0.6f_P$. Lopez-Poveda and Alves-Pinto (2008) have suggested that the latter may still undergo as much as 2:1 compression; hence they could lead to an underestimate of the degree of on-frequency compression. The agreement between the present compression estimates obtained from TMCs and DPOAEs at 4 kHz provides circumstantial support to the conclusion of Lopez-Poveda and Alves-Pinto (2008).

2.4.2 DPOAE notches and plateaus

Notches and plateaus are common in the present DPOAE I/O curves (Figs. 2.4–2.7, especially at the lower f_2, Table 2.2). This contrasts with the conclusion of Kummer *et al.* (1998), who reported that notches and plateaus were less common when DPOAE I/O curves were measured with their proposed primary-level rule than with

other rules. The Kummer *et al.* level rule is based on a group average and thus in some cases it may deviate considerably from the individual optimal (the optimal rule would be the one that evokes the strongest possible DPOAEs at all levels). Therefore, one possible explanation for the present observations is simply that the rule of the Kummer *et al.* (1998) was not optimal for the subjects used in the present study. This explanation is supported by Neely *et al.* (2005), who reported that L_1 should be systematically higher than prescribed by Kummer *et al.* (1998) and that the L_1– L_2 relationship should vary with f_2. Johnson *et al.* (2006) confirmed the latter and further suggested that the ratio f_2 / f_1 should vary slightly with f_2. Interestingly, it is for low frequencies and moderate–high stimulus levels where the rule of Kummer *et al.* deviates most from the rules of Neely *et al.* (2005) and Johnson *et al.* (2006). These are also the conditions where plateaus and dips are most common in the present data, which suggests that the rule of Kummer *et al.* is not optimal for the subjects considered in this study at low frequencies. On the other hand, Kummer *et al.* (2000) verified their original paradigm [derived from data of Gaskill and Brown (1990)] with a larger sample and still found it to be independent of frequency. This explanation is explored further in Chapter 5.

Among the possible reasons for these notches and plateaus is that another DP generation source (see Mills, 1997) at high stimulus levels create destructive interference with the normal mid-low level source. Notches and plateaus could also be due to fine structure despite the precautions taken to minimize its effects. However, this explanation may seem less likely as the fine structure is more frequent and has higher magnitude at low stimulus levels (Mauermann and Kollmeier, 2004) and the notches and plateaus in this study occurred at mid-high stimulus levels. These explanations are discussed with more detail in Section 6.4 and 6.6.

2.4.3 The level at which the minimum compression exponent occurred

Low correlation was found between estimates of this parameter obtained with the two methods (DPOAEs and TMCs). The reason for this is uncertain. Maybe DPOAEs grow faster with increasing stimulus level at high stimulus levels because of the contribution from the second high level DP generation source mentioned in the preceding sections.

2.4.4 Compression threshold

A moderately high correlation was found between compression-threshold estimates inferred with the two methods [Fig. 2.9(B)]. DPOAE-based estimates were, however, on average 7 dB lower than TMC-based estimates. A total of 22 I/O curves were measured in the ten subjects, but the compression threshold could be estimated in only 14 of these 22 cases. This could be interpreted as an argument against the apparent equivalence of the two methods with respect to estimating this parameter. Further analysis reveals, however, that there are good reasons why a compression threshold could not be estimated in the remaining eight cases. Four of them corresponded to DPOAE I/O curves at 0.5 kHz (Fig. 2.4) that extended over a range of input levels above 45 dB SPL that was too narrow to reveal a compression threshold. Another case was S4 at 4 kHz (Fig. 2.7), who did not have sufficiently strong DPOAE despite her hearing threshold being normal at that frequency. For the three remaining cases, the TMCs did not show a clear compression threshold or did not reach the criterion slope of 0.4 dB/dB (see Sec. 2.3.4). These cases were S1 at 2 kHz in Fig. 2.6; S3 and S4 at 1 kHz in Fig. 2.5. Three of the 14 cases where both methods demonstrated a compression threshold were considered outliers and excluded from the comparison of compression threshold: S1 at 4 kHz, and S5 and S8 at 1 kHz [depicted as circles in Fig. 2.9(B)]. The data for S5 and S8 were excluded because these listeners had great difficulties performing the TMC task (they needed six to eight attempts at high masker levels to obtain three measurements each having a SD below 6 dB; Sec. 2.2.3). There was no obvious

reason to exclude the data point of S1. This data point corresponded to the first condition on which the subject was tested and he might not have been sufficiently trained. In any case, it is noteworthy that DPOAE I/O curves allow estimating a compression threshold much more easily than do TMC-based I/O curves.

The reason why DPOAE-based compression threshold estimates were lower than corresponding TMC-based estimates is uncertain. Maybe the DPOAE response was influenced by mutual suppression of the primaries (Sec. 6.2). This could have linearized the I/O curve (e.g., by decreasing cochlear gain) and thus *increased* the compression threshold suggested by DPOAE I/O curves. This explanation does not fit the data as the actual compression thresholds estimated from DPOAEs were *lower* than those estimated from TMCs. Perhaps the difference in compression threshold estimate for the two methods is caused by the use of suboptimal DPOAE parameters (see Sec. 2.4.2).

The moderately high correlation between the compression-threshold estimates of the two methods also at low frequencies [Figs. 2.9(B), 2.10(E), and 2.10(F)] may seem surprising given the low correlation between the estimates of minimum compression exponent. One possible explanation could be that the DPOAE parameters were adequate for lower L_2, where the compression threshold occurs, but not for higher L_2 levels, at which, coincidentally, plateaus and notches occur.

The average compression threshold estimates at 4 kHz found in the present study were 47 and 40 dB SPL for TMC and DPOAE, respectively. These values are comparable to those (average 37 dB SPL at 3–4 kHz) found psychophysically by Yasin and Plack (2003) based on a three-line segment fitting procedure (Plack *et al.*, 2004). A value of ~35 dB SPL is obtained when applying the present definition of compression threshold to the DPOAE data of Neely *et al.* (2003). A compression threshold was also estimated from the DPOAEs reported by Gorga *et al.* (2007). After discarding their data for lowest stimulus levels because they were most likely contaminated by noise and applying the definition used in the present study (level at which I/O curve slope equals 0.4 dB/dB) to their data in their Fig. 3, the resulting

compression thresholds were 30 and 45 dB SPL at 0.5 and 4 kHz, respectively. The present values are reasonably in accordance with their results at 4 kHz but not at 0.5 kHz. Our results indicate that the compression threshold is approximately constant across frequencies [Figs. 2.10(E) and 2.10(F)].

2.4.5 Gain and the level of return to linearity

It is still controversial that BM I/O functions become linear at very high input levels in healthy cochleae (Robles and Ruggero, 2001). Assuming, however, that this is a truly physiological characteristic, DPOAEs are not considered a reliable predictor of the threshold level of return to linearity at high levels. First, there may be another mechanism involved in generation of DPOAE at the higher stimulus levels, as explained above (Sec. 2.4.2). Second, the best frequency of any BM site shifts with increasing level, hence the BM site where the two primaries cause maximum excitation is likely to shift accordingly with level. This may change the DPOAE response as the region of overlap of the two primaries changes with level. Because of this, the gain and return to linearity thresholds were not estimated based on DPOAE I/O curves and not compared with those inferred from TMCs.

The TMC-based I/O curves suggested that the cochlear gain increased with increasing CF because of a parallel increase in the threshold of return to linearity at high levels (Fig. 2.11). This is in agreement with physiological data (Robles and Ruggero, 2001). Gain estimates based on tip-to-tail level differences of DPOAE suppression tuning curves (Gorga *et al.*, 2008) showed the same tendency for the gain to decrease with decreasing frequency.

2.4.6 On the merits of the DPOAEs and TMCs for estimating cochlear I/O curves

The presence of plateaus and notches in the DPOAE I/O curves at 0.5 and 1 kHz results in zero or negative compression-exponent estimates which do not occur in corresponding TMC-based curves. While deep notches in apical BM I/O functions

have been reported [e.g., Fig. 7(a) of Rhode and Cooper (1996)], they typically occur for stimulation frequencies higher than the CF (Rhode and Cooper, 1996). Therefore, the notches reported here are unlikely to reflect notches in the underlying BM responses (see also Sec. 2.4.2).

It would be wrong to conclude, however, that it is inappropriate to use DPOAEs to infer cochlear I/O functions at low frequencies. As discussed earlier (Sec. 2.4.2), the present notches are possibly due to using suboptimal primary levels at low frequencies and it might be possible to find DPOAE stimulus parameters that would lead to higher correlations between TMC-based and DPOAE I/O curves.

It would also be wrong to conclude that I/O curves inferred from TMCs are more correct (i.e., reflect more closely the underlying BM responses) than DPOAE I/O functions as the TMC method rests on several assumptions, some of which have been questioned (see also Sec. 6.1).

In summary, the lack of correlation between the results of the two methods at low frequencies is uninformative at present of their relative accuracy for inferring cochlear I/O curves. Future studies should investigate the reason for the low correlation at low frequencies and whether higher correlations would be obtained using different DPOAE parameters and/or different psychophysical methods or assumptions.

The present study did not evaluate the merit of DPOAE I/O curves as a (*clinical*) tool for assessing residual cochlear compression characteristics in hearing-impaired listeners. The present results, however, suggest that they might be useful to assess residual compression in listeners with presbyacusis, who are mostly affected by high-frequency loss.

2.5 Conclusions

(1) The correlation between individual compression exponent estimates inferred from TMCs and DPOAEs is reasonably high at 4 kHz, but low at 0.5 and 1 kHz.

Both methods suggest that maximum compression is approximately 10:1 and constant across the frequency range from 0.5 to 4 kHz.

(2) The low correlation at low frequencies cast doubts on the postulates and interpretation of I/O curves inferred with either or both of the two methods. The most likely reason for the lack of correlation at low frequencies (0.5–1 kHz) is the presence of notches and plateaus in the DPOAE I/O curves. This suggests that the DPOAE stimulus paradigm of Kummer *et al.* (1998) may not be optimal (i.e., does not produce maximum DP magnitude) at low frequencies.

(3) A high correlation was found between estimates of compression threshold inferred from DPOAEs and TMC-based I/O curves between 1 and 4 kHz. The DPOAE and the TMC methods indicate that the compression threshold equals 40 and 47 dB SPL, respectively, and is approximately constant across the range of frequencies from 0.5 to 4 kHz for TMCs and from 1 to 4 kHz for DPOAEs.

(4) Cochlear gain and return-to-linearity thresholds were inferred from the TMCs only. Both parameters increased by ~16 dB with increasing characteristic frequency from 0.5 to 4 kHz.

(5) It seems reasonable to use TMCs and DPOAE I/O curves interchangeably to infer cochlear I/O curves at 4 kHz but doubts exist that the same applies to lower frequencies of 0.5 and 1 kHz.

CHAPTER 3

SIMULATING THE DEPENDENCY OF DISTORTION PRODUCT INPUT/OUTPUT CURVES ON STIMULUS PRIMARY RULE[2]

3.1 Introduction

In the present study, as in many previous ones (e.g., Dorn *et al*., 2001; Williams and Bacon, 2005; Johannesen and Lopez-Poveda, 2008; Neely *et al*., 2009), it is implicitly assumed that the I/O curve for the $2f_1-f_2$ DPOAE would be a reasonable description of the BM I/O curve (as measured using on-CF pure tones) for a cochlear region with a CF equal to f_2. In other words, that it would be possible to infer on-CF BM I/O curves by measuring the growth of the $2f_1-f_2$ DPOAE component with increasing L_2, for f_2 = CF. This chapter provides evidence that this assumption stands a better chance of being reasonable when optimal L_1 levels are considered. Evidence is also shown that sub-optimal L_1 levels may produce non-monotonic $2f_1-f_2$ DP I/O curves even in the absence of secondary DP sources (fine structure).

Ideally, the assumption in question should be demonstrated in analytical mathematical form for the nonlinearity underlying BM responses. This, however,

[2] This chapter is adapted from the appendix A of the published article: Johannesen, P. T., and Lopez-Poveda, E. A. **2010**. "Correspondence between behavioral and individually "optimized" otoacoustic emission estimates of human cochlear input/output curves," J. Acoust. Soc. Am. **127**, 3602-3613. Some portions of the article are reproduced here with permission from the Acoustical Society of America.

is not possible because nonlinear systems do not have exact analytic transfer functions and because the actual form of the BM nonlinearity is, to the authors' knowledge, yet to be confirmed. The present section provides only a crude numerical demonstration for two kinds of time-invariant compressive nonlinearities: a broken-stick nonlinear (Meddis *et al.*, 2001) and a double-Boltzmann gain function (e.g., Lukashkin and Russell, 1998, 2001). These nonlinearities have been successfully used to account for BM and OHC in computational models. For simplicity, no concurrent filtering effects are considered here.

3.2 Methods

Let $x(t)$ and $y(t)$ be the time-domain input and output waveforms to/from the nonlinearity. The broken-stick nonlinearity has the form:

$$y(y) = \min[ax(t), bx^c(t)], \tag{3.1}$$

where a and b are gain parameters and c is the compression exponent. Likewise, the double-Boltzmann nonlinearity has the form:

$$y(t) = G_M\left\{1 + e\left(\frac{x_1 - x(t)}{s_1}\right)\left[1 + e\left(\frac{x_2 - x(t)}{s_2}\right)\right]\right\}^{-1}, \tag{3.2}$$

where G_M, x_1, x_2, s_1 and s_2 are parameters. Here, the input was arbitrarily assumed in units of μPa and the parameters of the two functions were set to the following arbitrary values: a = 1; b = 437; c = 0.2; x_1 = 41; x_2 = 24; s_1 = 62.5; and s_2 = 15.38.

The I/O response characteristics of these two nonlinearities were obtained considering one- and two-sinusoid digital input waveforms (sampling frequency = 48 kHz). These will be referred to as pure tone (PT) and DP I/O curves, respectively. PT I/O curves were obtained using single sinusoids of a fixed frequency (f_2) and amplitudes (A_2) that varied in steps of 2 dB (re 20 μPa). A Fast Fourier Transform (FFT) was applied to the output waveforms and the amplitude at the f_2 frequency was noted and plotted as a function of A_2 in a log-log scale to obtain the I/O curve. DP I/O curves were obtained presenting two simultaneous sinusoids of frequencies

f_2 and $f_1 = f_2/1.2$. The amplitude of the f_2 sinusoid, A_2, was the same as in the one-sinusoid evaluations. For each value of A_2, the amplitude of the f_1 sinusoid, A_1, varied in 1-dB steps. An FFT was applied to the output waveform and the amplitude of the $2f_1–f_2$ (DP) frequency component was noted.

3.3 Results

The top and bottom panels of Fig. 3.1 illustrate the results for the broken-stick and Boltzmann nonlinearity, respectively. The left and right panels illustrate I/O curves and A_1-A_2 amplitude rules, respectively. PT I/O curves are shown as bold continuous lines. DP I/O curves are shown for three A_1-A_2 combinations: equal-amplitude primaries (A_1=A_2) (crosses); optimal A_1 (circles); and A_1 5 dB lower than optimal (triangles). Optimal A_1 are defined here as the A_1 amplitudes that produced the highest DP amplitudes for each value of A_2.

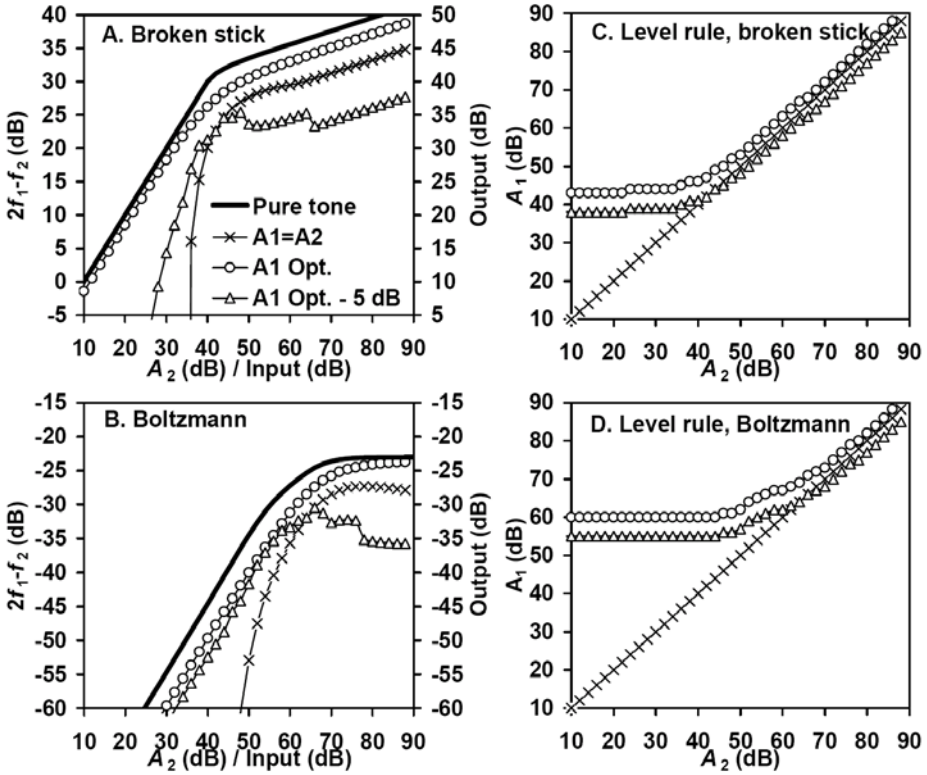

Figure 3.1. Simulated I/O curves (left panels) for different amplitude rules (right panels) and two non-linearities: a broken stick (upper panels) and a Boltzmann function (lower panels). (A) and (B) Bold continuous lines illustrate I/O curves based on single-sinusoid inputs. Symbols illustrate I/O curves for the $2f_1–f_2$ DP for 3 different amplitude rules: $A_1 = A_2$, A_1 = optimal, and $A_1 = A_1$ optimal – 5 dB (as indicated in the inset). (C) and (D) A_1-A_2 amplitude rules corresponding to the I/O curves of A and B.

Simple visual inspection to the I/O curves of Fig. 3.1 reveals that PT I/O curves are most similar to the DP I/O curve obtained with optimal A_1 amplitudes. The figure also reveals that suboptimal A_1 amplitudes can produce non-monotonic DP I/O curves (e.g., triangles) (see also Lukashkin and Russell, 2001). Furthermore, the broken-stick nonlinearity DP I/O curve obtained with slightly suboptimal A_1 amplitudes shows shallow notches. These results support the claims made at the beginning of this chapter.

Several things are to be noted. First, the present results do not imply that the maximal correspondence between DP and PT I/O curves occurs for optimal primary amplitudes; they only show that optimal primary amplitudes lead to a reasonable

correspondence between DP and PT I/O curves. This result is useful in practice because the A_1-A_2 combination that maximizes the correspondence between the two I/O curves will typically be unknown *a priori*. Second, the main conclusions of this exercise are independent of nonlinearity parameter values. Third, the present models are very simplistic and do not consider concurrent filtering. For this reason, the present simulations do not depend on the actual frequency of the sinusoids or the parameters of the nonlinearities. Filtering, however, would almost certainly produce different optimal amplitude combinations than those illustrated in Fig. 3.1(C)-(D).

3.4 Conclusions

Despite its being an oversimplification, this modelling exercise demonstrates that:

(1) Primary levels strongly influence the shape of the DP I/O curve.

(2) Suboptimal levels can produce non-monotonic DP I/O curves with plateaus and wide, shallow notches.

(3) The use of optimal levels may enhance the similarity between PT and DP I/O curves

CHAPTER 4

OTOACOUSTIC EMISSION THEORIES AND BEHAVIORAL ESTIMATES OF HUMAN BASILAR MEMBRANE MOTION ARE MUTUALLY CONSISTENT[3]

4.1 Introduction

DPOAE levels are commonly used as an indicator of the physiological status of the active mechanism within the inner ear. The sensitivity of DPOAEs for this purpose is greatest when the primaries have levels that evoke maximal-level DPOAEs (Mills and Rubel 1994; Whitehead *et al*. 1995). We will use the term DPOAE optimal rule to refer to the combination of primary levels that evokes the highest level of DPOAEs. Up to now, efforts have been directed to obtain optimal rules *empirically* (Kummer *et al*. 1998) but the form of the optimal rule is still controversial (Johnson *et al*. 2006; Kummer *et al*. 2000).

The controversy could be clarified by elucidating the cochlear mechanical conditions that maximize DPOAE levels. The overriding view is that maximal-level DPOAEs occur when the primaries produce equal excitation at the cochlear region most sensitive to f_2 (Kummer *et al*. 2000; Neely *et al*. 2005). Concurrent DPOAE

[3] This chapter is adapted from the published article: Lopez-Poveda, E. A., and Johannesen, P. T. **2009**. "Otoacoustic emission theories and behavioral estimates of human basilar membrane motion are mutually consistent," J. Assoc. Res. Otolaryngol. **10**, 511–523. Some portions of the article are reproduced here with permission from the Journal of the Association for Research in Otolaryngology.

and BM recordings have revealed that this view is approximately true for rodents (Rhode 2007), but a confirmation in living human is still pending because it is not possible to directly record the motion of their BM while monitoring their ear canal DPOAE levels.

On the other hand, it has been claimed that it is possible to infer the levels of two equally-effective pure tones at a given cochlear site from behavioral forward masking thresholds. The technique is known as the TMC method and is arguably the favoured procedure to infer human BM I/O curves (Lopez-Poveda *et al*. 2003; Nelson *et al*. 2001). The TMC method would seem an appropriate tool to verify the DPOAE-generation conjecture of Kummer *et al*. (2000) in human. Unfortunately, its assumptions (described below) have been validated only indirectly, using computer models or other psychoacoustical methods, but lack direct physiological support.

A high correlation between DPOAE optimal level rules and corresponding behavioral rules inferred using the TMC method would provide strong support to both the conjecture of Kummer *et al*. (2000) on the generation of maximal-level DPOAEs *and* the assumptions of the TMC method of inferring human BM responses. The present study aimed at investigating such correlation. It will be shown that a high correspondence exists for frequencies of 1 and 4 kHz and for levels below around 65 dB SPL.

4.2 Methods

4.2.1 Rationale

A TMC is a plot of the levels of a pure tone (masker) required to just mask a brief following tone (probe) as a function of the time gap between the masker and the probe. The probe level is fixed just above the absolute threshold for the probe. The masker level at the masking threshold increases with increasing time gap and is thought to depend on two variables (Nelson *et al*. 2001). First, it depends on the

time gap: the amount of masking decreases as the masker-probe time gap increases (Duifhuis 1973; Moore and Glasberg 1983; Nelson and Freyman 1987). Second, it depends on the relative excitation produced by the masker and the probe at the BM place tuned at or close to the probe frequency (Nelson *et al*. 2001; Oxenham *et al*. 1997; Oxenham and Moore 1995; Oxenham and Plack 1997). Because the probe level is fixed at all times, a TMC is assumed to represent the masker levels required to generate a fixed level of excitation after decaying during the masker-probe time gap. This is why the resulting functions are referred to as isoresponse temporal masking curves or TMCs (Nelson *et al*. 2001).

There is strong evidence that the rate of recovery from forward masking is approximately the same for different masker frequencies over a wide range of masker levels (Wojtczak and Oxenham 2009). Although this evidence is for a probe frequency of 4 kHz, indirect evidence suggests that the same applies to probe frequencies as low as 0.5 kHz (Lopez-Poveda and Alves-Pinto 2008). Therefore, it seems reasonable to *assume* that for any given masker-probe time gap, two masker of slightly different frequencies (e.g., f and f/1.2) with levels at their masking thresholds produce *identical* degrees of excitation at a cochlear site tuned approximately to the probe frequency. This assumption is commonplace when inferring cochlear I/O curves and compression exponents from TMCs (Lopez-Poveda *et al*. 2003, 2005; Nelson *et al*. 2001; Plack *et al*. 2004; Wojtczak and Oxenham 2009).

Based on the above, our approach consisted in measuring two TMCs, both for a probe frequency equal to the DPOAE test frequency (f_2), and for masker frequencies equal to the DPOAE primary tones (f_1, f_2; with f_2/f_1=1.2). We then plotted the resulting levels for the f_1 masker (L_1) against those for the f_2 masker (L_2), paired according to masker-probe time gap. Based on the previously-explained interpretation of TMCs, the resulting plot should illustrate the combination of levels, L_1-L_2, for which two pure tones of frequencies f_1 and f_2 produce approximately comparable degrees of excitation at the f_2 cochlear site. If the current DPOAE generation model (as described by Kummer *et al*., 2000) and the

assumptions of the TMC method are both correct, then this behavioral rule should match with a DPOAE optimal rule obtained empirically.

All human procedures were approved by the human experimentation ethical committee of the University of Salamanca.

4.2.2 Subjects

A total of 14 subjects participated in the study. Their ages ranged from 20 to 39 years. Their hearing was audiometrically normal (i.e., absolute hearing thresholds < 20 dB HL) at the three tests frequencies considered in this study (0.5, 1 and 4 kHz). Table 2.1 details, for subjects S1 to S10, their behavioral absolute thresholds (in dB SPL) for pure tones of 0.5, 1 and 4 kHz and durations of 10, 110, and 300 ms. Table 4.1 shows the same but for subjects S11 to S14.

Table 4.1. Thresholds in dB SPL measured with Etymotic ER2 insert earphones for subjects S11 to S14 and for tone durations of 300 ms absolute threshold, 110 ms masker threshold, and 10 ms probe threshold, respectively. n.a. stands for not available.

Frequency (kHz)	*0.5*	*1*	*4*
Tone duration (ms)	*300 / 110 / 10*	*300 / 110 / 10*	*300 / 110 / 10*
S11	13 / 13 / 43	8 / 5 / 26	9 / n.a. / 25
S12	22 / n.a. / n.a.	14 / 14 / 36	21 / n.a. / 39
S13	3 / 7 / 28	1 / 3 / 26	2 / n.a. / 21
S14	17 / n.a. / n.a.	7 / 10 / 30	9 / n.a. / 26

4.2.3 Behavioral rules

TMCs were measured for probe frequencies (f_P) of 0.5, 1 and 4 kHz and for masker frequencies equal to f_P and $f_P/1.2$. These masker frequencies were equal to those of the primary tones (f_1 and f_2, respectively) used to measure DPOAEs (see below). The masker-probe time gaps, defined as the 0-Volt period from masker offset to probe onset, ranged from 5 to 100 ms in 5-ms steps with an additional gap of 2 ms. The durations of the masker and the probe were 110 and 10 ms, respectively, including 5-ms cosine-squared onset and offset ramps. The probe had no steady state portion. The level of the probe was fixed at 9 dB above the individual absolute threshold for the probe as shown in Table 2.1.

Stimuli were generated with a Tucker Davies Technologies Psychoacoustics Workstation (System 3) operating at a sampling rate of 48.8 kHz and with analog to digital conversion resolution of 24 bit. If needed, signals were attenuated with a programmable attenuator (PA-5) before being output through the headphone buffer (HB-7). Stimuli were presented to the listeners through Etymotic ER-2 insert earphones. TMC sound pressure levels (SPL) were calibrated by coupling the earphones to a sound level meter through a Zwislocki DB-100 coupler. Calibration was performed at 1 kHz only and the obtained sensitivity was used at all other

frequencies because the earphone manufacturer guarantees an approximately flat (± 2 dB) response between 200 Hz and 10 kHz.

Masker levels at masking threshold were measured using a two-interval, two-alternative, forced-choice adaptive procedure with feedback. Two sound intervals were presented to the listener in each trial. One of them contained the masker only and the other contained the masker followed by the probe. The interval containing the probe was selected randomly. The subject was asked to indicate the interval containing the probe. The inter-stimulus interval was 500 ms. The initial masker level was set sufficiently low that the listener always could hear both the masker and the probe. The masker level was then changed according to a two-up, one-down adaptive procedure to estimate the 71% point on the psychometric function (Levitt 1971). An initial step size of 6 dB was applied, which was decreased to 2 dB after three reversals. A total of 15 reversals were measured. Threshold was calculated as the mean of the masker levels at the last 12 reversals. A measurement was discarded if the SD of the last 12 reversals exceeded 6 dB. Three threshold estimates were obtained in this way and their mean was taken as the threshold. If the SD of these three measurements exceeded 6 dB, a fourth threshold estimate was obtained and included in the mean. Measurements were made in a double-wall sound attenuating booth. Listeners were given at least 2 h of training on the TMC task before data collection began.

The resulting TMCs were least-squares fitted with Eq. (1) of Lopez-Poveda *et al*. (2005). Individual behavioral level rules were obtained by plotting the fitted levels for the f_1 masker against those for the f_2 masker, paired according to the masker-probe time gaps.

4.2.4 DPOAE optimal rules

The magnitude (in dB SPL) of the $2f_1–f_2$ DPOAE was measured for f_2 frequencies of 1 and 4 kHz and for a fixed primary frequency ratio of $f_2/f_1 = 1.2$. Individual DPOAE optimal rules were obtained by systematically varying the levels of the two

primaries (L_1 and L_2 for f_1 and f_2, respectively) to find the L_1-L_2 combinations that produced the highest DPOAE response levels. L_2 was varied in 5-dB steps within the range from 35 to 75 dB SPL. For each fixed L_2, L_1 was varied in 3-dB steps and the individual optimal value (i.e., the level that produced the highest DPOAE level) was noted.

The DPOAE magnitude can vary rapidly by changing the test frequency only slightly (Gaskill and Brown 1990). These variations are most clearly seen in a DP-gram (i.e., the graphical representation of the DPOAE magnitude as a function of test frequency f_2). They are known as 'DPOAE fine structure' and can be as large as 20 dB for an f_2 change of 1/32 octave (He and Schmiedt 1993). The fine structure is thought to occur by vector summation of two DPOAE contributions: one that originates at the BM region of maximum overlap between the cochlear excitation patterns evoked by the two primaries (i.e., the f_2 region); and one that originates at the cochlear site with CF~$2f_1-f_2$, where the first contribution reflects back to the ear canal. The varying phases of these two contributions give rise to constructive and destructive interferences, thus to peaks and valleys in the DP-gram (Heitmann *et al*. 1998; Shera and Guinan 1999).

In an attempt to reduce the potential influence of the fine structure on the DPOAE optimal rules, three such rules were obtained for three f_2 frequencies close to the frequency of interest and their mean was taken as the actual DPOAE optimal rule. The three test frequencies in question were equal to 0.99*f*, *f*, and 1.01*f*, where *f* denotes the frequency of interest. For instance, the final DPOAE optimal rule at 4 kHz was the mean of three optimal rules for f_2 frequencies of 3960, 4000, and 4040 Hz. This procedure was inspired by earlier studies that showed that a 'clean' DP-gram (i.e., a DP-gram without the influence of the fine structure) resembled very closely a moving average of the original DP-gram with fine structure (Kalluri and Shera 2001; Mauermann and Kollmeier 2004).

4.2.5 DPOAE I/O curves

DPOAE I/O curves were measured for f_2 frequencies of 1 and 4 kHz with *individual* behavioral and DPOAE optimal level rules, as well as with the rule of Kummer *et al*. (1998) ($L_1 = 0.4L_2+39$). I/O curves were also measured for an f_2 of 500 Hz, but using only individual behavioral rules and the rule of Kummer *et al*. When the rule of Kummer *et al*. was applied, L_2 ranged from 20 to 75 dB SPL in 5-dB steps, except for $f_2 = 0.5$ kHz for which it ranged from 45 to 75 dB SPL. The primary frequency ratio was always fixed at $f_2/f_1 = 1.2$.

To reduce the potential influence of the fine structure on the I/O curves, five such curves were measured for five close f_2 frequencies around the frequency of interest, and the resulting I/O curves were averaged (Johannesen and Lopez-Poveda 2008). For instance, the final DPOAE I/O curve at 4 kHz was the mean of five I/O functions for f_2 frequencies of 3920, 3960, 4000, 4040, and 4080 Hz. Three I/O curves were obtained in this way for the behavioral and Kummer rules per f_2 frequency and subject, the mean of which was taken as the 'true' I/O curve. For the individual DPOAE optimal rules, only one such I/O curve was measured.

4.2.6 DPOAE measurement procedure

DPOAE measurements were obtained with an Intelligent Hearing System's Smart device (with SmartOAE software version 4.52) equipped with an Etymotic ER-10D probe. During the measurements, subjects sat comfortably in a double-wall sound attenuating chamber and were asked to remain as steady as possible. When seeking DPOAE optimal rules, a recording session consisted of measuring DPOAE responses for all possible primary-level combinations (L_1, L_2) for one of the three adjacent f_2 frequencies considered per frequency of interest (see above). When measuring DPOAE I/O curves, a recording session consisted of measuring I/O curves for the five adjacent frequencies considered for each frequency of interest (see above).

The probe fit was checked before and after each recording session. The probe remained in the subject's ear throughout the whole measurement session to avoid measurement variance from probe fit. DPOAEs were measured for a preset measurement time, which ranged from 12 s for high L_2 to 1 min for low L_2. A DPOAE measurement was considered valid when it was 2SD above the measurement noise floor (defined as the mean level over 10 frequency bins adjacent to the $2f_1-f_2$ component in the OAE spectrum). When a response did not meet this criterion, the measurement was repeated and the measurement time was increased if necessary. The probe remained in the same position during these re-measurements. If the required criterion was not met after successive tries, the measurement point was discarded.

DPOAE measurements were regarded as valid only when they were 6 dB above the system's artifact response. The rationale behind this rather strict criterion and the details of the procedure for controlling for system's artifacts can be found in Chapter 2.

4.3 Results

4.3.1 Temporal masking curves

Figures 4.1 to 4.3 illustrate TMCs for probe frequencies (f_P) of 0.5, 1, and 4 kHz, respectively. Each panel illustrates the TMCs for one subject (as indicated in the top-left corner of the panel) and for two masker frequencies at $f_1 = f_P/1.2$ (filled symbols) and $f_2 = f_P$ (open symbols).

The characteristics of the present TMCs were overall consistent with those reported elsewhere for similar stimuli (Lopez-Poveda *et al*. 2003; Nelson and Schroder 2004; Plack and Drga 2003). In broad terms, these are as follows. TMCs were overall steeper for the on-frequency masker (i.e., the masker whose frequency was equal to the probe frequency) than for the off-frequency masker (i.e., the masker whose frequency was below the probe frequency). This difference in slope

is interpreted to reflect the different rates of growth of the corresponding cochlear responses for stimulus frequencies at or below the characteristic frequency (CF) of the cochlear site tuned to the probe frequency, respectively. That is, the steeper portions of the TMCs are interpreted to reflect shallower growths of BM response with increasing masker level, hence greater degrees of BM compression.

At short masker-probe time gaps, higher levels are required for the off- than for the on-frequency masker to mask the fixed-level probe. This is consistent with the fact that, at low levels, the level of a pure tone below the CF must be higher than that of an on-CF tone for both tones to produce equal responses at the cochlear site in question. For moderate-to-long gaps, however, the levels of the lower, off-frequency masker overlap or are even lower than those of the on-frequency masker. This is interpreted to reflect that at high levels, below-CF tones produce comparable or more cochlear excitation than on-CF tones, which is consistent with broader tuning at high levels and with the well-reported basalward shift of cochlear excitation with increasing level for CFs above approximately 1 kHz (Robles and Ruggero 2001; Ruggero *et al*. 1997). Interestingly, in a few instances (e.g., S10 and S15 in Fig. 4.1, or S11 in Fig. 4.2) the TMCs for both maskers crossed again at very long gaps, suggesting that the on-frequency masker became more effective than the off-frequency masker again at very high levels. A similar 'rebound' effect can be observed in earlier reports (e.g., Fig. 2 of Lopez-Poveda *et al*., 2003). The explanation of this result is uncertain. It would be consistent with an apicalward shift of cochlear excitation at very high levels following the previously mentioned basalward shift at moderate levels. Direct BM responses suggest that this shift is possible but existing evidence only applies to apical cochlear regions [see the cross symbols (×) in Figs. 2.3 and 2.4 of (Cooper 2004)]. Another possibility would be that the rate of decay of the post-cochlear masker effect becomes slower for the on-frequency masker than for the lower, off-frequency one at very high levels. Hence, for long gaps, the required masker level at threshold would be lower for the on- than for the off-frequency masker. To our knowledge, there is no evidence that this

is the case. In fact, existing evidence suggests the opposite (Wojtczak and Oxenham, 2009). In any case, the ‘rebound’ effect was rare and occurred over a range of masker levels much higher than the maximum primary level for which DPOAEs could be measured reliably (80 dB SPL). Therefore, it had no effect on the conclusions of the present paper.

More detailed interpretations of TMC characteristics are provided elsewhere (Lopez-Poveda *et al*. 2003; Nelson *et al*. 2001).

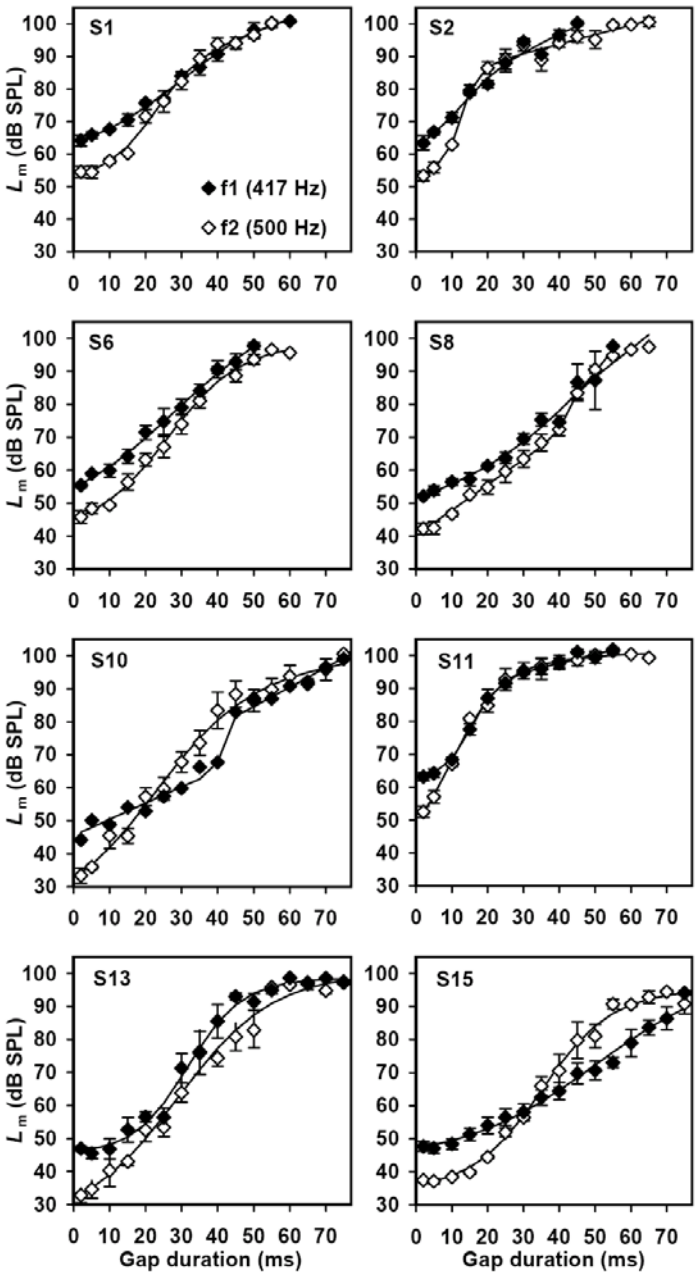

Figure 4.1. TMCs for all listeners for probe frequencies (f_2) of 500 Hz. Each panel shows data for one subject. Open symbols illustrate TMCs for a masker frequency equal to the probe frequency (f_P); filled symbols illustrate TMCs for a masker frequency below the probe frequency (f_P/1.2). Each data point is the mean of at least 3 measurements. Error bars illustrate 1SD of the mean (n > 3). Lines illustrate function fits to the data points. L_m stands for masker level.

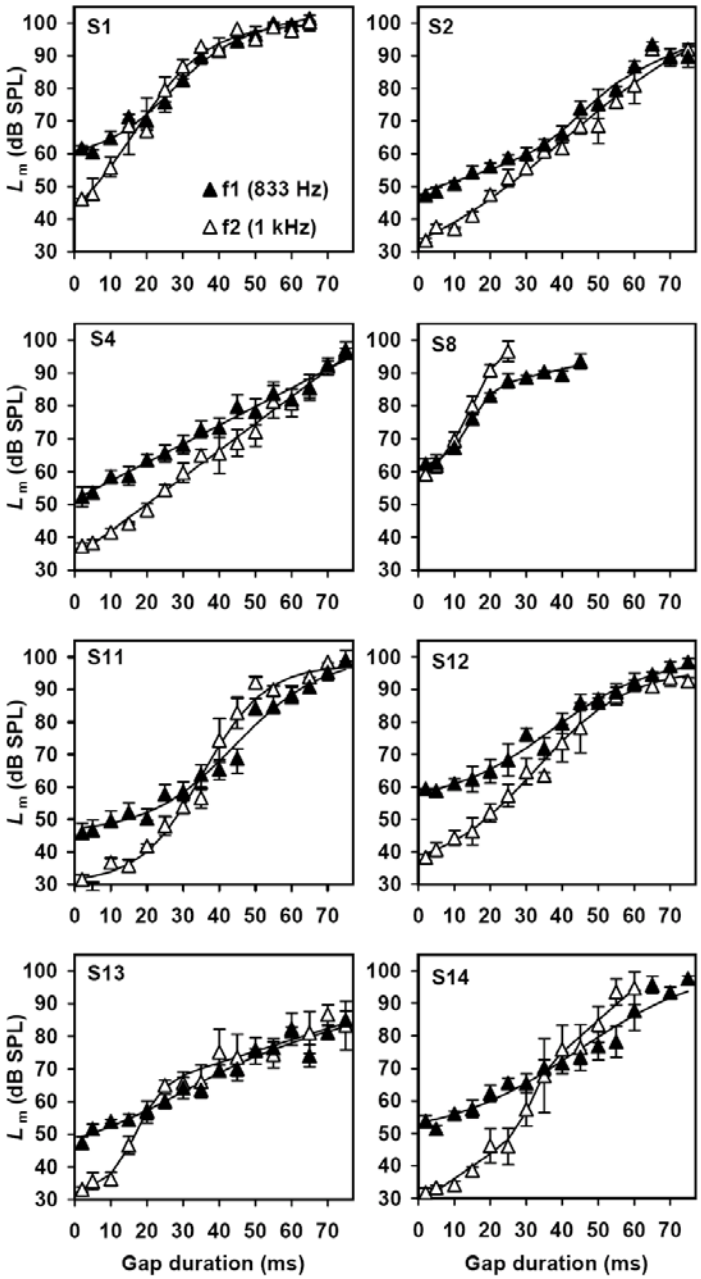

Figure 4.2. As Fig. 4.1 but for a probe frequency of 1 kHz.

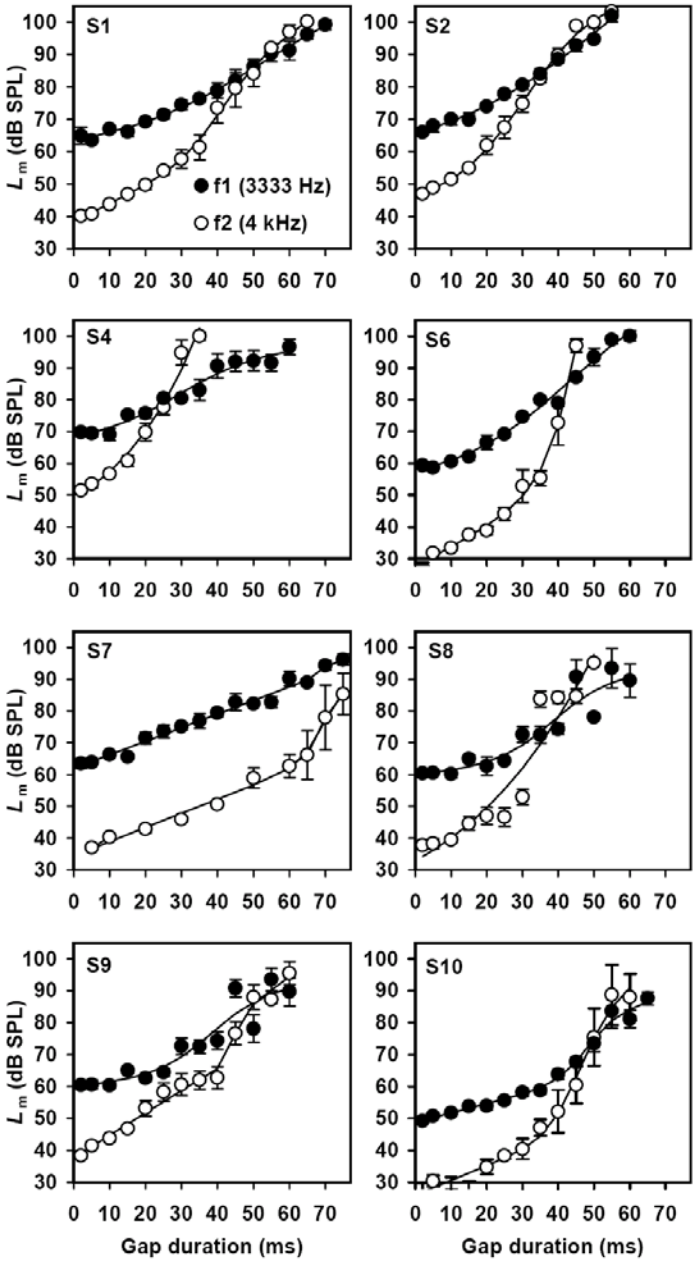

Figure 4.3. As Fig. 4.1 but for a probe frequency of 4 kHz.

4.3.2 The influence of the fine structure on DPOAE optimal rules

Figure 4.4 provides several illustrative examples of the influence of the DPOAE fine structure on the individual DPOAE optimal rules of several subjects for frequencies of 1 (left panels) and 4 kHz (right panels). Clearly, the level of f_1 (L_1)

that evoked the maximal DPOAE response for any given level of f_2 (L_2) changed by as much as 8 dB with a change in f_2 of only 1% for some conditions and subjects. The figure shows that the fine structure could have affected the DPOAE optimal rules, albeit only slightly, and thus justifies our approach to use the mean curve for three adjacent frequencies (illustrated with filled circles) as the DPOAE optimal rule.

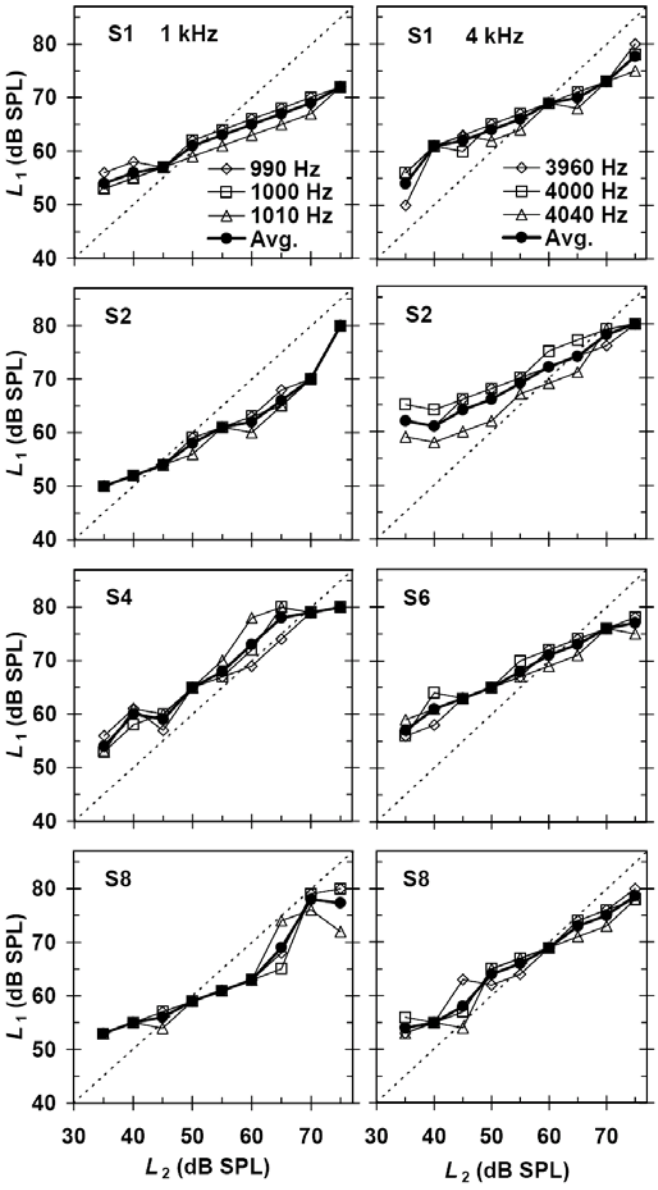

Figure 4.4. Examples of the influence of the fine structure on individual DPOAE optimal rules at 1 and 4 kHz (left and right panels, respectively). The listener identifier is shown on the top-left corner of each panel. Open symbols illustrate DPOAE optimal rules for a different test frequency at or around the frequency of interest (as indicated by the insets in the top panels). Filled circles illustrate the mean curves, which were regarded as the actual DPOAE optimal rules without the influence of the fine structure.

4.3.3 Behavioral *vs*. DPOAE optimal rules

Figures 4.5 and 4.6 illustrate plots of the mean levels of the f_1 masker at threshold against those of the f_2 masker, paired according to masker-probe time gap. Based on the interpretation of TMCs explained in the Methods section, these illustrate level combinations of two equally-effective maskers at the cochlear site tuned to the probe frequency (f_2). These behavioral rules are compared with individual DPOAE optimal rules (i.e., with primary level combinations, L_1-L_2, that produced maximal DPOAE levels) for primary tone frequencies equal to the masker frequencies. The match between the two rules varied from subject to subject and across frequencies. It was almost perfect at 1 kHz for several subjects (e.g., S1, S2, S4 or S14) over an L_2 range typically below 65 dB SPL. The degree of correspondence in the individual data was generally less at 4 kHz, but it was still apparent for several subjects (e.g., S2, S8 or S10). The reasons for the lower degree of correspondence above 65 dB SPL will be discussed later.

Individual behavioral rules were based on *mean* values of at least three independent estimates of L_1 and L_2, whereas the DPOAE optimal rules were based on a single L_1 for every L_2 (note that this L_1 was the mean of three estimates, each for a slightly different f_2 around the frequency of interest; see Methods). To test for the statistical significance of the difference between the two individual rules, we simply checked if the single DPOAE optimal rule estimate fell within the range of all possible combinations of L_1 and L_2 based on the available TMC data. The latter are illustrated as small gray dots in Figs. 4.5 and 4.6. Except, perhaps, for S7 at 4 kHz, *all* DPOAE optimal rules always fell within the variability of the behavioral combinations for $L_2 \leq 65$ dB SPL. Subject S7 was peculiar in that he repeatedly reported that the behavioral task was extremely difficult.

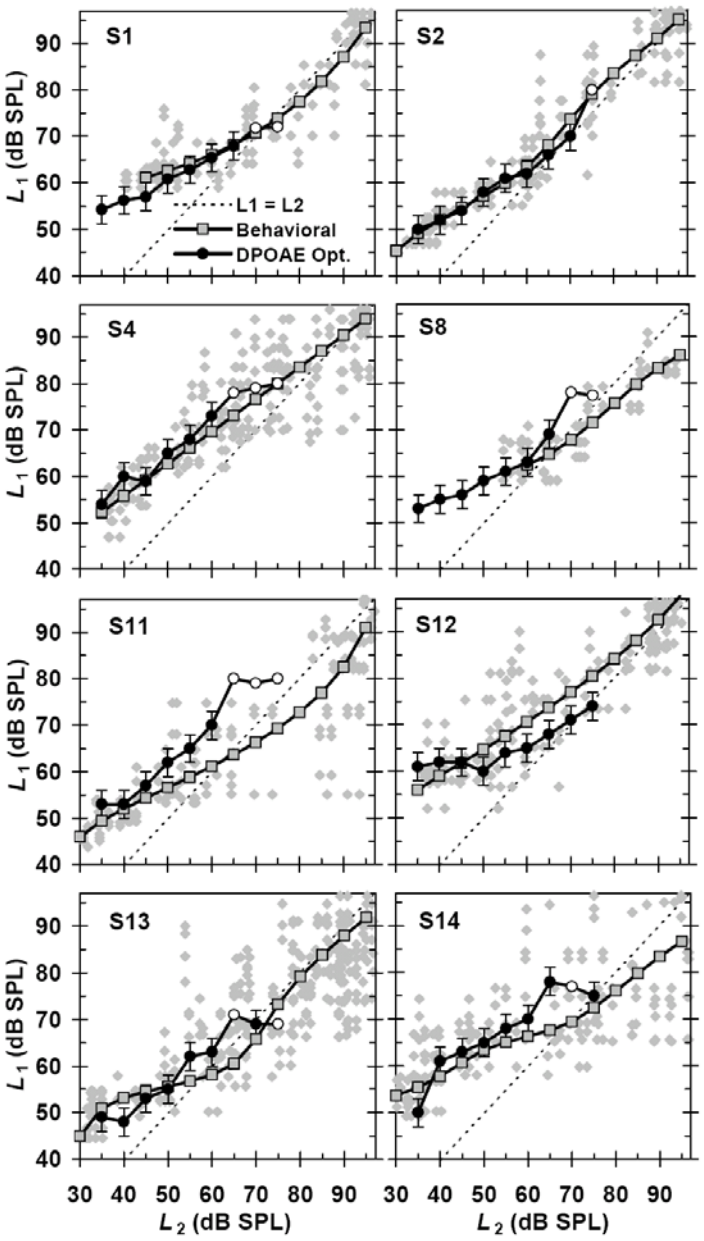

Figure 4.5. Comparison of behavioral and DPOAE optimal level rules for a test frequency of 1 kHz. Each panel illustrates results for a single subject. Gray squares illustrate behavioral rules based on *mean* TMCs. Small gray dots illustrate *all* possible L_1-L_2 combinations based on the available TMC data and thus inform of the variability of the behavioral rule based on the data available. Circles illustrate DPOAE optimal rules. Open circles indicate possible suboptimal L_1-L_2 combinations. That is, combinations that produced the strongest DPOAEs within the L_1 level range of the system (< 80 dB SPL), but it is possible that higher DPOAE levels might have been measured with higher L_1s, should the system had allowed them. Error bars are fixed at ±3 dB and indicate the precision of the optimal L_1 (see Methods for more details). They are not represented for the open circles. Dotted lines illustrate equal-primary level rules (L_1=L_2).

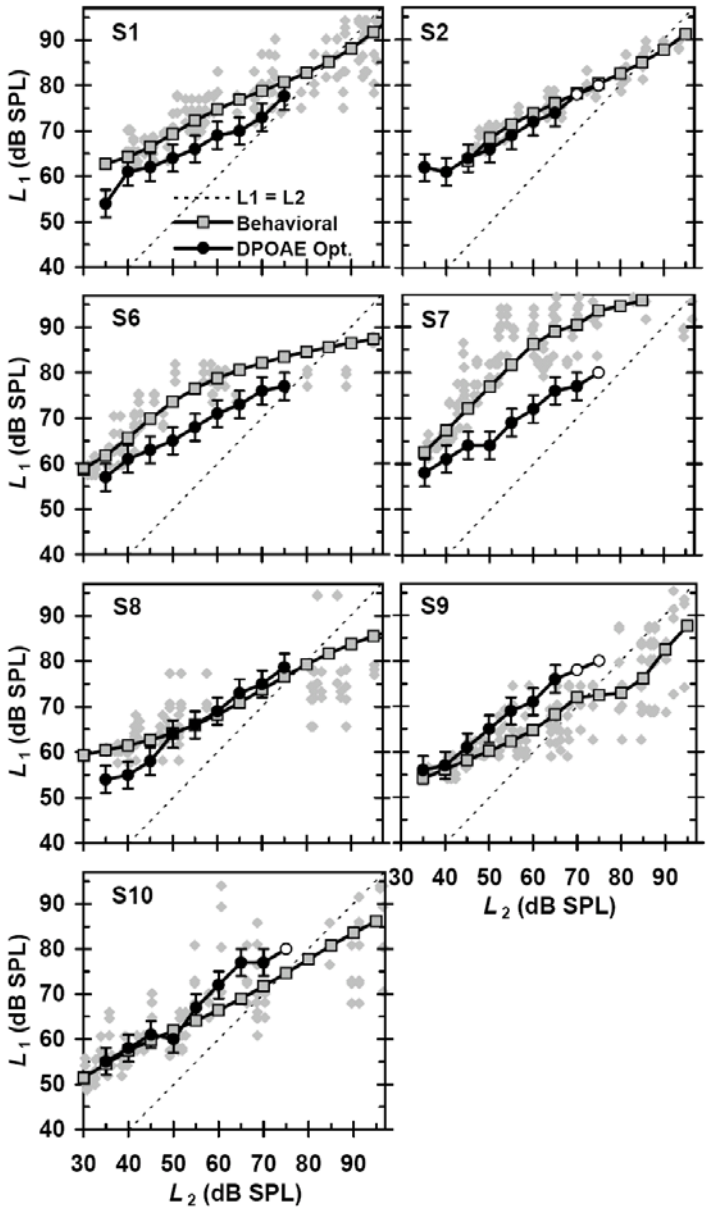

Figure 4.6. As Fig. 4.5 but for a test frequency of 4 kHz.

The reason for the variability in the behavioral L_1-L_2 combinations (gray dots in Figs. 4.5 and 4.6) is uncertain. Such variability reflects, by definition, the variability across TMC estimates (Figs. 4.2 and 4.3). Consistent with many previous studies (e.g., Lopez-Poveda *et al.*, 2003, 2005; Nelson *et al.*, 2001; Plack *et al.*, 2004), the variability in the present masker levels was noticeably larger over the

steeper portions of the TMCs (Figs. 4.2 and 4.3). The steeper portion of a TMC is assumed to reflect the range of levels where the masker is subject to greater cochlear compression (Nelson *et al*. 2001). Therefore, even small changes in cochlear responses or in the listener's sensitivity across measurement sessions would produce a large change in masker level at threshold and would explain such variability.

Figure 4.7 illustrates *mean* rules across listeners. Interestingly, the mean DPOAE optimal rules overlapped with behavioral rules at 1 kHz over the L_2 range for which both sets could be measured reliably (35-65 dB SPL). At 4 kHz, the two rules did not overlap but they were within 1SD of each other and their difference was not statistically significant for $L_2 \leq 65$ dB SPL ($p < 0.05$, point-by-point, two-tailed, paired *t*-test). Indeed, the difference was accentuated by the results of a single subject (S7).

4.3.4 The dependence of behavioral and DPOAE optimal rules on test frequency

Average behavioral rules (gray squares in Fig. 4.7) approached equal level for high L_2 levels (~75 dB SPL). The difference between L_1 and L_2 was larger at lower than at higher L_2s and increased gradually with increasing frequency. Straight lines were fitted to the data for $L_2 \leq 65$ dB SPL (thick continuous lines in Fig. 4.7). As indicated in Table 4.2, the lines for both behavioral and DPOAE optimal rules had similar slopes at 1 and 4 kHz, and these were shallower than the line fitted to the behavioral rule at 0.5 kHz. Behavioral rules, however, had slopes that were statistically indistinct across frequencies ($p > 0.05$, two-tailed, equal-variance, *t*-test).

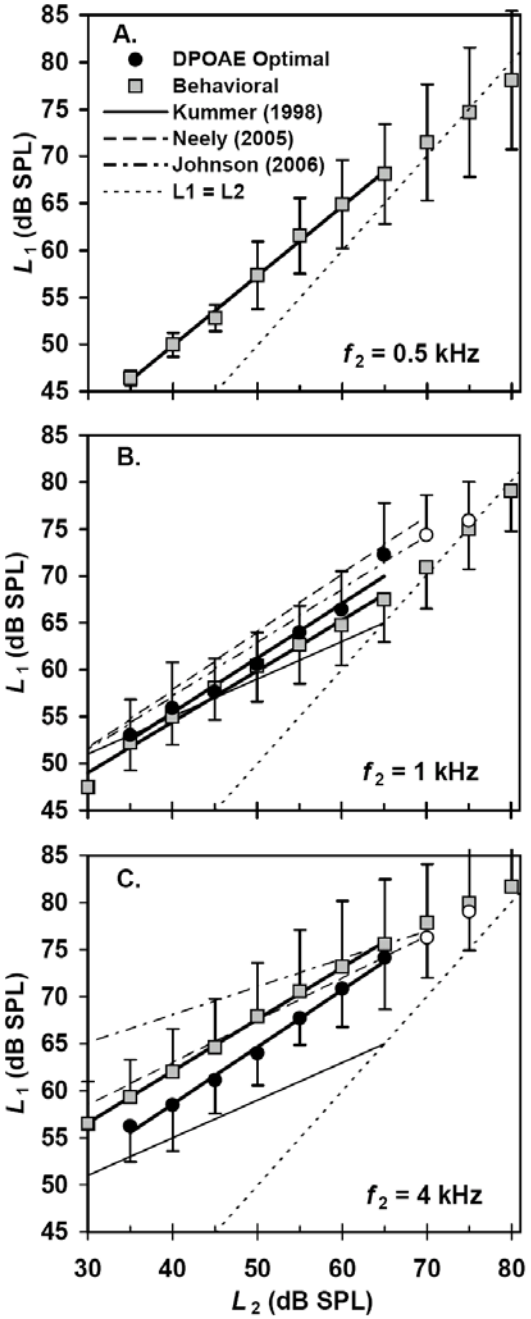

Figure 4.7. Average L_1-L_2 rules at different frequencies and as proposed by the present and earlier studies. (A) For f_2 = 0.5 kHz. (B) For f_2 = 1 kHz. (C) For f_2 = 4 kHz. Circles and squares illustrate average DPOAE optimal and behavioral rules obtained in the present study, respectively. Each data point is the mean of data for at least 3 subjects. Error bars illustrate 1SD (n ≥ 3, typically 5) Thick straight lines illustrate least-square fits to these experimental rules for $L_2 \leq 65$ dB SPL. Dotted lines illustrate equal-level rules. Primary-level rules proposed by earlier representative studies are also shown for comparison. These are plotted across the L_2 range originally proposed in their original reports. Open circles indicate possible suboptimal values (see the legend of Fig. 4.5).

Table 4.2. Regression parameters for linear relationships, $L_1 = aL_2 + b$, based on the present behavioral and DPOAE optimal rules, for an L_2 range between 30 and 65 dB SPL.

Rule	*n*	*Frequency (kHz)*	*b (dB SPL)*	*a*
DPOAE Optimal	8	1	31.1	0.61
	7	4	35.6	0.59
		Average	33.4	0.60
Behavioral	8	0.5	23.6	0.69
	8	1	35.1	0.49
	7	4	39.5	0.56
		Average	32.7	0.58

4.3.5 DPOAE I/O curves

The growth of the DPOAE magnitude as a function of L_2 was measured for each listener using his/her *individual* DPOAE-optimal and behavioral rules to obtain individual DPOAE I/O curves. The mean I/O curves are shown in Fig. 4.8. The figure also illustrates mean I/O curves measured with the rule of Kummer *et al*. (1998), which was identical across subjects and frequencies. For the three f_2 frequencies, DPOAEs grew with increasing L_2 at rates considerably lower than 1 dB/dB. DPOAE levels measured using DPOAE optimal rules (filled circles) were the highest and were consistently 3-5 dB higher than those measured with the behavioral rules (gray squares); that is, circles and squares run parallel to each other across the L_2 level range.

The DPOAE levels evoked by the rule of Kummer *et al*. were identical or lower than those evoked by the behavioral rule at 1 and 4 kHz across levels, except for $L_2 < 50$ dB SPL. The rule of Kummer *et al*. evoked slightly higher DPOAE levels than the behavioral rule at 500 Hz (Fig. 4.8A). The difference between the DPOAE levels measured with the optimal and the Kummer rules increased with increasing L_2 at 1 and 4 kHz.

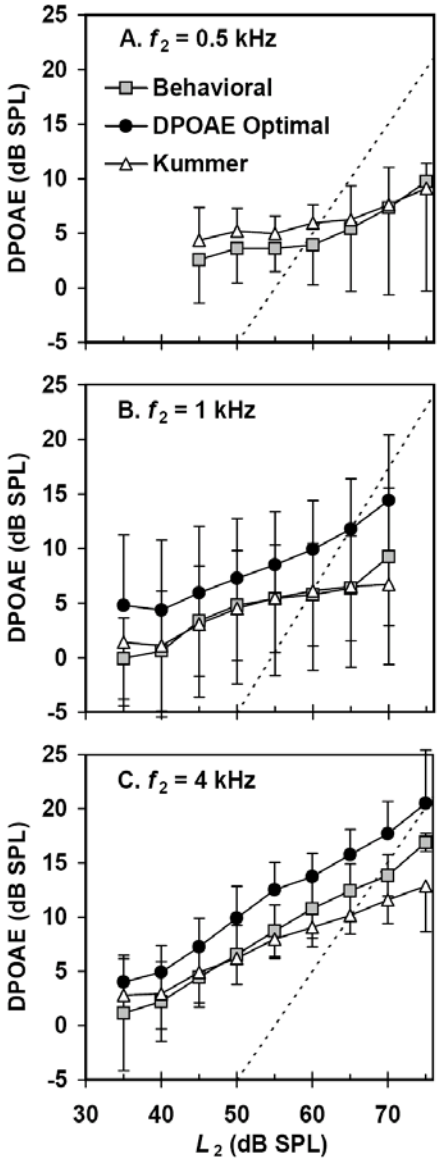

Figure 4.8. Average I/O curves for the $2f_1–f_2$ DPOAE for different primary level rules (as indicated in the inset of panel A). Different panels illustrate results for a different f_2 frequency. (A) f_2 = 0.5 kHz. (B) f_2 = 1 kHz. (C) f_2 = 4 kHz. Every point is the mean of data for between 3 and 7 subjects (typically n = 5). Error bars illustrate 1SD. Note that the rule of Kummer *et al*. was originally designed for $L_2 \leq 65$ dB SPL but was extrapolated here to higher L_2. Dotted lines indicate a linear growth with a slope of 1 dB/dB.

4.4 Discussion

The general overlap between behavioral and empirical DPOAE optimal rules at 1 and 4 kHz (Figs. 4.5-4.7) for L_2 below approximately 65 dB SPL supports the view that DPOAE levels are highest when the two primaries produce similar responses at the f_2 cochlear region (Kummer *et al*., 2000).

The present behavioral rules were derived on the assumption that the TMCs for the two maskers (f_1 and f_2) reflect *only* differences in the cochlear excitation evoked by the two tones at the cochlear site tuned to the probe frequency (f_2 in this case). That is, on the assumption that the post-cochlear interaction between the masker and the probe is linear and identical for the two maskers, for all masker-probe time gaps and masker levels. This assumption is commonplace when inferring cochlear I/O functions from TMCs (Lopez-Poveda *et al*. 2003; Nelson *et al*. 2001) and is supported by modeling (Oxenham and Moore 1994) and experimental studies, at least for levels below approximately 83 dB SPL (Lopez-Poveda and Alves-Pinto 2008; Wojtczak and Oxenham 2009). The post-cochlear interaction may be (slightly) different for masker frequencies that are an octave apart at higher masker levels (Lopez-Poveda and Alves-Pinto 2008; Wojtczak and Oxenham 2009), but this is unlikely to undermine the validity of the present approach because the present maskers were closer in frequency and conclusions are claimed to be valid only for $L_2 \leq$ ~65 dB SPL. Furthermore, given that the behavioral and DPOAE optimal rules were inferred using fundamentally different assumptions and methods, the correspondence between the two provides further support for the assumptions of the behavioral TMC method, at least below 65 dB SPL.

The level of the $2f_1-f_2$ DPOAE measured in the ear canal is almost certainly the sum of contributions from several cochlear sources and generation mechanisms. These include distortion generated by nonlinear interaction of the primaries in the f_2 cochlear region (Martin *et al*. 1998), reflection of this distortion at the $2f_1-f_2$ cochlear site (Kalluri and Shera 2001), and nonlinear interaction between f_2 and the

first harmonic of f_1 ($2f_1$) at a more basal ($2f_1$) cochlear site (Fahey *et al*. 2000). The relative weight of these contributions to the measured DPOAE is uncertain. The present behavioral rules were obtained from TMCs for probe frequencies equal to the DPOAE test frequencies (f_2). Based on the current interpretation of TMCs (Lopez-Poveda *et al*. 2003; Lopez-Poveda and Alves-Pinto 2008; Nelson *et al*. 2001), the behavioral rules thus reflect L_1-L_2 combinations for which the two maskers (f_1, f_2) produce equal responses at a cochlear site with a CF~f_2. Therefore, the match between behavioral and DPOAE optimal rules for levels below 65 dB SPL (Figs. 4.5-4.7) together with the evidence that DPOAEs originate at multiple cochlear locations suggest that the DPOAE contribution from the f_2 region is dominant and/or that the DPOAEs originated at the other sites are proportional to the contribution generated at the f_2 region.

Given the reasonable match between mean behavioral and DPOAE optimal rules (Fig. 4.7B-C), it is unclear why the DPOAE levels were on average 3-5 dB higher for the DPOAE optimal than for the behavioral rules (Fig. 4.8B-C). Recall that the DPOAE I/O curves of Fig. 4.8 represent the mean of the curves obtained with *individual* DPOAE optimal and behavioral rules. One possibility is that mutual suppression between the primaries may have affected the results (see also Sec. 6.2). One fundamental difference between the *individual* behavioral and DPOAE optimal rules is that the latter were obtained from DPOAE measurements that required the *simultaneous* presentation of the two primary tones, while behavioral rules were inferred from single-tone responses. In other words, DPOAE optimal rules implicitly take into account possible nonlinear interactions (e.g., suppression) between the primary tones that are disregarded by the behavioral rules. Concurrent recordings of DPOAE and BM responses in chinchilla suggest that DPOAE levels are submaximal for L_1-L_2 combinations that produce equal cochlear responses of simultaneously-presented primaries, and this possibly occurs because the primary tone f_1 suppresses the response of the BM to f_2 [Fig. 1 of Rhode (2007)]. Given that the behavioral rule reflects an equal-response criterion for non-simultaneous

primaries, this might explain why the mean DPOAE levels for the behavioral rules were consistently lower (on average) than those measured with the optimal rules. That said, the same suppression mechanism would have led to behavioral L_1 values being consistently lower than DPOAE optimal L_1 values and this was not the case (Figs. 4.5-4.7). Therefore, mutual suppression is unlikely to account for the difference in DPOAE levels evoked by the two rules [see also the discussion of (Kummer *et al*. 2000)].

An alternative simpler explanation is that individual DPOAE optimal rules were, by definition, those that evoked the highest DPOAE levels (within the 3-dB precision considered for L_1, see Methods). Therefore, any deviation, however small, of the individual behavioral rules from the individual DPOAE optimal rule (Figs. 4.5-4.6) would have always produced *lower* DPOAE levels for each subject and this would be reflected in the mean I/O curves (Fig. 4.8).

The correspondence between the present behavioral and DPOAE optimal rules tended to be less in the individual data (Figs. 4.5-4.6) for L_2 above around 65 dB SPL. In several cases (depicted by open circles in Figs. 4.5-4.6), the optimal L_1 levels were higher than the maximum output of our system and most of these were higher than the corresponding behavioral L_1 values. The reason for this result is uncertain, but it may reflect a shift of the DPOAE generation cochlear site towards the base of the cochlea with increasing L_2. It is well reported that the peak of the cochlear traveling wave shifts basally with increasing sound level (Robles and Ruggero 2001). Therefore, the region of maximum interaction between the traveling waves evoked by the two primaries is likely to shift from that with CF~f_2 to a more basal region as L_2 increases. This shift is illustrated in Fig. 4.9, where the region of maximal interaction at high L_2 is denoted f'_2. Figure 4.9 also illustrates how the increase in L_1 for a given increase in L_2 would be greater if the two primaries were to evoke equal responses at the f'_2 than at the f_2 cochlear regions.

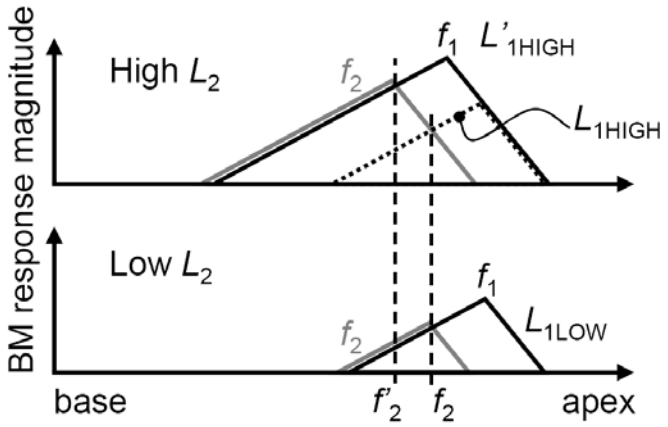

Figure 4.9. Schematic cochlear excitation patterns of the two DPOAE primary tones, f_1 and f_2, for low and high L_2 levels (lower and higher figures, respectively). For low L_2 levels, the maximum interaction between the two excitation patterns occurs at a cochlear region with CF~f_2. At high L_2 levels, however, the excitation patterns of the two primaries shift towards the cochlear base and the new region of maximum interaction, denoted f'_2, is basal to f_2. For the high L_2 condition, two excitation patterns are shown for the f_1 primary: one that produces approximately the same excitation as f_2 at the f_2 site (dashed dark line), and one that produces the same excitation as f_2 at the f'_2 region (continuous dark line). The level of f_1, L_1, would be lower in the former than in the latter case. These are denoted as L_{1HIGH} and L'_{1HIGH}, respectively, with $L'_{1HIGH} > L_{1HIGH} > L_{1LOW}$. The assumptions of the TMC method suggest that behavioral rules illustrate L_{1HIGH} *vs.* L_2, while DPOAE optimal rule might be revealing L'_{1HIGH} *vs.* L_2. See text for further details.

The present behavioral rules were unlikely affected by the shift in question because they were based on TMCs for *fixed*, low-level probes and thus presumably reflected cochlear responses at a *fixed* cochlear site with CF~f_2 at all L_2 (Nelson *et al*. 2001). *If the highest DPOAE levels occurred for equally-effective primaries near the peak of the f_2 traveling waves at each L_2 level*, then DPOAE optimal L_1 values would be higher than behavioral values at high L_2 levels (Fig. 4.9). This might explain why the behavioral L_1 were sometimes lower than the DPOAE optimal L_1 at high L_2 levels. Furthermore, if the level-dependent shift in question were gradual, then this would also explain why some of the DPOAE optimal L_1-L_2 rules appeared generally steeper than their behavioral counterparts over the range of L_2 levels where both of them could be measured (Figs. 4.5-4.6).

The theory that the *main* DPOAE generation site shifts basally with increasing L_2 might be tested by comparing the degree of correlation between

individual optimal level rules with behavioral rules measured with the present method and with the growth-of-masking (GOM) method for a signal frequency equal to f_2 and a masker frequency equal to f_1 (Nelson *et al*. 2001; Oxenham and Plack 1997). Based on the current interpretation of GOM functions (Oxenham and Plack 1997; Rosengard *et al*. 2005), the prediction would be that GOM functions would provide a more accurate behavioral correlate of optimal DPOAE level rules because they implicitly encompass potential level-dependent shifts of cochlear excitation.

4.4.1 On the controversy about DPOAE optimal primary-level rules

Kummer and colleagues have argued that "optimizing the L_1 level for any given L_2 is not a trivial DPOAE level maximization but rather appropriate for maximizing the sensitivity of DPOAE measurements," to discriminate between healthy and damaged cochleae [p. 54 of Kummer *et al*. (2000)]. It is uncertain that DPOAE optimal rules for normal-hearing subjects serve to maximize DPOAE levels of hearing impaired subjects. It is also unlikely that *average* DPOAE optimal rules account for *individual* DPOAE-level variations. Nevertheless, average DPOAE optimal rules for normal-hearing subjects may be regarded as the "best-guess" parameters for any given normal-hearing or hearing-impaired individual. Because of this and given its potential clinical implications, much effort has been spent on providing an accurate *average* DPOAE optimal rule. Unfortunately, different studies disagree in their conclusions. There exists consensus that the optimal L_1 should increase with increasing L_2 following a linear relationship ($L_1 = aL_2 + b$), but some studies have concluded that a and b should be constant across the f_2 range from 1 to 8 kHz (Kummer *et al*. 2000) while others have concluded that they should vary rather systematically with f_2 (Johnson *et al*. 2006; Neely *et al*. 2005).

Although the present study was not aimed at resolving this controversy, the behavioral and DPOAE data it produced incidentally support the view that optimal rules should be approximately similar at 1 and 4 kHz (Fig. 4.7B-C and Table 4.2).

The behavioral data suggest, however, that the optimal rule at 500 Hz is likely to be significantly different from those at higher frequencies (Fig. 4.7 and Table 4.2). Unfortunately, this could not be corroborated with empirical DPOAE optimal rules at this frequency. That said, the I/O curves of Fig. 4.8 suggest that the rule of Kummer *et al*. (2000), which was identical across test frequencies, evoked *average* DPOAE levels that were indistinguishable for most conditions from those evoked by the present behavioral rules, despite them being different, particularly at 500 Hz.

4.5 Conclusions

The agreement between the behavioral and DPOAE optimal rules supports the view that maximal DPOAE levels occur when the two primary tones produce approximately equal responses in the f_2 cochlear region as well as the assumptions of the popular TMC method of inferring human BM responses.

CHAPTER 5

CORRESPONDENCE BETWEEN BEHAVIORAL AND INDIVIDUALLY "OPTIMIZED" OTOACOUSTIC EMISSION ESTIMATES OF HUMAN COCHLEAR INPUT/OUTPUT CURVES[4]

5.1 Introduction

In Chapter 2, we assessed the within-subject correspondence between behaviorally inferred BM I/O curves and DPOAE I/O curves obtained with typical DPOAE parameters. We observed a reasonably high correspondence only for test frequencies above ~2 kHz. The correspondence was, however, poorer at lower frequencies (0.5 and 1 kHz) because DPOAE I/O curves showed notches and plateaus that did not occur in corresponding behavioral curves. The reason for the notches and plateaus was uncertain. Two explanations were suggested. First, notches and plateaus could be due to the DPOAE fine structure (Gaskill and Brown, 1990), despite our attempts to reduce it by spectral averaging (Kalluri and Shera 2001). Second, they could be due to our using suboptimal DPOAE stimulus parameters.

[4] This chapter is adapted from the published article: Johannesen, P. T., and Lopez-Poveda, E. A. **2010**. "Correspondence between behavioral and individually "optimized" otoacoustic emission estimates of human cochlear input/output curves," J. Acoust. Soc. Am. **127**, 3602-3613. Some portions of the article are reproduced here with permission from the Acoustical Society of America.

The term “fine structure” is used to refer to rapid and local variations present in the graphical representation of $2f_1-f_2$ DPOAE level against test frequency (f_2) (known as the DP-gram). These variations are thought to arise by vector summation of DP contributions generated by various mechanisms from spatially distributed regions within the cochlea, which give rise to peaks and notches in the DP-gram (for a review, see Shera and Guinan 2008). The same interference mechanism also influences DPOAE I/O curves (He and Schmiedt, 1993, 1997; Mauermann and Kollmeier, 2004).

The fine structure is equally common across test frequencies (Fig. 2 of Mauermann *et al.*, 1999; Fig. 3 of Dhar and Schaffer, 2004), although this evidence is based on rather small sample sizes (N = 4 and N = 10, respectively). The notches and plateaus of our I/O curves were, by contrast, present only for low test frequencies. Johnson *et al.* (2006b) reported a higher notch incidence at 2 than at 4 kHz based on a larger subject sample (N = 12-22), which might be taken as suggestive of greater incidence of plateaus and notches at low frequencies, consistent with the results of Chapter 2. The evidence of Johnson *et al*. (2006b), however, was based on low primary levels (L_2 = 30 dB SPL), for which the fine structure is known to be more pronounced (He and Schmiedt, 1993, 1997; Mauermann and Kollmeier, 2004; Johnson *et al*., 2006b). Therefore, it is uncertain that the results of Johnson *et al.* (2006b) may be generalized to 50-60 dB SPL, the range of levels where plateaus and notches were observed in the results of Chapter 2. Furthermore, it is also uncertain that the results of Johnson *et al*. (2006b) can be generalized to 0.5 and 1 kHz, the frequencies where plateaus and notches were more frequent in our data. Additionally, the plateaus and the majority of the notches in the I/O curves of our previous study extended over a wider input level range (approximately 20-30 dB) than that of the notches caused by interference from several DP source contributions [about 10 dB, as estimated from the I/O functions in Figs. 6-8 of He and Schmiedt (1993)]. Lastly, the DPOAE I/O curves reported in our previous study were the mean of several (typically five) I/O curves for adjacent

f_2 frequencies in an attempt to reduce the influence of the fine structure by spectral averaging (Kalluri and Shera, 2001). This number was sufficient to account for the fine structure for test frequencies above 2 kHz and so it seems reasonable to assume that it also accounted for the fine structure influence at low frequencies. Altogether, this suggests that the fine structure explanation for the poor correspondence between behavioral and DPOAE I/O curves at low frequencies is possible but may not be sufficient.

Indeed, a complementary explanation may be that the DPOAE I/O curves of Chapter 2 were measured with suboptimal stimulus parameters. DPOAE primaries had frequencies (f_1 and f_2) with a ratio of $f_2/f_1 = 1.2$ (Gaskill and Brown 1990) and their levels conformed to the rule of Kummer *et al.* (1998): $L_1 = 39 + 0.4L_2$, with L_1 and L_2 being the levels of primaries f_1 and f_2, respectively. These parameters were group-average optimal values and were constant across test frequencies; hence it is unlikely that they were optimal on an individual basis. Evidence exists that the primary frequency f_2/f_1 ratio has only a small effect on the shape and slope of I/O curves (Johnson *et al.,* 2006a). By contrast, there is significant evidence that primary levels have a stronger influence on the shape and slope of I/O curves. First, individualized optimal level rules vary significantly across listeners and test frequencies (Neely *et al.,* 2005). Second, only 10% of the subjects have non-monotonic I/O curves in the frequency range from 1 to 8 kHz when using individually optimized DPOAE primary levels (Kummer *et al.,* 2000). Third, numerical simulations shown in Chapter 3 and elsewhere (Lukashkin and Russell, 2001) suggest that (1) suboptimal primary levels may lead to non-monotonic DP I/O curves, even in the absence of secondary DP sources; and (2) the use of optimal primary levels improves the correspondence between DP I/O curves and I/O curves measured using single tones. Lastly, we have confirmed in Chapter 4 and elsewhere (Lopez-Poveda and Johannesen, 2010) that individualized optimal rules for the subjects of Chapter 2 differ from the rule of Kummer *et al*. Altogether this suggests that the correspondence between DPOAE and behaviorally-inferred BM I/O curves could improve by using individualized optimal primary levels.

The use of individualized DPOAE optimal levels may also reduce the fine structure. It is known that varying the primary levels can change the relative contribution from the various DP cochlear sources (He and Schmiedt, 1993). To the authors' knowledge, the details are uncertain but one possibility is that individualized optimal primary levels maximize the DPOAE levels by emphasizing the contribution from the "distortion" (f_2) source at the ear canal, which might yield a comparatively smaller "reflection" component and thus reduce the magnitude of the fine structure. If this were the case, then DPOAE I/O curves measured with individualized optimal primary levels may reflect more closely the underlying BM I/O curve at the f_2 cochlear site, and hence improve the match with its corresponding behaviorally inferred I/O curves.

In summary, the fine structure cannot be dismissed as an explanation for the poor correlation between DPOAE and behaviorally-inferred BM I/O curves at low frequencies, but the use of individualized DPOAE optimal primary levels would likely improve the correspondence between the results of the two methods.

This study is a re-analysis of the data shown in Chapter 2 and 4 aimed at testing this possibility for a fixed primary frequency ratio of $f_2/f_1 = 1.2$. An attempt is made to minimize fine structure effects by averaging DPOAE I/O curves for a number of f_2 frequencies around the test frequency of interest. If successful, such an approach would be more time consuming than DPOAE I/O curve measurement procedures with standard parameters but still advantageous over behavioral methods for estimating individualized, frequency-specific BM I/O curves. Furthermore, it could be put in practice with current advanced DPOAE clinical devices.

5.2 Methods

5.2.1 Approach

The approach involved within-subject comparisons of DPOAE I/O curves for optimal stimuli with BM I/O curves inferred behaviorally from temporal masking curves (TMCs).

A TMC is a graphical representation of the level of a pure tone forward masker required to just mask a fixed, low-level pure tone probe, as a function of the time interval between the masker and the probe. It is assumed that the slope of a TMC reflects the rate of increase of the BM response to the masker at the BM place tuned to the probe frequency *and* the post-cochlear decay rate of the internal masker effect (Nelson *et al.,* 2001). There is evidence that the latter is approximately constant across masker frequencies and over a wide range of masker levels (Lopez-Poveda and Alves-Pinto 2008; Wojtczak and Oxenham 2009). Hence, approximate BM I/O curves may be inferred by plotting the levels of a linear reference TMC (i.e., the TMC for a masker that evokes a linear BM response) against the levels for the TMC for the frequency of interest paired according to time interval (Nelson *et al.,* 2001).

Individualized optimal DPOAE stimuli may be found empirically; that is, by searching combinations of primary frequencies and levels that maximize the level of the $2f_1$-f_2 DPOAE (e.g., Kummer *et al.,* 2000; Johnson *et al.,* 2006a). Optimal stimuli differ across individuals and test frequencies. In the present study, the primary frequency ratio was fixed at $f_2/f_1 = 1.2$. Therefore, the terms “optimal DPOAE stimuli” and “optimal DPOAE level rule” are used here to refer to the individualized combination of primary levels, L_1 and L_2, that evokes the maximal level of the $2f_1$-f_2 DPOAE component for any test frequency, f_2.

It has been conjectured that primary levels are optimal when both primaries evoke comparable responses in the f_2 BM region (Kummer *et al.,* 2000). We have provided support for this conjecture by comparing empirical optimal levels with levels of equally effective primaries as inferred from TMCs (Chapter 4). We

measured TMCs for two masker frequencies, f_{m1} and f_{m2}, equal to the DPOAE primary frequencies ($f_{m1} = f_1$ and $f_{m2} = f_2$) for a probe frequency (f_P) equal to the DPOAE test frequency ($f_P = f_2$). Based on the TMC interpretation explained above, the levels of two equally effective pure tones were inferred by plotting the levels of the f_{m1} masker, L_{m1}, against those of the f_{m2} masker, L_{m2}, paired according to time interval. The resulting L_{m1}-L_{m2} functions overlapped reasonably well with corresponding plots of empirical DPOAE optimal L_1-L_2 levels for the same subject, thus supporting the conjecture of Kummer *et al.* (2000). The overlap was, however, restricted to L_2 levels below ~65 dB SPL.

The previous Chapters 2 and 4 did not address the main question of the present study but contained a great deal of the data necessary to do it. Therefore, for convenience, the approach here consisted of reanalyzing the data from Chapters 2 and 4 and from Lopez-Poveda and Johannesen (2010) with the aim of testing if the correspondence between behavioral and DPOAE estimates of BM I/O curves improves when DPOAEs are measured with individualized optimal primary levels. For completeness, the present study extended to DPOAE I/O curves measured using individualized TMC-based primary level rules [from Chapter 4] and the group-average level rule of Kummer *et al.* (1998) [from Chapter 2].

The absolute thresholds for the maskers and the probes can be found in Table 2.1 and 4.1. Linear reference (f_m=0.4f_p) and on-frequency TMCs ($f_m = f_p$) for subjects S1-S10 were shown in Figure 2.1. The linear reference TMCs for subjects S11-S15 have not been shown before. On-frequency TMCs ($f_m = f_p$) for subjects S11-S15 were taken from Chapter 4. Behavioral I/O curves have only been shown before for subjects S1-S10 in Figure 2.4 to 2.7 and are reproduced here for completeness (Figs. 5.1-5.3). The TMCs for masker frequencies equal to DPOAE primary frequencies f_1 and f_2 ($f_2 = f_P$ and $f_1 = f_p$/1.2, respectively) were taken from Chapter 4, and were used to infer equal-BM-excitation DPOAE level rules.

DPOAE I/O curves measured with the rule of Kummer *et al*. can be found in Chapter 2; TMC-based DPOAE primary level rules in Chapter 4; individualized optimal primary level rules for 1 and 4 kHz in Chapter 4; individualized optimal

primary level rules for 0.5 kHz in Lopez-Poveda and Johannesen (2010). In the case of TMC-based and individualized optimal rules, only *mean* DPOAE I/O curves have been shown previously in Chapter 4.

Methodological details have been amply described in Chapters 2 and 4 and only a brief description is provided here.

5.2.2 Subjects

Fifteen normal-hearing listeners participated in the study. Their ages ranged from 20 to 39 years. All of them had thresholds within 20 dB HL (ANSI, 1996) at the frequencies considered in this study (0.5, 1 and 4 kHz). They are identified here as in Chapters 2 and 4.

5.2.3 TMC stimuli

TMCs were measured for probe frequencies (f_P) of 0.5, 1 and 4 kHz and for masker frequencies (f_2) equal to the f_P. TMCs were also measured for a probe frequency of 4 kHz and a masker frequency of 1.6 kHz. The latter were regarded as the linear reference (Lopez-Poveda and Alves-Pinto, 2008). These TMCs were used to infer BM I/O curves for all probe frequencies (Lopez-Poveda *et al.*, 2003). Additional TMCs were measured for probe frequencies of 0.5, 1 and 4 kHz and for masker frequencies (f_1) equal to $f_P/1.2$. These TMCs, together with those for masker frequencies f_2, were used to infer individualized DPOAE primary rules so that the two primaries evoked equal BM excitation at the f_2 site (see Sec. 5.2.1 and Sec. 5.2.6).

Probe level was fixed at 9 dB sensation level (SL) (i.e., 9 dB above the individual's absolute threshold for the probe).

5.2.4 TMC procedure

Masker levels at threshold were measured using a two-interval, two-alternative, forced-choice, procedure. Feedback was provided to the listener. Masker level was

changed according to a two-up, one-down adaptive procedure to estimate the 71% point on the psychometric function (Levitt 1971). The initial step size was 6 dB. The step size was decreased to 2 dB after three reversals. A total of 15 reversals were measured. Threshold was calculated as the mean of the masker levels for the last 12 reversals. A measurement was discarded if the associated SD exceeded 6 dB. Three threshold estimates were obtained in this way and their mean was taken as the masker level at masked threshold. If the SD of these three measurements exceeded 6 dB, a fourth threshold estimate was obtained and included in the mean.

Listeners were trained in the forward-masking task for several hours; at first with a higher probe level of 15 dB SL, and later with a probe level of 9 dB SL, until performance became stable.

5.2.5 Inferring BM I/O functions from TMCs

BM I/O functions were inferred from TMCs by plotting the masker levels for the linear reference TMC against the levels for a masker equal in frequency to the probe and paired according to masker-probe time interval (Nelson *et al.*, 2001). The linear-reference TMCs were fitted (using a least-squares procedure) with a double exponential function and extrapolated to longer masker-probe time intervals to infer BM I/O functions over the wider possible range of levels.

5.2.6 Inferring DPOAE primary level rules from TMCs

A least-squares procedure was used to fit TMCs for the f_1 and f_2 maskers with an *ad-hoc* function (see Lopez-Poveda *et al.*, 2005). Individualized DPOAE level rules were then obtained by plotting the fitted levels for the f_1 masker against the fitted levels for the f_2 masker paired according to the masker-probe time intervals (Chapter 4). Under the assumption that the post-cochlear recovery from masking is independent of masker frequency, and given that the two masker frequencies were equal to the DPOAE primary frequencies, the resulting curves provided the level relationships for two pure tones of frequencies f_1 and f_2 that evoke comparable

excitation levels at the f_2 region of the BM (for a full justification, see Chapter 4). These will be referred to as TMC-based primary levels.

5.2.7 DPOAE Stimuli

DPOAE I/O curves were obtained by plotting the magnitude (in dB SPL) of the $2f_1-f_2$ DPOAE as a function of the level, L_2, of primary tone f_2. I/O curves were obtained for a fixed primary frequency ratio of $f_2/f_1 = 1.2$ and three different level rules.

- Individualized optimal levels. These were obtained by searching the (L_1, L_2) space to find the value of L_1 that produced the highest DPOAE response level for each value of L_2. L_2 was varied in 5-dB steps within the range 35-75 dB SPL. For each fixed L_2, L_1 was varied in 3-dB steps and the individual's optimal value was found.
- TMC-based levels. These were derived from the TMCs for maskers f_1 and f_2 as explained in Sec. 5.2.6. Recall that with these levels the two primaries are presumed to evoke comparable responses in the f_2 cochlear region (Chapter 4).
- The level rule of Kummer *et al.* (1998): $L_1 = 0.4L_2+39$, with L_1 and L_2 in dB SPL. This group average rule was originally designed for $L_2 \leq 65$ dB SPL, but here it was extrapolated to $L_2 = 75$ dB SPL. When this rule was applied, L_2 ranged from 20 to 75 dB SPL in 5-dB steps, except for $f_2 = 0.5$ kHz for which it ranged from 45 to 75 dB SPL.

DPOAE I/O curves were measured for test frequencies of 0.5, 1 and 4 kHz for all three level rules. Not all rules were applied to all participants. Indeed, at 0.5 kHz, I/O curves for optimal rules were measured for only 3 participants due to time restrictions.

A spectral averaging approach (e.g., Kalluri and Shera, 2001; Mauermann and Kollmeier, 2004) was used in an attempt to reduce variability of the DPOAE I/O

curves due to the fine structure. DPOAE I/O curves were measured for five close f_2 frequencies around the text frequency of interest, and the resulting I/O curves were averaged (details may be found in Chapter 2 and 4). For instance, the final DPOAE I/O curve at 4 kHz was the mean of five I/O functions for f_2=3920, 3960, 4000, 4040, 4080 Hz. When optimal levels rules were considered, the I/O curves of only three adjacent frequencies (e.g., f_2=3960, 4000, 4040 Hz) were averaged due to time constraints.

5.2.8 DPOAE stimulus calibration and system artifacts

DPOAE primary levels were calibrated with a Zwislocki DB-100 coupler for each pair of primary frequencies (f_1, f_2). No further *in situ* adjustment of this calibration was applied.

Instrument artifactual DP responses were controlled for by prolonged measurements in a DB-100 Zwislocki coupler and a plastic syringe with a volume of ~1.5 cc. Tests were performed for high L_2 levels (> 50 dB) and under the same conditions as real ear-canal measurements. Ear-canal measurements were rejected if they were less than 6 dB above the coupler DP response. This is a stricter criterion than commonly used in clinical contexts (for a comprehensive justification see Chapter 2).

5.2.9 DPOAE Procedure

DPOAE measurements were made with an IHS Smart system (with SmartOAE software version 4.52) equipped with an Etymotic ER-10D probe. During the measurements, subjects sat comfortably in a double-wall sound attenuating chamber and were asked to remain as steady as possible.

The probe fit was checked before and after each measurement session. The probe remained in the subject's ear throughout the whole measurement session to minimize measurement variance from altering the position of the probe in the ear canal. DPOAEs were measured for a fixed measurement time ranging from 10 s to

60 s. A DPOAE measurement was considered valid when it was 2SD above the measurement noise floor (defined as the mean level over 10 adjacent frequency bins in the spectrum). When a response did not meet this criterion, the measurement was repeated and the measurement time increased if necessary. The probe remained in the same position during these re-measurements. If the required criterion was not met after successive tries, the measurement point was discarded.

When measuring optimal level rules, system artifacts sometimes occurred for high L_2 levels (70-75 dB SPL) and some of the higher L_1 values. A data point for a certain L_2 level was discarded when an optimal L_1 level could not be found within the range of L_1 levels whose DPOAE responses passed the artifact criterion. In other cases, DPOAE responses passed the artifact criterion for the whole range of L_1 levels but no optimal L_1 value could be found because the instrument limits it to 80 dB SPL. That is, the optimal L_1 would have been almost certainly above 80 dB SPL. In these cases, the true DPOAE response would be higher. These points were noted and included in the correspondence analyses (see below).

5.2.10 Analysis of the correspondence between behavioral and DPOAE I/O curves

The degree of within-subject correspondence between BM I/O curves inferred from TMCs and DPOAEs was assessed by least-squares fitting 3rd-order polynomials to all I/O curves (e.g., Chapter 2). The first derivative of the polynomials was calculated analytically and evaluated for the range of input (behavioral) or L_2 (DPOAEs) levels for which experimental data were available. The similarity between behavioral and DPOAE I/O curves was then assessed by the root mean square (RMS) difference between the *slopes* of the behavioral and DPOAE I/O curves *for a corresponding range of input levels*. Considering the difference between the first derivatives instead of the polynomials themselves has the advantage of accounting for the large disparity between behavioral and DPOAE output levels while preserving the information about the shape of the I/O curves. In

other words, a first-derivative RMS difference of zero indicates I/O curves that may be vertically shifted from each other but are otherwise identical.

Two additional measures were employed to assess the similarity between the degree of BM *compression* suggested by behavioral and DPOAE I/O curves. First, a LR analysis was applied to the minimum value of slope of the 3rd-order polynomials fitted to the I/O curves. Second, a LR analysis was applied to the slope of straight lines fitted by least-squares to I/O curves segments for input (or L_2) levels between 40 and 65 dB SPL. This level range was considered because it typically covered the compressive portion of the I/O curves (Chapter 2).

These LR analyses were applied to the three sets of DPOAE I/O curves obtained with the three primary levels rules considered in the present study (i.e., optimal, TMC-based, and Kummer).

5.3 Results

5.3.1 TMCs

Except for the linear references of subjects S11 to S15, all other TMCs have been reported in Chapters 2 and 4. Detailed interpretations of their characteristics can be found in the Sec. 2.3.1 and 4.3.1.

5.3.2 DPOAE primary level rules

Individualized TMC-based primary level rules were derived from the TMCs for the f_1 and f_2 maskers as described in Sec. 5.2.6. Individualized optimal levels were also found as described in Sec. 5.2.7. These level rules have been reported and described in great detail in Sec. 4.3.3.

5.3.3 DPOAE I/O curves

Figures 5.1-5.3 illustrate DPOAE I/O curves for test frequencies of 0.5, 1 and 4 kHz, respectively. Curves are shown for individualized optimal primary levels

(open triangles), individualized TMC-based primary levels (open circles), and the group-average level rule of Kummer *et al.* (1998) (open squares). Each panel shows the results for a different listener. DPOAE noise levels (crosses) plus two SDs are shown for participant S2 (top-right panels) as a representative example of the noise levels for all other participants. Solid lines illustrate 3rd-order polynomials fitted to the DPOAE I/O curves. Dotted lines illustrate linear responses with a slope of 1 dB/dB. Gray-filled symbols indicate responses whose associated L_1 level could still have been suboptimal (included in fitted curves).

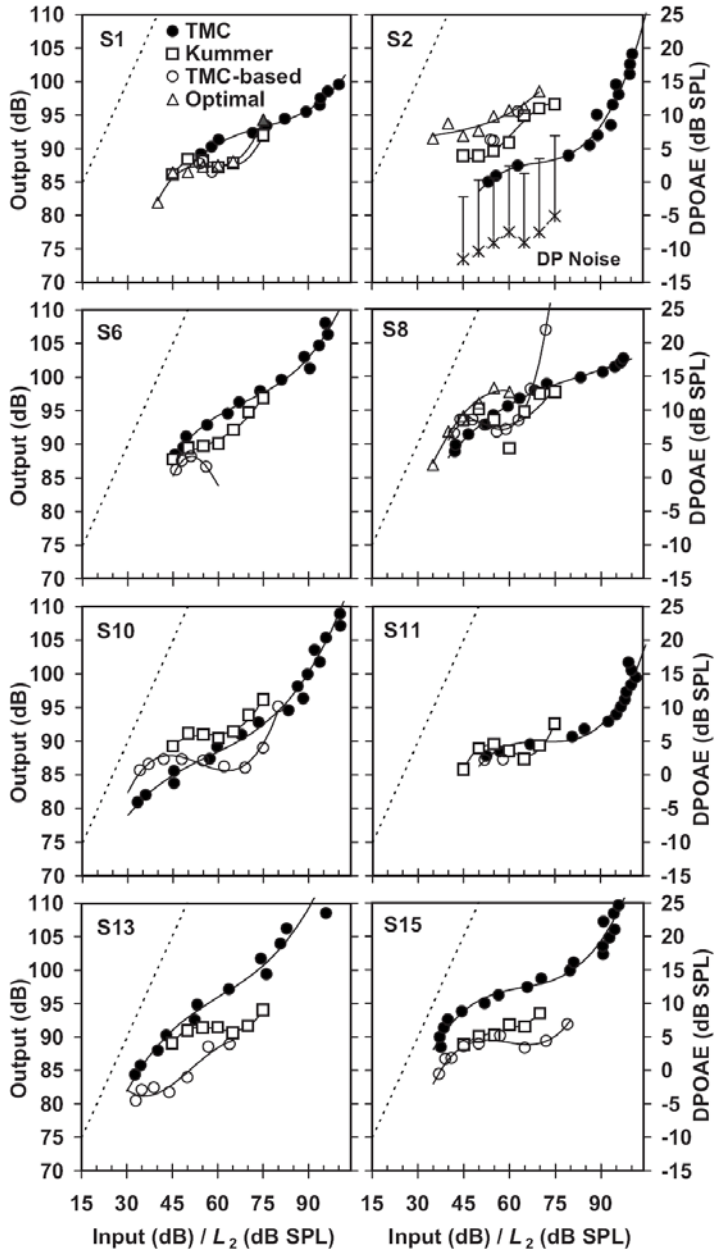

Figure 5.1. Behavioral (filled circles) and DPOAE (open symbols) I/O curves at 0.5 kHz for three different primary level rules, as indicated by the inset. Continuous lines illustrate 3rd-order polynomial fits to the experimental points. Each panel illustrates the result for a different participant. The panel for participant S2 also illustrates the mean DP noise floor and its corresponding 2SDs. Thin dotted lines illustrate a linear response for comparison. Gray symbols illustrate conditions for which primary level L_1 was likely suboptimal because the optimal value would have been higher than the maximum value allowed by the OAE system (80 dB SPL).

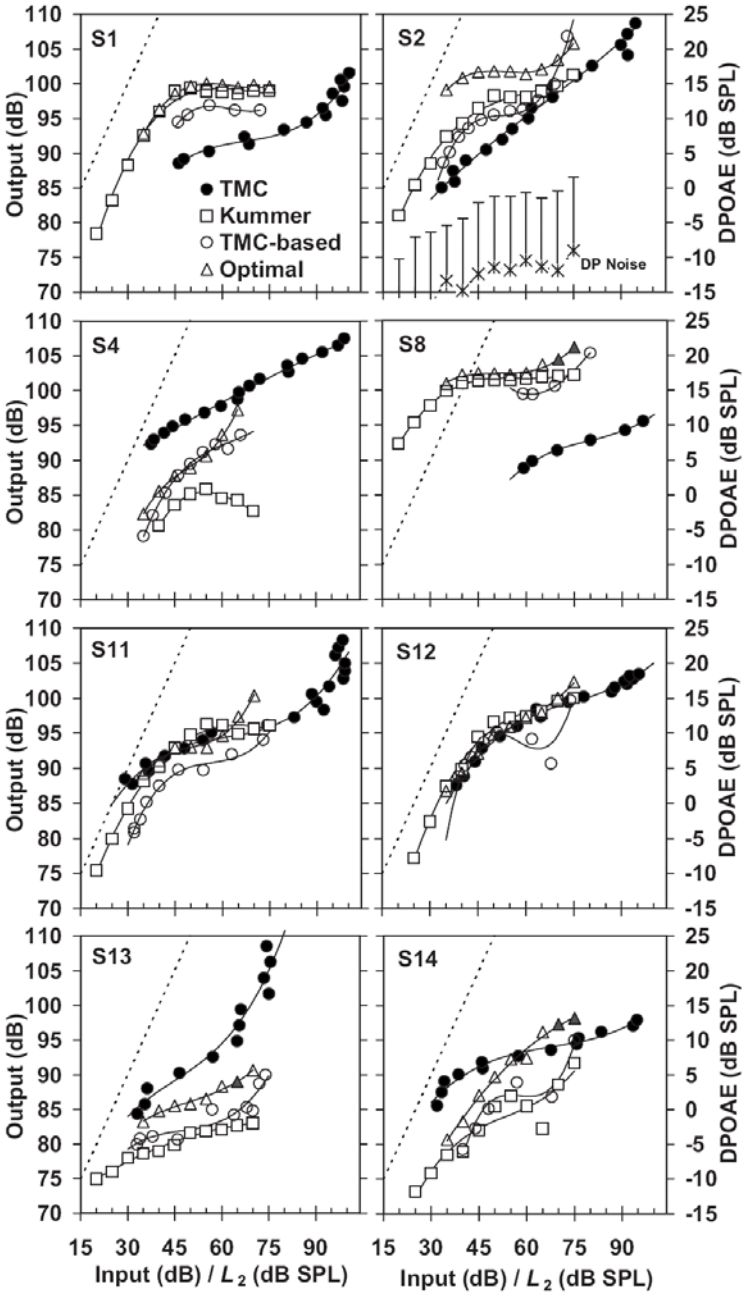

Figure 5.2. As Fig. 5.1 but for a frequency of 1 kHz.

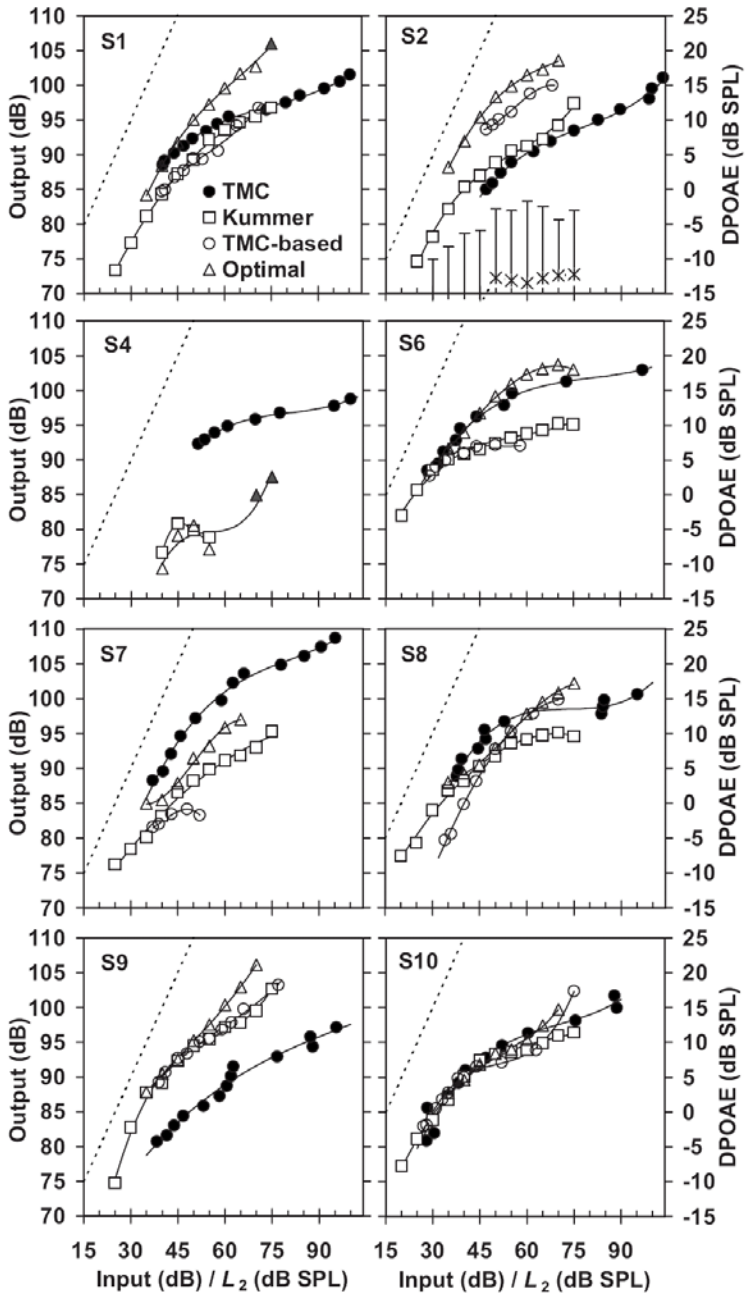

Figure 5.3. As Fig. 5.1 but for a frequency of 4 kHz.

As expected, DPOAE levels for the optimal stimuli were, with very rare exceptions, comparable or higher than DPOAE levels for TMC-based or Kummer stimuli. The very few exceptions (e.g., S11 at 1 kHz and L_2 = 50 dB SPL in Fig.

5.2) were possibly due to changes in the position of the probe across DPOAE measurements for the three level rules.

DPOAE I/O curves could be typically (but not always) described as having a steep segment (approaching linearity) at low L_2 levels, followed by a shallower, segment at midrange levels. Some curves showed a steeper segment at high L_2 levels that approached linearity. This general trend is characteristic of BM I/O curves (e.g., Robles and Ruggero 2001). It is noteworthy that the shallower segments of many curves showed plateaus (i.e., regions with a slope of ~0 dB/dB) or notches (i.e., regions with negative slopes), some of which were very sharp. These features were very common at 0.5 (e.g., S1, S6, S8, S10, S11, S13, S15 in Fig. 5.1) and 1 kHz (e.g., S1, S2, S8, S14 in Fig. 5.2) but virtually nonexistent at 4 kHz. Table 5.1 summarizes the number of I/O curves having plateaus or negative-slope segments for each of the three level rules and test frequencies. Rather surprisingly, notches and plateaus occurred at 0.5 and 1 kHz not only for the group-average level rule of Kummer *et al.* (1998) but also for individualized levels, even for optimal levels (e.g., S1, S2, or S8 at 1 kHz).

Table 5.1. Incidence of plateaus and notches in DPOAE I/O curves across test frequencies and primary level rules. The numbers between brackets indicate the total number of I/O curves measured for each condition.

Frequency (kHz)	0.5	1	4
Kummer rule	5 (8)	5 (8)	2 (8)
TMC-based	5 (8)	4 (8)	2 (8)
Optimal	2 (3)	4 (8)	1 (8)

It is interesting to compare the I/O curves obtained with the three primary level rules. At 0.5 kHz (Fig. 5.1), I/O curves for optimal levels were available for three participants only (S1, S2 and S8) and individualized TMC-based levels often did not cover a sufficient level range to allow a useful comparison. At 1 and 4 kHz (Figs. 5.2 and 5.4), where data points were more numerous, visual inspection

revealed that I/O curves for the three level rules had reasonably similar shapes (e.g. S1, S8, S11, S13 at 1 kHz; S1, S2, S7, S9, S10 at 4 kHz) but with some exceptions (S4, S14 at 1 kHz; S6, S8 at 4 kHz). For most of these exceptions, I/O curves for optimal (open triangles) and TMC-based levels (open circles) generally shared many characteristics and they both differed from the curves measured with Kummer's rule (open squares). Optimal levels and to a lesser extent also TMC-based levels tended to produce a rapid increase of DPOAE levels for levels L_2 > ~65 dB (e.g., S2, S8, S11 at 1 kHz; S9, S10 at 4 kHz), although there were exceptions (S1 at 1 kHz; S6 at 4 kHz). Such rapid increases were much rarer for I/O curves measured with Kummer's rule (S2 at 4 kHz).

5.3.4 Correspondence between behavioral and DPOAE I/O curves

Behavioral I/O curves (filled circles in Figs. 5.1-5.3) showed the same general trend as DPOAE I/O curves, having steep, shallow (compressive), and steep segments at low, moderate, and high input levels, respectively. Unlike for DPOAE I/O curves, however, behavioral I/O curves rarely showed plateaus or notches.

Visual inspection revealed a reasonable correspondence between behavioral and DPOAE I/O curves for the three level rules at 4 kHz (Fig. 5.3). This was expected for the Kummer rule because the present data had already been reported in Chapter 2. The question is whether the match improved for any of the individualized primary levels (optimal or TMC-based). Optimal levels did improve the correspondence sometimes (e.g., S6 and S7). Other times, however, the correspondence decreased (e.g., S8 and S9). TMC-based levels did not improve the correspondence between DPOAE and behavioral I/O curves.

At 0.5 and 1 kHz (Figs. 5.1 and 5.2), however, the correspondence between behavioral and DPOAE I/O curves was disappointing even for DPOAEs measured with individualized optimal or TMC-based levels. With a few exceptions (S11 and S12 at 1 kHz in Fig. 5.2), behavioral I/O curves were strikingly different from all

DPOAE I/O curves. The difference could generally be attributed to the above mentioned plateaus and notches that only occurred for DPOAE I/O curves.

Figure 5.4 provides quantitative support to the preceding qualitative description. It illustrates RMS differences between the slopes of DPOAE and behavioral I/O curves (see Sec. 5.2) against test frequencies, with primary level rules as the parameter (as indicated by the inset). Each open symbol illustrates the RMS difference for a given participant. Filled symbols illustrate mean values across participants. Mean differences (filled symbols) were statistically identical across test frequencies and rules. Mean differences tended to be smaller for the Kummer and optimal level rules at 4 kHz than at 0.5 kHz (unpaired *t*-test, $0.05 < p < 0.10$). Indeed, the mean difference for all three rules tended to decrease with increasing frequency. The *range* of individual differences was comparable for DPOAE I/O curves measured with the rule of Kummer *et al.* (1998) and the optimal rule and they were always narrower than the TMC-based rule.

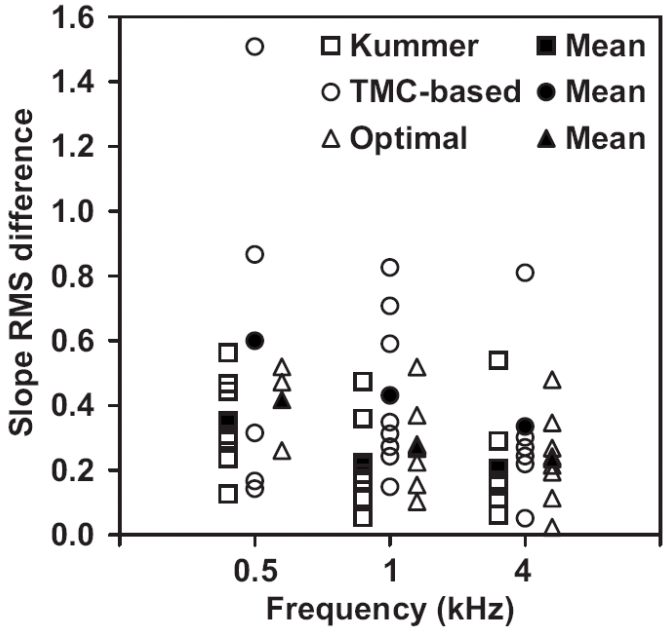

Figure 5.4. Root mean square differences between the slope of behavioral and DPOAE I/O curves for different primary level rules, as indicated by the inset. Open and filled symbols illustrate individual and mean values across listeners, respectively.

5.3.5 Correspondence of compression estimates

Some authors regard the slope of I/O curves over their nonlinear (compressive) segments as a useful description of the physiological status of the cochlea (reviewed by Robles and Ruggero 2001). Others suggest, by contrast, that the compression exponent does not change with amount of hearing loss (Plack *et al.*, 2004). In any case, given that DPOAE I/O curves sometimes show plateaus and/or notches, which may be absent in their behavioral counterparts, the slope of I/O curves over their compressive segments constitutes a useful variable for assessing their correspondence between behavioral and DPOAE I/O curves. Slopes have been often calculated as the average value over the I/O curve compressive region (e.g., Lopez-Poveda *et al.*, 2003; Plack *et al.*, 2004). Sometimes, however, the slope changes smoothly over the compressive region (e.g., Nelson *et al.*, 2001) and so the minimum, instead of the average, slope is used (e.g., Chapter 2). Another

consideration is that average slope may prove less sensitive to sharp notches, like those present in DPOAE I/O curves. Here, the two descriptors (minimum and average slope values) were considered for assessing the correspondence between DPOAE and behavioral I/O curves over their compressive regions (see Sec. 5.2).

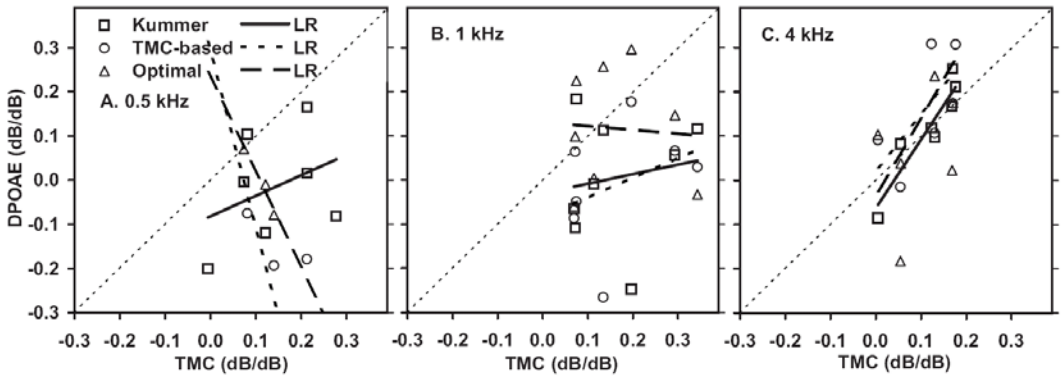

Figure 5.5. Correspondence between minimum compression-exponent estimates obtained from 3rd-order polynomials fitted to behavioral and DPOAE I/O curves at 0.5, 1 and 4kHz. Each point represents data for one subject and one level rule, as indicated by the inset in panel A. Thick lines show LR fits to the data points for each of the stimulus rules. Diagonal thin-dotted lines illustrate perfect correspondence.

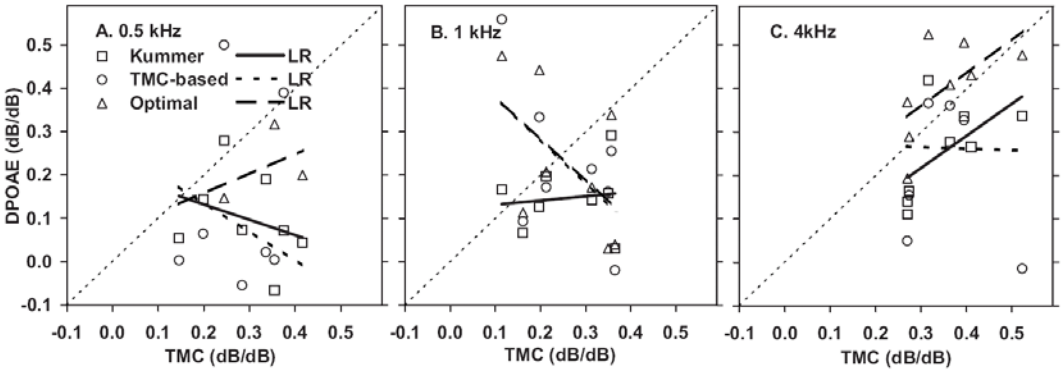

Figure 5.6. As Fig. 5.5 but for the average slope of behavioral and DPOAE I/O curves over their compressive region (input levels from 40 to 65 dB SPL).

Figure 5.5 and 5.6 show, respectively, plots of minimum and average slopes of DPOAE I/O curves against corresponding slopes for behavioral curves. Each panel is for a different test frequency. Each point illustrates results for a single

participant. Different symbols illustrate results for DPOAEs measured with different levels rules, as indicated by the inset in the top panels. Thick lines illustrate LR functions fitted by least squares to the experimental data points, as indicated by the inset. The diagonal represents perfect correspondence (illustrated by thin dotted lines). Tables 5.2 and 5.3 summarize the results of LR analysis on these compression estimates. The correspondence can be regarded as good when the LR slope (*a*) and intercept (*b*) are close to one and zero, respectively, and the *F*-statistics allows rejecting the null-hypothesis (there is no statistical association between the two data sets).

Correspondence between DPOAE and behavioral compression estimates was clearly poor at 0.5 and 1 kHz regardless of the compression descriptor (i.e., minimum or average slope) or the level rule used to measure DPOAEs. For both compression descriptors (Tables 5.2 and 5.3), the LR slope and intercept were far from 1 and 0, respectively, as would be expected if perfect correlation occurred between behavioural and DPOAE results.

Table 5.2. Results of LR analysis between the minimum slopes of 3rd-order polynomials fitted to behavioral and DPOAE I/O curves for the three level rules. *a*: slope; *b*: intercept; r^2: predicted variance; *N*: number of data points; *p* = probability of observed regression occurring by chance (i.e., $p < 0.05$ indicates a statistically-significant regression); n.a.: not applicable (LR statistics were not calculated if $N < 4$).

Frequency			*0.5 kHz*		
Level rule	*a*	*b*	r^2	*N*	*p*
Kummer	0.47	-0.08	0.13	7	0.42
TMC-based	-3.99	0.29	0.38	4	0.39
Optimal	n.a.	n.a.	n.a.	3	n.a.
			1 kHz		
Kummer	0.22	-0.03	0.03	8	0.70
TMC-based	0.45	-0.09	0.13	7	0.43
Optimal	-0.09	0.13	0.00	8	0.87
			4 kHz		
Kummer	1.55	-0.06	0.86	7	0.0025
TMC-based	1.25	0.02	0.43	6	0.16
Optimal	1.72	-0.03	0.27	8	0.18

Table 5.3. As Table 5.2, but for LR analysis between the slopes of straight lines fitted to the compressive segments of behavioral and DPOAE I/O curves (Fig. 5.6).

Frequency			*0.5 kHz*		
Level rule	*a*	*b*	r^2	*N*	*p*
Kummer	-0.35	0.20	0.10	8	0.45
TMC-based	-0.67	0.27	0.06	8	0.57
Optimal	n.a.	n.a.	n.a.	3	n.a.
			1 kHz		
Kummer	0.10	0.12	0.01	8	0.78
TMC-based	-1.01	0.48	0.33	8	0.13
Optimal	-0.97	0.48	0.31	8	0.15
			4 kHz		
Kummer	0.74	0.00	0.36	8	0.11
TMC-based	-0.03	0.28	0.00	7	0.98
Optimal	0.78	0.12	0.38	8	0.11

At 4 kHz, reasonably high correspondence was observed for the minimum slope estimate (Fig. 5.5) for DPOAEs measured with the Kummer rule (slope = 1.55; intercept = -0.06; r^2 = 0.86; p < 0.01). Correspondence was lower for individualized level rules and was not significant (Table 5.2). For the average slope estimates (Fig. 5.6), there was moderate correspondence for DPOAEs measured with the Kummer rule and the optimal rule, but it did not reach significance (both p = 0.11). The two individualized rules (optimal and TMC-based) yielded reasonably similar minimum compression estimates (slope = 1.72; intercept = -0.08; r^2 = 0.89; p < 0.005; data not shown). This is not surprising given that both individualized rules were very similar, presumably because they reflect levels of equally effective primaries at the f_2 cochlear site (Chapter 4).

5.4 Discussion

The long-term goal of the present research is to define DPOAE stimuli and conditions that would allow using DPOAEs as a universal and faster alternative to behavioral methods to infer individualized, frequency-specific BM I/O curves. In Chapter two it was shown that this is possible using typical DPOAE stimuli (specifically, the rule of Kummer *et al*. and a frequency ratio of $f_2/f_1 = 1.2$) for test frequencies above ~2 kHz but not for lower frequencies (Chapter 2). The main reason for the lack of correspondence at these lower frequencies seemed to be the presence of notches and plateaus in DPOAE I/O curves that did not have a counterpart in behavioral curves. The present study aimed at testing whether the correspondence between the I/O curves inferred with the two methods would improve by using individualized optimal primary levels. Two individualized level rules have been considered: one optimized to maximize DPOAE levels; and one intended for primaries to evoke comparable BM responses at the f_2 cochlear site. It has been shown that neither approach improves the correspondence between behavioral and DPOAE I/O curves at low frequencies with respect to the group average levels of Kummer *et al.* (1998). Indeed, it has been shown that plateaus and notches are equally common in DPOAE I/O curves for individualized and group-average DPOAE primary levels at 0.5 and 1 kHz (Table 5.1).

The similar morphology of DPOAE I/O curves measured with different primary levels (Figs. 5.1-5.3) suggests a lower dependency on individualized levels than expected (Chapter 3). For example, all DPOAE I/O curves suggested similar compression thresholds, as well as similar slopes below and above the compression threshold (Fig. 5.1-5.3). This was true for most participants (the only exceptions were S4 at 1 kHz, and S6 and S8 at 4 kHz). However, the tendency to a rapid increase in the DPOAE response at high levels (L_2 > ~70 dB SPL) seemed peculiar to individualized (optimal and TMC-based) levels. Such a tendency was rarely observed in I/O curves measured with the rule of Kummer *et al.* (1998). The slight influence of the level rule on slope may also explain the rather surprising result that

the match between behavioral and DPOAE I/O curves was closest for the Kummer rule even though it is a group-average estimate that disregards individual idiosyncrasies. Despite their modest influence on the shape of DPOAE I/O curves, individualized primary levels may nevertheless be important to maximize the DPOAE absolute levels, and thus to maximize the sensitivity of DPOAEs as a potential clinical diagnostic tool (Kummer *et al.*, 2000).

5.4.1 The causes of low-frequency plateaus of notches

It has been shown that plateaus and notches are equally common in DPOAE I/O curves for individualized and group-average DPOAE primary levels at 0.5 and 1 kHz (Table 5.1). The present study was not intended to provide an explanation for the low-frequency notches and plateaus, but the present results, in combination with existing evidence, provide interesting insights. As argued in the Introduction to this Chapter, plateaus and notches could be due to the fine structure and the use of suboptimal DPOAE stimuli. Further, it was conjectured that these two explanations need not be independent of each other in so far as the fine structure changes with changing primary levels (He and Schmiedt, 1993). However, the present data do not support the relationship between the two explanations. In principle, varying DPOAE primary levels could change the relative contribution from the various DP sources (as measured in the ear canal) and hence the magnitude of the fine structure. It was hypothesized in the Introduction that individualized primary levels could maximize DPOAE levels by emphasizing the contribution from the "distortion" (f_2) source relative to that of the "reflection" ($2f_1–f_2$) source, which would contribute to reduce the fine structure magnitude. The present results show that the incidence and magnitude of notches and plateaus is not reduced by using individualized optimal primary levels and so they do not support the conjecture in question. If anything, the present data actually suggest that the magnitude of the contributions from the "reflection" ($2f_1–f_2$) source is proportional to that of the "distortion" (f_2) source.

5.5 Conclusions

(1) For a fixed primary frequency ratio of f_2/f_1 = 1.2, DPOAE levels vary moderately depending on primary level rule but the fundamental morphology of DPOAE I/O curves hardly changes.

(2) The incidence of plateaus and notches in DPOAE I/O curves was comparable for individualized primary level rules and for the group-average rule of Kummer *et al.* These features were more frequent at low (0.5 and 1 kHz) than at high (4 kHz) frequencies and remain the most likely reason for the low correspondence between behavioral and DPOAE I/O curves at low frequencies.

(3) The correspondence between behavioral and DPOAE I/O curves was reasonably high at 4 kHz. The group average primary level rule of Kummer *et al.* (1998) is sufficient to estimate individualized BM I/O curves at high frequencies.

(4) It is uncertain which of two approaches DPOAEs or behavioral methods is more appropriate to infer individualized BM I/O responses at 0.5 and 1 kHz, but DPOAEs may *not* be used as an alternative to behavioral methods, even considering individualized optimal primary levels.

CHAPTER 6

GENERAL DISCUSSION

This thesis was intended to test the main hypothesis that behavioural (TMC) and DPOAE methods can be used indistinctly to infer human cochlear I/O curves. The approach was to compare I/O curves as measured by DPOAE and inferred from TMCs in normal hearing listeners. The results suggest that normal hearing human cochlear I/O curves estimated with TMC and DPOAE are mutually consistent for frequencies above ~2 kHz but less so for lower frequencies. The results also suggest that the match between I/O curves obtained with both methods hardly improves at low frequencies by using DPOAE primaries individually optimized for maximum distortion. It remains uncertain which of the two methods, if any, best reflects the BM I/O curve at low frequencies. Several factors may influence the estimated I/O curves of both methods and can contribute to the mismatch between I/O curves obtained with both methods. These are discussed in the following Sections.

6.1 TMC method assumptions

The TMC method of inferring behavioral BM I/O curves relies on several assumptions (see Sec. 5.2 and 2.4). The majority of them have received experimental support for a wide range of conditions (e.g., Chapter 4; Lopez-Poveda and Alves-Pinto 2008; Wojtczak and Oxenham 2009). Nevertheless, behavioral curves cannot be taken as an undisputed “golden standard” (e.g., Stainsby and Moore, 2006; Wojtczak and Oxenham, 2009), particularly at low frequencies where

there is a lack of correspondence between behavioral and DPOAE I/O curves (Chapter 2).

The TMC method is an indirect, psychophysical method, thus its results may be influenced by retro-cochlear mechanisms unknown to date. Indeed, there exist within-subject differences between I/O functions inferred with different psychophysical methods (e.g., Rosengard *et al.*, 2005). One of the main assumptions of the TMC method is that the rate of decay of the internal (post-compression) masker effect is identical across frequencies and levels (Nelson *et al.*, 2001; Lopez-Poveda *et al.*, 2003). The validity of this assumption is still controversial. Stainsby and Moore (2006) have argued that the decay rate is faster for low probe frequencies or equivalently, low CFs, at least for hearing-impaired listeners. By contrast, Lopez-Poveda and Alves-Pinto (2008) have provided indirect evidence for frequency-independent decay rates, at least for normal hearing listeners. Additionally, there is evidence that for any given frequency, the decay rate is slower at high levels (Lopez-Poveda and Alves-Pinto, 2008; Wojtczak and Oxenham, 2007). These issues complicate the selection of the linear reference TMC and thus cast doubts on the corresponding I/O curves, particularly at low frequencies (Stainsby and Moore, 2006; Lopez-Poveda and Alves-Pinto, 2008). If both concerns were true, the compression exponent would be higher than suggested by the present data (Fig. 2.8), particularly for low frequencies, and this would increase the difference between compression exponent estimates inferred from TMCs and DPOAEs (Fig. 2.8, 5.5 and 5.6).

6.2 Efferent system

Activation of the ipsilateral efferent reflex may affect the I/O curves inferred from both TMCs and from DPOAEs. The efferent system can be activated by ipsilateral and/or contralateral sounds above around 40 dB SL (Hood *et al.* 1996) and has an activation latency in the range 40-80 ms (Backus and Guinan, 2006; Roverud and

Strickland, 2010). At least for a contralateral elicitor, the efferent effect is strongest for low frequencies (Lilaonitkul and Guinan, 2009; Lopez-Poveda *et al.*, 2013; Aguilar *et al.*, 2013), which is coincidentally also the frequency region where the discrepancy between TMC and DPOAE I/O curves is largest. The masker of the TMC stimuli is often above this activation threshold and has a sufficient duration (100 ms) to activate the efferent system. The activation of the efferent system reduces cochlear gain and linearizes the I/O curve (Cooper and Guinan, 2006; Lopez-Poveda *et al.*, 2013). If the efferent system was activated during the present TMC measurements, the "true" I/O curves would be more compressive than the present data suggest (Fig. 2.8, 5.5 and 5.6) and the difference between compression exponent estimates inferred from TMCs and DPOAEs would decrease (Fig. 2.8, 5.5 and 5.6).

The higher L_1-L_2 levels of the DPOAE stimuli are also above the efferent activation threshold and the stimulus is long enough (400 ms) to activate the efferent system. The DPOAE response level is typically reduced ~2 dB by the efferent effect (Abdala *et al.* 2009) and this suppression effect would likely increase gradually with the stimulus level. The DPOAE I/O response magnitude would therefore be reduced only at higher stimulus levels and this would decrease the compression exponent estimates. If the efferent system was activated during the present DPOAE measurements, then the "true" I/O curve would be less compressive than suggested by the present DPOAE data and the difference between compression exponent estimates inferred from TMCs and DPOAEs would decrease (Fig. 2.8, 5.5 and 5.6). Given the moderate magnitude of the efferent effect on the DPOAEs, the efferent effect might contribute to explain the discrepancies at low frequencies between TMC and DPOAE data but seems insufficient to explain all the difference.

The efferent system can also change the magnitude of the distortion and reflection components of the DPOAE response and therefore affect DPOAE fine structure (Abdala *et al.* 2009). Their results suggest that fine structure magnitude

decrease due to the efferent effect but the magnitude of the changes (~2 dB) seems too small to have a significant contribution.

6.3 Suboptimal DPOAE primary frequency ratio

In Section 2.4.2 and 5.1 it was suggested that the a poor correspondence at low frequencies between I/O curves inferred from TMCs and DPOAEs could be due to suboptimal primary levels, which was supported by the simulations in Chapter 3. Another aspect of this suboptimal parameter explanation, however, is that a fixed primary frequency ratio of $f_2/f_1 = 1.2$ has been used in the present study based on early evidence that it maximizes (on average) DPOAE levels (Gaskill and Brown 1990). According to more recent reports, however, the optimal f_2/f_1 ratio increases slightly with decreasing f_2 frequency and with increasing L_2 level (Johnson *et al.*, 2006a), particularly for low frequencies. To the best of the authors' knowledge, the optimal ratio for a test frequency of 0.5 kHz is yet to be determined. There is evidence, however, that human cochlear processing in apical BM regions may be significantly different from that of basal zones (e.g. Lopez-Poveda *et al.*, 2003; Plack *et al.*, 2004), which suggests that the optimal frequency ratio at 0.5 kHz could differ from 1.2. Therefore, a better correspondence between behavioral and DPOAE I/O curves might have been obtained by considering not only optimal primary levels but also optimal primary frequencies. On the other hand, variations of the f_2/f_1 ratio seem to have only a small effect on the slope of average I/O curves below 65 dB SPL (e.g., Fig. 3 of Johnson *et al.*, 2006a), at least in the frequency range from 1 to 8 kHz. This casts doubts that optimizing the frequency ratio would improve the correspondence between behavioral and DPOAE I/O curves, but the benefit it could have on an individual basis is yet to be investigated.

6.4 Mutual suppression of DPOAE stimuli

There is one fundamental difference between the behavioral and the DPOAE stimulus paradigms: the stimuli of the behavioral TMC method are non-simultaneous whereas measuring DPOAEs, requires presenting the two primaries f_1 and f_2 *simultaneously*. One possible cause for difference between TMC-based and DPOAE I/O curves is hence that the f_1 primary may have suppressed the BM response to the f_2 primary at the f_2 site. Indeed, Rhode (2007) measured BM excitation and DPOAEs simultaneously in the same preparation and showed that the f_1 primary suppresses the BM response to the f_2 with a fixed level of 60 dB SPL for L_1 above 60 dB SPL (see his Fig. 1). As a result, DPOAE I/O curves may not correspond directly to single-tone BM I/O curves, as is commonly assumed.

Unlike DPOAE I/O curves, the I/O curves inferred from TMCs would not be affected by suppression because the masker and the probe tones were not presented simultaneously, in fact, this is one of the reasons that the TMC method is so widely used to infer BM I/O curves. Therefore, one might think that mutual suppression between the primary tones may have influenced DPOAE but not TMC-based I/O curves and that the effect would be more pronounced at low CFs because the nonlinear effects extend to a wider bandwidth (Rhode and Cooper, 1996; Lopez-Poveda *et al.*, 2003). This explanation, however, is unlikely to account for the low correlation between compression estimates obtained with the two methods at low frequencies. There is physiological and psychophysical evidence that suppression leads to I/O curves steeper than single-tone I/O curves (Nuttall and Dolan, 1993; Rhode, 2007; Yasin and Plack, 2007). This is true particularly for suppressor/suppressee combinations similar to the primary-tone combinations used here. If suppression had affected the DPOAE I/O curves, they should indicate less compression than TMC-based I/O curves and this has been found *not* to be the case (Fig. 2.8). Therefore, a more likely explanation for the low correlation at low CFs between I/O curves estimates obtained with the two methods is the presence of

notches and plateaus in the DPOAE I/O curves, which occur more frequently at low CFs (Figs. 2.4–2.7, Table 2.2).

6.5 Secondary DPOAE sources at high stimulus level

A first possible explanation for the notches and plateaus in the DPOAE I/O curves might be that another DP generation mechanism starts playing a role at high stimulus levels and notches reflect destructive interference between the DPs generated by this "new" high-level source and the normal source (see Mills, 1997). The measurement system is unlikely to be the source in question because a rather strict exclusion criterion was applied in the present study to eliminate system generated DPs. Liberman *et al.* (2004) showed that genetically modified mice without the necessary prestin protein to drive OHC electromotility still generated attenuated DP responses and thus supported the existence of a second possible DP-generation mechanism at high levels. Some of their DPOAE I/O curves showed notches at similar stimulus levels to the notches found in the current study. On the other hand, Avan *et al.* (2003) attributed low and high stimulus level DPOAEs to the same nonlinear mechanism. Also, Lukashkin *et al.* (2002) and Lukashkin and Russell (2002) have shown that a single saturating nonlinearity is sufficient to explain a notch in a DP I/O function (see also Fig. 3.1).

6.6 DPOAE fine structure

A further explanation for the plateaus and notches at low frequencies is that some of the DPOAE I/O curves still could have been influenced by the fine structure despite the precautions taken to minimize its effects. Plateaus occur for levels around 45–50 dB SPL (e.g., S1 and S8 at 1 kHz in Fig. 2.5; S3 at 4 kHz in Fig. 2.7), where the fine structure certainly can have an influence. It is, however, unclear if

this is also the explanation for the notches because they always occurred at moderate-to-high levels (60–70 dB SPL) and the fine structure has a higher influence at low levels (Mauermann and Kollmeier, 2004).

Arguments have been given in the Introduction to Chapter 5 that the fine-structure explanation is possible but may not be sufficient. The arguments provided were based on existing evidence for test frequencies equal to or greater than 2 kHz. The importance of the fine structure for frequencies below 1 kHz is still uncertain but the present results suggest that it could more significant than was thought at the outset of the present study. If this were the case, it would explain why the spectral averaging method employed here to minimize the fine structure seemed more efficient for high (4 kHz) than for low (0.5 and 1 kHz) test frequencies. The reason for this is uncertain. If the interference between the DP contributions from various cochlear regions (Shera and Guinan, 2008) was more pronounced and less sensitive to primary levels and f_2 frequencies for low than for high frequency DPOAEs, this might explain the relatively higher incidence of plateaus and notches at low frequencies and for all level rules considered. By extension, it might also explain the observed discrepancies between behavioral and DPOAE I/O curves. It is uncertain why destructive interference would be more pronounced or less sensitive to primary levels at low than at high frequencies. Interestingly, however, both OHC disorganization (e.g., Lonsbury-Martin *et al.,* 1988) and compression bandwidth appear comparatively greater for apical than for basal cochlear regions (e.g., Rhode and Cooper 1996; Lopez-Poveda *et al.*, 2003; Plack and Drga 2003). As a result, the generation region of DPOAEs by both “reflection” and “distortion” mechanisms is likely broader in the apex than in the base and hence potential interactions between DPOAE contributions from adjacent regions could be more significant for low test frequencies. This, however, is only a conjecture, whose detailed formulation and verification remain to be developed.

6.7 Ideas for future work

One important direction of future research should be aimed at testing the common notion that the DPOAE I/O curves reflect the BM I/O curves. Chapter 3 provided only crude evidence so one possibility could be to simulate BM I/O curves using a realistic BM model or conduct physiological experiments where DPOAE I/O curves and BM I/O curves are measured simultaneously.

DPOAE fine structure turns out to be a likely reason for the lack of correlation between DPOAE I/O curves and I/O curves inferred from TMCs at low frequency in normal hearing subjects. The studies described in Chapters 2-5 applied a simple method to reduce the influence of fine structure caused by the interference between "distortion" and "reflection" DPOAE sources. There do, however, exist better methods to minimize the contribution from the "reflection" source but they were not employed partly due to practical issues but also due to reasons explained as follows. A good method for avoiding the "reflection" source contribution is probably the inverse FFT paradigm (Stover *et al*., 1996). The major drawback of this method is that it is extremely time consuming, as it requires data points with a very high frequency resolution (~20 Hz) and that would need to be repeated for each L_2 stimulus level. It is therefore considered impractical for the present purpose.

Another way to minimize the effect of DPOAE fine structure would be the suppression method, where a third tone is placed close to the frequency of the "reflection" source ($2f_1$-f_2) and suppresses the contribution from the "reflection" source. It is as fast as the normal DPOAE procedure. The major problem with this approach is that it is difficult to set the correct suppressor tone level as it varies a lot between individuals (Johnson *et al*., 2006): if it is too low, the "reflection" source is not suppressed and if it is too high, it also suppresses the "distortion" source (f_2), which would affect the I/O curve. Despite this problem, it might be a good alternative as the suppressor level could be fixed at a relative low level (e.g 50 dB SPL) where it would still suppress somewhat the "reflection" component but without altering the "distortion" component.

Lastly, a swept primary paradigm has been suggested (e.g. Long *et al*., (2008) that separates the "distortion" and "reflection" components by time windowing and that supposedly is not time consuming. It would be interesting to investigate if the correspondence between TMC-based and DPOAE I/O curves would improve at low frequencies by measuring DPOAE using either the suppression or the swept method.

It might also be worthwhile investigating the influence of the optimal DPOAE stimulus parameters at low frequencies (f_2 < 1 kHz). In fact, optimal f_2/f_1 ratio and L_1-L_2 level rules are yet to be defined for low frequencies. This would be interesting also for clinical purposes.

There exist other psychophysical methods for inferring cochlear I/O curves like the growth of masking method (Oxenham and Plack 1997; Rosengard *et al*. 2005) or the variant TMC method (Lopez-Poveda and Alves-Pinto, 2008). The I/O curves inferred from the GOM method might correspond better to DPOAE I/O curves when the BM excitation shifts basally at high stimulus levels as discussed in Section 4.4. The variant TMC method does not rely on the assumptions of the linear reference condition discussed in Section 6.1, but on different assumptions (Lopez-Poveda and Alves-Pinto, 2008). An interesting project could therefore be to compare individual DPOAE I/O curves with variant TMC and/or GOM inferred cochlear I/O curves (Lopez-Poveda and Johannesen, 2010).

Also the present TMC method might be improved in at least two directions. First, by minimizing the eventual impact of the efferent system on the inferred I/O curve as discussed in Section 6.2 (Yasin *et al.*, 2013). Second, by reducing the measurement time and the need for training to make the TMC method more applicable in clinical contexts at least for cooperative subjects.

One limitation of the present work is that it only includes normal hearing listeners. Evidently, it should be extended to hearing impaired (HI) subjects and in fact we are currently undertaking similar studies in HI subjects. It is uncertain if I/O curves inferred from TMCs and DPOAEs would be more or less consistent with each other for HI listeners. The shape of average DPOAE I/O curves (Dorn *et al.*,

2001) for different degrees of hearing loss (supposedly attributed to OHC loss) is incompatible with the model of BM I/O curves for HI listeners by Plack *et al.* (2004) which suggested that the compression threshold increases with increasing hearing loss and that the compression exponent in the compressive segment is unaltered. Therefore the match between I/O curves inferred from TMCs and DPOAEs above ~2 kHz found in the present study might not apply to HI listeners. On the other hand, this does not preclude finding a good correlation between e.g. cochlear gain loss and some metric based on DPOAE I/O curves. Probably the DPOAE fine structure remains a problem also for HI subjects (He and Schmiedt, 1996). The fine structure is reduced when the "reflection" source ($2f_1$-f_2) coincides with an impaired frequency region (Mauermann *et al.*, 1999) but the opposite can be the case when the "distortion" source (f_2) coincides with an impaired region and therefore the influence of fine structure will depend on hearing loss configuration (Konrad-Martin *et al.*, 2002).

The long-term aim is to devise a method to estimate BM I/O curves in a clinically feasible way (Müller and Janssen, 2004; Lopez-Poveda *et al.*, 2009). The hypothesis is that, utilizing BM I/O curve information, hearing aid fitting could be made more individualized and the number of fitting sessions reduced and/or that the hearing aid performance could be improved.

CONCLUSIONS

The present studies aimed at adapting DPOAE measurement procedures with a view to using DPOAE I/O curves as a faster, reliable alternative to behavioral TMC methods for inferring cochlear I/O curves. The main conclusions are:

- TMC-based cochlear I/O curves and DPOAE I/O curves are mutually consistent for frequencies above around 2 kHz. Given that the two methods are based on different assumptions, this suggests that both sets of I/O curves depend on the same underlying mechanism and probably reflect the "true" BM I/O curve.
- The poorer correspondence between TMC-based and DPOAE I/O curves at lower frequencies may be due to pitfalls in the assumptions of either method and/or to using suboptimal stimuli.
- The poorer correspondence at lower frequencies is, nonetheless, associated to the occurrence of notches and plateaus in the DPOAE I/O curves. These features are unlikely due to using population average rather than individualized optimal DPOAE primary levels, even though simulations suggest the opposite. The experimental results may be thus caused by DPOAE fine structure.
- Maximal DPOAE levels occur when the two primaries tones produce approximately identical BM excitation at the cochlear site tuned to f_2.

REFERENCES

Abdala, C., Mishra, S. K. and Williams, T. L. (**2009**) "Considering distortion product otoacoustic emission fine structure in measurements of the medial olivocochlear reflex," J. Acoust. Soc. Am. **125**, 1584-1594.

Aguilar, E., Eustaquio-Martin, A. and Lopez-Poveda, E. A. (**2013**) "Contralateral efferent reflex effects on threshold and suprathreshold psychoacoustical tuning curves at low and high frequencies," J. Assoc. Res. Otolaryngol. **14**, 341-57.

ANSI (**1996)**. *S3.6 Specification for Audiometers* American National Standards Institute, New York.

Avan, P., Bonfils, P., Gilain, L., and Mom, T. (**2003)**. "Physiopathological significance of distortion-product otoacoustic emissions at 2f1-f2 produced by high-versus low-level stimuli," J. Acoust. Soc. Am. **113**, 430–441.

Backus, B. C., and Guinan, J. J., Jr., (**2006**). "Time-course of the human medial olivocochlear reflex," J. Acoust. Soc. Am. **119**, 2889–2904.

Bacon, S. P., and Oxenham, A. J. (**2004)**. "Psychophysical manifestations of compression: Normal-hearing listeners," in *Compression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper Springer, New York, Chap. 3, pp. 107–152.

Bacon, S. P. (**2004)**. "Overview of auditory compression," in *Compression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper Springer, New York, Chap. 1, pp. 1–17.

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

Cooper, N. P., and Guinan, J. J., Jr., (**2006**). "Efferent-mediated control of basilar membrane motion," J. Physiol. **576.1**, 49-54.

Cooper, N. P., and Rhode, W. S. (**1997)**. "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," J. Neurophysiol. **78**, 261–270.

Cooper, N. P. (**2004)**. "Compression in the peripheral auditory system," in C*ompression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper Springer, New York, Chap. 2, pp. 18–61.

Dhar, S., and Schaffer, 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**, 573–85.

Dorn, P. A., Konrad-Martin, D., Neely, S. T., Keefe, D. H., Cyr, E., and Gorga, M. P. (**2001)**. "Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears," J. Acoust. Soc. Am. **110**, 3119–3131.

Duifhuis, H. (**1973)**. "Consequences of peripheral frequency selectivity for non-simultaneous masking," J. Acoust. Soc. Am. **54,** 1471–1488.

Fahey, P. F., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (**2000)**. "Nonlinear interactions that could explain distortion product interference response areas," J. Acoust. Soc. Am. **108**, 1786–1802.

Gaskill, S. A., and Brown, A. M. (**1990)**. "The behavior of the acoustic distortion product, 2f1-f2, from the human ear and its relation to auditory sensitivity," J. Acoust. Soc. Am. **88**, 821–839.

Gaskill, S. A., and Brown, A. M. (**1996)**. "Suppression of human acoustic distortion product: Dual origin of 2f1-f2," J. Acoust. Soc. Am. **100**, 3268–3274.

Goldstein, J. L. (**1967)**. „Auditory nonlinearity," J. Acoust. Soc. Am. **41**, 676–689.

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

Gorga, M. P., Neely, S. T., Dierking, D. M., Kopun, J., Jolkowski, K., Groenenboom, K., Tan, H., and Stiegemann, B. (**2007)**. "Low-frequency and high-frequency cochlear nonlinearity in humans," J. Acoust. Soc. Am. **122**, 1671–1680.

Gorga, M. P., Neely, S. T., Dierking, D. M., Kopun, J., Jolkowski, K., Groenenboom, K., Tan, H., and Stiegemann, B. (**2008)**. "Low-frequency and high-frequency distortion product otoacoustic emission suppression in humans," J. Acoust. Soc. Am. **123**, 2172–2190.

He, N.-J., and Schmiedt, R. A. (**1993)**. "Fine structure of the 2*f*1-*f*2 acoustic distortion product: Changes with primary level," J. Acoust. Soc. Am. **94**, 2659–2669.

He, N.-J., and Schmiedt, R. A. (**1996)**. "Effects of aging on the fine structure of the 2*f*1-*f*2 acoustic distortion product," J. Acoust. Soc. Am. **99**, 1002-1015.

He, N.-J., and Schmiedt, R. A. (**1997)**. "Fine structure of the 2*f*1-*f*2 acoustic distortion product: Effects of primary level and frequency ratios," J. Acoust. Soc. Am. **101**, 3554–3565.

Heinz, M. G., and Young, E. D. (**2004)**. "Response growth with sound level in auditory-nerve fibers after noise induced hearing loss," J. Neurophysiol. **91**, 784–795.

Heitmann, J., Waldmann, B., Schnitzler, H., Plinkert, P. K., and Zenner, H. (**1998)**. "Suppression of distortion product otoacoustic emissions DPOAE near 2f1-f2 removes DP-gram fine structure-Evidence for a secondary generator," J. Acoust. Soc. Am. **103**, 1527–1531.

Hood, L. J., Berlin, C. I., Hurley, A., Cecola, R. P., Bell, B. (**1996**). "Contralateral suppression of transient-evoked otoacoustic emissions in humans: intensity effects," Hear. Res. **101**:113-118.

Janssen, T., and Müller, J. (**2008)**. "Otoacoustic emissions as a diagnostic tool in a clinical context," in *Active Processes and Otoacoustic Emissions*, edited by G.

A. Manley, R. R. Fay, and A. N. Popper Springer, New York, Chap. 13, pp. 421–460.

Johannesen, P. T., and Lopez-Poveda, E. A. (**2008)**. "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion product otoacoustic emissions," J. Acoust. Soc. Am. **124**, 2149–2163.

Johannesen, P. T., and Lopez-Poveda, E. A. (**2010)**. "Correspondence between behavioral and individually "optimized" otoacoustic emission estimates of human cochlear input/output curves," J. Acoust. Soc. Am. **127**, 3602-3613.

Johnson, T. A., Neely, S. T., Garner, C. A., and Gorga, M. P. (**2006a)**. "Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions," J. Acoust. Soc. Am. **119**, 418–428.

Johnson, T. A., Neely, S. T., Kopun, J. G., and Gorga, M. P. (**2006b)**. "Reducing reflected contributions to ear-canal distortion product otoacoustic emissions in humans," J. Acoust. Soc. Am. **119**, 3896–3907.

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

Kemp, D. T. (**1978)**. "Stimulated acoustic emissions from within the human auditory system," J. Acoust. Soc. Am. **64**, 1386–1391.

Konrad-Martin, D., Neely, S. T., Keefe, D. H., Dorn, P. A., Cyr., E., and Gorga, M. P. (**2002).** "Sources of DPOAEs revealed by suppression experiments, inverse fast Fourier transforms, and SFOAEs in impaired ears," J. Acoust. Soc. Am. **111**, 1800-1809.

Kummer, P., Janssen, T., and Arnold, W. (**1995)**. "Suppression tuning characteristics of the 2f1-f2 distortion product otoacoustic emission in humans," J. Acoust. Soc. Am. **98**, 197–210.

Kummer, P., Janssen, T., and Arnold, W. (**1998)**. "The level and growth behavior of the 2F1-F2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," J. Acoust. Soc. Am. **103**, 3431–3444.

Kummer, P., Janssen, T., Hulin, P., and Arnold, W. (**2000)**. "Optimal L1-L2 primary tone level separation remains independent of test frequency in humans," Hear. Res. **146**, 47–56.

Levitt, H. (**1971)**. "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R., and Coats, A. C. (**1988)**. "Spontaneous otoacoustic emissions in nonhuman primate. II. Cochlear anatomy," Hear. Res. **33**, 69–93.

Liberman, M. C., Zuo, J., and Guinan, J. J., Jr. (**2004)**. "Otoacoustic emissions without somatic motility: Can stereocilia mechanics drive the mammalian cochlea?," J. Acoust. Soc. Am. **116**, 1649–1655.

Lilaonitkul, W., and Guinan, J. J., Jr. (**2009**). "Human Medial Olivocochlear Reflex: Effects as Functions of Contralateral, Ipsilateral, and Bilateral Elicitor Bandwidths," J. Assoc. Res. Otolaryngol. **10**, 459-470.

Lonsbury-Martin, B. L., and Martin, G. K. (**1990)**. "The clinical utility of distortion product otoacoustic emissions," Ear Hear. **11**, 144–154.

Long, G., Talmadge, C. L., and Lee, J. (**2008)**. "Measuring distortion product otoacoustic emissions using continuously sweeping primaries," J. Acoust. Soc. Am. **124,** 1613-1626.

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (**2003)**. "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," J. Acoust. Soc. Am. **113**, 951–960.

Lopez-Poveda, E. A., Plack, C. J., Meddis, R., and Blanco, J. L. (**2005)**. "Cochlear compression in listeners with moderate sensorineural hearing loss," Hear. Res. **205**, 172–183.

Lopez-Poveda, E. A., and Alves-Pinto, A. (**2008)**. "A variant temporal masking-curve method for inferring peripheral auditory compression," J. Acoust. Soc. Am. **123**, 1544–1554.

Lopez-Poveda, E. A., Johannesen, P. T., and Merchán, M. A. (**2009)**. "Estimation of the degree of inner and outer hair cell dysfunction from distortion product otoacoustic emission input/output functions," Audiological Med. **7**, 22–28.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2009)**. "Otoacoustic emission theories and behavioral estimates of human basilar membrane motion are mutually consistent," J. Assoc. Res. Otolaryngol. **10**, 511–523.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2010)**. "Otoacoustic emission theories can be tested with behavioral methods," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, R. Meddis, and A. R. Palmer Springer, New York, pp. 3–14.

Lopez-Poveda, E. A., Aguilar, E., Johannesen, P. T. and Eustaquio-Martín, A. (**2013**) "Contralateral efferent regulation of human cochlear tuning: behavioural observations and computer model simulations," Adv. Exp. Med. Biol. **787**, 47-54.

Lukashkin, A. N., and Russell, I. J. (**1998)**. "A descriptive model of the receptor potential nonlinearities generated by the hair cell mechanoelectrical transducer," J. Acoust. Soc. Am. **103**, 973–980.

Lukashkin, A. N., and Russell, I. J. (**2001)**. "Origin of the bell-like dependence of the DPOAE amplitude on primary frequency ratio," J. Acoust. Soc. Am. **110**, 3097–3106.

Lukashkin, A. N., Lukashkina, V. A., and Russell, I. J. (**2002)**. "One source for distortion product otoacoustic emissions generated by low- and high level primaries," J. Acoust. Soc. Am. **111**, 2740–2748.

Lukashkin, A. N., and Russell, I. J. (**2002)**. "Modifications of a single saturating non-linearity account for post-onset changes in 2f1-f2 distortion product otoacoustic emission," J. Acoust. Soc. Am. **112**, 1561–1568.

Martin, G. K., Jassir, D., Stagner, B. B., Whitehead, M. L., and Lonsbury-Martin, B. L. (**1998)**. "Locus of generation for the 2f1–f2 vs 2f2–f1 distortion product otoacoustic emissions in normal-hearing humans revealed by suppression

tuning, onset latencies, and amplitude correlations," J. Acoust. Soc. Am. **103**, 1957–1971.

Mauermann, M., Uppenkamp, S., van Hengel, P. W., and Kollmeier, B. (**1999)**. "Evidence for the distortion product frequency place as a source of distortion product emission DPOAE fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio *f*2/*f*1," J. Acoust. Soc. Am. **106**, 3473–3483.

Mauermann, M., and 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.

Mauermann, M., and Kollmeier, B. (**2004)**. "Distortion product otoacoustic emission DPOAE input/output functions and the influence of the second DPOAE source," J. Acoust. Soc. Am. **116**, 2199–2212.

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001)**. "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.

Meddis, R., and O'Mard, L. P. (**2005)**. "A computer model of the auditory nerve response to forward-masking stimuli," J. Acoust. Soc. Am. **117**, 3787–3798.

Mills, D. M., and Rubel, E. W. (**1994)**. "Variation of distortion product otoacoustic emissions with furosemide injection," Hear Res. **77**, 183–199.

Mills, D. (**1997)**. "Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones," J. Acoust. Soc. Am. **102**, 413–429.

Moore, B. C., and Glasberg, B. R. (**1983)**. „Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise," J. Acoust. Soc. Am. **73**, 1249–1259.

Moore, B. C. J. (**2003)**. *An Introduction to the Psychology of Hearing*, 5th ed. Academic, London.

Moore, B. C. J. (**2007)**. *Cochlear Hearing Loss*, 2nd ed. Wiley, Chichester, UK.

Müller, J., and Janssen, T. (**2004)**. "Similarity in loudness and distortion product otoacoustic emission input/output functions: Implications for an objective hearing aid adjustment," J. Acoust. Soc. Am. **115**, 3081–3091.

Neely, S. T., Gorga, M. P., and Dorn, P. A. (**2003)**. "Cochlear compression estimates from measurements of distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **114**, 1499–1507.

Neely, S. T., Johnson, T. A., and Gorga, M. P. (**2005)**. "Distortion-product otoacoustic emission measured with continuously varying stimulus level," J. Acoust. Soc. Am. **117**, 1248–1259.

Neely, S. T., Johnson, T. A., Kopun, J., Dierking, D. M., and Gorga, M. P. (**2009)**. "Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears," J. Acoust. Soc. Am. **126**, 728–738.

Nelson, D. A., Schroder, A. C., and Wojtczak, M. (**2001)**. "A new procedure for measuring peripheral compression in normal-hearing and hearing impaired listeners," J. Acoust. Soc. Am. **110**, 2045–2064.

Nelson, D. A., and Freyman, R. L. (**1987)**. "Temporal resolution in sensorineural hearing-impaired listeners," J. Acoust. Soc. Am. **81**, 709–720.

Nelson, D. A., and Schroder, A. C. (**2004)**. "Peripheral compression as a function of stimulus level and frequency region in normal-hearing listeners," J. Acoust. Soc. Am. **115**, 2221–2233.

Nuttall, A. L., and Dolan, D. F. (**1993)**. "Two-tone suppression of inner hair cell and basilar membrane responses in the guinea pig," J. Acoust. Soc. Am. **93**, 390–400.

Oxenham, A. J., and Moore, B. C. J. (**1994)**. "Modeling the additivity of nonsimultaneous masking," Hear. Res. **80**, 105–118.

Oxenham, A. J., and Moore, B. C. (**1995)**. "Additivity of masking in normally hearing and hearing-impaired subjects," J. Acoust. Soc. Am. **98**, 1921–1934.

Oxenham, A. J., Moore, B. C., and Vickers, D. A. (**1997)**. "Short-term temporal ntegration: evidence for the influence of peripheral compression," J. Acoust. Soc. Am. **101**, 3676–3687.

Oxenham, A. J., and Plack, C. J. (**1997)**. "A behavioral measure of basilar membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666–3675.

Oxenham, A. J., and Bacon, S. P. (**2003)**. "Cochlear compression, perceptual measures and implications for normal and impaired hearing," Ear Hear. **24**, 352–366.

Oxenham, A. J., and Bacon, S. P. (**2004)**. "Psychophysical manifestations of compression: Hearing-impaired listeners," in *Compression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper Springer, New York, Chap. 4, pp. 62–106.

Plack, C. J., and Oxenham, A. J. (**1998)**. "Basilar-membrane nonlinearity and the growth of forward masking," J. Acoust. Soc. Am. **103**, 1598–1608.

Plack, C. J., and Oxenham, A. J. (**2000)**. "Basilar-membrane nonlinearity estimated by pulsation threshold," J. Acoust. Soc. Am. **107**, 501–507.

Plack, C. J., and Drga, V. (**2003)**. "Psychophysical evidence for auditory compression at low characteristic frequencies," J. Acoust. Soc. Am. **113**, 1574–1586.

Plack, C. J., Drga, V., and Lopez-Poveda, E. A. (**2004)**. "Inferred basilar membrane response functions for listeners with mild to moderate sensorineural hearing loss," J. Acoust. Soc. Am. **115**, 1684–1695.

Rhode, W. S., and Cooper, N. P. (**1996)**. "Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo," Aud. Neurosci. **3**, 101–121.

Rhode, W. S. (**2007)**. "Distortion product otoacoustic emissions and basilar membrane vibration in the 6–9 kHz region of sensitive chinchilla cochleae," J. Acoust. Soc. Am. **122**, 2725–2737.

Robles, L., and Ruggero, M. A. (**2001)**. "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1997)**. "Two-tone distortion on the basilar membrane of the chinchilla cochlae," J. Neurophysiol. **77**, 2385-2399.

Rosengard, P. S., Oxenham, A. J., and Braida, L. D. (**2005)**. "Comparing different estimates of cochlear compression in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **117**, 3028–3041.

Roverud, E., and Strickland, E. A. (**2010**) "The time course of cochlear gain reduction measured using a more efficient psychophysical technique," J. Acoust. Soc. Am. **128,** 1203-1214.

Ruggero, M. A. (**1993)**. „Distortion in those good vibrations. Curr. Biol. **3**, 755–758.

Ruggero, M. A., Rich, N. C.; Recio, A., Narayan, S. S., and Robles, L. (**1997)**. "Basilar membrane responses to tones at the base of the chinchilla cochlea," J. Acoust. Soc. Am. **101**, 2151–2163.

Shaffer, L. A., Withnell, R. H., Dhar, S., Lilly, D. J., Goodman, S. S., and Harmon, K. M. (**2003)**. "Sources and mechanisms of DPOAE generation: implications for the prediction of auditory sensitivity," Ear Hear. **24**, 367–379.

Shera, C. A., and 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.

Shera, C. A., and Guinan, J. J, Jr. (**2007)**. Cochlear traveling-wave amplification, suppression, and beamforming probed using noninvasive calibration of intracochlear distortion sources. J. Acoust. Soc. Am. **121**, 1003–1016.

Shera, C. A., and Guinan, J. J., Jr. (**2008)**. "Mechanisms of mammalian otoacoustic emissions," in *Active Processes and Otoacoustic Emissions*, edited by G. A. Manley, R. R. Fay, and A. N. Popper Springer, New York, Chap. 9, pp. 305–342.

Siegel, J. H. (**1994)**. "Ear-canal standing waves and high frequency calibration using otoacoustic emission probes," J. Acoust. Soc. Am. **95**, 2589–2597.

Stainsby, T. H., and Moore, B. C. J. (**2006)**. "Temporal masking curves for hearing-impaired listeners," Hear. Res. **218**, 98–111.

Stover, J. S., Neely, S. T., and Gorga, M. P. (**1996)**. "Latency and multiple sources of distortion product otoacoustic emissions," J. Acoust. Soc. Am. **99**, 1016–1024.

Talmadge, C., Tubis, A., Long, G. R., and Pisorski, P. (**1998)**. "Modeling otoacoustic emission and hearing threshold fine structure in humans," J. Acoust. Soc. Am. **104**, 1517–1543.

Talmadge, C. L., Long, G. R., Tubis, A., and 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.

Whitehead, M. L., Stagner, B. B., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (**1995)**. "Dependence of distortion-product otoacoustic emissions

on primary levels in normal and impaired ears. II. Asymmetry in L1, L2 space," J. Acoust. Soc. Am. **97**, 2359–2377.

Whitehead, M. L., Stagner, B., Lonsbury-Martin, B. L., and Martin, G. K. (**1995)**. "Effects of ear-canal standing waves on measurements of distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **98**, 3200–3214.

Williams, E. J., and Bacon, S. P. (**2005)**. "Compression estimates using behavioral and otoacoustic emission measures," Hear. Res. **201**, 44–54.

Wojtczak, M., and Oxenham, A. J. (**2007)**. "Verification of the assumption of frequency-independent recovery from forward masking," J. Acoust. Soc. Am. **121**, 3133.

Wojtczak, M., and Oxenham, A. J. (**2009)**. "Pitfalls in behavioral estimates of basilar-membrane compression in humans," J. Acoust. Soc. Am. **125**, 270–281.

Yasin, I., and Plack, C. J., (**2003)**. "The effect of a high-frequency suppressor on tuning curves and derived basilar-membrane response functions," J. Acoust. Soc. Am. **114**, 322–332.

Yasin, I., and Plack, C. J., (**2007)**. "The effect of low- and high-frequency suppressors on psychophysical estimates of basilar-membrane compression and gain," J. Acoust. Soc. Am. **121**, 2832–2841.

Yasin, I., Drga, V., and Plack, C. J., (**2013)**. "Improved psycophysical methods to estimate peripheral gain and compression," Adv. Exp. Med. Biol. **787**, 39-46.

APPENDIX A

PUBLICATIONS AND CONFERENCE COMMUNICATIONS

Publications

- Johannesen, P. T., and Lopez-Poveda, E. A. (**2008)**. "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **124**, 2149–2163.
- Lopez-Poveda, E. A., and Johannesen, P. T. (**2009)**. "Otoacoustic emission theories and behavioral estimates of human basilar membrane motion are mutually consistent," J. Assoc. Res. Otolaryngol. **10**, 511–523.
- Johannesen, P. T., and Lopez-Poveda, E. A. (**2010)**. "Correspondence between behavioral and individually "optimized" otoacoustic emission estimates of human cochlear input/output curves," J. Acoust. Soc. Am. **127**, 3602-3613.

Published conference abstracts

- Johannesen, P. T., and Lopez-Poveda, E. A. (**2008)**. "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions," 155th Meeting of the Acoustical Society of America, Journal of the Acoustical Society of America **123**, 3855.

Conference communications

- Johannesen, P. T., and Lopez-Poveda, E. A. (**2008)**. "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions," poster, 155th Meeting of the Acoustical Society of America, Acoustics´08, Paris, France.
- Lopez-Poveda, E. A. and Johannesen, P. T. (**2008)**. "A psychophysical test of the Kummer primary-level rule for measuring distortion-product otoacoustic emission input/output functions," poster, 155th Meeting of the Acoustical Society of America, Acoustics´08, Paris, France.
- Lopez-Poveda, E. A. and Johannesen, P. T. (**2009)**. "Otoacoustic emissions theories can be tested with behavioral methods," oral communication, 15th Meeting of the International Symposium on Hearing, Salamanca, Spain.
- Johannesen, P. T. and Lopez-Poveda, E. A. (**2009)**. "Estimating residual compression using otoacoustic emissions," oral communication, III Annual meeting "Jornadas de Audiologia", Salamanca, Spain.
- Lopez-Poveda, E. A. and Johannesen, P. T. (**2009)**. "Psychoacoustical verification of current theories on the generation of human distortion product otoacoustic emissions," oral communication, 3rd Iberian Conference on Perception, Guimaraes, Portugal.
- Lopez-Poveda, E. A. and Johannesen, P. T. (**2010)**. "Towards designing a clinical method for estimating basilar membrane input/output response characteristics in listeners with normal and impaired hearing," poster, 33rd Midwinter meeting of the Association for Research in Otolaryngology, Anaheim California, USA.
- Lopez-Poveda, E. A. and Johannesen, P. T. (**2012)**. "Estimates of inner and outer hair cell loss from behavioral and otoacoustic emission measures," poster, 35th Midwinter meeting of the Association for Research in Otolaryngology, San Diego, California, USA.

APPENDIX B

REPRINTS OF PUBLISHED ARTICLES

Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions

Peter T. Johannesen and Enrique A. Lopez-Poveda[a)]
Unidad de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, 37007 Salamanca, Spain

(Received 28 April 2008; revised 14 July 2008; accepted 14 July 2008)

The aim was to investigate the correlation between compression exponent, compression threshold, and cochlear gain for normal-hearing subjects as inferred from temporal masking curves (TMCs) and distortion-product otoacoustic emission (DPOAEs) input–output (I/O) curves. Care was given to reduce the influence of DPOAE fine structure on the DPOAE I/O curves. A high correlation between compression exponent estimates obtained with the two methods was found at 4 kHz but not at 0.5 and 1 kHz. One reason is that the DPOAE I/O curves show plateaus or notches that result in unexpectedly high compression estimates. Moderately high correlation was found between compression threshold estimates obtained with the two methods, although DPOAE-based values were around 7 dB lower than those based on TMCs. Both methods show that compression exponent and threshold are approximately constant across the frequency range from 0.5 to 4 kHz. Cochlear gain as estimated from TMCs was found to be ~16 dB greater at 4 than at 0.5 kHz. In conclusion, DPOAEs and TMCs may be used interchangeably to infer precise individual nonlinear cochlear characteristics at 4 kHz, but it remains unclear that the same applies to lower frequencies.
 [DOI: 10.1121/1.2968692]

PACS number(s): 43.64.Jb, 43.66.Dc, 43.64.Ri, 43.64.Bt [BLM] Pages: 2149–2163

I. INTRODUCTION

It is almost certain that our ability to perceive sounds over a 120 dB level range is accomplished via a form of compression that takes place in the basilar membrane (BM) (Bacon, 2004). Listeners with sensorineural hearing loss show reduced auditory dynamic ranges and this is typically interpreted as an indication of reduced cochlear compression. It is controversial, however, that this is actually the case (e.g., Heinz and Young, 2004; Plack *et al.*, 2004; Lopez-Poveda *et al.*, 2005) and yet knowing the degree of residual peripheral compression might help in diagnosing the type of hearing loss as well as improving its treatment with hearing aids. The long-term goal of this work is to design a fast and reliable technique to estimate the degree of *individual* residual compression in listeners with sensorineural hearing loss. This paper describes a first effort to determine whether distortion-product (DP) otoacoustic emission (DPOAE) input/output (I/O) functions could be used for that purpose in normal-hearing listeners. The motivation of the present work is thus similar to that of Müller and Janssen (2004).

There exist several psychoacoustical methods to estimate the amount of peripheral compression in humans [reviewed in Bacon (2004)]. Of these, the temporal masking curve (TMC) method of Nelson *et al.* (2001) (see also Lopez-Poveda *et al.*, 2003) is perhaps the most accurate because it minimizes off-frequency listening effects that might occur when the probe level is varied. This method consists of measuring the level of a tonal forward masker required to just mask a *fixed* tonal probe as a function of the time interval between the masker and the probe. A TMC is a graphical representation of the resulting masker levels against the corresponding masker–probe intervals. Because the probe level is fixed, the masker level increases with increasing masker–probe time interval and hence TMCs have positive slopes. Nelson *et al.* (2001) argued that the slope of any given TMC depends simultaneously on the amount of BM compression affecting the masker at a cochlear place whose characteristic frequency (CF) equals approximately the probe frequency and on the rate of decay of the internal (postcochlear) masker effect. By assuming that the decay rate is the same across masker frequencies, BM I/O functions may be estimated by plotting the masker levels of a linear reference TMC (i.e., the TMC for a masker that is processed linearly by the BM) against the levels for any other masker frequency, paired according to masker–probe delays [Nelson *et al.* (2001) provide a full justification of these assumptions; see also Lopez-Poveda and Alves-Pinto (2008)].

The TMC method has been used to infer BM I/O curves in normal-hearing and hearing-impaired listeners in a number of studies (e.g., Lopez-Poveda *et al.*, 2003, 2005; Lopez-Poveda and Alves-Pinto, 2008; Nelson *et al.*, 2001; Nelson and Schroder, 2004; Plack *et al.*, 2004; Rosengard *et al.*, 2005). It is arguably a reliable method but time consuming and requires the active participation of the listener. This would make it inconvenient for clinical purposes, particularly for testing newborns and the elderly.

On the other hand, DPOAE I/O curves share many characteristics with BM I/O functions (Cooper and Rhode, 1997; Dorn *et al.*, 2001; Neely *et al.*, 2003). Specifically, both of

a)Author to whom correspondence should be addressed. Electronic mail: ealopezpoveda@usal.es

them are generally linear at low levels but become compressive above a certain compression threshold (Dorn *et al.*, 2001; Kummer *et al.*, 1998) and both of them are similarly labile to outer hair cell damage (Rhode, 2007). Furthermore, measuring DPOAEs does not require the active participation of the listener. This suggests that DPOAE I/O functions could be used as a faster and universal way to infer individual BM I/O functions in clinical conditions (Müller and Janssen, 2004). The aim of the present study is to test this hypothesis by measuring the degree of correlation between BM I/O functions as inferred from TMCs and from DPOAEs in the same subject. Clearly, only if the results of the two methods correlate well will it be possible to add support to the use of DPOAEs to infer *individual* cochlear response characteristics. If they disagree, however, it will be difficult to resolve which of the two methods is more appropriate to reveal the true nonlinear characteristics of the underlying BM responses. Although the long-term goal is to extend the study to hearing-impaired listeners, the focus here is on normal-hearing listeners.

Several earlier studies have addressed this or related questions. Müller and Janssen (2004) investigated the similarity of loudness and DPOAE I/O curves in the same subject sample [Neely *et al.* (2003) had done it previously using different subject samples and slightly different methods]. They found a high resemblance between the characteristics (gain and compression) of the two sets of *average* I/O curves in normal-hearing and hearing-impaired listeners. Müller and Janssen (2004) acknowledged, however, that loudness may be affected by retrocochlear mechanisms (see also Heinz and Young, 2004) and it is also thought that loudness is affected by off-frequency effects (e.g., different spreads of excitation at different levels), which make it difficult to establish a one-to-one relationship between loudness and underlying BM I/O curves (Moore, 2003). This undermines the conclusions of Müller and Janssen (2004). Furthermore, their conclusions applied to *average* I/O curves and frequencies of 2–4 kHz, and thus may not be valid individually or for other frequencies, particularly 0.5 and 1 kHz.

Gorga *et al.* (2007) measured the degree of cochlear compression in a very large sample ($N=103$) of normal-hearing listeners as estimated from DPOAE I/O functions at 0.5 and 4 kHz. As a consequence, the I/O functions they reported likely provide a good description of *average* normal responses. Their results supported the conclusion of earlier psychophysical studies that the degree of compression is similar for apical and basal cochlear sites (Lopez-Poveda *et al.*, 2003; Plack and Drga, 2003). However, their study did not include within-subject psychophysical/physiological comparisons.

Williams and Bacon (2005) inferred cochlear I/O curves from TMCs and DPOAEs in four listeners and for frequencies of 1, 2, and 4 kHz. The results revealed that both methods yield similar *average* compression estimates. Like the above-mentioned studies, this study was not intended to investigate within-subject correlations between the results of both methods. Further, their DPOAE I/O curves could have been influenced by the DP fine structure. Indeed, Gaskill and Brown (1990) showed rapid variations (known as "fine structure") of the magnitude of the $2f_1-f_2$ DPOAE with changing the frequencies of the primaries (f_1 and f_2, with $f_2/f_1=1.21$) only slightly. This fine structure is thought to be the result of constructive and destructive interference between DPs generated at two spatially distant sites (Kummer *et al.*, 1995; Stover *et al.*, 1996; Gaskill and Brown, 1996; Heitmann *et al.*, 1998; Talmadge *et al.*, 1998; 1999; Mauermann *et al.*, 1999; Mauermann and Kollmeier, 1999; Shera and Guinan, 1999). The principal generation site is the BM region of maximum overlap between the excitation caused by the two primaries, that is the BM site with CF=f_2 (e.g., Kummer *et al.*, 1995). This component propagates back toward the oval window but also to the cochlear site with CF=$2f_1-f_2$ where it excites a second generation source. The DP generated at this second source propagates back toward the oval window and is summed with the response of the first source. The fine structure is thought to originate from vector summation of these two components, whose varying phases give rise to constructive or destructive interference and thereby to peaks and valleys in the DP-gram. The f_2 generator site is the dominant source at high stimulus levels [Fig. 3 of Mauermann and Kollmeier (2004)], which explains why the fine structure is more pronounced at low levels.

The fine structure has a large influence on DPOAE I/O curves, especially at low levels. He and Schmiedt (1993) mentioned that the DPOAE magnitude can change by as much as 20 dB for a change in f_2 of 1/32 octave. Mauermann and Kollmeier (2004) reported that the response varied by 10–15 dB when f_2 varied over the interval from 2250 to 2610 Hz. The influence of the fine structure is greater for individual than for average (across subjects) I/O curves, but can also affect average curves when the sample size is small. The sample size was small ($N=4$) in the study of Williams and Bacon (2005). Thus, fine-structure effects may have complicated the interpretation of their results or even led to wrong conclusions.

The present report extends these earlier studies in several respects. First, the focus here is on within-subject as opposed to average psychophysical/physiological correlations. Second, psychophysical BM I/O curves were inferred using what is arguably the most accurate method available to date for this purpose [see Nelson *et al.* (2001) for a full justification; but see also Sec. IV of the present paper]. Third, special care was exercised to reduce the influence of the fine structure on individual DPOAE I/O curves by averaging the magnitude of the $2f_1-f_2$ DPOAE for five f_2 frequencies near the frequency of interest. Fourth, the frequency range considered (0.5–4 kHz) included low and high frequencies. Fifth, the physiological/psychophysical comparisons extended to parameters pertaining to cochlear nonlinearity other than compression magnitude; specifically, compression threshold and the level at which maximum compression occurs.

It will be shown that reasonable correlation exists between the characteristics of individual TMC-based and DPOAE I/O curves at 4 kHz but not at 0.5 and 1 kHz. Reasons for the observed discrepancies at low frequencies will be discussed.

TABLE I. Thresholds (in dB SPL) measured with Etymotic ER2 insert earphones for all subjects and for tone durations of 300 ms (absolute threshold), 110 ms (masker threshold), and 10 ms (probe threshold), respectively. n.a. stands for not available.

	Frequency (kHz)			
Tone duration (ms)	0.5 300/110/10	1 300/110/10	2 300/110/10	4 300/110/10
S1	13/15/39	6/8/30	12/16/34	13/13/33
S2	14/21/43	10/15/30	20/22/n.a.	21/21/40
S3	13/16/39	7/8/32	5/8/27	4/10/25
S4	17/20/n.a.	10/16/34	16/21/42	21/24/44
S5	23/24/n.a.	13/17/41	7/12/n.a.	0/3/23
S6	9/10/36	10/11/31	9/13/n.a.	4/4/24
S7	6/10/n.a.	10/11/n.a.	16/20/n.a.	12/15/31
S8	11/18/34	5/7/38	25/28/n.a.	11/10/30
S9	12/n.a./n.a.	9/n.a./n.a.	10/n.a./n.a.	10/15/31
S10	5/n.a./n.a.	−2/n.a./n.a.	10/n.a./n.a.	3/1/23

II. METHOD

A. Subjects

Ten normal-hearing subjects participated in the study. Their age ranged from 20 to 39 years. Their absolute thresholds were measured using a two-down, one-up adaptive procedure. Signal duration was 300 ms, including 5 ms cosine-squared onset and offset ramps. All subjects had thresholds within 20 dB HL at the frequencies considered in this study (0.5, 1, 2, and 4 kHz, see Table I).

B. TMC stimuli

TMCs were measured for probe frequencies (f_P) of 0.5, 1, 2, and 4 kHz and for masker frequencies equal (on-frequency) to the f_P. Additional TMCs were measured for a probe frequency of 4 kHz and a masker frequency of 1.6 kHz (f_M=0.4f_p). The latter were selected as the linear references (Lopez-Poveda and Alves-Pinto, 2008) and used to infer BM I/O curves for all probe frequencies (Lopez-Poveda *et al.*, 2003). The masker–probe time intervals ranged from 5 to 100 ms in 5 ms steps with an additional interval of 2 ms. The duration of the masker was 110 ms, including 5 ms cosine-squared onset and offset ramps. The probe duration was 10 ms, including 5 ms cosine-squared onset/offset ramps and no steady state portion. The level of the probe was fixed at 9 dB sensation level (SL) (i.e., 9 dB above the individual absolute threshold for the probe), except for subject S5 for whom it was 15 dB SL. Stimuli were generated with a Tucker Davies Technologies Psychoacoustics Workstation (System 3) operating at a sampling rate of 48.8 kHz and with analog to digital conversion resolution resolution of 24 bits. If needed, signals were attenuated with a programmable attenuator (PA-5) before being output through the headphone buffer (HB-7). Stimuli were presented to the listeners through Etymotic ER-2 insert earphones. TMC sound pressure levels (SPL) were calibrated by coupling the earphones to a sound level meter through a Zwislocki DB-100 coupler. Calibration was performed at 1 kHz only and the obtained sensitivity was used at all other frequencies because the earphone manufacturer guarantees an approximately flat (±2 dB) frequency response between 200 Hz and 10 kHz.

C. TMC procedure

The procedure was identical to that of Lopez-Poveda and Alves-Pinto (2008). Masker levels at threshold were measured using a two-interval, two-alternative forced-choice paradigm. Two sound intervals were presented to the listener in each trial. One of them contained the masker only and the other contained the masker followed by the probe. The interval containing the probe was selected randomly. The subject was asked to indicate the interval containing the probe. The initial masker level was set sufficiently low that the subject always could hear both the masker and the probe. The masker level was then changed according to a two-up, one-down adaptive procedure to estimate the 71% point on the psychometric function (Levitt, 1971). An initial step size of 6 dB was applied, which was decreased to 2 dB after three reversals. A total of 15 reversals were measured. Threshold was calculated as the mean of the masker levels at the last 12 reversals. A measurement was discarded if the standard deviation (s.d.) of the last 12 reversals exceeded 6 dB. Three threshold estimates were obtained in this way and their mean was taken as the threshold. If the s.d. of these three measurements exceeded 6 dB, a fourth threshold estimate was obtained and included in the mean.

The maximum SPL was set to 104 dB to prevent subject discomfort and/or temporary threshold shifts. A measurement run was stopped and discarded when the subject reached this limit on more than two consecutive trials over the last 12 reversals. Masker levels at threshold were measured for masker–probe time intervals in increasing order. This was done to minimize the possibility that the measurements would be affected by potential temporary thresholds shifts that might have occurred if intervals had been presented in random order and a long interval (high masker level) immediately preceded a short interval. An attempt was made to measure masker levels for all masker–probe time intervals. Missing data indicate that the maximum output level (104 dB SPL) was reached for the time interval in question or that it was impossible within six to ten attempts to obtain three threshold estimates with s.d. ⩽6 dB.

The listeners' absolute threshold for the maskers and probes were measured using the same equipment and conditions used to measure the TMCs. At least three measurements were obtained and averaged. Results are shown in Table I. Listeners were trained in the forward-masking task for several hours, at first with a higher probe level of 15 dB SL and later with a probe level of 9 dB SL, until performance became stable. Listeners sat in a double-wall sound attenuating chamber during all measurements.

D. Inferring BM I/O functions from TMCs

BM I/O functions were inferred from TMCs by plotting the levels for the linear reference TMC against the levels for any other masker frequency paired according to masker–probe time interval (Nelson *et al.*, 2001). The off-frequency

TMC for f_P=4 kHz was used as the linear reference to infer I/O curves for all other frequencies, as suggested by Lopez-Poveda *et al.* (2003). It was sometimes necessary to extrapolate the linear references to longer masker–probe time intervals to infer BM I/O functions over the wider possible range of levels. In similar situations, some authors have fitted the linear reference TMC with a straight line (e.g., Lopez-Poveda *et al.*, 2003; 2005; Nelson *et al.*, 2001; Plack *et al.*, 2004). There is strong evidence, however, that the decay of forward masking is better described with two time constants (e.g., Lopez-Poveda and Alves-Pinto, 2008; Meddis and O'Mard, 2005; Oxenham and Moore, 1994; Plack and Oxenham, 1998). Based on this, the individual linear-reference TMCs were fitted here (using a least-squares procedure) with a double exponential function of the form:

$$L_m(t) = L_0 - 20\log_{10}[\alpha e^{(-t/\tau_a)} + (1-\alpha)e^{(-t/\tau_b)}], \quad (1)$$

where $L_m(t)$ is the masker level required to mask the probe at masker–probe time interval t; L_0 is masker level for a masker–probe interval of zero; τ_a and τ_b are time constants; and α determines the level at which the second exponential takes over from the first one. α, τ_a, τ_b, and L_0 were fitting parameters and were allowed to vary freely within certain boundaries: α was restricted to the interval [0, 1]; L_0 was restricted to vary within [50,120] dB; and τ_a and τ_b to the interval [1,200] ms. The lowest correlation between actual and predicted masker levels was r=0.94, which shows that the goodness of fit was excellent.

E. DPOAE stimuli

DPOAE I/O curves were obtained by plotting the magnitude (in dB SPL) of the $2f_1-f_2$ DP emission as a function of the level, L_2, of the primary tone f_2. DPOAEs were measured only for f_2 frequencies equal to the probe frequencies for which TMCs had been previously measured (0.5, 1, 2, and/or 4 kHz). The f_2/f_1 ratio was fixed at 1.2. L_2 ranged from 20 to 75 dB SPL in 5 dB steps, except for f_2 =0.5 kHz for which it ranged from 45 to 75 dB SPL. L_1 and L_2 were related according to $L_1=0.4L_2+39$ dB, the rule proposed by Kummer *et al.* (1998) to obtain maximum-level DPOAEs for $L_2 \leqslant 65$ dB SPL. In the current study, this rule was extrapolated to $L_2 > 65$ dB SPL.

In an attempt to reduce the potential variability of the I/O curves caused by the DP fine structure, DPOAE I/O curves were measured for five f_2 frequencies near the frequency of interest, and the resulting I/O curves were averaged. This procedure is supported by Kalluri and Shera (2001) and Mauermann and Kollmeier (2004), who showed that a "cleaned" DP-gram (i.e., a DP-gram where the fine structure has been accounted for) resembles very closely a moving average of the original DP-gram with fine structure. The five adjacent frequencies were selected to differ by as much as 2% of the frequency of interest based on a suggestion of Mauermann and Kollmeier (2004) that this frequency spacing is appropriate to reveal (or to account for) the influence of the fine structure. That meant measuring DPOAE I/O curves for frequencies of $0.98f_2$, $0.99f_2$, f_2, $1.01f_2$, and $1.02f_2$. For instance, the final DPOAE I/O curve at 4 kHz was the mean of five I/O functions for f_2 ={3920,3960,4000,4040,4080 Hz}. To further assess the potential influence of the fine structure around the frequency of interest, DPOAEs were measured for four additional adjacent f_2 frequencies on each side of the five main frequencies. In the case of 4 kHz these were: f_2 ={3760,3800,3840,3880,…,4120,4160,4200,4240 Hz}. DPOAEs for these latter frequencies were measured only for L_2={70,60,50,40 dB} and the resulting I/O curves were not included in the final mean I/O curve.

The influence of standing waves must be taken into consideration when measuring DPOAE at high frequencies. Siegel (1994) and Whitehead *et al.* (1995) found that restricting DPOAE measurements to $f_2 < 6$ kHz avoids the majority of the idiosyncratic variations of the sound pressure at the eardrum due to standing waves. For this reason, it was decided to restrict DPOAE measurements to $f_2 \leqslant 4$ kHz. Measuring DPOAEs for $f_2 < 1$ kHz is also problematic due to increased physiological subject noise. The moderate-to-high level part of the DPOAE I/O curve for f_2=0.5 kHz could still be measured for most listeners, however, by considering L_2 above 45 dB SPL only.

F. DPOAE stimulus calibration

DPOAE stimuli were calibrated with a Zwislocki DB-100 coupler for each f_1,f_2 frequency. In some studies, calibrated levels are further adjusted with the probe *in situ* to account for the acoustic effects associated with ear-canal resonances. This *in situ* adjustment, however, was not applied here for two reasons. First, Siegel (1994) has shown that it does not always work for frequencies above 2–3 kHz because of the errors in predicting the level at the eardrum from measurements made at the plane of the probe (where the standing waves interact). Second, as this is a comparison study, it was deemed important that the two methods applied the same stimulus level control. Given that the psychophysical equipment did not allow easy *in situ* level adjustment, this option was disabled in the OAE instrument.

G. DPOAE system artifacts

When measuring DPOAE I/O curves at high primary levels it is necessary to control for cubic distortion produced by the measurement instrument. This system artifact was assessed by measuring the magnitude of the $2f_1-f_2$ DP in two different couplers: a DB-100 Zwislocki coupler and a plastic syringe having a volume of approximately 1.5 cc. The test was performed for L_2 from 50 to 80 dB SPL with the same equipment and in the same conditions used to measure DPOAEs. The measurement time was prolonged to maximize the chances of discovering any artifacts. The magnitude of the cubic DP would be $-\infty$ dB SPL for an ideal OAE system. The system artifact limit was set to the higher of the responses for the two couplers that was also 2 s.d. above the mean level of ten adjacent frequency bins in its corresponding spectrum. This procedure was repeated for each of the f_2s considered when measuring DPOAEs.

It is common to accept DPOAE measurements when they are above the system artifact limit. This may be toler-

able in a clinical context where it comes to making a "pass/refer" decision. The magnitude of the measured DPOAE is the *vector* sum of DP contributions from any nonlinearity along the signal path, be it from the instrument or from the subject. If the clinical rule was applied, then the true physiological response would be any value within the range $(-\infty, +6]$ dB around the measured DPOAE in the worst possible case (i.e., when the measured DPOAE magnitude just exceeds the artifact limit and the physiological DP has opposite or equal phase to the system's DP, respectively). This uncertainty range, however, seems too broad for the present study where the slope of the I/O curve is of interest. Therefore, a more restrictive rule was applied. A DPOAE measurement was accepted as valid only when it exceeded the artifact limit by 6 dB or more. This guaranteed that the physiological DP contribution was within the range $[-6, +3.5]$ dB in the worst possible case (i.e., when the DPOAE measurement just met the present criterion). Therefore measurements were rejected if they were less than 6 dB above the system artifact limit. This was the case for subject S4 at 4 kHz, for whom most data points were discarded based on this criterion, and also for subject S3 at 1 kHz at high stimulus levels (75 dB SPL).

H. DPOAE procedure

DPOAE measurements were obtained with an IHS Smart system (with SmartOAE software version 4.52) equipped with an Etymotic ER-10D probe. During the measurements, subjects sat comfortably in a double-wall sound attenuating chamber and were asked to remain as steady as possible.

The probe fit was checked before and after each measurement session. The probe remained in the subject's ear throughout the whole measurement session to avoid measurement variance from probe fit. DPOAEs were measured for a preset measurement time. For $f_2=4$ kHz, the measurement time ranged from 60 s at $L_2=20$ dB SPL to 8 s at high L_2; for $f_2=0.5$ kHz, it ranged between 60 and 30 s for L_2 between 45 and 75 dB SPL, respectively. A DPOAE measurement was considered valid when it was 2 s.d. above the measurement noise floor (defined as the mean level over ten adjacent frequency bins in the spectrum). When a response did not meet this criterion, the measurement was repeated and the measurement time was increased if necessary. The probe remained in the same position during these remeasurements. If the required criterion was not met after successive attempts, the measurement point was discarded.

Each recording session consisted of measuring DPOAE I/O curves for one frequency of interest (consisting of five adjacent frequencies) (see Sec. II E) and was allowed to take up to 1 h. Three DPOAE measurements were obtained per condition (i.e., per f_2 and L_2) and averaged, except for subject S2 for whom only one measurement was obtained. Therefore, each point in the final I/O curves was the mean of 15 (3 measurements $\times$ 5 adjacent frequencies) measurements for each L_2 level. Occasionally, it was not possible to obtain all 15 points, particularly for the lower L_2 levels. In those cases, the direct mean of the available points would have been biased toward the DPOAE values for the frequencies giving the stronger responses. To minimize this potential bias, a mean was calculated only when two (out of the three possible) measurements were available per frequency and when eight of the new ten possible measurements were available. Otherwise, the corresponding point was neglected for further analysis.

III. RESULTS

A. TMCs

Figure 1 shows the TMCs for the ten subjects for probe frequencies of 0.5, 1, 2, and 4 kHz. Note that TMCs were not measured for all frequencies and for all subjects. Subject S5 was unable to perform the TMC task using a probe level of 9 dB SL, thus a probe level of 15 dB SL was used in that case. Even so, it was not possible to measure high masker levels at 4 kHz for subject S5 and the resulting TMC did not allow estimating the true degree of compression (open circles in panel S5 of Fig. 1). Therefore these data were discarded from further analysis.

The shapes of the present TMCs are generally consistent with those of previous studies (e.g., Nelson *et al.*, 2001; Lopez-Poveda *et al.*, 2003; Lopez-Poveda and Alves-Pinto, 2008; Plack *et al.*, 2004; Rosengard *et al.*, 2005). The linear-reference TMCs (open squares) can be described either by a straight line (e.g., S4, S5) or by a shallow and gradually saturating function (e.g., S3, S8). The latter justifies the decision to fit the linear-reference TMCs with a double-exponential function. The continuous, thick lines illustrate these fits. Several on-frequency TMCs show a shallow segment for short masker–probe time intervals (or gaps) followed by a steeper segment for moderate intervals. Others, however, are better described by a segment steeper than the linear reference followed sometimes by a shallower section at high masker levels (e.g., S1, S2, S6, and S8). In any case, all of the on-frequency TMCs have segments that are much steeper than the linear reference TMC. Assuming that the off-frequency masker condition used to generate the linear reference is processed linearly by the BM and that the rate of decay of the internal masker effect is identical across frequencies (Lopez-Poveda *et al.*, 2003; Lopez-Poveda and Alves-Pinto, 2008), the steeper segments may be interpreted to indicate BM compression (Nelson *et al.*, 2001). The validity of these assumptions is discussed in Sec. IV F.

B. DPOAEs

Figure 2 shows a typical example of the amount of data measured to estimate one DPOAE I/O curve (2 kHz in this particular case). Each data point is the average of three measurements. Figure 2 also serves to illustrate the influence of the DP fine structure on the resulting I/O function. The inset indicates L_2 (in dB SPL). Note that DPOAEs were measured for a wider range of stimulus levels (from 20 to 75 dB SPL) for the five central adjacent frequencies only. A narrower range of levels (40, 50, 60, and 70 dB SPL) was considered for frequencies outside this frequency range. This was suffi-

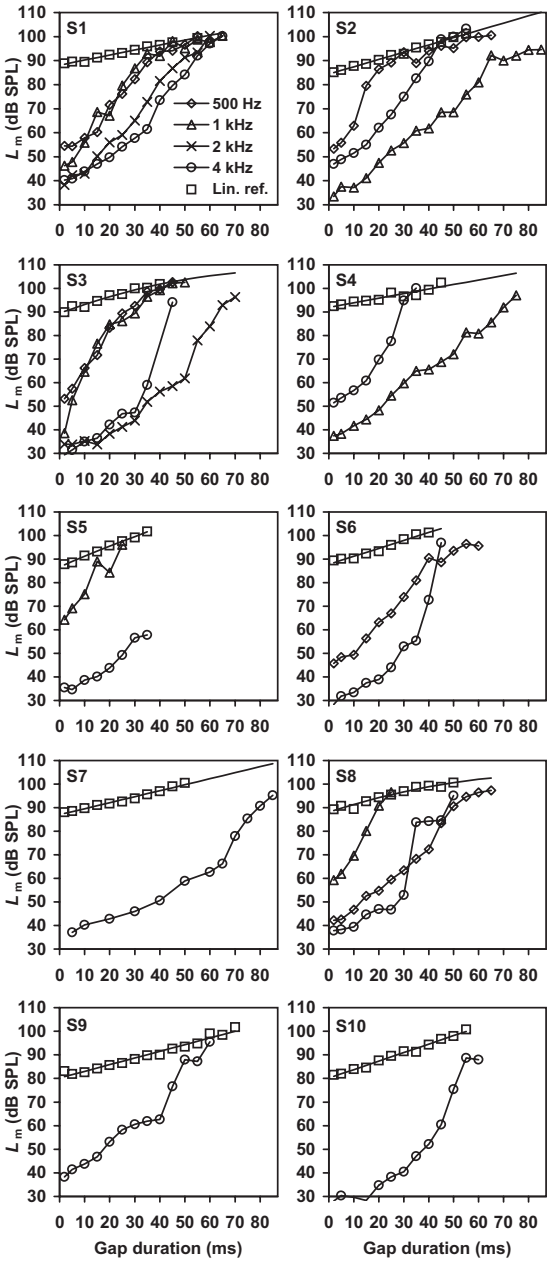

FIG. 1. TMCs for all subjects and probe and masker frequencies. Each panel illustrates the data for one subject. Open squares illustrate the linear reference TMC; i.e., the TMC for a probe frequency of 4 kHz and a masker frequency ($0.4f_P$) of 1.6 kHz. The smoother continuous lines illustrate fits to the linear reference TMC with a double-exponential function. Other symbols illustrate on-frequency TMCs for different probe frequencies (as indicated by the inset in the top-left panel). The probe level was 9 dB SL except for subject S5 for whom it was 15 dB SL.

cient, however, to get an idea of the surrounding DPOAE fine structure and its potential influence on the DPOAE I/O at the frequency of interest.

Obviously the DPOAE I/O curve would change considerably by changing f_2 within a narrow frequency range of 100 Hz, which emphasizes the need to take into account the effect of fine structure when estimating the actual I/O curve. The final I/O curve, representing the BM I/O function at 2 kHz, would be the mean of the five I/O curves for the five central frequencies. Figure 3 illustrates these five I/O curves for subjects S1 and S2 and for all frequencies considered in

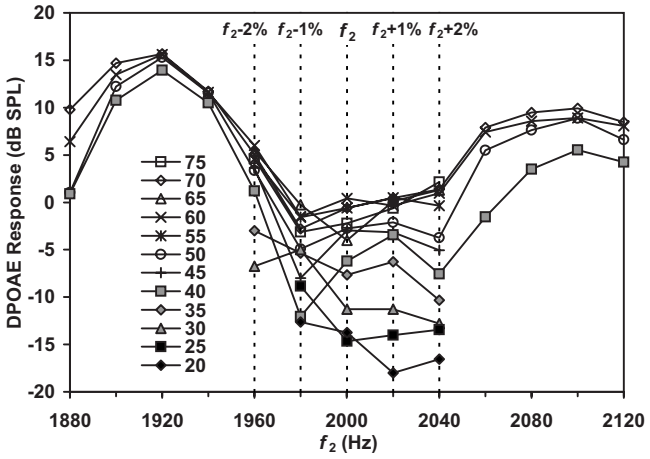

FIG. 2. An example of the influence of the DPOAE fine structure at 2 kHz. An example data set recorded to estimate every I/O curve is also illustrated. The DPOAE magnitude is shown for 12 f_2 frequencies around the frequency of interest (2 kHz) and for different L_2 levels (in dB SPL) as indicated in the inset.

the present study. As can be seen, there is variability across subjects and across frequencies. Nevertheless, variability appears to be greater for the lower stimulus levels. This seems reasonable given that the fine structure is more obvious at low stimulus levels (Mauermann and Kollmeier, 2004).

C. Comparison of BM I/O curves inferred from TMCs and DPOAEs

Figures 4–7 allow within-subject comparisons of DPOAE I/O curves (open squares) with BM I/O curves inferred from TMCs (closed circles) for corresponding cochlear sites with CFs of 0.5, 1, 2, and 4 kHz, respectively. Each panel illustrates the results for one subject. The associated solid lines are third-order polynomials fitted to the data. Error bars denote 1 standard error (s.e.) of the mean. For DPOAE I/O curves, the s.e. was based on up to 15 measurement points for each stimulus level, which explains why error bars are so short. The average error across all subjects and levels was 0.7 dB. The vertical and horizontal error bars of TMC-based I/O curves illustrate the s.e. of the linear reference or the TMC in question, respectively, based on at least three measurements. The average errors across all subjects and time gaps for the linear reference and the on-frequency TMCs were 0.9 and 2.3 dB, respectively.

The open triangles in the top-right panel of Figs. 4–7 and their associated error bars illustrate the mean DPOAE noise floor plus 2 s.d. (based on three measurements for the five adjacent frequencies considered) for one example subject. The noise levels were similar for the other subjects. In general, the noise has the effect of increasing the DPOAE magnitude, particularly at low levels. However, the strict criteria used here should have avoided this influence.

In general, both DPOAE and TMC-based I/O curves are similar in that they are linear at low levels and become gradually compressive with increasing level. There is a tendency in both sets for I/O curves to become linear again at the highest levels tested (e.g., S6 at 0.5 kHz in Fig. 4; or S2 and S5 at 4 kHz in Fig. 7). The degree of similarity between

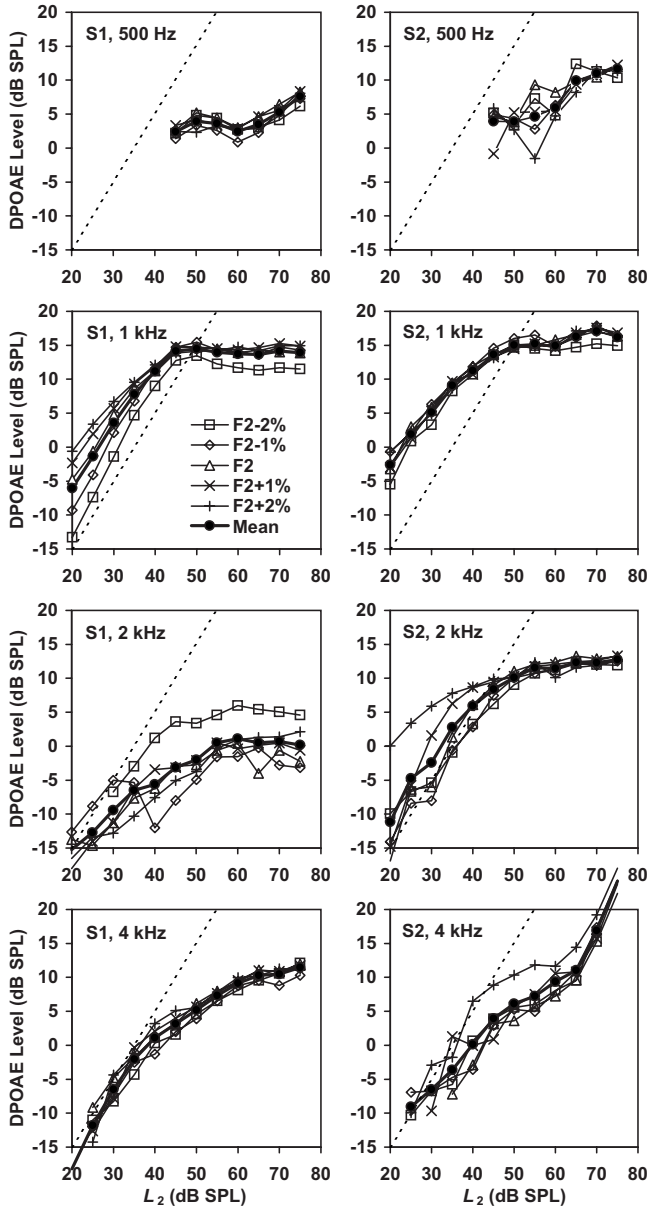

FIG. 3. Example DPOAE I/O curves for subjects S1 (left panels) and S2 (right panels), for frequencies from 0.5 (top) to 4 kHz (bottom) in octave steps. Each final I/O curve (closed circles, thick line) was obtained as the average of five I/O curves for frequencies of $0.98f_2$, $0.99f_2$, f_2, $1.01f_2$, and $1.02f_2$, illustrated with different symbols according to the inset. The dashed line illustrates a linear response for comparison.

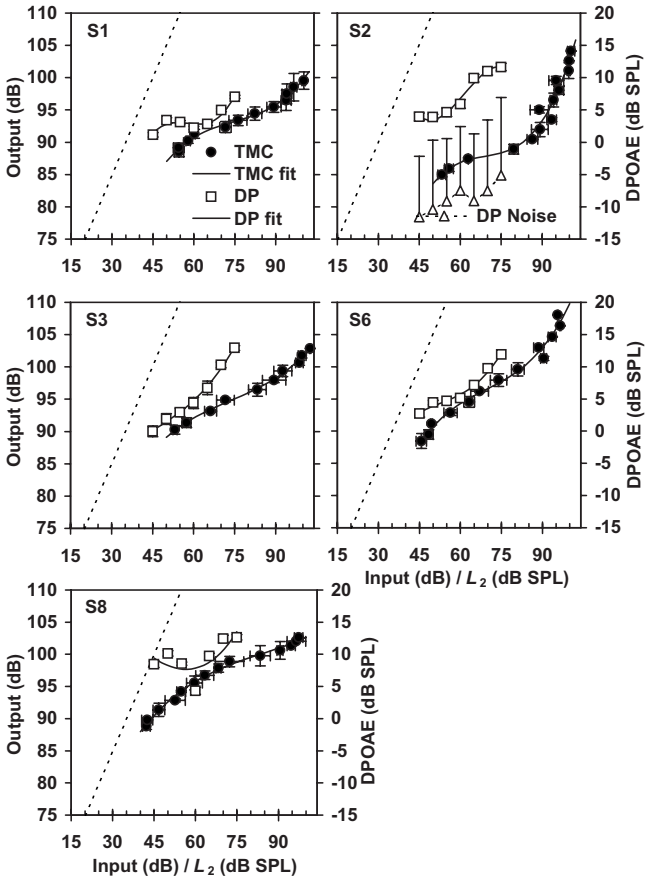

FIG. 4. Experimental DPOAE (open squares) and TMC-based (closed circles) I/O curves at 0.5 kHz ($=fp=f_2$). Continuous lines illustrate third-order polynomial fits to the experimental I/O curves. Error bars denote 1 s.e. of the mean. Horizontal bars (only for TMC-based curves) represent the standard error of the input level (i.e., the standard error for the on-frequency masker level). Each panel illustrates the result for a different subject. The panel for subject S2 also illustrates the mean DP noise floor and its corresponding 2 s.d. Thin dashed lines illustrate a linear response for comparison.

the two sets of I/O curves is greatest at 4 kHz (Fig. 7). At this frequency, both sets of I/O curves indicate equally mild compression for subjects S1, S2, S5, and S7, and equally strong compression for S3 and S8. The strongest disagreement between the two sets of I/O curves at 4 kHz occurs for subject S9 (Fig. 7). S4 deserves a special mention because her DPOAEs could not be measured for levels outside the 40–50 dB SPL level range despite her having normal hearing at 4 kHz (her DPOAE readings for $L_2<40$ dB SPL were below the physiological noise level and for $L_2>50$ dB SPL they did not meet the instrument artifact criterion, see Sec. II G).

The degree of similarity between the shapes of the two sets of I/O curves is, however, much lower for 0.5 and 1 kHz. It is noteworthy that, for these frequencies, some DPOAE I/O curves show plateaus and notches that do not have a clear correlate with TMC-based I/O curves (e.g., S1, S2, and S8 at 0.5 kHz in Fig. 4; or S3, S4, and S5 at 1 kHz in Fig. 5). Possible explanations for the notches and plateaus are discussed in Sec. IV B.

It was not possible to draw conclusive results regarding the degree of correlation between the two sets of I/O curves at 2 kHz because data at this frequency were collected for two subjects only (Fig. 6). Good correspondence (akin to what was observed at 4 kHz) was found for one of the two subjects (S3). The DPOAE I/O curve for the other subject (S1), however, exhibited a negative gradient at moderate-to-high levels that was not present in the TMC-based I/O curves. This negative gradient was typical of the I/O curves at frequencies of 0.5 and 1 kHz.

D. Comparison of derived cochlear nonlinearity parameters

Third-order polynomials were fitted by least squares to the TMC-based and DPOAE I/O curves (continuous lines in Figs. 4–7) and used to derive the following parameters pertaining to cochlear nonlinearity: minimum compression ex-

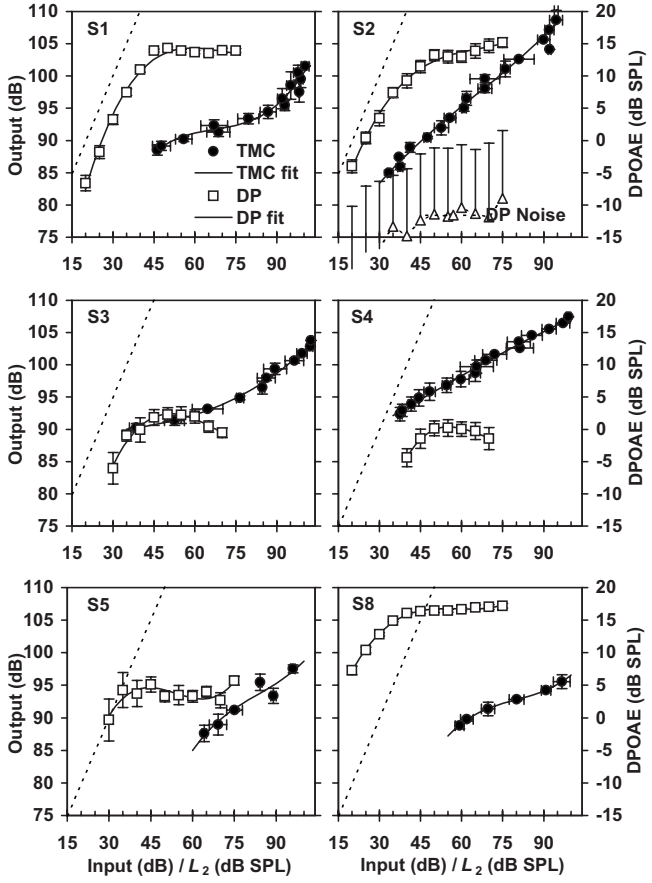

FIG. 5. The same as Fig. 4, but for a frequency of 1 kHz.

ponent, the level at which it occurred, compression threshold, cochlear gain, and the level at which the I/O curves returned to linearity at high levels.

The minimum compression exponent was estimated as the minimum slope of the fitted I/O curves over the measured range of input levels. Figure 8(A) compares the minimum compression exponent as inferred from the I/O curves obtained with the two methods. Closed symbols show the results for 4 kHz separately. Open symbols illustrate the results for the other frequencies (0.5, 1, and 2 kHz), as indicated by the inset. The solid and dashed lines illustrate linear regression fits to the 4 kHz data only or to all data points (including also the 4 kHz points), respectively. Both regression lines were forced to pass through the origin of the graph. Pearson's correlation coefficients were (high) 0.92 for the 4 kHz subsample and (low) 0.19 when all data points (all

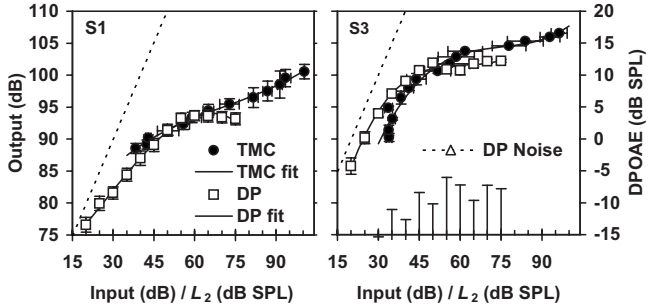

FIG. 6. The same as Fig. 4, but for a frequency of 2 kHz.

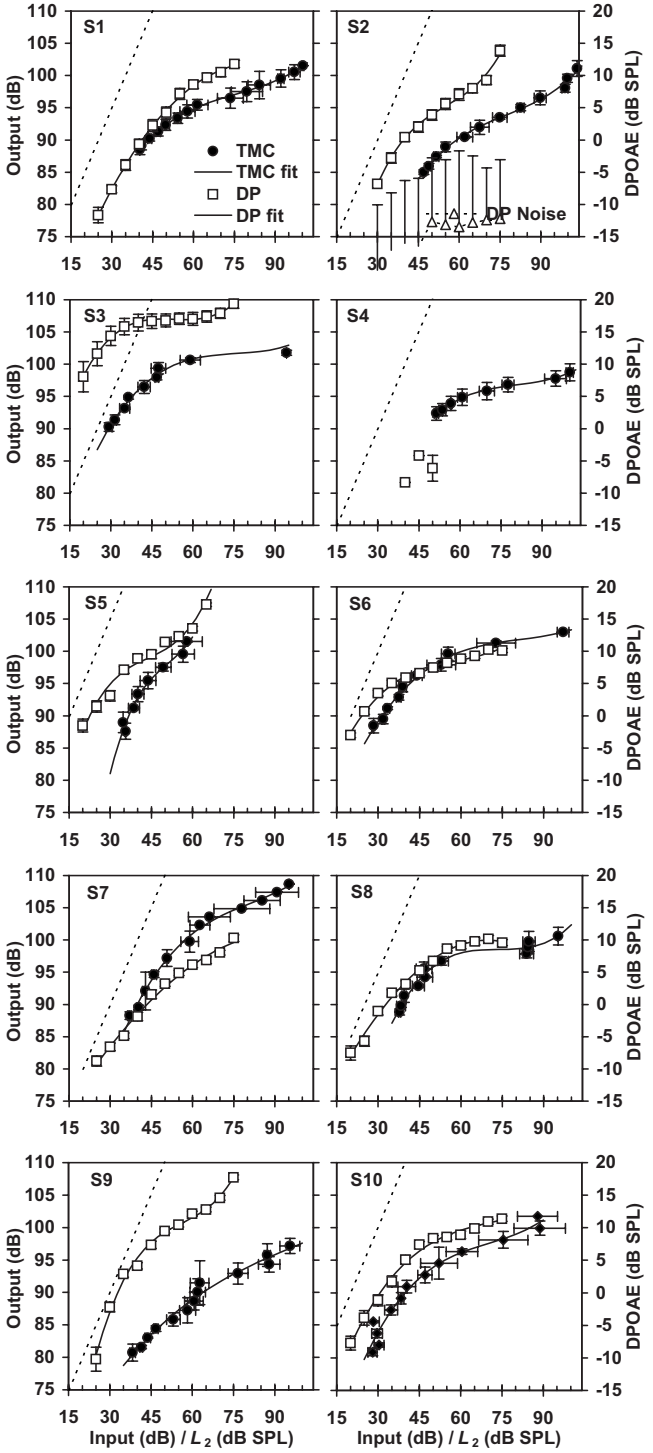

FIG. 7. The same as Fig. 4, but for a frequency of 4 kHz.

frequencies) were considered. The mean compression exponents across frequencies were 0.13 and 0.04 for the TMC-based and the DPOAE data sets, respectively. They, however, were 0.10 and 0.11, respectively, considering only the 4 kHz data points. The mean difference was significant when all data points were considered ($p<0.005$ for a paired two-tailed Student's t-test) but not for the 4 kHz subsample ($p=0.81$). This is clearly seen from Fig. 8(A) in that DPOAE-based compression estimates are lower overall than TMC-based ones for 0.5 and 1 kHz.

The reason for the disagreement between the results at

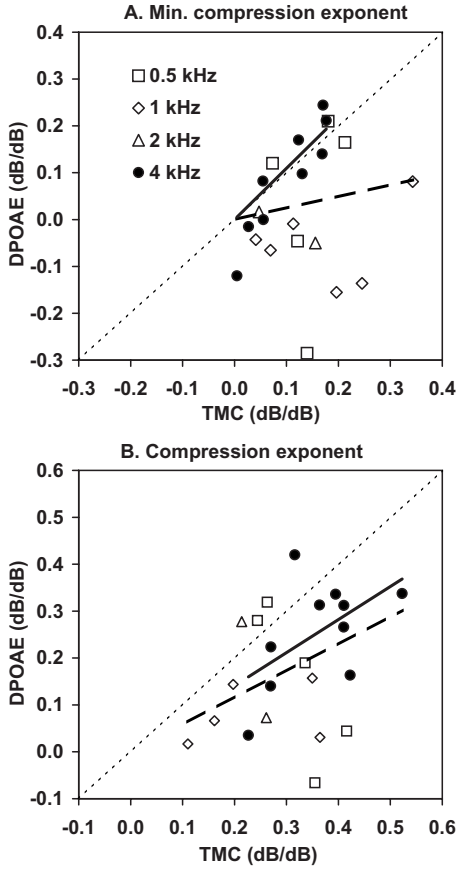

FIG. 8. Correlation between compression-exponent estimates obtained from DPOAE- and TMC-based I/O curves (i.e., each point represents data for one subject). (A) Minimum compression exponent based on third-order polynomial fits to the I/O curves. (B). Compression exponent based on linear-regression fits to the I/O curves over the input-level range 40–65 dB SPL. Closed and open symbols illustrate the 4 kHz data and the data at other frequencies (0.5, 1, and 2 kHz), respectively. Solid lines show linear regression fits to the 4 kHz data points constrained to cross through the origin. Thick dashed lines are linear regression fits considering all data points (including 4 kHz) constrained to cross through the origin. Diagonal thin-dashed lines illustrate perfect correlation.

0.5 and 1 kHz is, almost certainly, that many DPOAE I/O curves for the lower frequencies showed plateaus or notches, which resulted in unexpectedly low compression exponents (slopes <0 dB/dB). Table II summarizes the number of cases across frequencies when the fitted I/O curve exhibited plateaus or negative compression exponents.

These results contrast with those of Williams and Bacon (2005), who reported a high correlation between compression estimates inferred from TMC-based and DPOAE I/O curves also at low frequencies. The compression estimates of Williams and Bacon (2005), however, were calculated differently. They were calculated as the slope of linear segments fitted to the DPOAE and TMC-based I/O curves over the range of input levels from 40 to 65 dB approximately (they determined the actual level range by visual inspection). This fitting method was also employed here as it could prove less sensitive to notches. Figure 8(B) shows the correlation between compression exponent estimates for the two methods derived by linear regression fits to the I/O curves over the level range 40–65 dB. The 4 kHz data are shown by the closed symbols; open symbols illustrate the results for frequencies of 0.5, 1, and 2 kHz. The solid and dashed lines are linear regression fits to the 4 kHz data and to all data points (including also the 4 kHz data), respectively. Pearson's correlation coefficient considering all data points was 0.32 and Fig. 8(B) illustrates that DPOAE-based compression exponents were generally lower (most points are below the diagonal) than those inferred from TMCs. Group mean compression exponents were 0.31 and 0.18 for TMCs and DPOAEs, respectively. This difference was statistically significant ($p<0.0005$). A moderately higher correlation was found when analyzing the 4 kHz data separately ($r=0.53$), but the associated mean difference (0.36 and 0.25 for TMCs and DPOAEs, respectively) was still statistically significant ($p<0.01$).

TABLE II. Incidence of plateaus and notches in DPOAE I/O curves as a function of f_2.

	Frequency (kHz)			
	0.5	1	2	4
Number of I/O measured curves	5	6	2	10
Number of I/O curves with plateaus or notches	3	5	1	2

Another characteristic of cochlear nonlinearity is the level at which maximum compression (or, equivalently, minimum compression exponent) occurs. Correlations were sought between estimates of this parameter for the I/O curves obtained with the two methods. Figure 9(A) reveals that high correlation occurs in a few cases only (points close to or on the diagonal). Most points are below the diagonal, which means that the level at which minimum compression occurred was lower for DPOAE I/O curves than for I/O curves inferred from TMCs. The average input level for minimum compression exponent was 61 and 76 dB for DPOAEs and TMC-based I/O curves, respectively. The difference was statistically significant ($p<0.001$). Results were similar when the data for each frequency were analyzed separately.

Compression threshold was defined as the input level at which the slope of the fitted polynomial decreased from a value close to one at low levels to 0.4 dB/dB at a higher level. This is an arbitrary definition, but seems reasonable for our purpose. When the slope of the I/O curve at the lowest level for which a data point existed was reasonably close to 0.4 dB/dB, the fitted polynomial was extrapolated up to 5 dB to identify the compression threshold and the value was noted for further analysis. An extrapolated compression estimate was thus included in the analysis only if it was less than 5 dB from an existing data point. Figure 9(B) illustrates the correlation between the compression threshold estimates inferred from DPOAE and TMC-based I/O curves. Three compression estimates were considered outliers [symbols surrounded by an open circle in Fig. 9(B)] and thus excluded from the statistical analysis. The possible reasons for these outliers will be discussed in Sec. IV D. The regression line was constrained to cross through the graph origin. The line indicated a high degree of correlation between the estimates of the two methods and, indeed, Pearson's correlation coefficient was reasonably high ($r=0.8$). The average compression threshold estimates were lower for the DPOAE than for

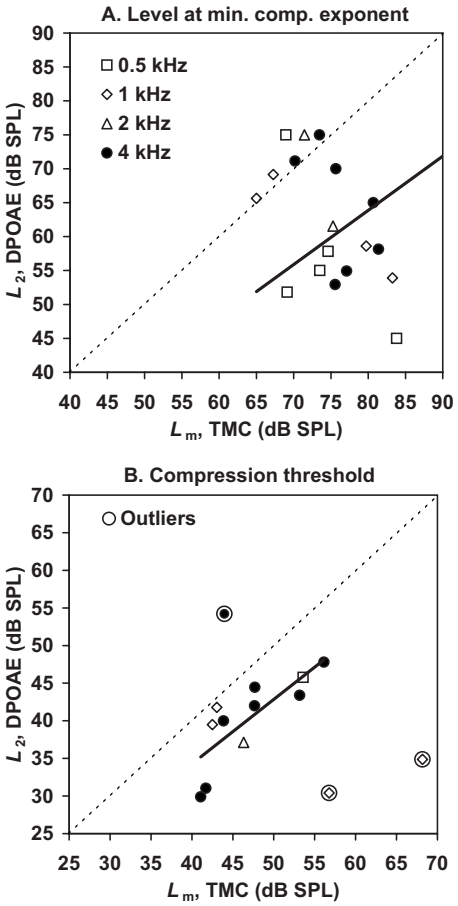

FIG. 9. (A) Correlation between the level at which maximum compression occurs based on the I/O curves inferred from DPOAE and TMCs (i.e., each point represents data for one subject). The solid line shows a linear regression fit to the data constrained to cross through the origin. (B) Correlation between compression-threshold estimates obtained with the two methods. Threshold was defined as the input level at which the I/O curve slope decreased from linearity to 0.4 dB/dB. The solid line shows the linear regression constrained to cross through the origin. Symbols surrounded by open circles illustrate three data points regarded as outliers. Diagonal thin-dashed lines illustrate perfect correlation.

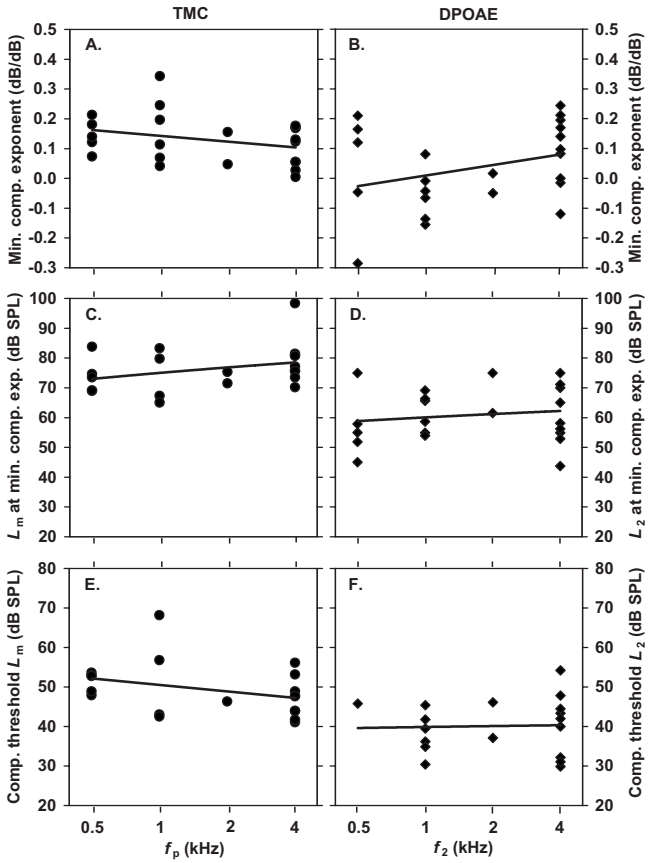

FIG. 10. Frequency dependency of cochlear nonlinearity parameters as estimated from TMCs (left panels) and DPOAEs (right panels). (A), (B) Minimum compression exponent. (C), (D) The level at which maximum compression occurs. (E), (F) Compression threshold.

the TMC-based I/O curves (40 and 47 dB SPL, respectively), which is also clearly seen in Fig. 9(B). This difference was statistically significant ($p<0.0001$).

E. Cochlear nonlinearity dependency on characteristic frequency

This section addresses whether the parameters considered in the preceding section vary with frequency similarly when they are inferred from TMCs or DPOAEs. Two additional parameters were considered based on TMCs, namely gain and the threshold of return to linearity at high levels. I/O curves may be generally described as having a linear segment (slope of 1 dB/dB) at low levels, followed by a compressive segment (slope <1 dB/dB) at moderate levels, followed by linear segment at high levels (e.g., Lopez-Poveda *et al.*, 2003). Perhaps, the clearest examples of this pattern were the TMCs for S6 at 0.5 kHz and S2 at 4 kHz. Few TMC-based I/O curves showed a linear segment at high levels (e.g., S2 at 0.5 kHz or S1 at 1 kHz). The slope of the I/O curves, however, always increased with increasing level beyond the inflection point of the curve (defined as the level at which maximum compression occurs or, equivalently, the level at which the second derivative of the I/O curve equals zero). This suggests that the I/O curves might approach linearity at very high levels. Since the minimum slopes were always <0.4 dB/dB [Fig. 8(A)], the return-to-linearity threshold was arbitrarily defined here as the level at which the slope of the fitted I/O equated to 0.4 dB/dB for levels above the inflection point. Gain was defined as the difference between the return-to-linearity and the compression thresholds (in decibels).

Figure 10 shows the frequency dependency of the minimum compression exponent (top panels), the level at which it occurs (middle panels), and the compression threshold (lower panels), respectively. Left and right panels show the value of these parameters as inferred from TMCs and DPOAE I/O curves, respectively. Figure 10 illustrates that all these parameters remain approximately constant across frequencies. Indeed, no significant differences were found between the mean values of every parameter across frequencies. The same applied to parameters estimated with both methods.

Figure 11 shows the gain and the return-to-linearity threshold as inferred from TMCs only. Both parameters tend to increase with increasing frequency. Since the compression threshold remained approximately constant across frequencies [Fig. 10(E)], the frequency dependency of the gain is

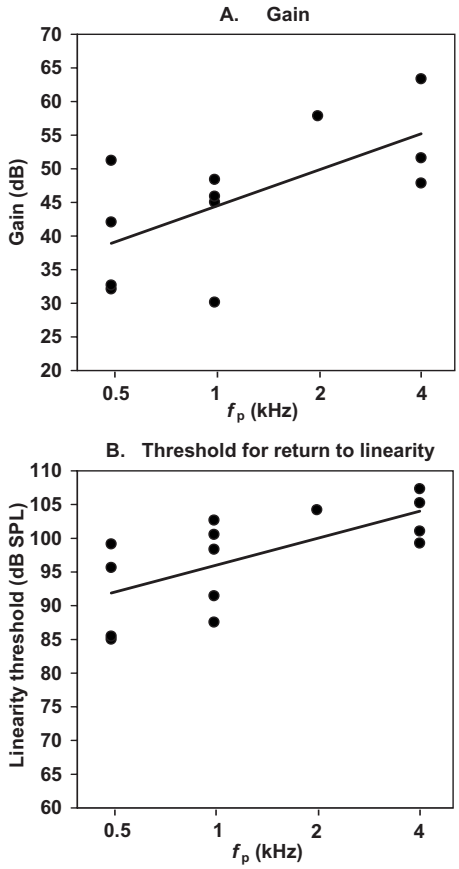

FIG. 11. Frequency dependency of gain (A) and the threshold of return to linearity (B) derived from TMC I/O curves. The latter was defined as the input level at which the I/O curve reached a slope of 0.4 dB/dB from a lower value for increasing input level.

fully attributed to the increase of the return-to-linearity threshold with increasing frequency. In any case, a one-to-one correspondence between these two parameters should not be expected because they are based on data for different subjects. Indeed, it was not possible to estimate the return-to-linearity threshold in several cases (one different subject for each frequency of 0.5, 1, and 2 kHz, and six subjects at 4 kHz). The I/O curves for these subjects may show a return to linearity at input levels higher than those considered in the present study. That the majority of these cases occurred at 4 kHz supports the idea that the return to linearity at higher frequencies occurs at higher input levels than those considered here.

IV. DISCUSSION

The goal of this study was threefold. The first objective was to compare cochlear nonlinearity parameters inferred from TMCs and DPOAEs, and, if coinciding, add support to the notion that they are two equivalent manifestations of cochlear nonlinearity. A second aim was to evaluate the feasibility of using DPOAE I/O curves as a fast tool for estimating individual parameters of cochlear nonlinearity. A third objective was to investigate the frequency dependency of parameters describing cochlear nonlinearity, as inferred from DPOAEs and TMCs.

A. Equivalence between DPOAE and TMC-based I/O curves

The degree of correlation between compression estimates inferred with the two methods was high (0.92) at 4 kHz but much lower at 0.5 and 1 kHz. The assumption has been made that DPOAE I/O curves reflect the characteristics of the BM response to single tones at the f_2 place. Measuring DPOAEs, however, requires presenting the two primaries (f_1 and f_2) *simultaneously*; hence the f_1 primary may have suppressed the BM response to the f_2 primary at the f_2 site. Indeed, Rhode (2007) measured BM excitation and DPOAEs simultaneously in the same preparation and showed that the f_1 primary suppresses the BM response to the f_2 (with a fixed level of 60 dB SPL) for L_1 above 60 dB SPL (see his Fig. 1). As a result, DPOAE I/O curves may not correspond directly to single-tone BM I/O curves, as is commonly assumed.

Unlike DPOAE I/O curves, the I/O curves inferred from TMCs would not be affected by suppression because the masker and the probe tones were not presented simultaneously (in fact, this is one of the reasons that they are so widely used to infer BM I/O curves). Therefore, one might think that mutual suppression between the primary tones may have influenced DPOAE but not TMC-based I/O curves and that the effect would be more pronounced at low CFs because the nonlinear effects extend to a wider bandwidth (Rhode and Cooper, 1996; Lopez-Poveda *et al.*, 2003). This explanation, however, is unlikely to account for the low correlation between compression estimates obtained with the two methods at low frequencies. There is physiological and psychophysical evidence that suppression leads to I/O curves steeper than single-tone I/O curves (Nuttall and Dolan, 1993; Rhode, 2007; Yasin and Plack, 2007). This is true particularly for suppressor/suppressee combinations similar to the primary-tone combinations used here. If suppression had affected the DPOAE I/O curves, they should indicate less compression than TMC-based I/O curves and this has been found *not* to be the case (Fig. 8). Therefore, the most likely explanation for the low correlation at low CFs between the compression estimates obtained with the two methods is the presence of notches and plateaus in the DPOAE I/O curves, which occur more frequently at low CFs (Table II).

The group mean compression-exponent estimates at 4 kHz obtained in the present study (0.10 and 0.11 dB/dB for TMC and DPOAE, respectively) are in agreement with those from Gorga *et al.* (2007), who reported a minimum slope value of ~0.12 dB/dB at 4 kHz based on DPOAEs (estimated from their Fig. 4 at L_2=60 dB SL). As for compression estimated using linear regression over the midlevel range [Fig. 8(B)], the group mean exponent values obtained in the present study (0.36 and 0.25 for TMC and DPOAE, respectively) were moderately higher than those reported by Williams and Bacon (2005) (0.26 and 0.15 for TMC and DPOAE at 4 kHz, respectively). In any case, the present study shows that the estimated degree of compression differs considerably depending on the method used to infer it (e.g., polynomial versus linear regression fits), which emphasizes the need to specify clearly the method used in every study.

The minimum compression exponent found here at 4 kHz (based on polynomial fits) is lower than previously

reported values obtained with the same (or different) methods [e.g., 0.14, third-order polynomial, Nelson and Schroder (2004), 0.20, third-order polynomial, Plack and Drga (2003), 0.13, sum of linear and sigmoidal function, and 0.23, straight line, Rosengard *et al.* (2005), 0.20, straight line, Plack *et al.* (2004), 0.25, straight line, Lopez-Poveda *et al.* (2003)]. The difference in TMC-based estimates may relate to differences in the linear reference used by different studies. Here, the linear reference was the TMC for a masker frequency of $0.4f_P$, whereas the above-mentioned studies used the TMC for a masker frequency between $0.5f_P$ and $0.6f_P$. Lopez-Poveda and Alves-Pinto (2008) have suggested that the latter may still undergo as much as 2:1 compression; hence they could lead to an underestimate of the degree of on-frequency compression. The agreement between the present compression estimates obtained from TMCs and DPOAEs at 4 kHz provides circumstantial support to the conclusion of Lopez-Poveda and Alves-Pinto (2008).

B. DPOAE notches and plateaus

Notches and plateaus are common in the present DPOAE I/O curves (Figs. 4–7, especially at the lower f_2 (Table II). This contrasts with the conclusion of Kummer *et al.* (1998), who reported that notches and plateaus were less common when DPOAE I/O curves were measured with their proposed primary-level rule than with other rules. The Kummer *et al.* level rule is based on a group average and thus in some cases it may deviate considerably from the individual optimal (the optimal rule would be the one that evokes the strongest possible DPOAEs at all levels). Therefore, one possible explanation for the present observations is simply that the rule of the Kummer *et al.* (1998) was not optimal for the subjects used in the present study. This explanation is supported by Neely *et al.* (2005), who reported that L_1 should be systematically higher than prescribed by Kummer *et al.* (1998) and that the L_1–L_2 relationship should vary with f_2. Johnson *et al.* (2006) confirmed the latter and further suggested that the f_2/f_1 ratio should vary slightly with f_2. Interestingly, it is for low frequencies and moderate–high stimulus levels where the rule of Kummer *et al.* deviates most from the rules of Neely *et al.* (2005) and Johnson *et al.* (2006). These are also the conditions where plateaus and dips are most common in the present data, which suggests that the rule of Kummer *et al.* is not optimal for the subjects considered in this study at low frequencies. On the other hand, Kummer *et al.* (2000) verified their original paradigm [derived from data of Gaskill and Brown (1990)] with a larger sample and still found it to be independent of frequency.

A second explanation for the plateaus and notches is that some of the DPOAE I/O curves still could have been influenced by the fine structure despite the precautions taken to minimize its effects. Plateaus occur for levels around 45–50 dB SPL (e.g., S1 and S8 at 1 kHz in Fig. 5; S3 at 4 kHz in Fig. 7), where the fine structure certainly can have influence. This is, however, an unlikely explanation for the notches because they always occurred at moderate-to-high levels (60–70 dB SPL) and the fine structure has a higher influence at low levels (Mauermann and Kollmeier, 2004).

A third explanation might be that another DP generation mechanism starts playing a role at high stimulus levels and notches reflect destructive interference between the DPs generated by this "new" high-level source and the normal source (see Mills, 1997). The measurement system is unlikely to be the source in question because a rather strict exclusion criterion was applied in the present study to eliminate system-generated DPs. Liberman *et al.* (2004) showed that genetically modified mice without the necessary prestin protein to drive outer-hair-cell electromotility still generated attenuated DP responses and thus supported the existence of a second possible DP-generation mechanism at high levels. Some of their DPOAE I/O curves showed notches at similar stimulus levels to the notches found in the current study. On the other hand, Avan *et al.* (2003) attributed low and high stimulus level DPOAEs to the same nonlinear mechanism. Also, Lukashkin *et al.* (2002) and Lukashkin and Russell (2002) have shown that a single saturating nonlinearity is sufficient to explain a notch in a DP I/O function.

C. The level at which the minimum compression exponent occurred

Low correlation was found between estimates of this parameter obtained with the two methods (DPOAEs and TMCs). The reason for this is uncertain. Maybe DPOAEs grow faster with increasing stimulus level at high stimulus levels because of the contribution from the second high level DP generation source discussed in the preceding sections.

D. Compression threshold

A moderately high correlation was found between compression-threshold estimates inferred with the two methods [Fig. 9(B)]. DPOAE-based estimates were, however, on average 7 dB lower than TMC-based estimates. A total of 22 I/O curves were measured in the ten subjects, but the compression threshold could be estimated in only 14 of these 22 cases. This could be interpreted as an argument against the apparent equivalence of the two methods with respect to estimating this parameter. Further analysis reveals, however, that there are good reasons why a compression threshold could not be estimated in the remaining eight cases. Four of them corresponded to DPOAE I/O curves at 0.5 kHz (Fig. 4) that extended over a range of input levels above 45 dB SPL that was too narrow to reveal a compression threshold. Another case was S4 at 4 kHz (Fig. 7), who did not have sufficiently strong DPOAE despite her hearing threshold being normal at that frequency. For the three remaining cases, the TMCs did not show a clear compression threshold or did not reach the criterion slope of 0.4 dB/dB (see Sec. III D). These cases were S1 at 2 kHz in Fig. 6; S3 and S4 at 1 kHz in Fig. 5. Three of the 14 cases where both methods demonstrated a compression threshold were considered outliers and excluded from the comparison of compression threshold: S1 at 4 kHz, and S5 and S8 at 1 kHz [depicted as circles in Fig. 9(B)]. The data for S5 and S8 were excluded because these listeners had great difficulties performing the TMC task (they needed six to eight attempts at high masker levels to obtain three measurements each having a standard deviation below

6 dB; Sec. II C). There was no obvious reason to exclude the data point of S1. This data point corresponded to the first condition on which the subject was tested and he might not have been sufficiently trained. In any case, it is noteworthy that DPOAE I/O curves allow estimating a compression threshold much more easily than do TMC-based I/O curves.

The reason why DPOAE-based compression threshold estimates were lower than corresponding TMC-based estimates is uncertain. Maybe the DPOAE response was influenced by mutual suppression of the primaries. This could have linearized the I/O curve (e.g., by decreasing cochlear gain) and thus *increased* the compression threshold suggested by DPOAE I/O curves. This explanation does not fit the data as the actual compression thresholds estimated from DPOAEs were *lower* than those estimated from TMCs. Perhaps the difference in compression threshold estimate for the two methods is caused by the use of suboptimal DPOAE parameters (see Sec. IV B).

The moderately high correlation between the compression-threshold estimates of the two methods also at low frequencies [Figs. 9(B), 10(E), and 10(F)] may seem surprising given the low correlation between the estimates of minimum compression exponent. One possible explanation could be that the DPOAE parameters were adequate for lower L_2, where the compression threshold occurs, but not for higher L_2 levels, at which, coincidentally, plateaus and notches occur.

The average compression threshold estimates at 4 kHz found in the present study were 47 and 40 dB SPL for TMC and DPOAE, respectively. These values are comparable to those (average 37 dB SPL at 3–4 kHz) found psychophysically by Yasin and Plack (2003) based on a three-line segment fitting procedure (Plack *et al.*, 2004). A value of ~35 dB SPL is obtained when applying the present definition of compression threshold to the DPOAE data of Neely *et al.* (2003). A compression threshold was also estimated from the DPOAEs reported by Gorga *et al.* (2007). After discarding their data for lowest stimulus levels because they were most likely contaminated by noise and applying the definition used in the present study (level at which I/O curve slope equals 0.4 dB/dB) to their data in their Fig. 3, the resulting compression thresholds were 30 and 45 dB SPL at 0.5 and 4 kHz, respectively. The present values are reasonably in accordance with their results at 4 kHz but not at 0.5 kHz. Our results indicate that the compression threshold is approximately constant across frequencies [Figs. 10(E) and 10(F)].

E. Gain and the level of return to linearity

It is still controversial that BM I/O functions become linear at very high input levels in healthy cochleae (Robles and Ruggero, 2001). Assuming, however, that this is a truly physiological characteristic, DPOAEs are not considered a reliable predictor of the threshold level of return to linearity at high levels. First, there may be another mechanism involved in generation of DPOAE at the higher stimulus levels, as explained above (Sec. IV B). Second, the best frequency of any BM site shifts with increasing level, hence the BM site where the two primaries cause maximum excitation is likely to shift accordingly with level. This may change the DPOAE response as the region of overlap of the two primaries changes with level. Because of this, the gain and return-to-linearity thresholds were not estimated based on DPOAE I/O curves and not compared with those inferred from TMCs.

The TMC-based I/O curves suggested that the cochlear gain increased with increasing CF because of a parallel increase in the threshold of return to linearity at high levels (Fig. 11). This is in agreement with physiological data (Robles and Ruggero, 2001). Gain estimates based on tip-to-tail level differences of DPOAE suppression tuning curves (Gorga *et al.*, 2008) showed the same tendency for the gain to decrease with decreasing frequency.

F. On the merits of the DPOAEs and TMCs for estimating cochlear I/O curves

The presence of plateaus and notches in the DPOAE I/O curves at 0.5 and 1 kHz results in zero or negative compression-exponent estimates which do not occur in corresponding TMC-based curves. While deep notches in apical BM I/O functions have been reported [e.g., Fig. 7(a) of Rhode and Cooper (1996)], they typically occur for stimulation frequencies higher than the CF (Rhode and Cooper, 1996). Therefore, the notches reported here are unlikely to reflect notches in the underlying BM responses (see also Sec. IV B).

It would be wrong to conclude, however, that it is inappropriate to use DPOAEs to infer cochlear I/O functions at low frequencies. As discussed earlier (Sec. IV B), the present notches are possibly due to using suboptimal primary levels at low frequencies and it might be possible to find DPOAE stimulus parameters that would lead to higher correlations between TMC-based and DPOAE I/O curves.

It would also be wrong to conclude that I/O curves inferred from TMCs are more correct (i.e., reflect more closely the underlying BM responses) than DPOAE I/O functions at low frequencies. The TMC method is an indirect, psychophysical method, thus its results may be influenced by retrocochlear mechanisms unknown to date. Indeed, there exist within-subject differences between I/O functions inferred with different psychophysical methods (e.g., Rosengard *et al.*, 2005). Furthermore, the TMC method rests on several assumptions, the main of which is that the rate of decay of the internal masker effect is identical across frequencies and levels (Nelson *et al.*, 2001; Lopez-Poveda *et al.*, 2003). The validity of these assumptions is still controversial. Stainsby and Moore (2006) have argued that the decay rate is faster for low probe frequencies (or equivalently, low CFs), at least for hearing-impaired listeners. By contrast, Lopez-Poveda and Alves-Pinto (2008) have provided indirect evidence for frequency-independent decay rates, at least for normal-hearing listeners. Additionally, there is evidence that for any given frequency, the decay rate is slower at high levels (Lopez-Poveda and Alves-Pinto, 2008; Wojtczak and Oxenham, 2007). These issues complicate the selection of the linear reference TMC and thus cast doubts on the correspond-

ing I/O curves, particularly at low frequencies (Stainsby and Moore, 2006; Lopez-Poveda and Alves-Pinto, 2008).

In summary, the lack of correlation between the results of the two methods at low frequencies is uninformative at present of their relative accuracy for inferring cochlear I/O curves. Future studies should investigate the reason for the low correlation at low frequencies and whether higher correlations would be obtained using different DPOAE parameters and/or different psychophysical methods or assumptions.

The present study did not evaluate the merit of DPOAE I/O curves as a (*clinical*) tool for assessing residual cochlear compression characteristics in hearing-impaired listeners. The present results, however, suggest that they might be useful to assess residual compression in listeners with presbyacusis, who are mostly affected by high-frequency loss.

V. CONCLUSIONS

(1) The correlation between individual compression exponent estimates inferred from TMCs and DPOAEs is reasonably high at 4 kHz, but low at 0.5 and 1 kHz. Both methods suggest that maximum compression is approximately 10:1 and constant across the characteristic frequency range from 0.5 to 4 kHz.
(2) The low correlation at low frequencies cast doubts on the postulates and interpretation of I/O curves inferred with either (or both) of the two methods. The most likely reason for the lack of correlation at low frequencies (0.5–1 kHz) is the presence of notches and plateaus in the DPOAE I/O curves. This suggests that the DPOAE stimulus paradigm of Kummer *et al.* (1998) may not be optimal (i.e., does not produce maximum DP magnitude) at low frequencies.
(3) A high correlation was found between estimates of compression threshold inferred from DPOAEs and TMC-based I/O curves between 1 and 4 kHz. The DPOAE and the TMC methods indicate that the compression threshold equals 40 and 47 dB SPL, respectively, and is approximately constant across the range of frequencies from 0.5 to 4 kHz for TMCs and from 1 to 4 kHz for DPOAEs.
(4) Cochlear gain and return-to-linearity thresholds were inferred from the TMCs only. Both parameters increased by ~16 dB with increasing characteristic frequency from 0.5 to 4 kHz.
(5) It seems reasonable to use TMCs and DPOAE I/O curves interchangeably to infer cochlear I/O curves at 4 kHz but doubts exist that the same applies to lower frequencies of 0.5 and 1 kHz.

ACKNOWLEDGMENTS

The authors are grateful for the comments and suggestions provided by two anonymous reviewers to improve an earlier version of this paper. This work was supported by IMSERSO 131/06, PROFIT CIT-390000-2005-4, MEC BFU-2006-07536, and The Oticon Foundation.

Avan, P., Bonfils, P., Gilain, L., and Mom, T. (**2003**). "Physiopathological significance of distortion-product otoacoustic emissions at 2f1-f2 produced by high-versus low-level stimuli," J. Acoust. Soc. Am. **113**, 430–441.

Bacon, S. P. (**2004**). "Overview of auditory compression," in *Compression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper (Springer, New York), Chap. 1, pp. 1–17.

Cooper, N. P., and Rhode, W. S. (**1997**). "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," J. Neurophysiol. **78**, 261–270.

Dorn, P. A., Konrad-Martin, D., Neely, S. T., Keefe, D. H., Cyr, E., and Gorga, M. P. (**2001**). "Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears," J. Acoust. Soc. Am. **110**, 3119–3131.

Gaskill, S. A., and Brown, A. M. (**1990**). "The behavior of the acoustic distortion product, 2f1-f2, from the human ear and its relation to auditory sensitivity," J. Acoust. Soc. Am. **88**, 821–839.

Gaskill, S. A., and Brown, A. M. (**1996**). "Suppression of human acoustic distortion product: Dual origin of 2f1-f2," J. Acoust. Soc. Am. **100**, 3268–3274.

Gorga, M. P., Neely, S. T., Dierking, D. M., Kopun, J., Jolkowski, K., Groenenboom, K., Tan, H., and Stiegemann, B. (**2007**). "Low-frequency and high-frequency cochlear nonlinearity in humans," J. Acoust. Soc. Am. **122**, 1671–1680.

Gorga, M. P., Neely, S. T., Dierking, D. M., Kopun, J., Jolkowski, K., Groenenboom, K., Tan, H., and Stiegemann, B. (**2008**). "Low-frequency and high-frequency distortion product otoacoustic emission suppression in humans," J. Acoust. Soc. Am. **123**, 2172–2190.

He, N. J., and Schmiedt, R. A. (**1993**). "Fine structure of the 2f1-f2 acoustic distortion product: Changes with primary level," J. Acoust. Soc. Am. **94**, 2659–2669.

Heinz, M. G., and Young, E. D. (**2004**). "Response growth with sound level in auditory-nerve fibers after noise induced hearing loss," J. Neurophysiol. **91**, 784–795.

Heitmann, J., Waldmann, B., Schnitzler, H., Plinkert, P. K., and Zenner, H. (**1998**). "Suppression of distortion product otoacoustic emissions (DPOAE) near 2f1-f2 removes DP-gram fine structure—Evidence for a secondary generator," J. Acoust. Soc. Am. **103**, 1527–1531.

Johnson, T. A., Neely, S. T., Garner, C. A., and Gorga, M. P. (**2006**). "Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions," J. Acoust. Soc. Am. **119**, 418–428.

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

Kummer, P., Janssen, T., and Arnold, W. (**1995**). "Suppression tuning characteristics of the 2f1-f2 distortion product otoacoustic emission in humans," J. Acoust. Soc. Am. **98**, 197–210.

Kummer, P., Janssen, T., and Arnold, W. (**1998**). "The level and growth behavior of the 2F1-F2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," J. Acoust. Soc. Am. **103**, 3431–3444.

Kummer, P., Janssen, T., Hulin, P., and Arnold, W. (**2000**). "Optimal L1-L2 primary tone level separation remains independent of test frequency in humans," Hear. Res. **146**, 47–56.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 466–477.

Liberman, M. C., Zuo, J., and Guinan, J. J., Jr. (**2004**). "Otoacoustic emissions without somatic motility: Can stereocilia mechanics drive the mammalian cochlea?," J. Acoust. Soc. Am. **116**, 1649–1655.

Lopez-Poveda, E. A., and Alves-Pinto, A. (**2008**). "A variant temporal-masking-curve method for inferring peripheral auditory compression," J. Acoust. Soc. Am. **123**, 1544–1554.

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (**2003**). "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," J. Acoust. Soc. Am. **113**, 951–960.

Lopez-Poveda, E. A., Plack, C. J., Meddis, R., and Blanco, J. L. (**2005**). "Cochlear compression in listeners with moderate sensorineural hearing loss," Hear. Res. **205**, 172–183.

Lukashkin, A. N., Lukashkina, V. A., and Russell, I. J. (**2002**). "One source for distortion product otoacoustic emissions generated by low- and high-level primaries," J. Acoust. Soc. Am. **111**, 2740–2748.

Lukashkin, A. N., and Russell, I. J. (**2002**). "Modifications of a single saturating non-linearity account for post-onset changes in 2f1-f2 distortion product otoacoustic emission," J. Acoust. Soc. Am. **112**, 1561–1568.

Mauermann, M., and 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.

Mauermann, M., and Kollmeier, B. (**2004**). "Distortion product otoacoustic emission (DPAOE) input/output functions and the influence of the second DPOAE source," J. Acoust. Soc. Am. **116**, 2199–2212.

Mauermann, M., Uppenkamp, S., van Hengel, P. W. J., and Kollmeier, B. (**1999**). "Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of frequency ratio f2/f1," J. Acoust. Soc. Am. **106**, 3473–3483.

Meddis, R., and O'Mard, L. P. (**2005**). "A computer model of the auditory-nerve response to forward-masking stimuli," J. Acoust. Soc. Am. **117**, 3787–3798.

Mills, D. (**1997**). "Interpretation of distortion product otoacoustic emission measurements. I. Two stimulus tones," J. Acoust. Soc. Am. **102**, 413–429.

Moore, B. C. J. (**2003**). *An Introduction to the Psychology of Hearing*, 5th ed. (Academic, London).

Müller, J., and Janssen, T. (**2004**). "Similarity in loudness and distortion product otoacoustic emission input/output functions: Implications for an objective hearing aid adjustment," J. Acoust. Soc. Am. **115**, 3081–3091.

Neely, S. T., Gorga, M. P., and Dorn, P. A. (**2003**). "Cochlear compression estimates from measurements of distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **114**, 1499–1507.

Neely, S. T., Johnson, T. A., and Gorga, M. P. (**2005**). "Distortion-product otoacoustic emission measured with continuously varying stimulus level," J. Acoust. Soc. Am. **117**, 1248–1259.

Nelson, D. A., and Schroder, A. C. (**2004**). "Peripheral compression as a function of stimulus level and frequency region in normal-hearing listeners," J. Acoust. Soc. Am. **115**, 2221–2233.

Nelson, D. A., Schroder, A. C., and Wojtczak, M. (**2001**). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **110**, 2045–2064.

Nuttall, A. L., and Dolan, D. F. (**1993**). "Two-tone suppression of inner hair cell and basilar membrane responses in the guinea pig," J. Acoust. Soc. Am. **93**, 390–400.

Oxenham, A. J., and Moore, B. C. J. (**1994**). "Modeling the additivity of nonsimultaneous masking," Hear. Res. **80**, 105–118.

Plack, C. J., and Drga, V. (**2003**). "Psychophysical evidence for auditory compression at low characteristic frequencies," J. Acoust. Soc. Am. **113**, 1574–1586.

Plack, C. J., Drga, V., and Lopez-Poveda, E. A. (**2004**). "Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss," J. Acoust. Soc. Am. **115**, 1684–1695.

Plack, C. J., and Oxenham, A. J. (**1998**). "Basilar-membrane nonlinearity and the growth of forward masking," J. Acoust. Soc. Am. **103**, 1598–1608.

Rhode, W. S. (**2007**). "Distortion product otoacoustic emissions and basilar membrane vibration in the 6–9 kHz region of sensitive chinchilla cochleae," J. Acoust. Soc. Am. **122**, 2725–2737.

Rhode, W. S., and Cooper, N. P. (**1996**). "Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*," Aud. Neurosci. **3**, 101–121.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Rosengard, P. S., Oxenham, A. J., and Braida, L. D. (**2005**). "Comparing different estimates of cochlear compression in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **117**, 3028–3041.

Shera, C. A., and 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.

Siegel, J. H. (**1994**). "Ear-canal standing waves and high frequency calibration using otoacoustic emission probes," J. Acoust. Soc. Am. **95**, 2589–2597.

Stainsby, T. H., and Moore, B. C. J. (**2006**). "Temporal masking curves for hearing-impaired listeners," Hear. Res. **218**, 98–111.

Stover, J. S., Neely, S. T., and Gorga, M. P. (**1996**). "Latency and multiple sources of distortion product otoacoustic emissions," J. Acoust. Soc. Am. **99**, 1016–1024.

Talmadge, C., Tubis, A., Long, G. R., and Pisorski, P. (**1998**). "Modeling otoacoustic emission and hearing threshold fine structure in humans," J. Acoust. Soc. Am. **104**, 1517–1543.

Talmadge, C. L., Long, G. R., Tubis, A., and 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.

Whitehead, M. L., Stagner, B., Lonsbury-Martin, B. L., and Martin, G. K. (**1995**). "Effects of ear-canal standing waves on measurements of distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **98**, 3200–3214.

Williams, E. J., and Bacon, S. P. (**2005**). "Compression estimates using behavioral and otoacoustic emission measures," Hear. Res. **201**, 44–54.

Wojtczak, M., and Oxenham, A. J. (**2007**). "Verification of the assumption of frequency-independent recovery from forward masking," J. Acoust. Soc. Am. **121**, 3133.

Yasin, I., and Plack, J. P. (**2003**). "The effect of a high-frequency suppressor on tuning curves and derived basilar-membrane response functions," J. Acoust. Soc. Am. **114**, 322–332.

Yasin, I., and Plack, J. P. (**2007**). "The effect of low- and high-frequency suppressors on psychophysical estimates of basilar-membrane compression and gain," J. Acoust. Soc. Am. **121**, 2832–2841.

JARO 10: 511–523 (2009)
DOI: 10.1007/s10162-009-0176-9

JARO
Journal of the Association for Research in Otolaryngology

Otoacoustic Emission Theories and Behavioral Estimates of Human Basilar Membrane Motion Are Mutually Consistent

Enrique A. Lopez-Poveda[1] and Peter T. Johannesen[1]

[1]*Unidad de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, C/Pintor Fernando Gallego 1, 37007, Salamanca, Spain*

Received: 2 March 2009; Accepted: 26 May 2009; Online publication: 13 June 2009

ABSTRACT

When two pure tones (or primaries) of slightly different frequencies (f_1 and f_2) are presented to the ear, new frequency components are generated by nonlinear interaction of the primaries within the cochlea. These new components can be recorded in the ear canal as otoacoustic emissions (OAE). The level of the $2f_1-f_2$ OAE component is known as the distortion product otoacoustic emission (DPOAE) and is regarded as an indicator of the physiological state of the cochlea. The current view is that maximal level DPOAEs occur for primaries that produce equal excitation at the f_2 cochlear region, but this notion cannot be directly tested in living humans because it is impossible to record their cochlear responses while monitoring their ear canal DPOAE levels. On the other hand, it has been claimed that the temporal masking curve (TMC) method of inferring human basilar membrane responses allows measurement of the levels of equally effective pure tones at any given cochlear site. The assumptions of this behavioral method, however, lack firm physiological support in humans. Here, the TMC method was applied to test the current notion on the conditions that maximize DPOAE levels in humans. DPOAE and TMC results were mutually consistent for frequencies of 1 and 4 kHz and for levels below around 65 dB sound pressure level. This match supports the current view on the generation of maximal level DPOAEs as well as the assumptions of the behavioral TMC method.

Correspondence to: Enrique A. Lopez-Poveda · Unidad de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León · Universidad de Salamanca · C/Pintor Fernando Gallego 1, 37007, Salamanca, Spain. Telephone: +34-923-294500; fax: +34-923-294750; email: ealopezpoveda@usal.es

Keywords: cochlear nonlinearity, DPOAE, auditory masking, psychoacoustics, human physiology, basilar membrane

INTRODUCTION

The human ear is not a high-fidelity system. It distorts acoustic signals within the cochlea (Ruggero 1993). The distortions can be perceived as audible sounds (Goldstein 1967) and are emitted from the cochlea back to the ear canal as otoacoustic emissions (Kemp 1978). Indeed, emitted distortions are a sign of a healthy ear: the weaker the emission, the greater the cochlear damage (Dorn et al. 2001; Lonsbury-Martin and Martin 1990). The level of these emissions also depends on the parameters of the sounds used to evoke them. Typically, two pure tones (or primaries) of slightly different frequencies (f_1 and f_2; $f_2/f_1 \sim 1.2$) are used and the level of the $2f_1-f_2$ emitted distortion at the ear canal is regarded as an indicator of the physiological state of the cochlea (Gorga et al. 1997). We will refer to this indicator as the distortion product otoacoustic emission (DPOAE). The sensitivity of this measure is greatest when the primaries have levels that evoke maximal level DPOAEs (Mills and Rubel 1994; Whitehead et al. 1995). We will use the term DPOAE optimal rule to refer to the combination of primary levels that evokes the highest level of DPOAEs. Up to now, efforts have been directed to obtain optimal rules *empirically* (Kummer et al. 1998) but the form of the optimal rule is still controversial (Johnson et al. 2006; Kummer et al. 2000).

The controversy could be clarified by elucidating the cochlear mechanical conditions that maximize

DPOAE levels. The overriding view is that maximal level DPOAEs occur when the primaries produce equal excitation at the cochlear region most sensitive to f_2 (Kummer et al. 2000; Neely et al. 2005; Shera and Guinan 2007). Concurrent DPOAE and basilar membrane (BM) recordings have revealed that this view is approximately true for rodents (Rhode 2007), but a confirmation in living humans is not currently feasible because it is not possible to directly record the motion of their BM while monitoring their ear canal DPOAE levels.

On the other hand, it has been claimed that it is possible to infer the levels of two equally effective pure tones at a given cochlear site from behavioral forward masking thresholds. The technique is known as the temporal masking curve (TMC) method and is arguably the most powerful procedure to infer human BM input/output (I/O) curves (Lopez-Poveda et al. 2003; Nelson et al. 2001). The TMC method would seem an appropriate tool to verify the DPOAE generation conjecture of Kummer et al. (2000) in humans. Unfortunately, its assumptions (described below) have been validated only indirectly, using computer models or other psychoacoustical methods, and lack direct physiological support.

A high correlation between DPOAE optimal level rules and corresponding behavioral rules inferred using the TMC method would provide strong support to both the conjecture of Kummer et al. (2000) on the generation of maximal level DPOAEs and the assumptions of the TMC method of inferring human BM responses. The present study aimed at investigating such correlation. It will be shown that a high correspondence exists for frequencies of 1 and 4 kHz and for levels below around 65 dB sound pressure level (SPL).

METHODS

Rationale

A TMC is a plot of the levels of a pure tone (masker) required to just mask a brief following tone (probe) as a function of the time gap between the masker and the probe. The probe level is fixed just above the absolute threshold for the probe. The masker level at the masking threshold increases with increasing time gap and is thought to depend on two variables (Nelson et al. 2001). First, it depends on the time gap: the amount of masking decreases as the masker–probe time gap increases (Duifhuis 1973; Moore and Glasberg 1983; Nelson and Freyman 1987). Second, it depends on the relative excitation produced by the masker and the probe at the BM place tuned at or close to the probe frequency (Nelson et al. 2001; Oxenham et al. 1997; Oxenham and Moore 1995; Oxenham and Plack 1997). Because the probe level is fixed at all times, a TMC is assumed to represent the masker levels required to generate a fixed level of excitation after decaying during the masker–probe time gap. This is why the resulting functions are referred to as isoresponse temporal masking curves or TMCs (Nelson et al. 2001).

There is strong evidence that the rate of recovery from forward masking is approximately the same for different masker frequencies over a wide range of masker levels (Wojtczak and Oxenham 2009). Although this evidence is for a probe frequency of 4 kHz, indirect evidence suggests that the same applies to probe frequencies as low as 0.5 kHz (Lopez-Poveda and Alves-Pinto 2008). Therefore, it seems reasonable to *assume* that for any given masker–probe time gap, two maskers of slightly different frequencies (e.g., f and $f/1.2$) with levels at their masking thresholds produce *identical* degrees of excitation at a cochlear site tuned approximately to the probe frequency. This assumption is commonplace when inferring cochlear I/O curves and compression exponents from TMCs (Lopez-Poveda et al. 2003, 2005; Nelson et al. 2001; Plack et al. 2004; Wojtczak and Oxenham 2009).

Based on the above, our approach consisted in measuring two TMCs, both for a probe frequency equal to the DPOAE test frequency (f_2) and for masker frequencies equal to the DPOAE primary tones (f_1, f_2; with f_2/f_1=1.2). We then plotted the resulting levels for the f_1 masker (L_1) against those for the f_2 masker (L_2), paired according to masker–probe time gap. Based on the previously explained interpretation of TMCs, the resulting plot should illustrate the combination of levels, L_1–L_2, for which two pure tones of frequencies f_1 and f_2 produce approximately comparable degrees of excitation at the f_2 cochlear site. If the current DPOAE generation model (as described by Kummer et al. 2000) and the assumptions of the TMC method are both correct, then this behavioral rule should match with a DPOAE optimal rule obtained empirically.

All human procedures were approved by the human experimentation ethical committee of the University of Salamanca.

Subjects

A total of 14 subjects participated in the study. Their ages ranged from 20 to 39 years. Their hearing was audiometrically normal (i.e., absolute hearing thresholds <20 dB HL) at the three tests frequencies considered in this study (0.5, 1, and 4 kHz). Table 1 details their behavioral absolute thresholds (in decibel sound pressure level) for pure tones of 0.5, 1, and 4 kHz and durations of 10, 110, and 300 ms.

TABLE 1

Thresholds (in decibel sound pressure level) measured with Etymotic ER2 insert earphones for all subjects and for tone durations of 300 ms (absolute threshold), 110 ms (masker threshold), and 10 ms (probe threshold), respectively

Frequency (kHz)	*0.5*	*1*	*4*
Tone duration (ms)	*300/110/10*	*300/110/10*	*300/110/10*
S1	13/15/39	6/8/30	13/13/33
S2	14/21/43	10/15/30	21/21/40
S3	13/16/39	7/8/32	4/10/25
S4	17/20/n.a.	10/16/34	21/24/44
S5	23/24/n.a.	13/17/41	0/3/23
S6	9/10/36	10/11/31	4/4/24
S7	6/10/n.a.	10/11/n.a.	12/15/31
S8	11/18/34	5/7/38	11/10/30
S9	12/n.a./n.a.	9/n.a./n.a.	10/15/31
S10	5/n.a./n.a.	−2/n.a./n.a.	3/1/23
S11	13/13/43	8/5/26	9/n.a./25
S12	22/n.a./n.a.	14/14/36	21/n.a./39
S13	3/7/28	1/3/26	2/n.a./21
S14	17/n.a./n.a.	7/10/30	9/n.a./26

n.a. not available

Behavioral rules

TMCs were measured for probe frequencies (f_P) of 0.5, 1, and 4 kHz and for masker frequencies equal to f_P and $f_P/1.2$. These masker frequencies were equal to those of the primary tones (f_1 and f_2, respectively) used to measure DPOAEs (see below). The masker–probe time gaps, defined as the 0-V period from masker offset to probe onset, ranged from 5 to 100 ms in 5-ms steps with an additional gap of 2 ms. The durations of the masker and the probe were 110 and 10 ms, respectively, including 5-ms cosine-squared onset and offset ramps. The probe had no steady-state portion. The level of the probe was fixed at 9 dB above the individual absolute threshold for the probe as shown in Table 1.

Stimuli were generated with a Tucker Davies Technologies Psychoacoustics Workstation (System 3) operating at a sampling rate of 48.8 kHz and with analog to digital conversion resolution of 24 bits. If needed, signals were attenuated with a programmable attenuator (PA-5) before being output through the headphone buffer (HB-7). Stimuli were presented to the listeners through Etymotic ER-2 insert earphones. TMC SPLs were calibrated by coupling the earphones to a sound level meter through a Zwislocki DB-100 coupler. Calibration was performed at 1 kHz only, and the obtained sensitivity was used at all other frequencies because the earphone manufacturer guarantees an approximately flat (±2 dB) response between 200 Hz and 10 kHz.

Masker levels at masking threshold were measured using a two-interval, two-alternative, forced-choice adaptive procedure with feedback. Two sound intervals were presented to the listener in each trial. One of them contained the masker only and the other contained the masker followed by the probe. The interval containing the probe was selected randomly. The subject was asked to indicate the interval containing the probe. The inter-stimulus interval was 500 ms. The initial masker level was set sufficiently low that the listener always could hear both the masker and the probe. The masker level was then changed according to a two-up, one-down adaptive procedure to estimate the 71% point on the psychometric function (Levitt 1971). An initial step size of 6 dB was applied, which was decreased to 2 dB after three reversals. A total of 15 reversals were measured. Threshold was calculated as the mean of the masker levels at the last 12 reversals. A measurement was discarded if the standard deviation (SD) of the last 12 reversals exceeded 6 dB. Three threshold estimates were obtained in this way and their mean was taken as the threshold. If the SD of these three measurements exceeded 6 dB, a fourth threshold estimate was obtained and included in the mean. Measurements were made in a double-wall sound attenuating booth. Listeners were given at least 2 h of training on the TMC task before data collection began.

The resulting TMCs were least-squares fitted with Eq. (1) of Lopez-Poveda et al. (2005). Individual behavioral level rules were obtained by plotting the fitted levels for the f_1 masker against those for the f_2 masker, paired according to the masker–probe time gaps.

DPOAE optimal rules

The magnitude (in decibel sound pressure level) of the $2f_1-f_2$ DPOAE was measured for f_2 frequencies of 1 and 4 kHz and for a fixed primary frequency ratio of $f_2/f_1=1.2$. Individual DPOAE optimal rules were obtained by systematically varying the levels of the two primaries (L_1 and L_2 for f_1 and f_2, respectively) to find the L_1-L_2 combinations that produced the highest DPOAE response levels. L_2 was varied in 5-dB steps within the range from 35 to 75 dB SPL. For each fixed L_2, L_1 was varied in 3-dB steps and the individual optimal value (i.e., the level that produced the highest DPOAE level) was noted.

The DPOAE magnitude can vary rapidly by changing the test frequency only slightly (Gaskill and Brown 1990). These variations are most clearly seen in a DP gram (i.e., the graphical representation of the DPOAE magnitude as a function of test frequency f_2). They are known as "DPOAE fine structure" and can be as large as 20 dB for an f_2 change of 1/32 octave (He and Schmiedt 1993). The fine structure is thought to occur by vector summation of two DPOAE

contributions: one that originates at the BM region of maximum overlap between the cochlear excitation patterns evoked by the two primaries (i.e., the f_2 region), and one that originates at the cochlear site with characteristic frequency (CF) $\sim 2f_1-f_2$, where the first contribution reflects back to the ear canal. The varying phases of these two contributions give rise to constructive and destructive interference, thus to peaks and valleys in the DP gram (Heitmann et al. 1998; Shera and Guinan 1999).

In an attempt to reduce the potential influence of the fine structure on the DPOAE optimal rules, three such rules were obtained for three f_2 frequencies close to the frequency of interest and their mean was taken as the actual DPOAE optimal rule. The three test frequencies in question were equal to $0.99f$, f, and $1.01f$, where f denotes the frequency of interest. For instance, the final DPOAE optimal rule at 4 kHz was the mean of three optimal rules for f_2 frequencies of 3,960, 4,000, and 4,040 Hz. This procedure was inspired by earlier studies that showed that a "clean" DP gram (i.e., a DP gram without the influence of the fine structure) resembled very closely a moving average of the original DP gram with fine structure (Kalluri and Shera 2001; Mauermann and Kollmeier 2004).

DPOAE I/O curves

DPOAE I/O curves were measured for f_2 frequencies of 1 and 4 kHz with *individual* behavioral and DPOAE optimal level rules, as well as with the rule of Kummer et al. (1998) ($L_1=0.4L_2+39$). I/O curves were also measured for an f_2 of 500 Hz, but using only individual behavioral rules and the rule of Kummer et al. When the rule of Kummer et al. was applied, L_2 ranged from 20 to 75 dB SPL in 5-dB steps, except for $f_2=0.5$ kHz for which it ranged from 45 to 75 dB SPL. The primary frequency ratio was always fixed at $f_2/f_1=1.2$.

To reduce the potential influence of the fine structure on the I/O curves, five such curves were measured for five close f_2 frequencies around the frequency of interest, and the resulting I/O curves were averaged (Johannesen and Lopez-Poveda 2008). For instance, the final DPOAE I/O curve at 4 kHz was the mean of five I/O functions for f_2 frequencies of 3,920, 3,960, 4,000, 4,040, and 4,080 Hz. Three I/O curves were obtained in this way for the behavioral and Kummer rules per f_2 frequency and subject, the mean of which was taken as the "true" I/O curve. For the individual DPOAE optimal rules, only one such I/O curve was measured.

DPOAE measurement procedure

DPOAE measurements were obtained with an Intelligent Hearing System's Smart device (with SmartOAE software version 4.52) equipped with an Etymotic ER-10D probe. During the measurements, subjects sat comfortably in a double-wall sound attenuating chamber and were asked to remain as steady as possible. When seeking DPOAE optimal rules, a recording session consisted of measuring DPOAE responses for all possible primary level combinations (L_1, L_2) for one of the three adjacent f_2 frequencies considered per frequency of interest (see above). When measuring DPOAE I/O curves, a recording session consisted of measuring I/O curves for the five adjacent frequencies considered for each frequency of interest (see above).

The probe fit was checked before and after each recording session. The probe remained in the subject's ear throughout the whole measurement session to avoid measurement variance from probe fit. DPOAEs were measured for a preset measurement time, which ranged from 12 s for high L_2 to 1 min for low L_2. A DPOAE measurement was considered valid when it was 2 SD above the measurement noise floor (defined as the mean level over 10 frequency bins adjacent to the $2f_1-f_2$ component in the OAE spectrum). When a response did not meet this criterion, the measurement was repeated and the measurement time was increased if necessary. The probe remained in the same position during these remeasurements. If the required criterion was not met after successive tries, the measurement point was discarded.

DPOAE measurements were regarded as valid only when they were 6 dB above the system's artifact response. The rationale behind this rather strict criterion and the details of the procedure for controlling for system's artifacts can be found elsewhere (Johannesen and Lopez-Poveda 2008).

RESULTS

Temporal masking curves

Figures 1, 2, and 3 illustrate TMCs for probe frequencies (f_P) of 0.5, 1, and 4 kHz, respectively. Each panel illustrates the TMCs for one subject (as indicated in the top-left corner of the panel) and for two masker frequencies at $f_1=f_P/1.2$ (filled symbols) and $f_2=f_P$ (open symbols).

The characteristics of the present TMCs were overall consistent with those reported elsewhere for similar stimuli (Lopez-Poveda et al. 2003; Nelson and Schroder 2004; Plack and Drga 2003). In broad terms, these are as follows: TMCs were overall steeper for the on-frequency masker (i.e., the masker whose frequency was equal to the probe frequency) than for the off-frequency masker (i.e., the masker whose frequency was below the probe frequency). This difference in

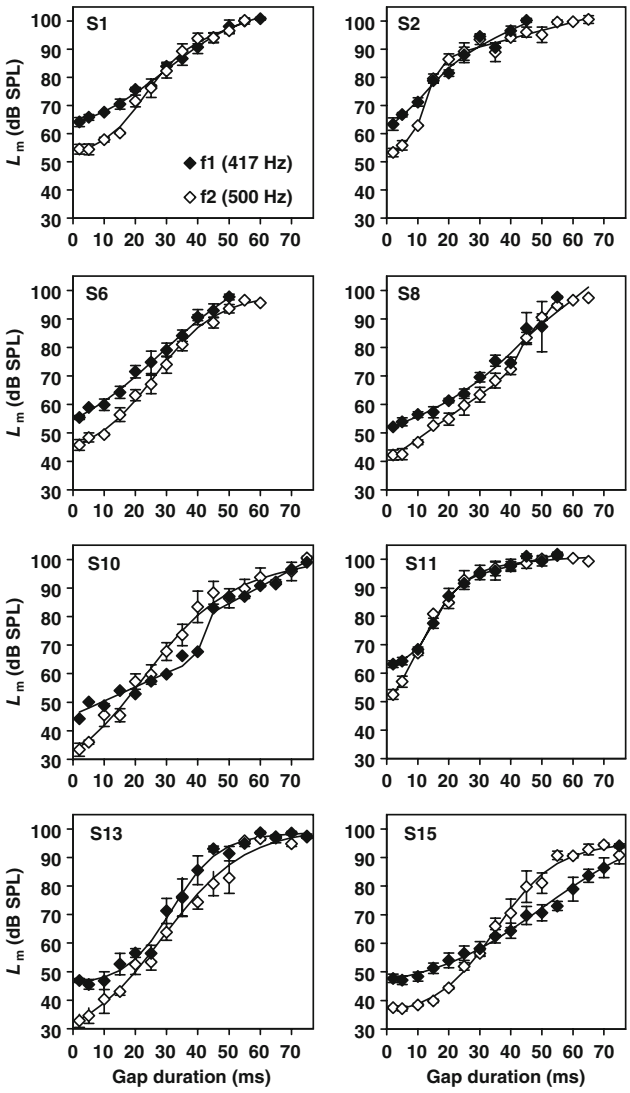

FIG. 1. TMCs for all listeners for probe frequencies (f_p) of 500 Hz. *Each panel* shows data for one subject. *Open symbols* illustrate TMCs for a masker frequency equal to the probe frequency ($f_2=f_P$); *filled symbols* illustrate TMCs for a masker frequency below the probe frequency ($f_1=f_P/1.2$). Each data point is the mean of at least three measurements. *Error bars* illustrate 1 SE of the mean ($n\geq3$). *Lines* illustrate function fits to the data points. L_m masker level.

slope is interpreted to reflect the different rates of growth of the corresponding cochlear responses for stimulus frequencies at or below the CF of the cochlear site tuned to the probe frequency, respectively. That is, the steeper portions of the TMCs are interpreted to reflect shallower growths of BM response with increasing masker level, hence greater degrees of BM compression.

At short masker–probe time gaps, higher levels are required for the off- than for the on-frequency masker to mask the fixed level probe. This is consistent with the fact that, at low levels, the level of a pure tone below the CF must be higher than that of an on-CF tone for both tones to produce equal responses at the cochlear site in question. For moderate-to-long gaps, however, the levels of the lower, off-frequency masker overlap or are even lower than those of the on-frequency masker. This is interpreted to reflect that at high levels, below-CF tones produce comparable or more cochlear excitation than on-CF tones, which is consistent with broader tuning at high levels and with the well-reported basalward shift of cochlear excitation with increasing level for CFs above approximately 1 kHz (Robles and Ruggero 2001; Ruggero et al. 1997). Interestingly, in a few instances (e.g., S10 and S15 in Fig. 1, or S11 in Fig. 2) the TMCs for both maskers crossed again at very long gaps, suggesting that the on-frequency masker became more effective than the off-frequency masker again at very high levels. A similar "rebound" effect can be observed in earlier reports (e.g., Fig. 2 of Lopez-Poveda et al. 2003). The explanation of this result is uncertain. It

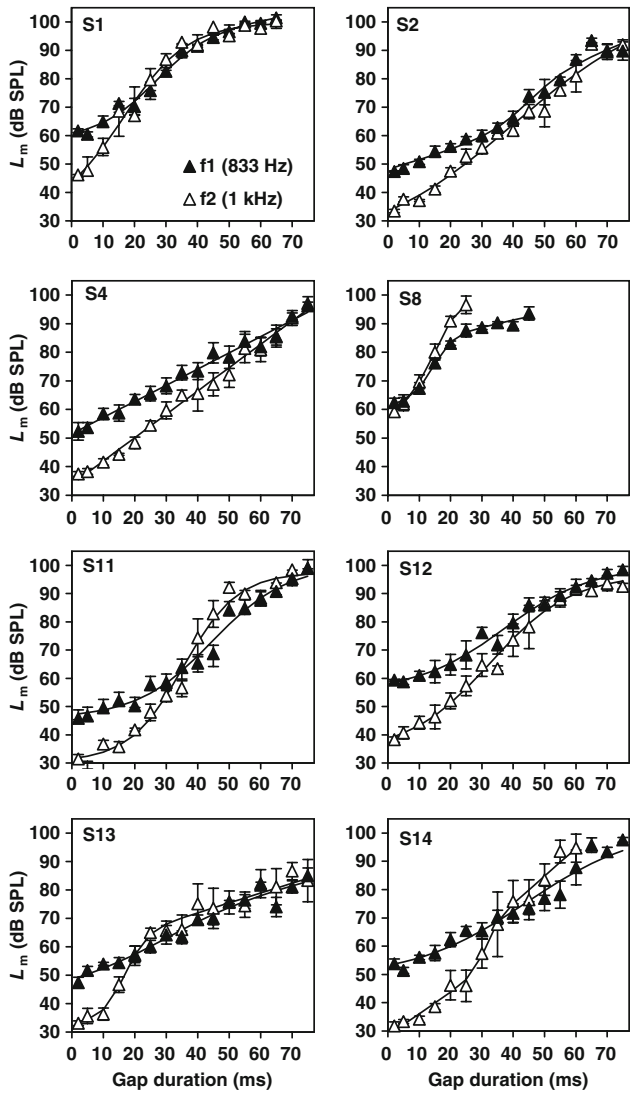

FIG. 2. As Figure 1 but for a probe frequency of 1 kHz.

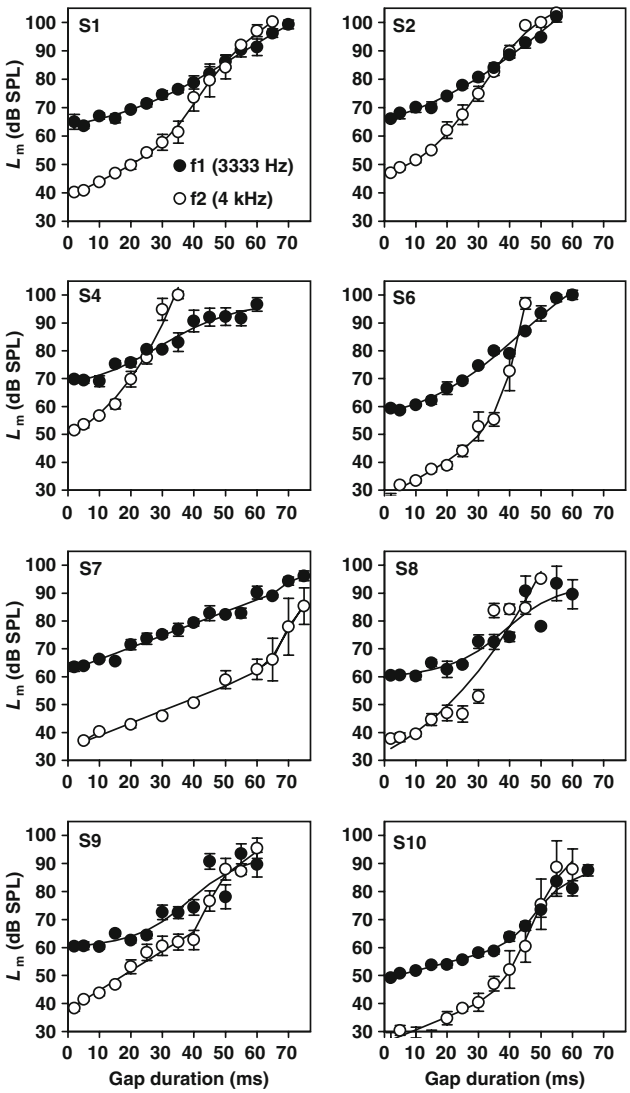

FIG. 3. As Figure 1 but for a probe frequency of 4 kHz.

would be consistent with an apicalward shift of cochlear excitation at very high levels following the previously mentioned basalward shift at moderate levels. Direct BM responses suggest that this shift is possible but existing evidence only applies to apical cochlear regions [see the cross symbols (×) in Figs. 2.3 and 2.4 of Cooper (2004)]. Another possibility would be that the rate of decay of the post-cochlear masker effect becomes slower for the on-frequency masker than for the lower, off-frequency one at very high levels. Hence, for long gaps, the required masker level at threshold would be lower for the on- than for the off-frequency masker. To our knowledge, there is no evidence that this is the case. In fact, existing evidence suggests the opposite (Wojtczak and Oxenham 2009). In any case, the "rebound" effect was rare and occurred over a range of masker levels much higher than the maximum primary level for which DPOAEs could be measured reliably (80 dB SPL). Therefore, it had no effect on the conclusions of the present paper.

More detailed interpretations of TMC characteristics are provided elsewhere (Lopez-Poveda et al. 2003; Nelson et al. 2001).

The influence of the fine structure on DPOAE optimal rules

Figure 4 provides several illustrative examples of the influence of the DPOAE fine structure on the individual DPOAE optimal rules of several subjects

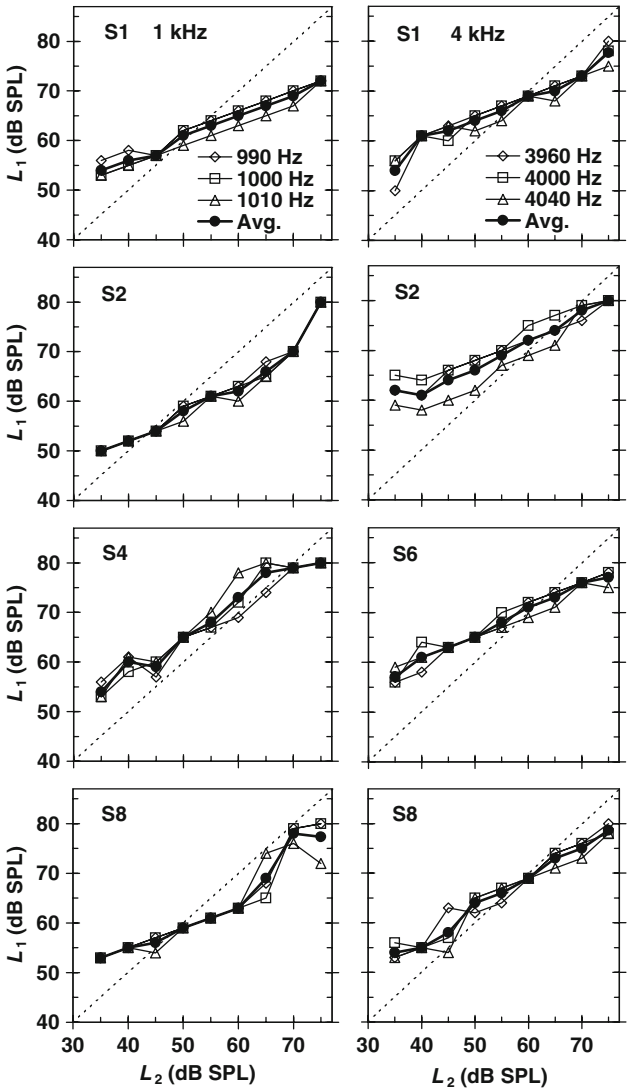

FIG. 4. Examples of the influence of the fine structure on individual DPOAE optimal rules at 1 and 4 kHz (*left and right panels*, respectively). The listener identifier is shown on the *top-left corner* of each panel. *Open symbols* illustrate DPOAE optimal rules for a different test frequency at or around the frequency of interest (as indicated by the *insets in the top panels*). *Filled circles* illustrate the mean curves, which were regarded as the actual DPOAE optimal rules without the influence of the fine structure.

for frequencies of 1 (left panels) and 4 kHz (right panels). Clearly, the level of f_1 (L_1) that evoked the maximal DPOAE response for any given level of f_2 (L_2) changed by as much as 8 dB with a change in f_2 of only 1% for some conditions and subjects. The figure shows that the fine structure could have affected the DPOAE optimal rules, albeit only slightly, and thus justifies our approach to use the mean curve for three adjacent frequencies (illustrated with filled circles) as the DPOAE optimal rule.

Behavioral vs. DPOAE optimal rules

Figures 5 and 6 illustrate plots of the mean levels of the f_1 masker at threshold against those of the f_2 masker, paired according to the masker–probe time gap. Based on the interpretation of TMCs explained in the "Methods" section, these illustrate level combinations of two equally effective maskers at the cochlear site tuned to the probe frequency (f_2). These behavioral rules are compared with individual DPOAE optimal rules (i.e., with primary level combinations, L_1–L_2, that produced maximal DPOAE levels) for primary tone frequencies equal to the masker frequencies. The match between the two rules varied from subject to subject and across frequencies. It was extremely close at 1 kHz for several subjects (e.g., S1, S2, S4, or S14) over an L_2 range typically below 65 dB SPL. The degree of correspondence in the individual data was generally less at 4 kHz, but for several subjects (e.g., S2, S8, or S10) the agreement was reasonably close even at this frequency. The reasons for the lower degree of correspondence above 65 dB SPL will be discussed later.

Individual behavioral rules were based on *mean* values of at least three independent estimates of L_1 and L_2, whereas the DPOAE optimal rules were based on a single L_1 for every L_2 (note that this L_1 was the mean of three estimates, each for a slightly different f_2 around the frequency of interest; see "Methods"). To test for the statistical significance of the difference between the two individual rules, we simply checked if the single DPOAE optimal rule estimate fell within the range of all possible combinations of L_1 and L_2 based on the available TMC data. The latter are illustrated as small gray dots in Figures 5 and 6. Except, perhaps, for S7 at 4 kHz, *all* DPOAE optimal rules always fell within the variability of the behavioral combinations for $L_2 \leq 65$ dB SPL. Subject S7 was peculiar in that he repeatedly reported that the behavioral task was extremely difficult.

The reason for the variability in the behavioral L_1–L_2 combinations (gray dots in Figs. 5 and 6) is uncertain. Such variability reflects, by definition, the variability across TMC estimates (Figs. 2 and 3). Consistent with many previous studies (e.g., Lopez-Poveda et al. 2003, 2005; Nelson et al. 2001; Plack et al. 2004), the variability in the present masker levels was noticeably larger over the steeper portions of the TMCs (Figs. 2 and 3). The steeper portion of a TMC is assumed to reflect the range of levels where the masker is subject to greater cochlear compression (Nelson et al. 2001). Therefore, even small changes in cochlear responses or in the listener's sensitivity

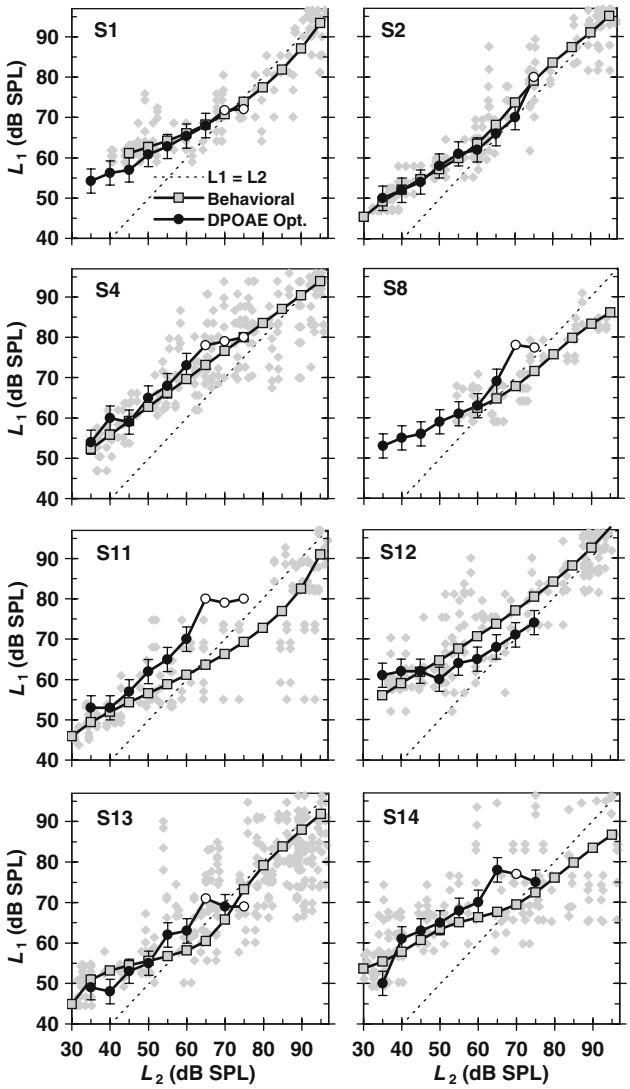

FIG. 5. Comparison of behavioral and DPOAE optimal level rules for a test frequency of 1 kHz. *Each panel* illustrates results for a single subject. *Gray squares* illustrate behavioral rules based on *mean* TMCs. *Small gray dots* illustrate *all* possible L_1–L_2 combinations based on the available TMC data and thus inform of the variability of the behavioral rule based on the data available. *Circles* illustrate DPOAE optimal rules. *Open circles* indicate possible suboptimal L_1–L_2 combinations, that is, combinations that produced the strongest DPOAEs within the L_1 level range of the system (<80 dB SPL), but it is possible that higher DPOAE levels might have been measured with higher L_1s, should the system had allowed them. Error bars are fixed at ±3 dB and indicate the precision of the optimal L_1 (see "Methods" for more details). They are not represented for the *open circles*. *Dotted lines* illustrate equal primary level rules ($L_1 = L_2$).

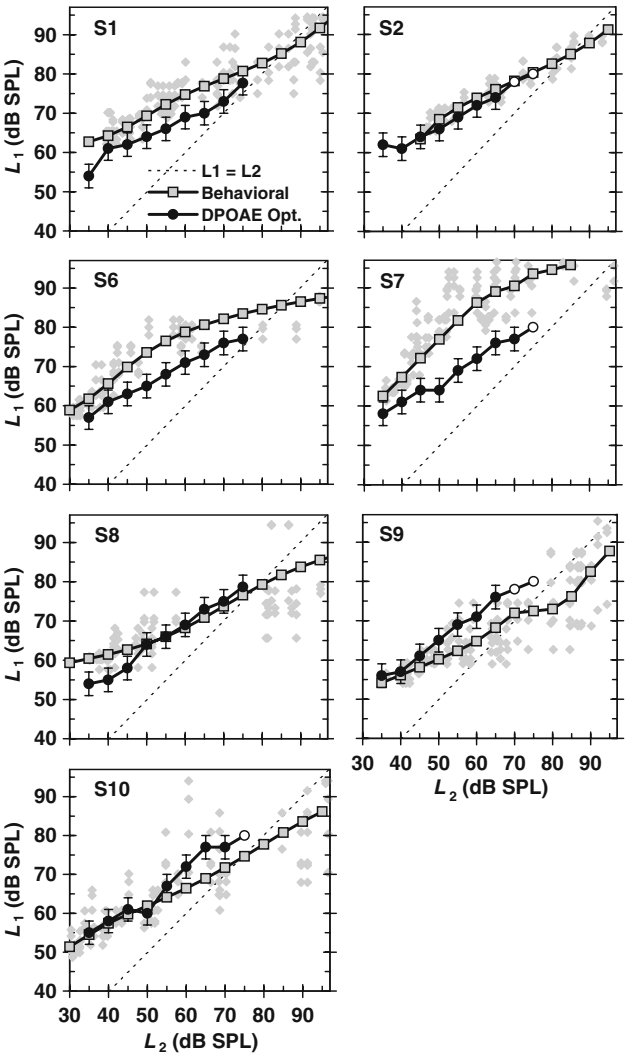

FIG. 6. As Figure 5 but for a test frequency of 4 kHz.

across measurement sessions would produce a large change in masker level at threshold and would explain such variability.

Figure 7 illustrates *mean* rules across listeners. Interestingly, the mean DPOAE optimal rules overlapped with behavioral rules at 1 kHz over the L_2 range for which both sets could be measured reliably (35–65 dB SPL). At 4 kHz, the two rules did not overlap but they were within 1 SD of each other and their difference was not statistically significant for $L_2 \leq$ 65 dB SPL ($p<0.05$, point-by-point, two-tailed, paired t test). Indeed, the difference was accentuated by the results of a single subject (S7).

The dependence of behavioral and DPOAE optimal rules on test frequency

Average behavioral rules (gray squares in Fig. 7) approached equal level for high L_2 levels (~75 dB SPL). The difference between L_1 and L_2 was larger at lower than at higher L_2s and increased gradually with increasing frequency. Straight lines were fitted to the data for $L_2 \leq 65$ dB SPL (thick continuous lines in

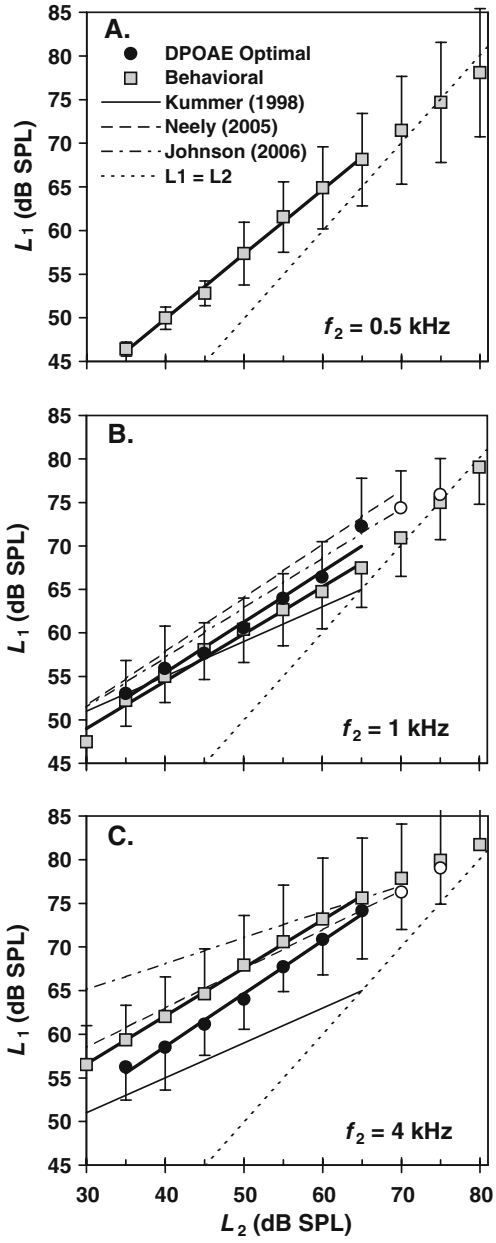

FIG. 7. Average L_1–L_2 rules at different frequencies and as proposed by the present and earlier studies. **A** For f_2=0.5 kHz. **B** For f_2=1 kHz. **C** For f_2=4 kHz. *Circles and squares* illustrate average DPOAE optimal and behavioral rules obtained in the present study, respectively. Each data point is the mean of data for at least three subjects. *Error bars* illustrate one SD ($n \geq 3$, typically 5) *Thick straight lines* illustrate least-square fits to these experimental rules for $L_2 \leq$ 65 dB SPL. *Dotted lines* illustrate equal level rules. Primary level rules proposed by earlier representative studies are also shown for comparison. These are plotted across the L_2 range originally proposed in their original reports. *Open circles* indicate possible suboptimal values (see the legend of Fig. 5).

TABLE 2

Regression parameters for linear relationships, $L_1 = aL_2 + b$, based on the present behavioral and DPOAE optimal rules, for an L_2 range between 30 and 65 dB SPL

Rule	*n*	*Frequency (kHz)*	*b (dB SPL)*	*a*
DPOAE optimal	8	1	31.1	0.61
	7	4	35.6	0.59
		Average	33.4	0.60
Behavioral	8	0.5	23.6	0.69
	8	1	35.1	0.49
	7	4	39.5	0.56
		Average	32.7	0.58

Fig. 7). As indicated in Table 2, the lines for both behavioral and DPOAE optimal rules had similar slopes at 1 and 4 kHz, and these were shallower than the line fitted to the behavioral rule at 0.5 kHz. Behavioral rules, however, had slopes that were statistically indistinct across frequencies ($p>0.05$, two-tailed, equal variance, t test).

DPOAE I/O curves

The growth of the DPOAE magnitude as a function of L_2 was measured for each listener using his/her *individual* DPOAE-optimal and behavioral rules to obtain individual DPOAE I/O curves. The mean I/O curves are shown in Figure 8. The figure also illustrates mean I/O curves measured with the rule of Kummer et al. (1998), which was identical across subjects and frequencies. For the three f_2 frequencies, DPOAEs grew with increasing L_2 at rates considerably lower than 1 dB/dB. DPOAE levels measured using DPOAE optimal rules (filled circles) were the highest and were consistently 3–5 dB higher than those measured with the behavioral rules (gray squares); that is, circles and squares run parallel to each other across the L_2 level range.

The DPOAE levels evoked by the rule of Kummer et al. were identical or lower than those evoked by the behavioral rule at 1 and 4 kHz across levels, except for $L_2<50$ dB SPL. The rule of Kummer et al. evoked slightly higher DPOAE levels than the behavioral rule at 500 Hz (Fig. 8A). The difference between the DPOAE levels measured with the optimal and the Kummer rules increased with increasing L_2 at 1 and 4 kHz.

DISCUSSION

The general overlap between behavioral and empirical DPOAE optimal rules at 1 and 4 kHz (Figs. 5, 6, and 7) for L_2 below approximately 65 dB SPL supports the view that DPOAE levels are highest when the two primaries produce similar responses at the f_2 cochlear region (Kummer et al. 2000).

The present behavioral rules were derived on the assumption that the TMCs for the two maskers (f_1 and f_2) reflect *only* differences in the cochlear excitation

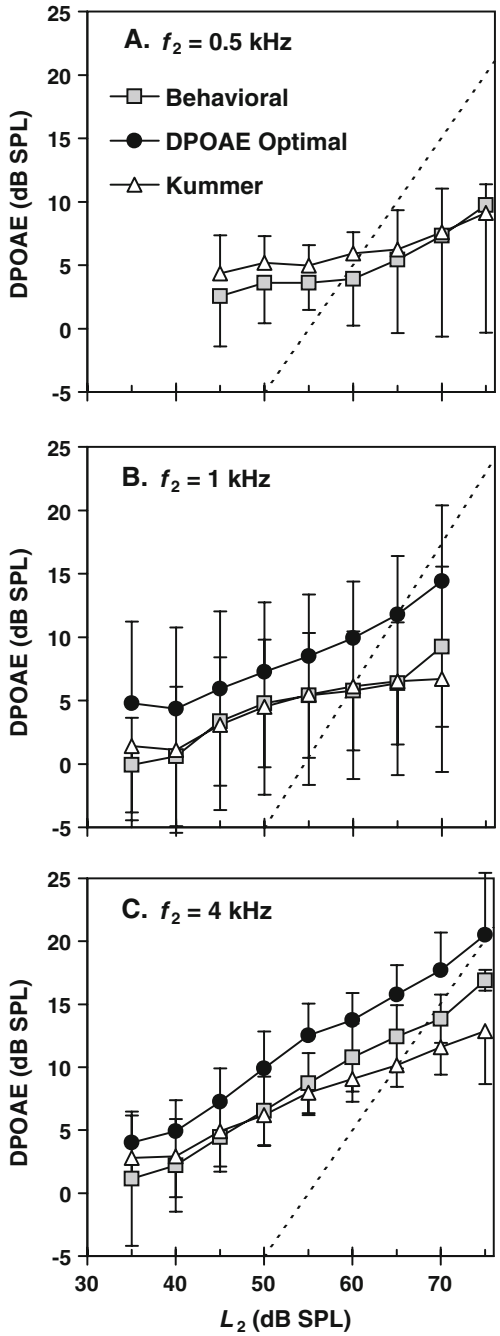

FIG. 8. Average I/O curves for the $2f_1-f_2$ DPOAE for different primary level rules (as indicated in the *inset* of **A**). *Different panels* illustrate results for a different f_2 frequency. **A** f_2=0.5 kHz. **B** f_2=1 kHz. **C** f_2=4 kHz. Every point is the mean of data for between three and seven subjects (typically n=5). *Error bars* illustrate 1 SD. Note that the rule of Kummer et al. was originally designed for $L_2 \leq 65$ dB SPL but was extrapolated here to higher L_2. *Dotted lines* indicate a linear growth with a slope of 1 dB/dB.

evoked by the two tones at the cochlear site tuned to the probe frequency (f_2 in this case). That is, on the assumption that the post-cochlear interaction between the masker and the probe is linear and identical for the two maskers, for all masker–probe time gaps and masker levels. This assumption is commonplace when inferring cochlear I/O functions from TMCs (Lopez-Poveda et al. 2003; Nelson et al. 2001) and is supported by modeling (Oxenham and Moore 1994) and experimental studies, at least for levels below approximately 83 dB SPL (Lopez-Poveda and Alves-Pinto 2008; Wojtczak and Oxenham 2009). The post-cochlear interaction may be (slightly) different for masker frequencies that are an octave apart at higher masker levels (Lopez-Poveda and Alves-Pinto 2008; Wojtczak and Oxenham 2009), but this is unlikely to undermine the validity of the present approach because the present maskers were closer in frequency and conclusions are claimed to be valid only for $L_2 \leq \sim$65 dB SPL. Furthermore, given that the behavioral and DPOAE optimal rules were inferred using fundamentally different assumptions and methods, the correspondence between the two provides further support for the assumptions of the behavioral TMC method, at least below 65 dB SPL.

The level of the $2f_1-f_2$ DPOAE measured in the ear canal is almost certainly the sum of contributions from several cochlear sources and generation mechanisms (Shaffer et al. 2003). These include distortion generated by nonlinear interaction of the primaries in the f_2 cochlear region (Martin et al. 1998), reflection of this distortion at the $2f_1-f_2$ cochlear site (Kalluri and Shera 2001), and nonlinear interaction between f_2 and the first harmonic of f_1 ($2f_1$) at a more basal ($2f_1$) cochlear site (Fahey et al. 2000). The relative weight of these contributions to the measured DPOAE is uncertain. The present behavioral rules were obtained from TMCs for probe frequencies equal to the DPOAE test frequencies (f_2). Based on the current interpretation of TMCs (Lopez-Poveda et al. 2003; Lopez-Poveda and Alves-Pinto 2008; Nelson et al. 2001), the behavioral rules thus reflect L_1–L_2 combinations for which the two maskers (f_1, f_2) produce equal responses at a cochlear site with a CF$\sim f_2$. Therefore, the match between behavioral and DPOAE optimal rules for levels below 65 dB SPL (Figs. 5, 6, and 7) together with the evidence that DPOAEs originate at multiple cochlear locations suggest that the DPOAE contribution from the f_2 region is dominant and/or that the DPOAEs originated at the other sites are proportional to the contribution generated at the f_2 region.

Given the reasonable match between mean behavioral and DPOAE optimal rules (Fig. 7B, C), it is unclear why the DPOAE levels were on average 3–5 dB higher for the DPOAE optimal than for the behavioral rules (Fig. 8B, C). Recall that the DPOAE I/O curves of Figure 8 represent the mean of the curves obtained with *individual* DPOAE optimal and behavioral rules. One possibility is that mutual suppression between the primaries may have affected the results. One fundamental difference between the *individual* behavioral and DPOAE optimal rules is that the latter were obtained from DPOAE measurements that required the *simultaneous* presentation of the two primary tones, while behavioral rules were inferred from single-tone responses. In other words, DPOAE optimal rules implicitly take into account possible nonlinear interactions (e.g., suppression) between the primary tones that are disregarded by the behavioral rules. Concurrent recordings of DPOAE and basilar membrane responses in chinchilla suggest that DPOAE levels are submaximal for L_1–L_2 combinations that produce equal cochlear responses of simultaneously presented primaries, and this possibly occurs because the primary tone f_1 suppresses the response of the basilar membrane to f_2 [Fig. 1 of Rhode (2007)]. Given that the behavioral rule reflects an equal response criterion for nonsimultaneous primaries, this might explain why the mean DPOAE levels for the behavioral rules were consistently lower (on average) than those measured with the optimal rules. That said, the same suppression mechanism would have led to behavioral L_1 values being consistently lower than DPOAE optimal L_1 values and this was not the case (Figs. 5, 6, and 7). Therefore, mutual suppression is unlikely to account for the difference in DPOAE levels evoked by the two rules [see also the discussion of Kummer et al. (2000)].

An alternative simpler explanation is that individual DPOAE optimal rules were, by definition, those that evoked the highest DPOAE levels (within the 3-dB precision considered for L_1, see "Methods"). Therefore, any deviation, however small, of the individual behavioral rules from the individual DPOAE optimal rule (Figs. 5 and 6) would have always produced *submaximal* DPOAE levels for each subject and this would be reflected in the mean I/O curves (Fig. 8).

The correspondence between the present behavioral and DPOAE optimal rules tended to be less in the individual data (Figs. 5 and 6) for L_2 above around 65 dB SPL. In several cases (depicted by open circles in Figs. 5 and 6), the optimal L_1 levels were higher than the maximum output of our system and most of these were higher than the corresponding behavioral L_1 values. The reason for this result is uncertain, but it may reflect a shift of the DPOAE generation cochlear site towards the base of the cochlea with increasing L_2. It is reported that the peak of the cochlear traveling wave shifts basally with increasing sound level (Robles and Ruggero 2001).

Therefore, the region of maximum interaction between the traveling waves evoked by the two primaries is likely to shift from that with CF~f_2 to a more basal region as L_2 increases. This shift is illustrated in Figure 9, where the regions of maximal interaction at low and high L_2 levels are denoted x_2 and x_2', respectively. Figure 9 also illustrates how the increase in L_1 for a given increase in L_2 would be greater if the two primaries were to evoke equal responses at the x_2' than at the x_2 cochlear regions. The present behavioral rules were unlikely affected by the shift in question because they were based on TMCs for *fixed*, low level probes and thus presumably reflected cochlear responses at a *fixed* cochlear site with CF~f_2 at all L_2 (Nelson et al. 2001). If the highest DPOAE levels occurred for equally effective primaries near the peak of the f_2 traveling waves at each L_2 level, then DPOAE optimal L_1 values would be higher than behavioral values at high L_2 levels (Fig. 9). This might explain why the behavioral L_1 were sometimes lower than the DPOAE optimal L_1 at high L_2 levels. Furthermore, if the level-dependent shift in question were gradual, then this would also explain why some of the DPOAE optimal L_1–L_2 rules appeared generally steeper than their behavioral counterparts over the range of L_2 levels where both of them could be measured (Figs. 5 and 6).

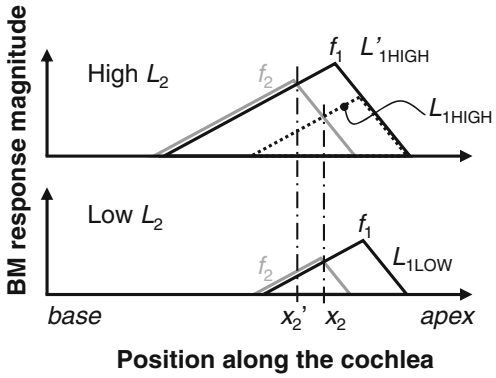

FIG. 9. Schematic cochlear excitation patterns of the two DPOAE primary tones, f_1 and f_2, for low and high L_2 levels (*lower and higher figures,* respectively). For low L_2 levels, the maximum interaction between the two excitation patterns occurs at the x_2 cochlear region whose CF is approximately equal to f_2. At high L_2 levels, however, the excitation patterns of the two primaries shift towards the cochlear base and the new region of maximum interaction, denoted x_2', is basal to x_2. For the high L_2 condition, two excitation patterns are shown for the f_1 primary: one that produces approximately the same excitation as f_2 at the x_2 site (*dotted dark line*), and one that produces the same excitation as f_2 at the x_2' region (*continuous dark line*). The level of f_1, L_1, would be lower in the former than in the latter case. These are denoted as L_{1HIGH} and $L'_{1HIGH}x_2'$, respectively, with $L'_{1HIGH} > L_{1HIGH} > L_{1LOW}$. The assumptions of the TMC method suggest that behavioral rules illustrate L_{1HIGH} vs. L_2, while DPOAE optimal rule might be revealing L'_{1HIGH} vs. L_2. See text for further details.

The theory that the *main* DPOAE generation site shifts basally with increasing L_2 might be tested by comparing the degree of correlation between individual optimal level rules with behavioral rules measured with the present method and with the growth-of-masking (GOM) method for a signal frequency equal to f_2 and a masker frequency equal to f_1 (Nelson et al. 2001; Oxenham and Plack 1997). Based on the current interpretation of GOM functions (Oxenham and Plack 1997; Rosengard et al. 2005), the prediction would be that GOM functions would provide a more accurate behavioral correlate of optimal DPOAE level rules because they implicitly encompass potential level-dependent shifts of cochlear excitation (Lopez-Poveda and Johannesen, in press).

On the controversy about DPOAE optimal primary level rules

Kummer and colleagues have argued that "optimizing the L_1 level for any given L_2 is not a trivial DPOAE level maximization but rather appropriate for maximizing the sensitivity of DPOAE measurements," to discriminate between healthy and damaged cochleae [p. 54 of Kummer et al. (2000)]. It is uncertain that DPOAE optimal rules for normal-hearing subjects serve to maximize DPOAE levels of hearing-impaired subjects. It is also unlikely that *average* DPOAE optimal rules account for *individual* DPOAE level variations. Nevertheless, average DPOAE optimal rules for normal-hearing subjects may be regarded as the "best-guess" parameters for any given normal-hearing or hearing-impaired individual. Because of this and given its potential clinical implications, much effort has been spent on providing an accurate *average* DPOAE optimal rule. Unfortunately, different studies disagree in their conclusions. There exists consensus that the optimal L_1 should increase with increasing L_2 following a linear relationship ($L_1 = aL_2 + b$), but some studies have concluded that a and b should be constant across the f_2 range from 1 to 8 kHz (Kummer et al. 2000) while others have concluded that they should vary rather systematically with f_2 (Johnson et al. 2006; Neely et al. 2005).

Although the present study was not aimed at resolving this controversy, the behavioral and DPOAE data it produced incidentally support the view that optimal rules should be approximately similar at 1 and 4 kHz (Fig. 7B, C and Table 2). The behavioral data suggest, however, that the optimal rule at 500 Hz is likely to be significantly different from those at higher frequencies (Fig. 7 and Table 2). Unfortunately, this could not be corroborated with empirical DPOAE optimal rules at this frequency. That said, the I/O curves of Figure 8 suggest that the rule of Kummer et al. (2000), which was identical across test frequencies, evoked *average* DPOAE levels that were

indistinguishable for most conditions from those evoked by the present behavioral rules, despite them being different, particularly at 500 Hz.

In summary, the agreement between the behavioral and DPOAE optimal rules supports the view that maximal DPOAE levels occur when the two primary tones produce approximately equal responses in the f_2 cochlear region as well as the assumptions of the popular TMC method of inferring human BM responses.

ACKNOWLEDGMENTS

We thank Almudena Eustaquio-Martin for technical support, and the associate editor, Barbara Shinn-Cunningham, and three anonymous reviewers for their thoughtful comments on an earlier version of this paper. This work was supported by the Spanish Ministry of Science and Education (refs. BFU2006-07536 and CIT-390000-2005-4) and by The William-Demant Oticon Foundation.

REFERENCE

COOPER NP. Compression in the peripheral auditoy system. In: Bacon SP, Fay RR, Popper AN (eds) Compression: From Cochlea to Cochlear Implants. New York, Springer, pp. 18–61, 2004.

DORN PA, KONRAD-MARTIN D, NEELY ST, KEEFE DH, CYR E, GORGA MP. Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears. J. Acoust. Soc. Am. 110:3119–3131, 2001.

DUIFHUIS H. Consequences of peripheral frequency selectivity for nonsimultaneous masking. J. Acoust. Soc. Am. 54:1471–1488, 1973.

FAHEY PF, STAGNER BB, LONSBURY-MARTIN BL, MARTIN GK. Nonlinear interactions that could explain distortion product interference response areas. J. Acoust. Soc. Am. 108:1786–1802, 2000.

GASKILL SA, BROWN AM. The behavior of the acoustic distortion product, 2f1–f2, from the human ear and its relation to auditory sensitivity. J. Acoust. Soc. Am. 88:821–839, 1990.

GOLDSTEIN JL. Auditory nonlinearity. J. Acoust. Soc. Am. 41:676–689, 1967.

GORGA MP, NEELY ST, OHLRICH B, HOOVER B, REDNER J, PETERS J. From laboratory to clinic: a large scale study of distortion product otoacoustic emissions in ears with normal hearing and ears with hearing loss. Ear Hear. 18:440–455, 1997.

HE NJ, SCHMIEDT RA. Fine structure of the 2f1–f2 acoustic distortion product: changes with primary level. J. Acoust. Soc. Am. 94:2659–2669, 1993.

HEITMANN J, WALDMANN B, SCHNITZLER H, PLINKERT PK, ZENNER H. Suppression of distortion product otoacoustic emissions (DPOAE) near 2f1–f2 removes DP-gram fine structure—evidence for a secondary generator. J. Acoust. Soc. Am. 103:1527–1531, 1998.

JOHANNESEN PT, LOPEZ-POVEDA EA. Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions. J. Acoust. Soc. Am. 124:2149–2163, 2008.

JOHNSON TA, NEELY ST, GARNER CA, GORGA MP. Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions. J. Acoust. Soc. Am. 119:418–428, 2006.

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

KEMP DT. Stimulated acoustic emissions from within the human auditory system. J. Acoust. Soc. Am. 64:1386–1391, 1978.

KUMMER P, JANSSEN T, ARNOLD W. The level and growth behavior of the 2 f1–f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss. J. Acoust. Soc. Am. 103:3431–3444, 1998.

KUMMER P, JANSSEN T, HULIN P, ARNOLD W. Optimal L(1)–L(2) primary tone level separation remains independent of test frequency in humans. Hear Res. 146:47–56, 2000.

LEVITT H. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49(Suppl 2)1971.

LONSBURY-MARTIN BL, MARTIN GK. The clinical utility of distortion-product otoacoustic emissions. Ear Hear. 11:144–154, 1990.

LOPEZ-POVEDA EA, ALVES-PINTO A. A variant temporal-masking-curve method for inferring peripheral auditory compression. J. Acoust. Soc. Am. 123:1544–1554, 2008.

LOPEZ-POVEDA EA, JOHANNESEN PT. Otoacoustic emission theories can be tested with behavioral methods. In: Lopez-Poveda EA, Palmer AR, Meddis R (eds) Advances in Auditory Research: Physiology, Psychophysics, and Models. New York, Springer.

LOPEZ-POVEDA EA, PLACK CJ, MEDDIS R. Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing. J. Acoust. Soc. Am. 113:951–960, 2003.

LOPEZ-POVEDA EA, PLACK CJ, MEDDIS R, BLANCO JL. Cochlear compression in listeners with moderate sensorineural hearing loss. Hear Res. 205:172–183, 2005.

MARTIN GK, JASSIR D, STAGNER BB, WHITEHEAD ML, LONSBURY-MARTIN BL. Locus of generation for the 2f1–f2 vs 2f2–f1 distortion-product otoacoustic emissions in normal-hearing humans revealed by suppression tuning, onset latencies, and amplitude correlations. J. Acoust. Soc. Am. 103:1957–1971, 1998.

MAUERMANN M, KOLLMEIER B. Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source. J. Acoust. Soc. Am. 116:2199–2212, 2004.

MILLS DM, RUBEL EW. Variation of distortion product otoacoustic emissions with furosemide injection. Hear Res. 77:183–199, 1994.

MOORE BC, GLASBERG BR. Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise. J. Acoust. Soc. Am. 73:1249–1259, 1983.

NEELY ST, JOHNSON TA, GORGA MP. Distortion-product otoacoustic emission measured with continuously varying stimulus level. J. Acoust. Soc. Am. 117:1248–1259, 2005.

NELSON DA, FREYMAN RL. Temporal resolution in sensorineural hearing-impaired listeners. J. Acoust. Soc. Am. 81:709–720, 1987.

NELSON DA, SCHRODER AC. Peripheral compression as a function of stimulus level and frequency region in normal-hearing listeners. J. Acoust. Soc. Am. 115:2221–2233, 2004.

NELSON DA, SCHRODER AC, WOJTCZAK M. A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 110:2045–2064, 2001.

OXENHAM AJ, MOORE BC. Modeling the additivity of nonsimultaneous masking. Hear Res 80:105–118, 1994.

OXENHAM AJ, MOORE BC. Additivity of masking in normally hearing and hearing-impaired subjects. J. Acoust. Soc. Am. 98:1921–1934, 1995.

OXENHAM AJ, MOORE BC, VICKERS DA. Short-term temporal integration: evidence for the influence of peripheral compression. J. Acoust. Soc. Am. 101:3676–3687, 1997.

OXENHAM AJ, PLACK CJ. A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. J. Acoust. Soc. Am. 101:3666–3675, 1997.

Plack CJ, Drga V. Psychophysical evidence for auditory compression at low characteristic frequencies. J. Acoust. Soc. Am. 113:1574–1586, 2003.

Plack CJ, Drga V, Lopez-Poveda EA. Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss. J. Acoust. Soc. Am. 115:1684–1695, 2004.

Rhode WS. Distortion product otoacoustic emissions and basilar membrane vibration in the 6–9 kHz region of sensitive chinchilla cochleae. J. Acoust. Soc. Am. 122:2725–2737, 2007.

Robles L, Ruggero MA. Mechanics of the mammalian cochlea. Physiol. Rev. 81:1305–1352, 2001.

Rosengard PS, Oxenham AJ, Braida LD. Comparing different estimates of cochlear compression in listeners with normal and impaired hearing. J. Acoust. Soc. Am. 117:3028–3041, 2005.

Ruggero MA. Distortion in those good vibrations. Curr. Biol. 3:755–758, 1993.

Ruggero MA, Rich NC, Recio A, Narayan SS, Robles L. Basilar-membrane responses to tones at the base of the chinchilla cochlea. J. Acoust. Soc. Am. 101:2151–2163, 1997.

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

Shera CA, Guinan JJ, Jr. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J. Acoust. Soc. Am. 105:782–798, 1999.

Shera CA, Guinan JJ, Jr. Cochlear traveling-wave amplification, suppression, and beamforming probed using noninvasive calibration of intracochlear distortion sources. J. Acoust. Soc. Am. 121:1003–1016, 2007.

Whitehead ML, Stagner BB, McCoy MJ, Lonsbury-Martin BL, Martin GK. Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. II. Asymmetry in L1,L2 space. J. Acoust. Soc. Am. 97:2359–2377, 1995.

Wojtczak M, Oxenham AJ. Pitfalls in behavioral estimates of basilar-membrane compression in humans. J. Acoust. Soc. Am. 125:270–281, 2009.

Correspondence between behavioral and individually "optimized" otoacoustic emission estimates of human cochlear input/output curves[a)]

Peter T. Johannesen and Enrique A. Lopez-Poveda[b)]
Unidad de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, 37007 Salamanca, Spain

(Received 14 December 2009; revised 10 March 2010; accepted 10 March 2010)

Previous studies have shown a high within-subject correspondence between distortion product otoacoustic emission (DPOAE) input/output (I/O) curves and behaviorally inferred basilar membrane (BM) I/O curves for frequencies above ~2 kHz. For lower frequencies, DPOAE I/O curves contained notches and plateaus that did not have a counterpart in corresponding behavioral curves. It was hypothesized that this might improve by using individualized optimal DPOAE primary levels. Here, data from previous studies are re-analyzed to test this hypothesis by comparing behaviorally inferred BM I/O curves and DPOAE I/O curves measured with well-established group-average primary levels and two individualized primary level rules: one optimized to maximize DPOAE levels and one intended for primaries to evoke comparable BM responses at the f_2 cochlear region. Test frequencies were 0.5, 1, and 4 kHz. Behavioral I/O curves were obtained from temporal (forward) masking curves. Results showed high within-subject correspondence between behavioral and DPOAE I/O curves at 4 kHz only, regardless of the primary level rule. Plateaus and notches were equally common in low-frequency DPOAE I/O curves for individualized and group-average DPOAE primary levels at 0.5 and 1 kHz. Results are discussed in terms of the adequacy of DPOAE I/O curves for inferring individual cochlear nonlinearity characteristics.
 [DOI: 10.1121/1.3377087]

PACS number(s): 43.64.Jb, 43.66.Dc [BLM] Pages: 3602–3613

I. INTRODUCTION

Humans can perceive sounds over a wide range of sound pressure levels (SPLs) most likely thanks to the functioning of the outer hair cells and their effect on the basilar membrane (BM) vibrations (e.g., Oxenham and Bacon 2003). Indeed, the healthy BM is sensitive to very low-level sounds and the magnitude of its response grows compressively with increasing sound level, thus accommodating a range of ~120 dB SPL to a narrower range of physiological responses (e.g., Robles and Ruggero, 2001). The damaged BM, by contrast, shows reduced sensitivity and linearized responses, which likely explains why hearing-impaired listeners show abnormally high thresholds and narrower dynamic ranges (e.g., Moore, 2007).

Despite its importance to hearing, the input/output (I/O) characteristics of the human BM response are not completely understood. Human BM responses cannot be measured directly so indirect techniques are used to infer I/O curves. There exist various methods to infer BM I/O curves from perceptual data (e.g., Lopez-Poveda and Alves-Pinto, 2008; Nelson *et al.*, 2001; Oxenham and Plack, 1997; Plack and Oxenham, 2000). These behavioral methods are reasonably well grounded (e.g., Bacon and Oxenham, 2004; Oxenham and Bacon, 2004). The different methods yield similar within-subject results (Lopez-Poveda and Alves-Pinto, 2008; Nelson *et al.*, 2001; Rosengard *et al.*, 2005) and they are replicable across different measurement sessions. Unfortunately, they are time consuming and require active participation from the listeners and their training. Therefore, they are unsuitable in clinical contexts or for non-cooperative patients (e.g., infants and the elderly).

Distortion product otoacoustic emission (DPOAE) I/O curves are on average broadly similar to BM I/O curves, both for healthy and damaged cochleae (e.g., Dorn *et al.*, 2001; Neely *et al.*, 2009). Their measurement does not require active participation from the listeners and so it has been suggested that DPOAEs could provide a useful alternative to behavioral methods to infer BM I/O curves (e.g., Janssen and Müller, 2008; Johannesen and Lopez-Poveda, 2008; Lopez-Poveda *et al.*, 2009; Müller and Janssen, 2004). Unfortunately, it is still uncertain that DPOAE I/O curves constitute reasonable estimates of BM I/O curves on an *individual* basis (e.g., Johannesen and Lopez-Poveda, 2008; Williams and Bacon, 2005). The long-term goal of the present research is to define DPOAE stimuli and conditions that would allow using DPOAE I/O curves as a reliable alternative to behavioral methods for inferring individualized, frequency-specific BM I/O curves.

In an earlier study (Johannesen and Lopez-Poveda, 2008), we assessed the within-subject correspondence between behaviorally inferred BM I/O curves and DPOAE I/O curves obtained with typical DPOAE parameters (see be-

[a)]Portions of this work were presented in poster form at Acoustics'08 Paris, 29 June–4 July 2008, Paris, France. Abstract published: J. Acoust. Soc. Am. 123, 3854 (2008).

[b)]Author to whom correspondence should be addressed. Electronic mail: ealopezpoveda@usal.es

 0001-4966/2010/127(6)/3602/12/$25.00

low). We observed a reasonably high correspondence only for test frequencies above ~2 kHz. The correspondence was, however, poorer at lower frequencies (0.5 and 1 kHz) because DPOAE I/O curves showed notches and plateaus that did not occur in corresponding behavioral curves. The reason for the notches and plateaus was uncertain. Two explanations were suggested. First, notches and plateaus could be due to the DPOAE fine structure (Gaskill and Brown, 1990), despite our attempts to reduce it by spectral averaging (Kalluri and Shera, 2001). Second, they could be due to our using suboptimal DPOAE stimulus parameters.

The term "fine structure" is used to refer to rapid and local variations present in the graphical representation of $2f_1-f_2$ DPOAE level against test frequency (f_2) (known as the DP-gram). These variations are thought to arise by vector summation of DP contributions generated by various mechanisms from spatially distributed regions within the cochlea, which give rise to peaks and notches in the DP-gram (for a review, see Shera and Guinan, 2008). The same interference mechanism also influences DPOAE I/O curves (He and Schmiedt, 1993, 1997; Mauermann and Kollmeier, 2004).

The fine structure is equally common across test frequencies (Fig. 2 of Mauermann *et al.*, 1999; Fig. 3 of Dhar and Schaffer, 2004), although this evidence is based on rather small sample sizes ($N=4$ and $N=10$, respectively). The notches and plateaus of our I/O curves were, by contrast, present only for low test frequencies. Johnson *et al.* (2006b) reported a higher notch incidence at 2 than at 4 kHz based on a larger subject sample ($N=12-22$), which might be taken as suggestive of greater incidence of plateaus and notches at low frequencies, consistent with the results of our previous study. The evidence of Johnson *et al.* (2006b), however, was based on low primary levels ($L_2=30$ dB SPL), for which the fine structure is known to be more pronounced (He and Schmiedt, 1993, 1997; Mauermann and Kollmeier, 2004; Johnson *et al.*, 2006b). Therefore, it is uncertain that the results of Johnson *et al.* (2006b) may be generalized to 50–60 dB SPL, the range of levels where plateaus and notches were observed in our previous study (Johannesen and Lopez-Poveda, 2008). Furthermore, it is also uncertain that the results of Johnson *et al.* (2006b) can be generalized to 0.5 and 1 kHz, the frequencies where plateaus and notches were more frequent in our data. Additionally, the plateaus and the majority of the notches in the I/O curves of our previous study extended over a wider input level range (approximately 20–30 dB) than that of the notches caused by interference from several DP source contributions [about 10 dB, as estimated from the I/O functions in Figs. 6–8 of He and Schmiedt (1993)]. Lastly, the DPOAE I/O curves reported in our previous study were the mean of several (typically five) I/O curves for adjacent f_2 frequencies in an attempt to reduce the influence of the fine structure by spectral averaging (Kalluri and Shera, 2001). This number was sufficient to account for the fine structure for test frequencies above 2 kHz and so it seems reasonable to assume that it also accounted for the fine structure influence at low frequencies. Altogether, this suggests that the fine structure explanation for the poor correspondence between behavioral and DPOAE I/O curves at low frequencies is possible but may not be sufficient.

Indeed, a complementary explanation may be that the DPOAE I/O curves of our previous study were measured with suboptimal stimulus parameters (Johannesen and Lopez-Poveda, 2008). DPOAE primaries had frequencies (f_1 and f_2) with a ratio of $f_2/f_1=1.2$ (Gaskill and Brown, 1990) and their levels conformed to the rule of Kummer *et al.* (1998): $L_1=39+0.4L_2$, with L_1 and L_2 being the levels of primaries f_1 and f_2, respectively. These parameters were group-average optimal values and were constant across test frequencies; hence it is unlikely that they were optimal on an individual basis. Evidence exists that the primary frequency f_2/f_1 ratio has only a small effect on the shape and slope of I/O curves (Johnson *et al.*, 2006a). By contrast, there is significant evidence that primary levels have a stronger influence on the shape and slope of I/O curves. First, individualized optimal level rules vary significantly across listeners and test frequencies (Neely *et al.*, 2005). Second, only 10% of the subjects have non-monotonic I/O curves in the frequency range from 1 to 8 kHz when using individually optimized DPOAE primary levels (Kummer *et al.*, 2000). Third, numerical simulations shown in Appendix A and elsewhere (Lukashkin and Russell, 2001) suggest that (1) suboptimal primary levels may lead to non-monotonic DP I/O curves, even in the absence of secondary DP sources; and (2) the use of optimal primary levels improves the correspondence between DP I/O curves and I/O curves measured using single tones. Lastly, Lopez-Poveda and Johannesen (2009, 2010) have confirmed that individualized optimal rules for the subjects of their first study (Johannesen and Lopez-Poveda, 2008) differ from the rule of Kummer *et al.* (1998). Altogether this suggests that the correspondence between DPOAE and behaviorally inferred BM I/O curves could improve by using individualized optimal primary levels.

The use of individualized DPOAE optimal levels may also reduce the fine structure. It is known that varying the primary levels can change the relative contribution from the various DP cochlear sources (He and Schmiedt, 1993). To the authors' knowledge, the details are uncertain but one possibility is that individualized optimal primary levels maximize the DPOAE levels by emphasizing the contribution from the "distortion" (f_2) source at the ear canal, which might yield a comparatively smaller "reflection" component and thus reduce the magnitude of the fine structure. If this were the case, then DPOAE I/O curves measured with individualized optimal primary levels may reflect more closely the underlying BM I/O curve at the f_2 cochlear site, and hence improve the match with its corresponding behaviorally inferred I/O curves.

In summary, the fine structure cannot be dismissed as an explanation for the poor correlation between DPOAE and behaviorally inferred BM I/O curves at low frequencies, but the use of individualized DPOAE optimal primary levels would likely improve the correspondence between the results of the two methods.

This study is a re-analysis of previously published data aimed at testing this possibility for a fixed primary frequency

ratio of $f_2/f_1=1.2$. An attempt is made to minimize fine structure effects by averaging DPOAE I/O curves for a number of f_2 frequencies around the test frequency of interest. If successful, such an approach would be more time consuming than DPOAE I/O curve measurement procedures with standard parameters but still advantageous over behavioral methods for estimating individualized, frequency-specific BM I/O curves. Furthermore, it could be put in practice with current advanced DPOAE clinical devices.

II. METHODS

A. Approach

The approach involved within-subject comparisons of DPOAE I/O curves for optimal stimuli with BM I/O curves inferred behaviorally from temporal masking curves (TMCs).

A TMC is a graphical representation of the level of a pure tone forward masker required to just mask a fixed, low-level pure tone probe, as a function of the time interval between the masker and the probe. It is assumed that the slope of a TMC reflects the rate of increase of the BM response to the masker at the BM place tuned to the probe frequency *and* the post-cochlear decay rate of the internal masker effect (Nelson *et al.*, 2001). There is evidence that the latter is approximately constant across masker frequencies and over a wide range of masker levels (Lopez-Poveda and Alves-Pinto, 2008; Wojtczak and Oxenham, 2009). Hence, approximate BM I/O curves may be inferred by plotting the levels of a linear reference TMC (i.e., the TMC for a masker that evokes a linear BM response) against the levels for the TMC for the frequency of interest paired according to time interval (Nelson *et al.*, 2001).

Individualized optimal DPOAE stimuli may be found empirically; that is, by searching combinations of primary frequencies and levels that maximize the level of the $2f_1-f_2$ DPOAE (e.g., Kummer *et al.*, 2000; Johnson *et al.*, 2006a). Optimal stimuli differ across individuals and test frequencies. In the present study, the primary frequency ratio was fixed at $f_2/f_1=1.2$. Therefore, the terms "optimal DPOAE stimuli" and "optimal DPOAE level rule" are used here to refer to the individualized combination of primary levels L_1 and L_2 that evokes the maximal level of the $2f_1-f_2$ DPOAE component for any test frequency f_2.

It has been conjectured that primary levels are optimal when both primaries evoke comparable responses in the f_2 BM region (Kummer *et al.*, 2000). We have recently provided support for this conjecture by comparing empirical optimal levels with levels of equally effective primaries as inferred from TMCs (Lopez-Poveda and Johannesen, 2009). We measured TMCs for two masker frequencies, f_{m1} and f_{m2}, equal to the DPOAE primary frequencies ($f_{m1}=f_1$ and $f_{m2}=f_2$) for a probe frequency (f_P) equal to the DPOAE test frequency ($f_P=f_2$). Based on the TMC interpretation explained above, the levels of two equally effective pure tones were inferred by plotting the levels of the f_{m1} masker, L_{m1}, against those of the f_{m2} masker, L_{m2}, paired according to time interval. The resulting $L_{m1}-L_{m2}$ functions overlapped reasonably well with corresponding plots of empirical DPOAE optimal L_1-L_2 levels for the same subject, thus supporting the conjecture of Kummer *et al.* (2000). The overlap was, however, restricted to L_2 levels below ~65 dB SPL.

Our former reports did not address the main question of the present study but contained a great deal of the data necessary to do it. Therefore, for convenience, the approach here consisted of re-analyzing the data of our earlier reports (Johannesen and Lopez-Poveda, 2008; Lopez-Poveda and Johannesen, 2009, 2010) with the aim of testing if the correspondence between behavioral and DPOAE estimates of BM I/O curves improves when DPOAEs are measured with individualized optimal primary levels. For completeness, the present study extended to DPOAE I/O curves measured using individualized TMC-based primary level rules [from Lopez-Poveda and Johannesen (2009)] and the group-average level rule of Kummer *et al.* (1998). Methodological details have been amply described in the relevant earlier studies (summarized in Appendix B) and only a brief description is provided here.

B. Subjects

Fifteen normal-hearing listeners participated in the study. Their ages ranged from 20 to 39 years. All of them had thresholds within 20 dB hearing level (HL) (ANSI, 1996) at the frequencies considered in this study (0.5, 1, and 4 kHz). They are identified here as in our earlier studies (see Appendix B).

C. TMC stimuli

TMCs were measured for probe frequencies (f_P) of 0.5, 1, and 4 kHz and for masker frequencies (f_2) equal to the f_P. TMCs were also measured for a probe frequency of 4 kHz and a masker frequency of 1.6 kHz. The latter were regarded as the linear reference (Lopez-Poveda and Alves-Pinto, 2008). These TMCs were used to infer BM I/O curves for all probe frequencies (Lopez-Poveda *et al.*, 2003). Additional TMCs were measured for probe frequencies of 0.5, 1, and 4 kHz and for masker frequencies (f_1) equal to $f_P/1.2$. These TMCs, together with those for masker frequencies f_2, were used to infer individualized DPOAE primary rules so that the two primaries evoked equal-BM excitation at the f_2 site (see Secs. II A and II F).

Probe level was fixed at 9 dB sensation level (SL) (i.e., 9 dB above the individual's absolute threshold for the probe).

D. TMC procedure

Masker levels at threshold were measured using a two-interval, two-alternative, forced-choice, procedure. Feedback was provided to the listener. Masker level was changed according to a two-up, one-down adaptive procedure to estimate the 71% point on the psychometric function (Levitt, 1971). The initial step size was 6 dB. The step size was decreased to 2 dB after three reversals. A total of 15 reversals were measured. Threshold was calculated as the mean of the masker levels for the last 12 reversals. A measurement was discarded if the associated standard deviation (SD) exceeded 6 dB. Three threshold estimates were obtained in this way and their mean was taken as the masker level at masked

threshold. If the SD of these three measurements exceeded 6 dB, a fourth threshold estimate was obtained and included in the mean.

Listeners were trained in the forward-masking task for several hours; at first with a higher probe level of 15 dB SL, and later with a probe level of 9 dB SL, until performance became stable.

E. Inferring BM I/O functions from TMCs

BM I/O functions were inferred from TMCs by plotting the masker levels for the linear reference TMC against the levels for a masker equal in frequency to the probe and paired according to masker-probe time interval (Nelson *et al.*, 2001). The linear reference TMCs were fitted (using a least-squares procedure) with a double exponential function and extrapolated to longer masker-probe time intervals to infer BM I/O functions over the wider possible range of levels.

F. Inferring DPOAE primary level rules from TMCs

A least-squares procedure was used to fit TMCs for the f_1 and f_2 maskers with an *ad-hoc* function (see Lopez-Poveda *et al.*, 2005). Individualized DPOAE level rules were then obtained by plotting the fitted levels for the f_1 masker against the fitted levels for the f_2 masker paired according to the masker-probe time intervals (Lopez-Poveda and Johannesen, 2009). Under the assumption that the post-cochlear recovery from masking is independent of masker frequency, and given that the two masker frequencies were equal to the DPOAE primary frequencies, the resulting curves provided the level relationships for two pure tones of frequencies f_1 and f_2 that evoke comparable excitation levels at the f_2 region of the BM (for a full justification, see Lopez-Poveda and Johannesen, 2009). These will be referred to as TMC-based primary levels.

G. DPOAE stimuli

DPOAE I/O curves were obtained by plotting the magnitude (in dB SPL) of the $2f_1-f_2$ DPOAE as a function of the level L_2 of primary tone f_2. I/O curves were obtained for a fixed primary frequency ratio of $f_2/f_1=1.2$ and three different level rules.

- *Individualized optimal levels.* These were obtained by searching the (L_1,L_2) space to find the value of L_1 that produced the highest DPOAE response level for each value of L_2. L_2 was varied in 5-dB steps within the range 35–75 dB SPL. For each fixed L_2, L_1 was varied in 3-dB steps and the individual's optimal value was found.
- *TMC-based levels.* These were derived from the TMCs for maskers f_1 and f_2 as explained in Sec. II F. Recall that with these levels the two primaries are presumed to evoke comparable responses in the f_2 cochlear region (Lopez-Poveda and Johannesen, 2009).
- *The level rule of* Kummer *et al.* (1998)*:* $L_1=0.4L_2+39$, with L_1 and L_2 in dB SPL. This group-average rule was originally designed for $L_2\leq 65$ dB SPL, but here it was extrapolated to $L_2=75$ dB SPL. When this rule was applied, L_2 ranged from 20 to 75 dB SPL in 5-dB steps, except for $f_2=0.5$ kHz for which it ranged from 45 to 75 dB SPL.

DPOAE I/O curves were measured for test frequencies of 0.5, 1, and 4 kHz for all three level rules. Not all rules were applied to all participants. Indeed, at 0.5 kHz, I/O curves for optimal rules were measured for only three participants due to time restrictions.

A spectral averaging approach (e.g., Kalluri and Shera, 2001; Mauermann and Kollmeier, 2004) was used in an attempt to reduce DPOAE I/O variability due to the fine structure. DPOAE I/O curves were measured for five close f_2 frequencies around the text frequency of interest, and the resulting I/O curves were averaged (details may be found in Johannesen and Lopez-Poveda, 2008; Lopez-Poveda and Johannesen, 2009). For instance, the final DPAOE I/O curve at 4 kHz was the mean of five I/O functions for $f_2=3920$, 3960, 4000, 4040, 4080 Hz. When optimal levels rules were considered, the I/O curves of only three adjacent frequencies (e.g., $f_2=3960$, 4000, 4040 Hz) were averaged due to time constraints.

H. DPOAE stimulus calibration and system artifacts

DPOAE primary levels were calibrated with a Zwislocki DB-100 coupler for each pair of primary frequencies (f_1,f_2). No further *in-situ* adjustment of this calibration was applied.

Instrument artifactual DP responses were controlled for by prolonged measurements in a DB-100 Zwislocki coupler and a plastic syringe with a volume of $\sim$1.5 cc. Tests were performed for high L_2 levels ($>$50 dB) and under the same conditions as real ear-canal measurements. Ear-canal measurements were rejected if they were less than 6 dB above the coupler DP response. This is a stricter criterion than commonly used in clinical contexts (for a comprehensive justification, see Johannesen and Lopez-Poveda, 2008).

I. DPOAE procedure

DPOAE measurements were made with an IHS Smart system (with SMARTOAE software version 4.52) equipped with an Etymotic ER-10D probe. During the measurements, subjects sat comfortably in a double-wall sound attenuating chamber and were asked to remain as steady as possible.

The probe fit was checked before and after each measurement session. The probe remained in the subject's ear throughout the whole measurement session to minimize measurement variance from altering the position of the probe in the ear canal. DPOAEs were measured for a fixed measurement time ranging from 10 to 60 s. A DPOAE measurement was considered valid when it was 2 SD above the measurement noise floor (defined as the mean level over ten adjacent frequency bins in the spectrum). When a response did not meet this criterion, the measurement was repeated and the measurement time increased if necessary. The probe remained in the same position during these re-measurements. If the required criterion was not met after successive tries, the measurement point was discarded.

When measuring optimal level rules, system artifacts sometimes occurred for high L_2 levels (70–75 dB SPL) and some of the higher L_1 values. A data point for a certain L_2 level was discarded when an optimal L_1 level could not be found within the range of L_1 levels whose DPOAE responses passed the artifact criterion. In other cases, DPOAE responses passed the artifact criterion for the whole range of L_1 levels but no optimal L_1 value could be found because the instrument limits it to 80 dB SPL. That is, the optimal L_1 would have been almost certainly above 80 dB SPL. In these cases, the true DPOAE response would be higher. These points were noted and included in the correspondence analyses (see below).

J. Analysis of the correspondence between behavioral and DPOAE I/O curves

The degree of within-subject correspondence between BM I/O curves inferred from TMCs and DPOAEs was assessed by least-squares fitting third-order polynomials to all I/O curves (e.g., Johannesen and Lopez-Poveda, 2008). The first derivative of the polynomials was calculated analytically and evaluated for the range of input (behavioral) or L_2 (DPOAEs) levels for which experimental data were available. The similarity between behavioral and DPOAE I/O curves was then assessed by the root mean square (rms) difference between the *slopes* of the behavioral and DPOAE I/O curves *for a corresponding range of input levels*. Considering the difference between the first derivatives instead of the polynomials themselves has the advantage of accounting for the large disparity between behavioral and DPOAE output levels while preserving the information about the shape of the I/O curves. In other words, a first-derivative rms difference of zero indicates I/O curves that may be vertically shifted from each other but are otherwise identical.

Two additional measures were employed to assess the similarity between the degree of BM *compression* suggested by behavioral and DPOAE I/O curves. First, a linear regression (LR) analysis was applied to the minimum value of slope of the third-order polynomials fitted to the I/O curves. Second, a LR analysis was applied to the slope of straight lines fitted by least-squares to I/O curves segments for input (or L_2) levels between 40 and 65 dB SPL. This level range was considered because it typically covered the compressive portion of the I/O curves (Johannesen and Lopez-Poveda, 2008).

These LR analyses were applied to the three sets of DPOAE I/O curves obtained with the three primary levels rules considered in the present study (i.e., optimal, TMC-based, and Kummer).

III. RESULTS

A. TMCs

Except for the linear references of subjects S11–S15, all other TMCs have been reported elsewhere for different purposes. Detailed interpretations of their characteristics can be found in the relevant studies (summarized in Appendix B).

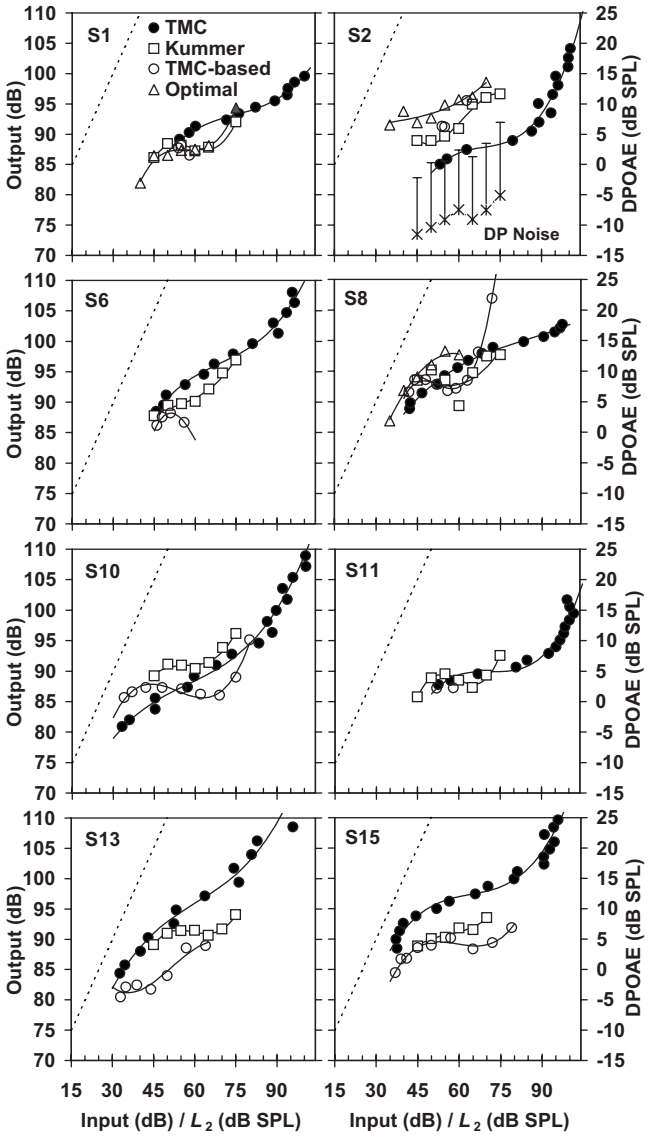

FIG. 1. Behavioral (filled circles) and DPOAE (open symbols) I/O curves at 0.5 kHz for three different primary level rules, as indicated by the inset. Continuous lines illustrate third-order polynomial fits to the experimental points. Each panel illustrates the result for a different participant. The panel for participant S2 also illustrates the mean DP noise floor and its corresponding 2 SDs. Thin dotted lines illustrate a linear response for comparison. Gray symbols illustrate conditions for which primary level L_1 was likely suboptimal because the optimal value would have been higher than the maximum value allowed by the OAE system (80 dB SPL).

B. DPOAE primary level rules

Individualized TMC-based primary level rules were derived from the TMCs for the f_1 and f_2 maskers as described in Sec. II F. Individualized optimal levels were also found as described in Sec. II G. These level rules have been reported and described in great detail elsewhere (see Appendix B).

C. DPOAE I/O curves

Figures 1–3 illustrate DPOAE I/O curves for test frequencies of 0.5, 1, and 4 kHz, respectively. Curves are shown for individualized optimal primary levels (open triangles), individualized TMC-based primary levels (open circles), and the group-average level rule of Kummer *et al.*,

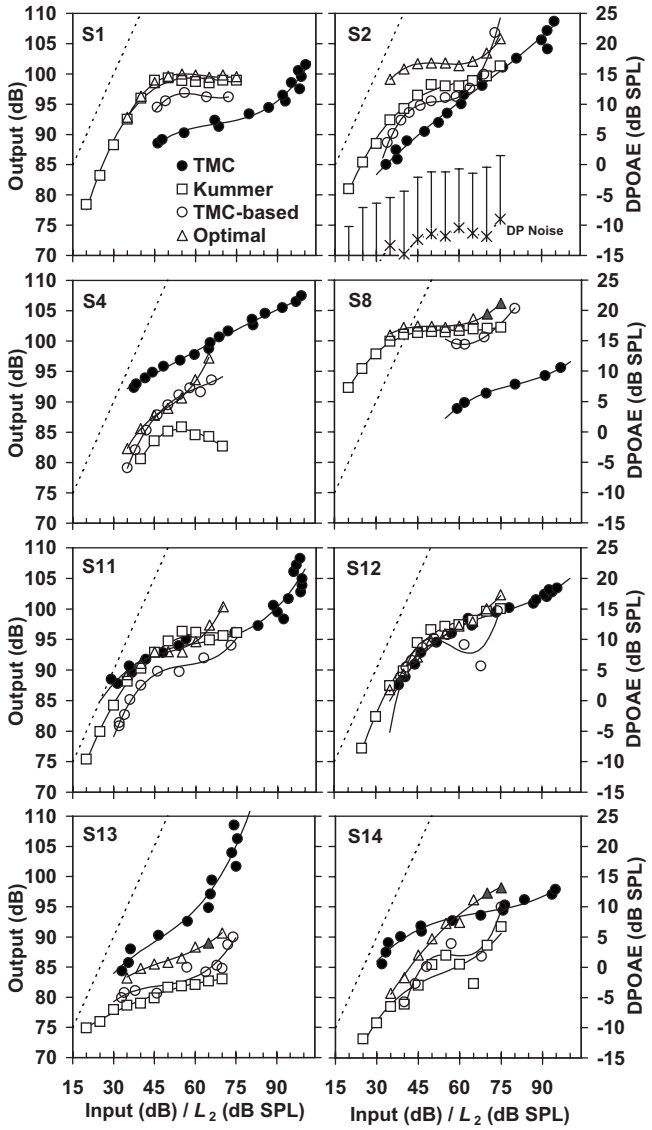

FIG. 2. As Fig. 1 but for a frequency of 1 kHz.

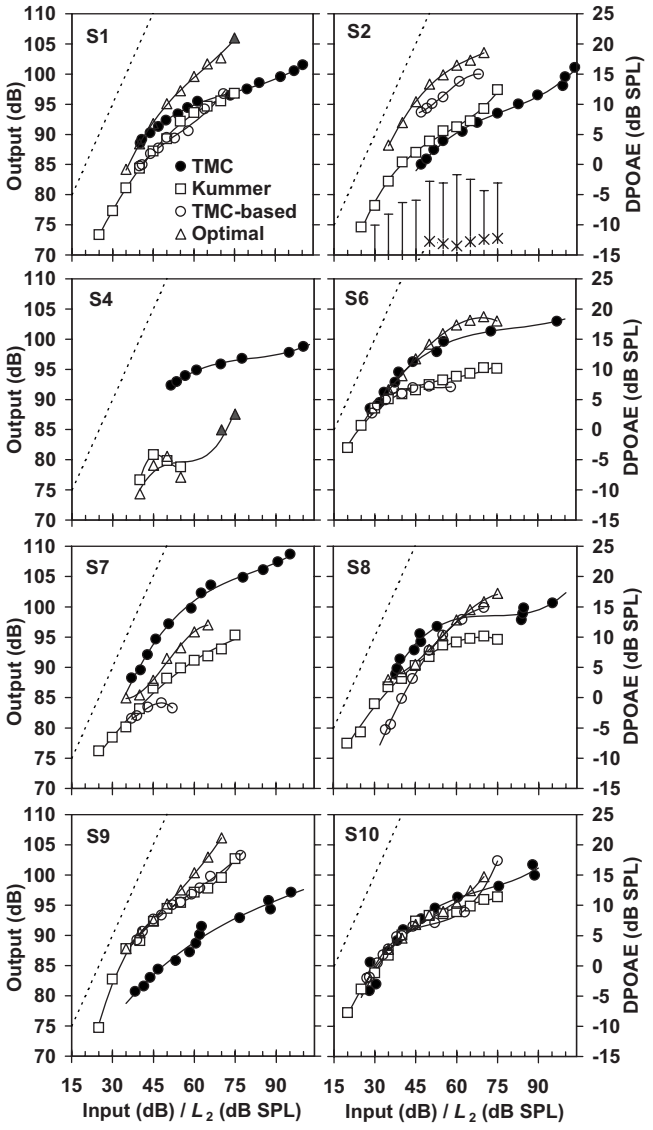

FIG. 3. As Fig. 1 but for a frequency of 4 kHz.

1998 (open squares). Each panel shows the results for a different listener. DPOAE noise levels (crosses) plus two SDs are shown for participant S2 (top-right panels) as a representative example of the noise levels for all other participants. Solid lines illustrate third-order polynomials fitted to the DPOAE I/O curves. Dotted lines illustrate linear responses with a slope of 1 dB/dB. Gray-filled symbols indicate responses whose associated L_1 level could still have been suboptimal (included in fitted curves).

As expected, DPOAE levels for the optimal stimuli were, with very rare exceptions, comparable or higher than DPOAE levels for TMC-based or Kummer stimuli. The very few exceptions (e.g., S11 at 1 kHz and L_2=50 dB SPL in Fig. 2) were possibly due to changes in the position of the probe across DPOAE measurements for the three level rules.

DPOAE I/O curves could be typically (but not always) described as having a steep segment (approaching linearity) at low L_2 levels, followed by a shallower segment at midrange levels. Some curves showed a steeper segment at high L_2 levels that approached linearity. This general trend is characteristic of BM I/O curves (e.g., Robles and Ruggero, 2001). It is noteworthy that the shallower segments of many curves showed plateaus (i.e., regions with a slope of ~0 dB/dB) or notches (i.e., regions with negative slopes), some of which were very sharp. These features were very common at 0.5 (e.g., S1, S6, S8, S10, S11, S13, and S15 in Fig. 1) and 1 kHz (e.g., S1, S2, S8, and S14 in Fig. 2) but virtually nonexistent at 4 kHz. Table I summarizes the number of I/O curves having plateaus or negative-slope segments for each of the three level rules and test frequencies. Rather surprisingly, notches and plateaus occurred at 0.5 and 1 kHz not only for the group-average level rule of Kummer *et al.* (1998) but also for individualized levels, even for optimal levels (e.g., S1, S2, or S8 at 1 kHz).

TABLE I. Incidence of plateaus and notches in DPOAE I/O curves across test frequencies and primary level rules. The numbers between parentheses indicate the total number of I/O curves measured for each condition.

Frequency (kHz)	0.5	1	4
Kummer rule	5 (8)	5 (8)	2 (8)
TMC-based	5 (8)	4 (8)	2 (8)
Optimal	2 (3)	4 (8)	1 (8)

It is interesting to compare the I/O curves obtained with the three primary level rules. At 0.5 kHz (Fig. 1), I/O curves for optimal levels were available for three participants only (S1, S2, and S8) and individualized TMC-based levels often did not cover a sufficient level range to allow a useful comparison. At 1 and 4 kHz (Figs. 2 and 4), where data points were more numerous, visual inspection revealed that I/O curves for the three level rules had reasonably similar shapes (e.g., S1, S8, S11, and S13 at 1 kHz; S1, S2, S7, S9, and S10 at 4 kHz) but with some exceptions (S4 and S14 at 1 kHz; S6 and S8 at 4 kHz). For most of these exceptions, I/O curves for optimal (open triangles) and TMC-based levels (open circles) generally shared many characteristics and they both differed from the curves measured with Kummer's rule (open squares). Optimal levels and to a lesser extent also TMC-based levels tended to produce a rapid increase of DPOAE levels for levels $L_2 > \sim 65$ dB (e.g., S2, S8, and S11 at 1 kHz; S9 and S10 at 4 kHz), although there were exceptions (S1 at 1 kHz; S6 at 4 kHz). Such rapid increases were much rarer for I/O curves measured with Kummer's rule (S2 at 4 kHz).

D. Correspondence between behavioral and DPOAE I/O curves

Behavioral I/O curves (filled circles in Figs. 1–3) showed the same general trend as DPOAE I/O curves, having steep, shallow (compressive), and steep segments at low, moderate, and high input levels, respectively. Unlike for DPOAE I/O curves, however, behavioral I/O curves rarely showed plateaus or notches.

Visual inspection revealed a reasonable correspondence between behavioral and DPOAE I/O curves for the three level rules at 4 kHz (Fig. 3). This was expected for the Kummer rule because the present data had already been reported by Johannesen and Lopez-Poveda (2008). The question is whether the match improved for any of the individualized primary levels (optimal or TMC-based). Optimal levels did improve the correspondence sometimes (e.g., S6 and S7). Other times, however, the correspondence decreased (e.g., S8 and S9). TMC-based levels did not improve the correspondence between DPOAE and behavioral I/O curves.

At 0.5 and 1 kHz (Figs. 1 and 2), however, the correspondence between behavioral and DPOAE I/O curves was disappointing even for DPOAEs measured with individualized optimal or TMC-based levels. With a few exceptions (S11 and S12 at 1 kHz in Fig. 2), behavioral I/O curves were strikingly different from all DPOAE I/O curves. The difference could generally be attributed to the above mentioned plateaus and notches that only occurred for DPOAE I/O curves.

Figure 4 provides quantitative support to the preceding qualitative description. It illustrates rms differences between the slopes of DPOAE and behavioral I/O curves (see Sec. II) against test frequencies, with primary level rules as the parameter (as indicated by the inset). Each open symbol illustrates the rms difference for a given participant. Filled symbols illustrate mean values across participants. Mean differences (filled symbols) were statistically identical across test frequencies and rules. Mean differences tended to be smaller for the Kummer and optimal level rules at 4 kHz than at 0.5 kHz (unpaired t-test, $0.05 < p < 0.10$). Indeed, the mean difference for all three rules tended to decrease with increasing frequency. The *range* of individual differences was comparable for DPOAE I/O curves measured with the rule of Kummer *et al.* (1998) and the optimal rule and they were always narrower than the TMC-based rule.

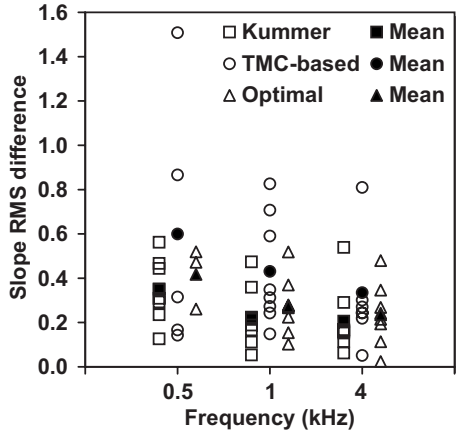

FIG. 4. Root mean square differences between the slope of behavioral and DPOAE I/O curves for different primary level rules, as indicated by the inset. Open and filled symbols illustrate individual and mean values across listeners, respectively.

E. Correspondence of compression estimates

Some authors regard the slope of I/O curves over their nonlinear (compressive) segments as a useful description of the physiological status of the cochlea [reviewed by Robles and Ruggero (2001)]. Others suggest, by contrast, that it does not change with amount of hearing loss (Plack *et al.*, 2004). In any case, given that DPOAE I/O curves sometimes show plateaus and/or notches, which may be absent in their behavioral counterparts, the slope of I/O curves over their compressive segments constitutes a useful variable for assessing their correspondence between behavioral and DPOAE I/O curves. Slopes have been often calculated as the average value over the I/O curve compressive region (e.g., Lopez-Poveda *et al.*, 2003; Plack *et al.*, 2004). Sometimes, however, the slope changes smoothly over the compressive region (e.g., Nelson *et al.*, 2001) and so the minimum, instead of the average, slope is used (e.g., Johannesen and Lopez-Poveda, 2008). Another consideration is that average slope may prove less sensitive to sharp notches, like those present in DPOAE I/O curves. Here, the two descriptors (minimum and average slope values) were considered for assessing the correspondence between DPOAE and behavioral I/O curves over their compressive regions (see Sec. II).

Figure 5 and 6 show, respectively, plots of minimum and average slopes of DPOAE I/O curves against corresponding slopes for behavioral curves. Each panel is for a different test frequency. Each point illustrates results for a single participant. Different symbols illustrate results for DPOAEs measured with different levels rules, as indicated by the inset in the top panels. Thick lines illustrate LR functions fitted by least-squares to the experimental data points, as indicated by the inset. The diagonal represents perfect correspondence (il-

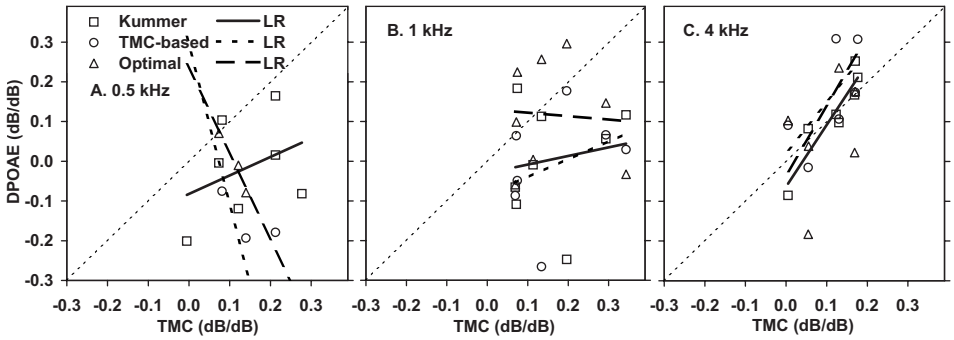

FIG. 5. Correspondence between minimum compression-exponent estimates obtained from third-order polynomials fitted to behavioral and DPOAE I/O curves at 0.5, 1, and 4 kHz. Each point represents data for one subject and one level rule, as indicated by the inset in panel (A). Thick lines show LR fits to the data points for each of the stimulus rules. Diagonal thin dotted lines illustrate perfect correspondence.

lustrated by thin dotted lines). Tables II and III summarize the results of LR analysis on these compression estimates. The correspondence can be regarded as good when the LR slope (a) and intercept (b) are close to one and zero, respectively, and the F-statistics allow rejecting the null-hypothesis (there is no statistical association between the two data sets).

Correspondence between DPOAE and behavioral compression estimates was clearly poor at 0.5 and 1 kHz regardless of the compression descriptor (i.e., minimum or average slope) or the level rule used to measure DPOAEs. For both compression descriptors (Tables II and III), the linear regression slope and intercept were far from 1 and 0.

At 4 kHz, reasonably high correspondence was observed for the minimum slope estimate (Fig. 5) for DPOAEs measured with the Kummer rule (slope=1.55; intercept=−0.06; r^2=0.86; $p<0.01$). Correspondence was lower for individualized level rules and was not significant (Table II). For the average slope estimates (Fig. 6), there was moderate correspondence for DPOAEs measured with the Kummer rule and the optimal rule, but it did not reach significance (both p =0.11). The two individualized rules (optimal and TMC-based) yielded reasonably similar minimum compression estimates (slope=1.72; intercept=−0.08; r^2=0.89; $p<0.005$; data not shown). This is not surprising given that both individualized rules were very similar, presumably because they reflect levels of equally effective primaries at the f_2 cochlear site (Lopez-Poveda and Johannesen, 2009).

IV. DISCUSSION

The long-term goal of the present research is to define DPOAE stimuli and conditions that would allow using DPOAEs as a universal and faster alternative to behavioral methods to infer individualized, frequency-specific BM I/O curves. Previous studies have shown that this is possible using typical DPOAE stimuli (specifically, the rule of Kummer *et al.*, 1998 and a frequency ratio of f_2/f_1=1.2) for test frequencies above ~2 kHz but not for lower frequencies (Johannesen and Lopez-Poveda, 2008). The main reason for the lack of correspondence at these lower frequencies seemed to be the presence of notches and plateaus in DPOAE I/O curves that did not have a counterpart in behavioral curves. The present study aimed at testing whether the correspondence between the I/O curves inferred with the two methods would improve by using DPOAE individualized optimal primary levels. Two individualized level rules have been considered: one optimized to maximize DPOAE levels and one intended for primaries to evoke comparable BM responses at the f_2 cochlear site. It has been shown that this approach does not improve the correspondence between behavioral and DPOAE I/O curves at low frequencies with respect to the group-average levels of Kummer *et al.* (1998). Indeed, it has been shown that plateaus and notches are equally common in DPOAE I/O curves for individualized and group-average DPOAE primary levels at 0.5 and 1 kHz (Table I).

The similar morphology of DPOAE I/O curves measured with different primary levels (Figs. 1–3) suggests a lower dependency on individualized levels than expected (see Introduction and Appendix A). For example, all DPOAE I/O curves suggested similar compression thresholds, as well as similar slopes below and above the compression threshold (Fig. 1–3). This was true for most participants (the only exceptions were S4 at 1 kHz, and S6 and S8 at 4 kHz). How-

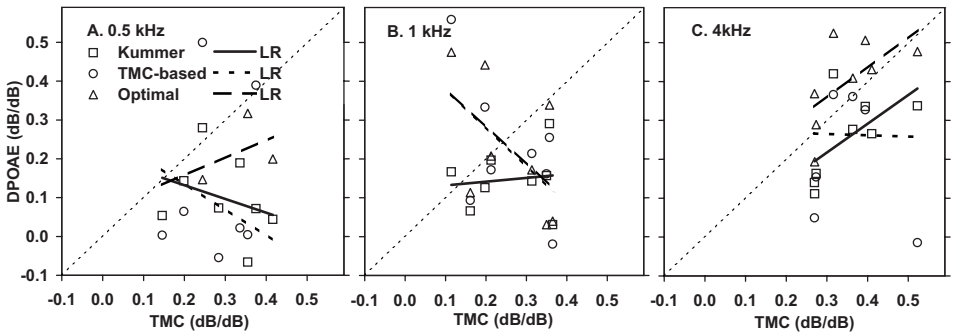

FIG. 6. As Fig. 5 but for the average slope of behavioral and DPOAE I/O curves over their compressive region (input levels from 40 to 65 dB SPL).

TABLE II. Results of linear regression analysis between the minimum slopes of third-order polynomials fitted to behavioral and DPOAE I/O curves for the three level rules. *a*: slope; *b*: intercept; r^2: predicted variance; *N*: number of data points; *p*=probability of observed regression occurring by chance (i.e., $p<0.05$ indicates a statistically significant regression); n.a.: not applicable (LR statistics were not calculated if $N<4$).

Frequency (kHz)			0.5		
Level rule	*a*	*b*	r^2	*N*	*p*
Kummer	0.47	−0.08	0.13	7	0.42
TMC-based	−3.99	0.29	0.38	4	0.39
Optimal	n.a.	n.a.	n.a.	3	n.a.
Frequency (kHz)			1		
Level rule	*a*	*b*	r^2	*N*	*P*
Kummer	0.22	−0.03	0.03	8	0.70
TMC-based	0.45	−0.09	0.13	7	0.43
Optimal	−0.09	0.13	0.00	8	0.87
Frequency (kHz)			4		
Level rule	*a*	*b*	r^2	*N*	*P*
Kummer	1.55	−0.06	0.86	7	0.0025
TMC-based	1.25	0.02	0.43	6	0.16
Optimal	1.72	−0.03	0.27	8	0.18

ever, the tendency to a rapid increase in the DPOAE response at high levels ($L_2 > \sim 70$ dB SPL) seemed peculiar to individualized (optimal and TMC-based) levels. Such a tendency was rarely observed in I/O curves measured with the rule of Kummer *et al.* (1998). The slight influence of the level rule on slope may also explain the rather surprising result that the match between behavioral and DPOAE I/O curves was closest for the Kummer rule even though it is a group-average estimate that disregards individual idiosyncrasies. Despite their modest influence on the shape of DPOAE I/O curves, individualized primary levels may nevertheless be important to maximize the DPOAE absolute levels, and thus to maximize the sensitivity of DPOAEs as a potential clinical diagnostic tool (Kummer *et al.*, 2000).

TABLE III. As Table II, but for linear regression analysis between the slopes of straight lines fitted to the compressive segments of behavioral and DPOAE I/O curves (Fig. 6).

Frequency (kHz)			0.5		
Level rule	*a*	*b*	r^2	*N*	*P*
Kummer	−0.35	0.20	0.10	8	0.45
TMC-based	−0.67	0.27	0.06	8	0.57
Optimal	n.a.	n.a.	n.a.	3	n.a.
Frequency (kHz)			1		
Level rule	*a*	*b*	r^2	*N*	*P*
Kummer	0.10	0.12	0.01	8	0.78
TMC-based	−1.01	0.48	0.33	8	0.13
Optimal	−0.97	0.48	0.31	8	0.15
Frequency (kHz)			4		
Level rule	*a*	*b*	r^2	*N*	*P*
Kummer	0.74	0.00	0.36	8	0.11
TMC-based	−0.03	0.28	0.00	7	0.98
Optimal	0.78	0.12	0.38	8	0.11

A. The causes of low-frequency plateaus of notches

It has been shown that plateaus and notches are equally common in DPOAE I/O curves for individualized and group-average DPOAE primary levels at 0.5 and 1 kHz (Table I). The present study was not intended to provide an explanation for the low-frequency notches and plateaus, but the present results, in combination with existing evidence, provide interesting insights. As argued in the Introduction, plateaus and notches could be due to the fine structure and the use of suboptimal DPOAE stimuli. Further, it was conjectured that these two explanations need not be independent of each other in so far as the fine structure changes with changing primary levels (He and Schmiedt, 1993).

Regarding the suboptimal parameter explanation, a fixed primary frequency ratio of $f_2/f_1=1.2$ has been used in the present study based on early evidence that it maximizes (on average) DPOAE levels (Gaskill and Brown, 1990). According to more recent reports, however, the optimal f_2/f_1 ratio increases slightly with decreasing f_2 frequency and with increasing L_2 level (Johnson *et al.*, 2006a), particularly for low frequencies. To the best of the authors' knowledge, the optimal ratio for a test frequency of 0.5 kHz is yet to be determined. There is evidence, however, that human cochlear processing in apical BM regions may be significantly different from that of basal zones (e.g., Lopez-Poveda *et al.*, 2003; Plack *et al.*, 2004), which suggests that the optimal frequency ratio at 0.5 kHz likely differs from 1.2. Therefore, a better correspondence between behavioral and DPOAE I/O curves might have been obtained by considering not only optimal primary levels but also optimal primary frequencies. On the other hand, variations of the f_2/f_1 ratio seem to have only a small effect on the slope of average I/O curves below 65 dB SPL (e.g., Fig. 3 of Johnson *et al.*, 2006a), at least in the frequency range from 1 to 8 kHz. This casts doubts that optimizing the frequency ratio would improve the correspondence between behavioral and DPOAE I/O curves, but the benefit it could have on an individual basis is yet to be investigated.

Regarding the fine structure explanation, arguments have been given in the Introduction to suggest that this explanation is possible but may not be sufficient. The arguments provided were based on existing evidence for test frequencies equal to or greater than 2 kHz. The importance of the fine structure for frequencies below 1 kHz is still uncertain but the present results suggest that it could more significant than was thought at the outset of the present study. If this were the case, it would explain why the spectral averaging method employed here to minimize the fine structure seemed more efficient for high (4 kHz) than for low (0.5 and 1 kHz) test frequencies. The reason for this is uncertain. If the interference between the DP contributions from various cochlear regions (Shera and Guinan, 2008) was more pronounced and less sensitive to primary levels and f_2 frequencies for low than for high-frequency DPOAEs, this might explain the relatively higher incidence of plateaus and notches at low frequencies and for all level rules considered. By extension, it might also explain the observed discrepancies between behavioral and DPOAE I/O curves. It is uncer-

tain why destructive interference would be more pronounced or less sensitive to primary levels at low than at high frequencies. Interestingly, however, both outer hair cell disorganization (e.g., Lonsbury-Martin *et al.*, 1988) and compression bandwidth appear comparatively greater for apical than for basal cochlear regions (e.g., Rhode and Cooper, 1996; Lopez-Poveda *et al.*, 2003; Plack and Drga, 2003). As a result, the generation region of DPOAEs by both "reflection" and "distortion" mechanisms is likely broader in the apex than in the base and hence potential interactions between DPOAE contributions from adjacent regions could be more significant for low test frequencies. This, however, is only a conjecture, whose detailed formulation and verification remain to be developed.

The present data do not support the relationship between the two explanations (fine structure and suboptimal parameters) that has been conjectured in the Introduction. In principle, varying DPOAE primary levels could change the relative contribution from the various DP sources (as measured in the ear canal) and hence the magnitude of the fine structure. It was hypothesized in the Introduction that individualized primary levels could maximize DPOAE levels by emphasizing the contribution from the "distortion" (f_2) source relative to that of the "reflection" ($2f_1-f_2$) source, which would contribute to reduce the fine structure magnitude. The present results show that the incidence and magnitude of notches and plateaus is not reduced by using individualized optimal primary levels and so they do not support the conjecture in question. If anything, the present data would suggest that the magnitude of the contributions from the "reflection" ($2f_1-f_2$) source is proportional to that of the "distortion" (f_2) source.

B. Remarks

The TMC method of inferring behavioral BM I/O relies on several assumptions (see Sec. II and the Discussion of Johannesen and Lopez-Poveda, 2008). The majority of them have received experimental support for a wide range of conditions (e.g., Lopez-Poveda and Johannesen, 2009; Lopez-Poveda and Alves-Pinto, 2008; Wojtczak and Oxenham, 2009). Nevertheless, behavioral curves cannot be taken as an undisputed "golden standard" (e.g., Stainsby and Moore, 2006; Wojtczak and Oxenham, 2009), particularly at low frequencies where there is a lack of correspondence between behavioral and DPOAE I/O curves. In other words, it remains unclear which of the two sets of I/O curves, behavioral or DPOAEs, best represents the underlying BM I/O curves at 0.5 and 1 kHz.

V. CONCLUSIONS

For a fixed primary frequency ratio of $f_2/f_1=1.2$, DPOAE levels vary moderately depending on primary level rule but the fundamental morphology of DPOAE I/O curves hardly changes.

The incidence of plateaus and notches in DPOAE I/O curves was comparable for individualized primary level rules and for the group-average rule of Kummer *et al.* (1998). These features were more frequent at low (0.5 and 1 kHz) than at high (4 kHz) frequencies and remain the most likely reason for the low correspondence between behavioral and DPOAE I/O curves at low frequencies.

The correspondence between behavioral and DPOAE I/O curves was reasonably high at 4 kHz. The group-average primary level rule of Kummer *et al.* (1998) is sufficient to estimate individualized BM I/O curves at high frequencies.

It is uncertain which of two approaches (DPOAEs or behavioral methods) is more appropriate to infer individualized BM I/O responses at 0.5 and 1 kHz, but DPOAEs may *not* be used as an alternative to behavioral methods, even considering individualized optimal primary levels.

ACKNOWLEDGMENTS

We thank the associate editor and two anonymous reviewers for their suggestions to improve the quality of the paper. Work supported by the Spanish Ministry of Science and Innovation (PROFIT CIT-390000-2005-4 and BFU-2006-07536), the Junta de Castilla y León (GR221) and the William Demant Oticon Foundation.

APPENDIX A

In the present study, as in many previous ones (e.g., Dorn *et al.*, 2001; Williams and Bacon, 2005; Johannesen and Lopez-Poveda, 2008; Neely *et al.*, 2009), it is implicitly assumed that the I/O curve for the $2f_1-f_2$ DPOAE would be a reasonable description of a BM site's I/O curve (as measured using pure tones at the site's characteristic frequency-CF) for a site with CF$\sim f_2$. In other words, that it would be possible to infer on-CF BM I/O curves by measuring the growth of the $2f_1-f_2$ DPOAE component with increasing L_2, for f_2=CF. This section provides evidence that this assumption stands a better chance of being reasonable when optimal L_1 levels are considered. Evidence is also shown that suboptimal L_1 levels may produce non-monotonic $2f_1-f_2$ DP I/O curves even in the absence of secondary DP sources (fine structure).

Ideally, the assumption in question should be demonstrated in analytical mathematical form for the nonlinearity underlying BM responses. This, however, is not possible because nonlinear systems do not have exact analytic transfer functions and because the actual form of the BM nonlinearity is, to the authors' knowledge, yet to be confirmed. The present section provides only a crude numerical demonstration of the claims made in the preceding paragraph for two kinds of time-invariant compressive nonlinearities: a broken-stick nonlinear (Meddis *et al.*, 2001) and a double-Boltzmann gain function (e.g., Lukashkin and Russell, 1998, 2001). These nonlinearities have been successfully used to account for BM and outer hair cell responses in computational models. For simplicity, no concurrent filtering effects are considered here.

Let $x(t)$ and $y(t)$ be the time-domain input and output waveforms to/from the nonlinearity. The broken-stick nonlinearity has the form

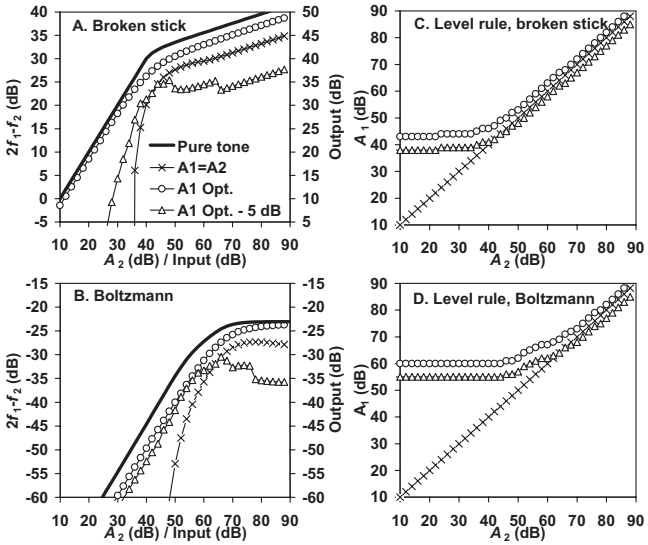

FIG. 7. Simulated I/O curves (left panels) for different amplitude rules (right panels) and two non-linearities: a broken stick (upper panels) and a Boltzmann function (lower panels). [(A) and (B)] Bold continuous lines illustrate I/O curves based on single-sinusoid inputs. Symbols illustrate I/O curves for the $2f_1-f_2$ DP for three different amplitude rules: $A_1=A_2$, A_1 =optimal, and $A_1=A_1$ optimal −5 dB (as indicated in the inset). [(C) and (D)] A_1-A_2 amplitude rules corresponding to the I/O curves of (A) and (B).

$$y(t) = \min[ax(t), bx^c(t)], \tag{A1}$$

where a and b are gain parameters and c is the compression exponent. Likewise, the double-Boltzmann nonlinearity has the form

$$y(t) = G_M\left\{1 + \exp\left(\frac{x_1 - x(t)}{s_1}\right)\left[1 + \exp\left(\frac{x_2 - x(t)}{s_2}\right)\right]\right\}^{-1}, \tag{A2}$$

where G_M, x_1, x_2, s_1, and s_2 are parameters. Here, the input was arbitrarily assumed in units of μPa and the parameters of the two functions were set to the following arbitrary values: $a=1$, $b=437$, $c=0.2$, $G_M=7$, $x_1=41$, $x_2=24$, $s_1=62.5$, and $s_2=15.38$.

The I/O response characteristics of these two nonlinearities were obtained considering one- and two-sinusoid digital input waveforms (sampling frequency=48 kHz). These will be referred to as pure tone (PT) and DP I/O curves, respectively. PT I/O curves were obtained using single sinusoids of a fixed frequency (f_2) and amplitudes (A_2) that varied from 20 to 20×10^5 μPa in steps of 2 dB. A fast Fourier transform (FFT) was applied to the output waveforms and the amplitude at the f_2 frequency was noted and plotted as a function of A_2 in a log-log scale to obtain the I/O curve. DP I/O curves were obtained presenting two simultaneous sinusoids of frequencies f_2 and $f_1=f_2/1.2$. The amplitude of the f_2 sinusoid, A_2, was the same as in the one-sinusoid evaluations. For each value of A_2, the amplitude of the f_1 sinusoid, A_1, varied in 1-dB steps. A FFT was applied to the output waveform and the amplitude of the $2f_1-f_2$ (DP) frequency component was noted.

The top and bottom panels of Fig. 7 illustrate the results for the broken-stick and Boltzmann nonlinearity, respectively. The left and right panels illustrate I/O curves and A_1-A_2 amplitude rules, respectively. PT I/O curves are shown as bold continuous lines. DP I/O curves are shown for three A_1-A_2 combinations: equal-amplitude primaries ($A_1=A_2$) (crosses), optimal A_1 (circles), and A_1 5 dB lower than optimal (triangles). Optimal A_1 are defined here as the A_1 amplitudes that produced the highest DP amplitudes for each value of A_2.

Simple visual inspection to the I/O curves of Fig. 7 reveals that PT I/O curves are most similar to DP I/O obtained with optimal A_1 amplitudes. The figure also reveals that suboptimal A_1 amplitudes can produce non-monotonic DP I/O curves (e.g., triangles). Furthermore, the broken-stick nonlinearity DP I/O curve obtained with slightly suboptimal A_1 amplitudes shows shallow notches. These results are consistent with those of Lukashkin and Russell (2001) and support the claims made at the beginning of this section.

Several things are to be noted. First, the present results do not imply that the maximal correspondence between DP and PT I/O occurs for optimal primary amplitudes; they show that optimal primary amplitudes lead to a reasonable correspondence between DP and PT I/O curves. This result is useful in practice because the A_1-A_2 combination that maximizes the correspondence between the two I/O curves will typically be unknown *a priori*. Second, the main conclusions of this exercise are independent of the nonlinearity parameter values. Third, the present models are very simplistic and do not consider concurrent filtering. For this reason, the present simulations do not depend on the actual frequency of the sinusoids or the parameters of the nonlinearities. Filtering, however, would almost certainly produce different optimal amplitude combinations than those illustrated in Figs. 7(C) and 7(D).

Despite its being an oversimplification, this modeling exercise demonstrates that (1) primary levels strongly influence the shape of the DP I/O curve; (2) suboptimal levels can produce non-monotonic DP I/O curves with plateaus and wide, shallow notches; and (3) the use of optimal levels may enhance the similarity between PT and DP I/O curves.

APPENDIX B

The original data can be found in the following references: the absolute thresholds for the maskers and the probes in Lopez-Poveda and Johannesen (2009); linear reference ($f_m=0.4f_p$) and on-frequency TMCs ($f_m=f_p$) for subjects S1–S10 in Johannesen and Lopez-Poveda (2008); the linear reference TMCs for subjects S11–S15 have not been published before. On-frequency TMCs ($f_m=f_p$) for subjects S11–S15 were taken from Lopez-Poveda and Johannesen (2009). Behavioral I/O curves have been published before for subjects S1–S10 only and are reproduced here for completeness (Figs. 1–3). The TMCs for masker frequencies equal to DPOAE primary frequencies f_1 and f_2 ($f_2=f_P$ and $f_1=f_p/1.2$, respectively) were taken from Lopez-Poveda and Johannesen (2009), and were used to infer TMC-based DPOAE level rules.

DPOAE I/O curves measured with the rule of Kummer *et al.* (1998) can be found in Johannesen and Lopez-Poveda (2008); TMC-based DPOAE primary level rules in Lopez-Poveda and Johannesen (2009); individualized optimal pri-

mary level rules for 1 and 4 kHz in Lopez-Poveda and Johannesen (2009); and individualized optimal primary level rules for 0.5 kHz in Lopez-Poveda and Johannesen (2010). In the case of TMC-based and individualized optimal rules, only *mean* DPOAE I/O curves have been published previously (Lopez-Poveda and Johannesen, 2009).

ANSI (**1996**). *S3.6 Specification for Audiometers* (American National Standards Institute, New York).

Bacon, S. P., and Oxenham, A. J. (**2004**). "Psychophysical manifestations of compression: Normal-hearing listeners," in *Compression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper (Springer, New York), Chap. 3, pp. 107–152.

Dhar, S., and Schaffer, 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.

Dorn, P. A., Konrad-Martin, D., Neely, S. T., Keefe, D. H., Cyr, E., and Gorga, M. P. (**2001**). "Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears," J. Acoust. Soc. Am. **110**, 3119–3131.

Gaskill, S. A., and Brown, A. M. (**1990**). "The behavior of the acoustic distortion product, 2*f*1-*f*2, from the human ear and its relation to auditory sensitivity," J. Acoust. Soc. Am. **88**, 821–839.

He, N.-J., and Schmiedt, R. A. (**1993**). "Fine structure of the $2f_1$-f_2 acoustic distortion product: Changes with primary level," J. Acoust. Soc. Am. **94**, 2659–2669.

He, N.-J., and Schmiedt, R. A. (**1997**). "Fine structure of the $2f_1$-f_2 acoustic distortion product: Effects of primary level and frequency ratios," J. Acoust. Soc. Am. **101**, 3554–3565.

Janssen, T., and Müller, J. (**2008**). "Otoacoustic emissions as a diagnostic tool in a clinical context," in *Active Processes and Otoacoustic Emissions*, edited by G. A. Manley, R. R. Fay, and A. N. Popper (Springer, New York), Chap. 13, pp. 421–460.

Johannesen, P. T., and Lopez-Poveda, E. A. (**2008**). "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion product otoacoustic emissions," J. Acoust. Soc. Am. **124**, 2149–2163.

Johnson, T. A., Neely, S. T., Garner, C. A., and Gorga, M. P. (**2006a**). "Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions," J. Acoust. Soc. Am. **119**, 418–428.

Johnson, T. A., Neely, S. T., Kopun, J. G., and Gorga, M. P. (**2006b**). "Reducing reflected contributions to ear-canal distortion product otoacoustic emissions in humans," J. Acoust. Soc. Am. **119**, 3896–3907.

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

Kummer, P., Janssen, T., and Arnold, W. (**1998**). "The level and growth behavior of the 2*f*1-*f*2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," J. Acoust. Soc. Am. **103**, 3431–3444.

Kummer, P., Janssen, T., Hulin, P., and Arnold, W. (**2000**). "Optimal L1-L2 primary tone level separation remains independent of test frequency in humans," Hear. Res. **146**, 47–56.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R., and Coats, A. C. (**1988**). "Spontaneous otoacoustic emissions in nonhuman primate. II. Cochlear anatomy," Hear. Res. **33**, 69–93.

Lopez-Poveda, E. A., and Alves-Pinto, A. (**2008**). "A variant temporal-masking-curve method for inferring peripheral auditory compression," J. Acoust. Soc. Am. **123**, 1544–1554.

Lopez-Poveda, E. A., Johannesen, P. T., and Merchán, M. A. (**2009**). "Estimation of the degree of inner and outer hair cell dysfunction from distortion product otoacoustic emission input/output functions," Audiological Med. **7**, 22–28.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2009**). "Otoacoustic emission theories and behavioral estimates of human basilar membrane motion are mutually consistent," J. Assoc. Res. Otolaryngol. **10**, 511–523.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2010**). "Otoacoustic emission theories can be tested with behavioral methods," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, R. Meddis, and A. R. Palmer (Springer, New York), pp. 3–14.

Lopez-Poveda, E. A., Plack, C. J., Meddis, R., and Blanco, J. L. (**2005**). "Cochlear compression in listeners with moderate sensorineural hearing loss," Hear. Res. **205**, 172–183.

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (**2003**). "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," J. Acoust. Soc. Am. **113**, 951–960.

Lukashkin, A. N., and Russell, I. J. (**1998**). "A descriptive model of the receptor potential nonlinearities generated by the hair cell mechanoelectrical transducer," J. Acoust. Soc. Am. **103**, 973–980.

Lukashkin, A. N., and Russell, I. J. (**2001**). "Origin of the bell-like dependence of the DPOAE amplitude on primary frequency ratio," J. Acoust. Soc. Am. **110**, 3097–3106.

Mauermann, M., Uppenkamp, S., van Hengel, P. W., and Kollmeier, B. (**1999**). "Evidence for the distortion product frequency place as a source of distortion product emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio *f*2/*f*1," J. Acoust. Soc. Am. **106**, 3473–3483.

Mauermann, M., and Kollmeier, B. (**2004**). "Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source," J. Acoust. Soc. Am. **116**, 2199–2212.

Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.

Moore, B. C. J. (**2007**). *Cochlear Hearing Loss*, 2nd ed. (Wiley, Chichester, UK).

Müller, J., and Janssen, T. (**2004**). "Similarity in loudness and distortion product otoacoustic emission input/output functions: Implications for an objective hearing aid adjustment," J. Acoust. Soc. Am. **115**, 3081–3091.

Neely, S. T., Johnson, T. A., and Gorga, M. P. (**2005**). "Distortion-product otoacoustic emission measured with continuously varying stimulus level," J. Acoust. Soc. Am. **117**, 1248–1259.

Neely, S. T., Johnson, T. A., Kopun, J., Dierking, D. M., and Gorga, M. P. (**2009**). "Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears," J. Acoust. Soc. Am. **126**, 728–738.

Nelson, D. A., Schroder, A. C., and Wojtczak, M. (**2001**). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **110**, 2045–2064.

Oxenham, A. J., and Bacon, S. P. (**2003**). "Cochlear compression, perceptual measures and implications for normal and impaired hearing," Ear Hear. **24**, 352–366.

Oxenham, A. J., and Bacon, S. P. (**2004**). "Psychophysical manifestations of compression: Hearing-impaired listeners," in *Compression. From Cochlea to Cochlear Implants*, edited by S. P. Bacon, R. R. Fay, and A. N. Popper (Springer, New York), Chap. 4, pp. 62–106.

Oxenham, A. J., and Plack, C. J. (**1997**). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666–3675.

Plack, C. J., and Drga, V. (**2003**). "Psychophysical evidence for auditory compression at low characteristic frequencies," J. Acoust. Soc. Am. **113**, 1574–1586.

Plack, C. J., and Oxenham, A. J. (**2000**). "Basilar-membrane nonlinearity estimated by pulsation threshold," J. Acoust. Soc. Am. **107**, 501–507.

Plack, C. J., Drga, V., and Lopez-Poveda, E. A. (**2004**). "Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss," J. Acoust. Soc. Am. **115**, 1684–1695.

Rhode, W. S., and Cooper, N. P. (**1996**). "Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo," Aud. Neurosci. **3**, 101–121.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Rosengard, P. S., Oxenham, A. J., and Braida, L. D. (**2005**). "Comparing different estimates of compression in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **117**, 3028–3041.

Shera, C. A., and Guinan, J. J., Jr. (**2008**). "Mechanisms of mammalian otoacoustic emissions," in *Active Processes and Otoacoustic Emissions*, edited by G. A. Manley, R. R. Fay, and A. N. Popper (Springer, New York), Chap. 9, pp. 305–342.

Stainsby, T. H., and Moore, B. C. J. (**2006**). "Temporal masking curves for hearing-impaired listeners," Hear. Res. **218**, 98–111.

Williams, E. J., and Bacon, S. P. (**2005**). "Compression estimates using behavioral and otoacoustic emission measures," Hear. Res. **201**, 44–54.

Wojtczak, M., and Oxenham, A. J. (**2009**). "Pitfalls in behavioral estimates of basilar-membrane compression in humans," J. Acoust. Soc. Am. **125**, 270–281.

APPENDIX C

SUMMARY IN SPANISH

VNiVERSiDAD
D SALAMANCA

ESTIMACIONES PSICOACÚSTICAS Y FISIOLÓGICAS DE CURVAS DE ENTRADA/SALIDA COCLEARES EN HUMANOS

Resumen en castellano de la tesis titulada
"Physiological and psycoacoustical estimation of human cochlear input/output curves".

Tesis doctoral de

Peter Tinggaard Johannesen

Noviembre 2013

Instituto de Neurociencias de Castilla y León

Universidad de Salamanca

RESUMEN

Las curvas de entrada/salida (E/S) de la cóclea humana no se conocen completamente dado que sólo pueden obtenerse con métodos indirectos. El método psicofísico de enmascaramiento temporal post-estimulatorio (CET) se emplea popularmente para inferir estas curvas pero es inviable en la clínica. Las curvas de E/S de otoemisiones acústicas de producto de distorsión (OEAPD) comparten muchas características con las curvas de E/S cocleares. Constituyen, además, un método clínicamente viable y, por tanto, podrían ser una alternativa a las curvas CET. Desgraciadamente, los mecanismos de generación de OEAPD no se conocen completamente y tampoco está clara la relación entre las curvas de E/S de OEAPD y las curvas E/S cocleares. El objetivo del presente trabajo es comprobar si se pueden utilizar los métodos CET y OEAPD de forma indistinta para inferir las curvas de E/S cocleares en sujetos con audición normal. El abordaje consiste en comparar curvas individuales inferidas con ambos métodos. Si los resultados fuesen consistentes entre sí, apoyarían los supuestos de ambos métodos.

En un primer estudio, se demuestra que la correspondencia entre las curvas de E/S obtenidas con los dos métodos es buena para frecuencias mayores de 2 kHz pero no para frecuencias menores. En bajas frecuencias, las curvas de E/S de OEAPD frecuentemente presentan mesetas y muescas que no aparecen en las curvas de E/S inferidas mediante el método CET. Se discute que esto podría deberse a que las curvas de E/S de OEAPD se midieron con tonos primarios cuyos niveles sonoros son genéricos (Kummer et al., 1998) y no individualizados.

Esta explicación se explora en un segundo estudio mediante simulaciones. Los resultados sugieren que, en efecto, el empleo de primarios con niveles sonoros optimizados para maximizar el producto de distorsión minimizaría las mesetas y las muescas y, además, maximizaría el parecido entre la curva de E/S de OEAPD y la curva de E/S de la no-linealidad subyacente. Las simulaciones también sugieren

que pequeñas desviaciones respecto a los niveles primarios óptimos bastarían para producir muescas en las curvas de E/S de OEAPD.

Es una hipótesis común que el máximo nivel de OEAPD se genera cuando los dos tonos primarios (f_1 y f_2, $f_2>f_1$) producen idéntica excitación en el sitio de la membrana basilar (MB) sintonizado a la frecuencia del primario f_2. Esta hipótesis se comprueba aquí, en un tercer estudio, mediante un novedoso abordaje psicofísico basado en curvas CET. Los resultados apoyan la hipótesis, ya que los niveles primarios inferidos a partir de estas curvas para obtener igual excitación coclear coinciden con los valores empíricos que producen la máxima OEAPD.

Finalmente, se comparan las curvas de E/S inferidas mediante CET con curvas de E/S de OEAPD obtenidas con niveles primarios optimizados de dos formas distintas: los inferidos mediante el método CET para garantizar que los primarios produzcan igual excitación en la MB y los individualizados empíricamente para generar niveles máximos de OEAPDs. Se concluye que la correspondencia entre las curvas de E/S de OEAPD y las inferidas con el método CET sigue siendo alta para frecuencias mayores de 2 kHz, independiente de los niveles sonoros primarios empleados, y que el empleo de niveles sonoros optimizados individualmente no mejora la correspondencia a bajas frecuencias.

La discusión general se centra en las posibles razones para la baja correspondencia entre las curvas de E/S de OEAPD y las inferidas mediante CET en bajas frecuencias y también en las razones para las mesetas y las muescas presentes en las curvas de E/S de OEAPD. Finalmente, se sugieren nuevas líneas de investigación.

LISTA DE ABREVIATURAS Y ACRÓNIMOS[1]

ANSI	Instituto nacional americano de estándares (*American National Standards Institute*)
MB	Membrana basilar
CCE	Células ciliadas externas
CET	Curva de enmascaramiento temporal (*temporal masking curve*)
dB	Decibelios
E/S	Entrada-salida
EE	Error estándar
FC	Frecuencia característica
f_1, f_2	Frecuencias de los tonos primarios de OEAPD
f_P, f_M	Frecuencias de sonda y máscara
HI	Hipoacúsico (*hearing impaired*)
HL	Nivel de audición (*hearing level*)
L_M	Nivel de máscara
L_1, L_2	Niveles de f_1 y de f_2
NH	Audición normal (*normal hearing*)
OEAPD	Otoemisiones acústicas de productos de distorsión
PD	Producto de distorsión
RL	Regresión lineal
RMS	Root Mean Square
SPL	Nivel presión sonora (*sound pressure level*)
DE	Desviación estándar
SL	Nivel de sensación sonora (*sensation level*)
TFR	Transformación rápida de Fourier
TP	Tono puro

[1] Se ha decidido mantener algunas abreviaciones con sus siglas en inglés dado que no tienen una buena equivalencia en castellano.

CONTENIDO

CAPÍTULO 1

INTRODUCCIÓN GENERAL

1.1 Motivación y estado de la cuestión

Los humanos perciben los sonidos dentro de un rango muy amplio de niveles sonoros, probablemente gracias al funcionamiento de las células ciliadas externas (CCE) y su efecto de amplificación sobre la vibración de la membrana basilar (MB) (i.e., Oxenham y Bacon 2003). Efectivamente, una MB sana es extremadamente sensible frente a sonidos muy suaves y la magnitud de su respuesta crece compresivamente al incrementar el nivel sonoro (Fig. 1.1 y 1.2). De este modo, acomoda un rango de ~120 dB SPL al rango fisiológico más estrecho (por ejemplo, Robles y Ruggero 2001). Por otro lado, una MB dañada muestra menor sensibilidad y respuestas linealizadas (Fig. 1.1), lo cual posiblemente explica porqué los hipoacúsicos muestran umbrales de percepción elevados y un rango dinámico auditivo reducido (véase por ejemplo, Moore 2007). Por eso las características de la MB son importantes para describir la audición y típicamente se cuantifican mediante la curva de entrada/salida (E/S) de la MB, que es una representación de la relación entre el nivel sonoro de estimulación y la magnitud de la respuesta mecánica de la MB (Fig. 1.1). En roedores, la curva de E/S de la MB se puede medir directamente y, cuando se mide con un tono de frecuencia igual a la frecuencia característica (FC) del sitio de la MB en cuestión, típicamente muestra tres secciones: una con tasa de crecimiento lineal (~1 dB/dB) en niveles de entrada bajos, seguida de otra sección compresiva (tasa de crecimiento < 1 dB/dB) en

niveles medianos y, finalmente, seguida, en ocasiones, de otra sección lineal en niveles muy altos de entrada (Robles and Ruggero 2001).

A pesar de su importancia para la audición, las características de E/S de la MB humana no se conocen completamente. La respuesta de la MB humana no se puede medir directamente y por tanto se usan métodos indirectos para inferir las curvas de E/S. Existen varios métodos para inferir estas curvas de E/S de la MB a partir de datos conductuales (por ejemplo, López-Poveda y Alves-Pinto 2008; Nelson *et al.,* 2001; Oxenham y Plack 1997; Plack y Oxenham 2000). Estos métodos conductuales están razonablemente bien fundamentados (por ejemplo, Bacon y Oxenham 2004; Oxenham y Bacon 2004). Los distintos métodos consiguen resultados intra-sujeto similares (López-Poveda y Alves-Pinto 2008; Nelson *et al.,* 2001; Rosengard *et al.,* 2005) y son replicables en diferentes sesiones de medición.

El método de enmascaramiento temporal (CET) de Nelson *et al.* (2001) (véase también López-Poveda *et al.,* 2003) es, quizás, el más utilizado de los métodos psicoacústicos [revisado en Bacon (2004)] porque minimiza los efectos de escucha "fuera de frecuencia" que podrían ocurrir cuando el nivel de la sonda varía. Este método consiste en medir el nivel sonoro umbral para que un tono (máscara) pueda enmascarar a otro tono con nivel sonoro fijo (sonda) en función del tiempo transcurrido entre la máscara y la sonda. La curva CET es una representación gráfica del nivel resultante de la máscara en función del tiempo transcurrido entre la máscara y la sonda. Dado que el nivel sonoro de la sonda es fijo, el nivel de la máscara se incrementa al aumentar el tiempo máscara-sonda y, por lo tanto, las CET tienen pendientes positivas. Nelson *et al.* (2001) argumentaron que la pendiente de cualquier CET depende simultáneamente de la tasa temporal de recuperación del efecto interno (post-coclear o post-compresión) de la máscara y del grado de compresión en la MB que afecta a la máscara en el sitio coclear con FC aproximadamente igual a la frecuencia de la sonda. Asumiendo que la tasa de recuperación del efecto interno de la máscara es independiente de la frecuencia y del nivel sonoro de la máscara, se pueden inferir las curvas de E/S de la MB

representando el nivel sonoro de la máscara de una referencia lineal (es decir, una CET procesada linealmente por la MB) en función del nivel sonoro de cualquier otra máscara, para idénticos intervalos de tiempo máscara-sonda. Nelson *et al.* (2001) proporcionan una justificación completa de estos supuestos (véase también López-Poveda y Alves-Pinto, 2008).

El método CET se ha empleado para inferir las curvas de E/S de la MB en normoyentes e hipoacúsicos (por ejemplo, López-Poveda *et al.*, 2003, 2005; López-Poveda y Alves-Pinto, 2008; Nelson *et al.*, 2001; Nelson y Schroder, 2004; Plack *et al.*, 2004; Rosengard *et al.*, 2005). Es un método fiable, pero requiere mucho tiempo de ejecución y la participación activa del sujeto. Esto lo hace inviable en la clínica en general y, en particular, en sujetos no-cooperativos como niños pequeños o ancianos.

El oído humano no es un sistema de alta fidelidad; distorsiona la señal acústica dentro de la cóclea (Ruggero 1993). La distorsión se puede percibir como un sonido audible (Goldstein 1967) y se emite desde la cóclea hacia el canal auditivo, donde se manifiesta como una otoemisión acústica (Kemp 1978). De hecho, la distorsión emitida es un indicador de un oído sano, mientras que a mayor daño de la cóclea, más débil será la distorsión emitida (Dorn *et al*. 2001; Lonsbury-Martin y Martin 1990). El nivel de la distorsión emitida depende de los parámetros de los sonidos usados para generarla. Típicamente se usan dos tonos puros (denominados primarios) con una ligera diferencia de frecuencia (f_1 y f_2; f_2/f_1~1.2). La magnitud de la distorsión con frecuencia $2f_1-f_2$ se considera el indicador del estado fisiológico de la cóclea. Nos referimos a este indicador como otoemisión acústica de producto de distorsión (OEAPD) y a la representación gráfica de su nivel sonoro en función del nivel sonoro, L_2, del tono primario f_2 como curva de E/S de la OEAPD (Fig. 1.2).

Las curvas de E/S de OEAPD comparten muchas características con las curvas de E/S de la MB (Cooper y Rhode, 1997; Dorn *et al.*, 2001; Neely *et al.*, 2003). Específicamente, ambas son generalmente lineales a niveles de entrada bajos

pero compresivas por encima de un cierto umbral de compresión (Dorn *et al.*, 2001; Kummer *et al.*, 1998) y ambas son sensibles al daño de las células ciliadas externas (véase la Fig. 1.1) (Rhode, 2007). Esto sugiere que se podría usar las curvas de E/S de OEAPD como una forma rápida y universal de inferencia de curvas individuales de E/S de la MB en condiciones clínicas (Müller y Janssen, 2004). Además, su medición no requiere la participación activa del sujeto, por lo que se ha sugerido que OEAPD podría ser una alternativa útil a los métodos conductuales para inferir las curvas de E/S de la MB (véase, por ejemplo, Jansen y Müller 2008; Müller y Janssen 2004). Desgraciadamente, no existe seguridad acerca de si las curvas de E/S de OEAPD constituyen una estimación razonable de las curvas de E/S de la MB a nivel *individual* (por ejemplo, Williams y Bacon 2005).

El objetivo de esta investigación es definir los estímulos de OEAPD y las condiciones que permitan usar OEAPD como alternativa a los métodos conductuales para inferir curvas individuales de E/S cocleares para cada frecuencia. Aunque el objetivo a largo plazo sería extender este estudio a hipoacúsicos, aquí nos centramos en personas normoyentes.

1.2 Hipótesis

La hipótesis general es que las curvas de E/S cocleares inferidas psicoacústicamente y las curvas de E/S de OEAPD son consistentes mutuamente, dado que ambas dependen de las características de E/S de la MB.

Las hipótesis específicas son:

1. El método conductual de CET y el método fisiológico de OEAPD se pueden utilizar indistintamente para inferior curvas de E/S cocleares en humanos.
2. Las curvas de E/S de producto de distorsión (PD) se parecen mas a la no-linealidad causal subyacente cuando se miden con primarios cuyos parámetros están optimizados para generar la máxima distorsión.

3. La respuesta de OEAPD es máxima cuando los tonos primarios evocan la misma excitación en el sitio de la MB sintonizado a la frecuencia del primario f_2.
4. La correspondencia entre curvas de E/S conductuales, inferidas a partir de CET, y de OEAPD es mejor cuando se miden las OEAPD con parámetros optimizados individualmente para generar la respuesta máxima distorsión.

1.3 Objetivos

El objetivo general de esta tesis es comprobar las hipótesis planteadas, para ello procederemos de la siguiente manera:

1. Compararemos las curvas de E/S medidas con OEAPD con las inferidas mediante el método psicofísico de CET.
2. Simularemos curva E/S de PD para dos funciones no lineales simples y con tonos primarios de diferentes niveles sonoros. En un caso, los primarios tendrán niveles sonoros optimizados para generar la máxima distorsión.
3. Compararemos los niveles sonoros primarios inferidos a partir de CET diseñados para producir la misma excitación en el sitio de la MB sintonizado a f_2, con los optimizados empíricamente para maximizar el nivel sonoro de OEAPD.
4. Compararemos las curvas de E/S inferidas de CET con curvas de E/S de OEAPD medidas con primarios de niveles sonoros optimizados individualmente.

1.4 Estructura de la tesis

Esta tesis está organizada de manera que cada capítulo corresponde a una de las hipótesis y de los objetivos propuestos.

El Capítulo 2 responde a la primera de las hipótesis planteadas y presenta una comparación entre curvas de E/S inferidas de CET y de OEAPD. Las curvas de E/S de OEAPD se obtuvieron usando primarios con niveles óptimos en promedio pero no individualmente.

En el Capítulo 3, se presentan simulaciones dirigidas a comprobar la segunda hipótesis. Las curvas de E/S de PD fueron simuladas para dos funciones no lineales distintas y para primarios con dos niveles sonoros distintos, y comparadas con la no-linealidad causal subyacente. Además, se analiza cuál es el impacto sobre la curva E/S cuando los estímulos se desvían ligeramente de los niveles sonoros óptimos.

El Capítulo 4 describe un abordaje psicofísico para inferir los niveles sonoros óptimos de los primarios para cada sujeto, y comprueba la hipótesis habitual que afirma que la máxima generación de OEAPD ocurre cuando los dos primarios evocan idéntica excitación en el sitio de la MB sintonizado a f_2.

En el Capítulo 5, se usan los niveles sonoros de primarios optimizados individualmente para comprobar si se mejora la correlación entre curvas individuales de E/S, basadas en OEAPD y CET (cuarta hipótesis).

Los Capítulos 2, 4 y 5 están cada uno basados en artículos ya publicados de forma independiente. Los artículos originales se han reproducido en el Apéndice B.

El Capítulo 6 discute de forma integrada los resultados descritos en los capítulos anteriores y describe futuras investigaciones.

Finalmente, el apartado Conclusiones presenta las principales conclusiones de esta investigación.

1.5 Contribuciones originales.

Las contribuciones originales de esta tesis son:

1. Una demostración experimental de que las curvas de E/S en personas normoyentes, inferidas a partir de CET y estimadas de OEAPD, son consistentes para frecuencias mayores que ~2 kHz pero no para frecuencias inferiores.
2. Un conjunto de simulaciones que sugieren que las curvas de E/S de PD son más similares a la no-linealidad subyacente cuando los parámetros están optimizados para producir la máxima distorsión.
3. Un novedoso abordaje psicoacústico para apoyar la hipótesis habitual que afirma que la respuesta de OEAPD es máxima cuando los primarios producen la misma excitación en el sitio de la MB sintonizado a al primario f_2.
4. Una demostración experimental de que la correspondencia entre curvas de E/S inferidas de CET y de OEAPD no mejora cuando las OEAPD se miden con parámetros optimizados para producir la máxima OEAPD.

1.6 Unidades de medida

Todas las variables y medidas se expresan en unidades del Sistema Internacional.

CAPÍTULO 2

LA NO-LINEALIDAD COCLEAR EN PERSONAS NORMOYENTES INFERIDA CON TÉCNICAS PSICOFÍSICAS Y DE OTOEMISIONES ACÚSTICAS DE PRODUCTOS DE DISTORSIÓN[2]

2.1 Introducción

Este capítulo describe, un primer intento, de determinar si las curvas de E/S de OEAPD podrían utilizarse para estimar el grado individual de compresión coclear en normoyentes. Esta hipótesis se comprueba midiendo el grado de correlación entre las curvas de E/S de la MB, inferidas mediante el método psicofísico CET y OEAPD en el mismo sujeto. Solo si los resultados de los dos métodos coinciden, será posible apoyar el uso de OEAPD para inferir características individuales de las respuestas cocleares. Si no coinciden, será difícil averiguar cuál de los dos métodos (si alguno) estima mejor las características no-lineales de la respuesta de la MB.

Williams y Bacon (2005) infirieron curvas de E/S cocleares a partir de CET y OEAPD en cuatro sujetos para frecuencias de 1, 2, y 4 kHz. Los resultados revelaron que ambos métodos consiguen estimaciones similares de compresión en promedio, pero no a nivel individual. Aunque el objetivo de dicho estudio no fue investigar el grado de correlación intra-sujeto entre los dos métodos. Además, sus

[2] Este capítulo se basa en el artículo: Johannesen, P. T., and López-Poveda, E. A. **2008**. "Cochlear nonlinearity in normal-hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions," publicado en J. Acoust. Soc. Am. **124**, 2149–2163. Se reproduce aquí parte del artículo con permiso de "the Acoustical Society of America".

curvas de E/S de OEAPD pudieron estar afectadas por la estructura fina de OEAPD. En efecto, Gaskill y Brown (1990) mostraron variaciones rápidas (conocidas como "estructura fina") de la magnitud del $2f_1-f_2$ OEAPD cambiando solo ligeramente la frecuencia de los primarios (f_1 y f_2, con f_2 / f_1 =1.21). Se piensa que esta estructura fina es el resultado de interferencia constructiva y destructiva entre PDs, generados en dos sitios espacialmente distantes de la MB (Kummer *et al.*, 1995; Stover *et al.*, 1996; Gaskill y Brown, 1996; Heitmann *et al.*, 1998; Talmadge *et al.*, 1998; 1999; Mauermann *et al.*, 1999; Mauermann y Kollmeier, 1999; Shera y Guinan, 1999). El sitio principal de generación de OEAPD es la zona de la MB con máximo solapamiento entre las excitaciones causadas por los dos primarios, es decir, el sitio de la MB con FC= f_2 (e.g., Kummer *et al.*, 1995). La distorsión generada en la zona de máximo solapamiento se propaga hacia atrás, hacia la ventana oval, pero también hacia el sitio coclear con FC=$2f_1-f_2$ donde excita una segunda fuente de generación, denominada "fuente de reflexión". El PD generado en este sitio también se propaga hacia atrás, hacia la ventana oval, y es sumado de forma vectorial a la respuesta del primer componente. La estructura fina ocurre cuando las fases de los dos componentes varían, y a veces se suman y a veces son opuestos, dando lugar de esta forma a interferencia constructiva o destructiva, respectivamente, y por lo tanto a picos y muescas en el PD-grama. El generador situado en f_2 domina con altos niveles de estimulación [Fig. 3 de Mauermann y Kollmeier (2004)], lo cual explica porqué la estructura fina es más pronunciada en bajos niveles.

La estructura fina también influye en las curvas de E/S de OEAPD, en especial a niveles bajos (Mauermann y Kollmeier, 2004). La estructura fina afecta en particular a nivel individual, pero también al promedio de una muestra pequeña, por lo tanto puede haber influido en las conclusiones de Williams y Bacon (2005).

El presente estudio amplía otros estudios en varios aspectos. Primero, el enfoque es la correlación psicofísica/fisiológica intra-sujeto y no el promedio de grupo. Segundo, las curvas de E/S de la MB se infirieron con el, probablemente, mejor método psicofísico (Nelson *et al.*, 2001). Tercero, se ha pretendido disminuir la influencia de la estructura fina promediando cinco frecuencias f_2, situadas

alrededor de la frecuencia de interés. Cuarto, el rango de frecuencias estudiadas incluye frecuencias bajas y altas (0.5–4 kHz). Quinto, la comparación entre psicofísica y fisiología es ampliada a otros parámetros, aparte de la magnitud de compresión, como por ejemplo; umbral de compresión, nivel al cual ocurre la compresión máxima, etc.

2.2 Metodología

2.2.1 Sujetos

Diez normoyentes participaron en este estudio. Todos tenían umbrales auditivos inferiores a 20 dB HL en las frecuencias estudiadas (0.5, 1, 2, y 4 kHz, véase tabla 2.1).

2.2.2 Estímulos de CET

Se midieron CET para frecuencias de sonda (f_P) de 0.5, 1, 2, y 4 kHz y con igual frecuencia de la máscara (frecuencia-igual). Una CET adicional fue medida para una frecuencia de sonda de 4 kHz y una frecuencia de máscara de 1.6 kHz (f_M=0.4f_p). Esta última fue seleccionada como referencia lineal (López-Poveda y Alves-Pinto, 2008) y usada para inferior curvas de E/S de la MB para todas las frecuencias de la sonda (López-Poveda *et al.*, 2003). La duración de la máscara fue de 110 ms, incluyendo rampas de 5 ms. La duración de la sonda fue de 10 ms. incluyendo rampas de 5 ms. El nivel de la sonda fue de 9 dB de nivel de sensación (SL) (es decir, 9 dB encima del umbral individual de la sonda).

2.2.3 Procedimiento CET

El procedimiento fue idéntico al de López-Poveda y Alves-Pinto (2008). Los umbrales de enmascaramiento fueron medidos usando un paradigma de dos intervalos, con dos alternativas y elección forzada. El nivel de la máscara fue modificado según un procedimiento adaptativo, dos-arriba y uno-abajo, que estima

el punto 71% de la función psicométrica (Levitt, 1971). Los umbrales fueron repetidos tres veces y su promedio considerado el umbral resultante. Si la desviación estándar superaba 6 dB se medía e incluía una cuarta repetición.

Los umbrales absolutos fueron medidos de la misma forma y se muestran los resultados en tabla 2.1.

2.2.4 Inferencia de curvas de E/S de la MB desde CET

Las curvas de E/S de la MB fueron inferidas a partir de CET dibujando los niveles de la referencia lineal, en función de los niveles de máscara (para todas sus frecuencias), pareado según intervalo máscara-sonda (Nelson *et al.*, 2001). La CET de f_P=4 kHz fue usada como referencia lineal para inferir las curvas de E/S para el resto de las frecuencias, como han sugerido López-Poveda *et al.* (2003). En ocasiones, fue necesario extrapolar la referencia lineal a intervalos más largos, para poder inferir las curvas de E/S del rango más amplio posible. Hay evidencia de que el decremento del efecto de enmascaramiento temporal, se describe mejor con dos constantes de tiempo (por ejemplo, López-Poveda y Alves-Pinto, 2008; Meddis y O'Mard, 2005; Oxenham y Moore, 1994; Plack y Oxenham, 1998) y, por lo tanto, se ha ajustado la referencia lineal utilizando una función doble exponencial de la siguiente forma:

$$L_m = L_0 - 20log10\left[\alpha e^{(-t/\tau_a)} + (1-\alpha)e^{(-t/\tau_b)}\right], \qquad (2.1)$$

donde $L_m(t)$ es el nivel de máscara del intervalo de tiempo t; L_0 es el nivel de máscara en el tiempo cero; τ_a y τ_b son constantes de tiempo; y α determina el nivel a partir del cual la segunda exponencial domina sobre la primera. Los parámetros α , τ_a, τ_b, y L_0 fueron parámetros de ajuste. El ajuste a los datos reales fue muy bueno (r=0.94).

2.2.5 Estímulos de OEAPD

Las curvas de E/S de OEAPD se obtuvieron dibujando la magnitud (en dB SPL) del PD $2f_1-f_2$, en función del nivel L_2 del primario f_2. Las OEAPD se midieron solo

para las frecuencias f_2 iguales a las frecuencias de la sonda para las que se midieron las CET (0.5, 1, 2 y 4 kHz). El ratio f_1/f_2 fue fijado a 1.2. L_2 variaba entre 20 y 75 dB SPL en pasos de 5 dB, excepto para f_2 =0.5 kHz que variaba entre 45 y 75 dB SPL. L_1 y L_2 se relacionaban según la regla sugerida por Kummer *et al.* (1998), L_1=0.4L_2+39 dB.

En un intento de reducir la variabilidad de las curvas de E/S, por la estructura fina del PD, se midieron cinco frecuencias cercanas a la frecuencia de interés, y las curvas de E/S resultantes fueron promediadas. Este procedimiento se apoya en Kalluri y Shera (2001) y Mauermann y Kollmeier (2004) que mostraron, que un promedio por frecuencia, se parece a un PD-grama, en el cual ha sido eliminado el efecto de la estructura fina. Por esa razón se midieron las curvas de E/S para frecuencias 0.98f_2, 0.99f_2, f_2, 1.01f_2, y 1.02f_2. Por ejemplo, la curva de E/S final para 4 kHz fue el promedio de las curvas de E/S para f_2 = {3920, 3960, 4000, 4040, 4080 Hz}.

Las OEAPD se midieron para frecuencias $f_2 \leq 4$ kHz, para evitar la influencia de ondas estacionarias (Siegel, 1994; Whitehead *et al.* 1995). Dado que es difícil medir OEAPD para frecuencias bajas (f_2<1 kHz), por el ruido fisiológico, solo se midieron niveles superiores a 45 dB SPL para la frecuencia f_2=0.5 kHz.

2.2.6 Calibración de los estímulos de OEAPD

Los estímulos de OEAPD se calibraron en un oído artificial, Zwislocki DB-100, para cada frecuencia f_1, f_2. Los niveles calibrados *no* fueron modificados posteriormente *in situ*.

2.2.7 Artefactos de sistema de OEAPD

El sistema de OEAPD puede producir respuestas propias, no fisiológicas (artefactos), con altos niveles de primarios. Los límites de estas respuestas artificiales se fijaron, midiendo en oídos artificiales, usando las mismas condiciones que en las mediciones en sujetos.

En este estudio se utilizó un criterio más exigente para aceptar una respuesta como fisiológica de lo que es común en la práctica clínica. Una respuesta tenía que superar el límite del artefacto más 6 dB, en caso contrario la medición era descartada.

2.2.8 Procedimiento de OEAPD

Las mediciones OEAPD se obtuvieron en una cabina insonorizada con un equipo "IHS Smart system" (con software SmartOAE versión 4.52) equipado con una sonda Etymotic ER-10D.

El tiempo de medición de OEAPD oscilaba entre los 8 y los 60 segundos para f_2=4 kHz y entre 30 y 60 segundos para f_2=0.5 kHz. La medición se consideraba válida cuando la respuesta era superior al ruido más 2 desviaciones estándar.

2.3. Resultados

2.3.1 CET

La Figura 2.1 muestra las curvas CET para los 10 sujetos y para frecuencias de sonda de 0.5, 1, 2 y 4 kHz. La forma de las curvas CET obtenidas es consistente con la de anteriores estudios (por ejemplo, Nelson *et al.*, 2001; López-Poveda *et al.*, 2003; López-Poveda y Alves-Pinto, 2008; Plack *et al.*, 2004; Rosengard *et al.*, 2005). La CET de la referencia lineal (cuadrados abiertos) se puede describir como una línea recta (por ejemplo, S4, S5) o como una función pendiente que gradualmente se satura (por ejemplo, S3, S8). Este último caso justifica ajustar una función doble exponencial a la CET de referencia lineal (línea continua y gruesa). Algunas CET, de la condición con igual frecuencia de sonda y máscara, muestran un segmento casi plano para intervalos cortos de máscara-sonda, seguido de un segmento más pendiente para intervalos medios. Otras se describen mejor con un segmento pendiente, seguido a veces de un segmento casi plano a altos niveles de máscara (por ejemplo, S1, S2, S6, y S8). Todas las curvas CET, de la condición con

igual frecuencia de máscara y sonda, tienen segmentos mucho más pendientes que la referencia lineal. Asumiendo, que la condición con desigual frecuencia de sonda y máscara, utilizada para la referencia lineal, se procesa de forma lineal en la MB, y que la recuperación del efecto interno de la máscara es idéntica para todas las frecuencias (López-Poveda *et al.*, 2003; López-Poveda y Alves-Pinto, 2008), los segmentos pendientes se puede interpretar como indicación de compresión en la MB (Nelson *et al.*, 2001). La validez de estos supuestos se discutirá en profundidad en la sección 6.1.

2.3.2 OEAPD

La Figura 2.2 muestra un ejemplo de la cantidad de datos necesarios para estimar una curva de E/S de OEAPD. Cada punto de dato es el promedio de 3 repeticiones. La Figura 2.2 también ilustra la influencia de la estructura fina del PD sobre la curva de E/S resultante. Evidentemente la curva de E/S cambiaría mucho con un pequeño cambio de f_2, lo cual subraya la necesidad de tener en cuenta el efecto de la estructura fina a la hora de estimar curvas de E/S. La curva final consiste en el promedio de las curvas de E/S de las 5 frecuencias centrales.

La Figura 2.3 ilustra estas 5 curvas de E/S para los sujetos S1 y S2. Como se aprecia, hay mucha variabilidad entre sujetos y entre frecuencias. Parece que hay menos variabilidad a altos niveles de estimulación, lo cual resulta razonable dado que la estructura fina es relativamente más importante a niveles bajos de estimulación (Mauerman y Kollmeier, 2004).

2.3.3 Comparación de curvas de E/S de la MB inferidas a partir de CET y OEAPD

Las Figuras 2.4 a 2.7 permiten comparaciones intra-sujeto de curvas de E/S de OEAPD (cuadrados abiertos) y curvas de E/S de la MB, inferidas a partir de CET (círculos llenos), para los correspondientes sitios cocleares, con FC de 0.5, 1, 2 y 4 kHz, respectivamente. Las líneas sólidas son polinomios de tercer orden ajustados

a los datos. Las barras de error muestran un error estándar (EE) del promedio. Las barras de errores verticales y horizontales, de las curvas de E/S basadas en CET, ilustran el E/S basado en tres repeticiones de la referencia lineal y de la condición de frecuencias iguales de mascara y sonda en cuestión, respectivamente.

Los triángulos, en el panel arriba-derecho de las Figuras 2.4–2.7, y sus barras de error asociadas, ilustran el umbral de ruido de las OEAPD más 2 desviaciones estándar.

En general, las curvas de E/S basadas en OEAPD y CET son similares, en el sentido de que son lineales a niveles bajos y, gradualmente, serán compresivas con incrementos del nivel de estimulación. Hay una tendencia a volver a la linealidad a niveles más altos. (por ejemplo, S6 en 0.5 kHz en Fig. 2.4; o S2 y S5 en 4 kHz en Fig. 2.7). El grado de semejanza de las curvas de E/S de los dos métodos es mayor en 4 kHz (Fig. 2.7 y sujetos S1, S2, S5, S7, S3 y S8).

El grado de semejanza entre las formas de las curvas de E/S es muy inferior en 0.5 y 1 kHz. Es destacable que en estas frecuencias, las curvas de E/S de OEAPD a menudo tienen muescas y mesetas, que no tienen una clara correlación con curvas de E/S basadas en CET (por ejemplo, S1, S2, y S8 en 0.5 kHz en Fig. 2.4; o S3, S4, y S5 en 1 kHz en Fig. 2.5).

2.3.4 Comparación de parámetros derivados de la no-linealidad coclear

Los polinomios de tercer orden fueron ajustados a las curvas de E/S de OEAPD, y a las basadas en CET (líneas continuas en Figs. 2.4–2.7), y utilizados para inferir los siguientes parámetros relativos a la no-linealidad coclear: exponente mínimo de compresión, el nivel en el que este ocurrió, umbral de compresión, ganancia coclear y el nivel en el cual la curva de E/S vuelve a linealidad a niveles altos.

El exponente mínimo de compresión fue estimado como la pendiente mínima de la función ajustada a los datos. La Figura 2.8(A) compara el exponente mínimo de compresión inferido de las curvas de E/S obtenidas con los dos métodos. Las

regresiones lineales (líneas gruesas continuas y punteadas) fueron obligadas a pasar por el origen de la gráfica e indican que hay una correlación alta entre los dos métodos a 4 kHz (coeficiente de correlación de Pearson de 0.92) y correlación baja (0.19) incluyendo todas las frecuencias. La razón para la disimilitud entre los resultados a 0.5 y 1 kHz es, casi seguramente, que muchas curvas de E/S de OEAPD muestran muescas y mesetas, lo cual resultó en exponentes de compresión bajos (pendientes < 0 dB/dB). La tabla 2.2 muestra, para cada frecuencia, el número de casos de curvas ajustadas exhibiendo mesetas o exponentes negativos de compresión.

Al contrario del resultado anterior, Williams y Bacon (2005), reportaron una correlación alta entre las estimaciones de compresión basadas en CET y OEAPD, lo cual se mostró también en baja frecuencia. Ellos ajustaron una línea recta a las curvas de E/S de CET y de OEAPD y este método podría ser menos sensible frente a muescas. La Figura 2.8 (B) muestra la correlación entre el exponente de compresión de ambos métodos, estimado a través de una regresión lineal aplicada a las curvas de E/S, entre los niveles 40 y 65 dB. En la Figura 2.8 (B) se puede apreciar que la mayoría de los exponentes de los OEAPD son inferiores a los de las CET (los datos están debajo de la diagonal), lo cual también se refleja en el coeficiente de correlación de Pearson, que se sitúa a 0.32 contemplando todas las frecuencias y a 0.53 limitándose solo a 4 kHz.

Otra característica de la no-linealidad coclear es el nivel donde ocurre la máxima compresión (o equivalente, mínimo exponente de compresión). La correlación entre estimaciones de este parámetro, obtenido a través de los dos métodos, es baja como revela la Figura 2.9(A). La mayoría de los datos están por debajo de la diagonal, y por lo tanto, el nivel donde ocurre la máxima compresión es inferior para las curvas de E/S de OEAPD (promedio de 61 dB SPL) que para las de CET (promedio de 76 dB SPL). La diferencia es estadísticamente significativa (p<0.001).

El umbral de compresión fue definido como el nivel de entrada en el cual la pendiente pasa de valores cerca de uno a valores inferiores de 0.4 dB/dB, yendo

hacia mayores niveles de entrada. La Figura 2.9 (B) ilustra la correlación entre umbrales de compresión estimados a partir de curvas de E/S basadas en CET y OEAPD. Las líneas de regresión indican un alto grado de correlación entre los dos métodos siendo el coeficiente de correlación de Pearson relativamente alto (r=0.8). El promedio del umbral de compresión fue inferior para las curvas de E/S de OEAPD que para las de CET, (40 y 47 dB SPL, respectivamente), lo cual también se aprecia en la Figura 2.9 (B). La diferencia es estadísticamente significativa (p<0.0001).

2.3.5 Dependencia de la no-linealidad coclear de la frecuencia característica

Esta sección trata acerca de si los parámetros descritos en la sección anterior varían con la frecuencia y si se infirieron a partir de CET u OEAPD. Además, se contemplan dos parámetros adicionales, basados solo en CET; ganancia y vuelta a linealidad a niveles altos. Las curvas de E/S se pueden describir generalmente como un segmento lineal (pendiente de 1 dB/dB) a niveles bajos, seguido de un segmento compresivo (pendiente < 1 dB/dB), a niveles moderados, y otro segmento lineal a niveles altos (por ejemplo, López-Poveda *et al.*, 2003). La pendiente de la curva de E/S, sin embargo, siempre se incrementa a medida que aumenta el nivel de entrada, a partir del punto de inflexión de la curva (definido como el nivel donde hay máxima compresión o equivalente del nivel al cual la segunda derivada de la curva de E/S es igual a cero). Esto sugiere que la curva de E/S podría alcanzar linealidad a niveles muy altos. Dado que las pendientes mínimas siempre fueron inferiores de 0.4 dB/dB [Fig. 2.8(A)], el nivel de vuelta a linealidad se ha definido, como el nivel, donde la pendiente de la curva de E/S ajustada, alcanza 0.4 dB/dB para niveles por encima del punto de inflexión. La ganancia se definió como la diferencia entre la vuelta a la linealidad y el umbral de compresión en decibelios.

La Figura 2.10 muestra que el exponente mínimo de compresión, el nivel donde ocurre y el umbral de compresión son aproximadamente constantes, no

cambian con la frecuencia y no dependen de si se estiman con CET u OEAPD.

La Figura 2.11 muestra que la ganancia y vuelta a la linealidad, inferidas solo a partir de CET, tienden a incrementarse con la frecuencia. Dado que el umbral de compresión se mantiene aproximadamente constante con la frecuencia, el incremento de la ganancia con la frecuencia se atribuye al incremento con la frecuencia de la vuelta a la linealidad. Una correspondencia exacta entre estos parámetros no es esperable, dado que no se basan en los mismos sujetos. En algunos sujetos no fue posible medir una vuelta a la linealidad, sobre todo en 4 kHz, y estos sujetos podrían tener su vuelta a la linealidad a niveles mayores de los considerados aquí. Esto también apoya la tendencia observada que la vuelta a la linealidad ocurre a niveles mayores en alta frecuencia.

2.4 Discusión

Fueron tres los objetivos de este estudio. El primero era comparar parámetros de no-linealidad coclear, inferidos a partir de CET y OEAPD, y en caso de coincidencia, apoyar la conjetura de que ambos métodos son manifestaciones equivalentes de la no-linealidad coclear. El segundo objetivo era evaluar la factibilidad de usar curvas de E/S de OEAPD, como una herramienta rápida, para estimar parámetros individuales de no-linealidad coclear. El tercer objetivo era investigar la dependencia respecto a la frecuencia de los parámetros inferidos de CET y OEAPD que describen la no-linealidad coclear.

2.4.1 Equivalencia entre curvas de E/S de OEAPD y de CET

El grado de correlación entre estimaciones de compresión inferidas con los dos métodos fue alta (0.92) en 4 kHz, pero mucho más baja en 0.5 y 1 kHz. Se ha supuesto que las curvas de E/S de OEAPD reflejan las características de la respuesta de la MB frente a un solo tono en el sitio f_2. Una posible explicación de la falta de coincidencia es que la supresión mutua entre los primarios f_1 y f_2 podría influir en

la forma de las curvas de E/S de OEAPD, en particular en baja FC (tabla 2.2). Esto se discutirá más ampliamente en la sección 6.2.

Los valores del promedio de grupo del exponente de compresión en 4 kHz (0.10 y 0.11 dB/dB para CET y OEAPD, respectivamente) son parecidos a los valores basados en OEAPD de Gorga *et al.* (2007). Los promedios de grupo del exponente de compresión (0.36 y 0.25 para CET y OEAPD, respectivamente), obtenidos con regresión lineal sobre el rango de niveles medios [Fig. 2.8 (B)], fueron ligeramente más altos que los reportados por Williams y Bacon (2005). En todo caso, el presente estudio subraya la necesidad de especificar claramente el método usado para inferir cada parámetro, visto que hay mucha diferencia entre ajustar con polinomios y la regresión lineal.

El exponente mínimo de compresión a 4 kHz (basado en CET y ajustado con polinomio) es inferior a los valores de otros estudios [por ejemplo, Nelson y Schroder (2004), Plack y Drga (2003), Rosengard *et al.* (2005), Plack *et al.* (2004), López-Poveda *et al.* (2003)]. La diferencia podría tener que ver con la referencia lineal usada. Aquí se ha usado una máscara con frecuencia de $0.4f_P$, mientras que los otros estudios han usado una frecuencia de máscara entre $0.5f_P$ y $0.6f_P$. López-Poveda y Alves-Pinto (2008) sugieren que este último podría estar comprimido tanto como 2:1; por lo que podría subestimarse la compresión real. La coherencia entre estimaciones obtenidas de CET y OEAPD apoya de forma circunstancial las conclusiones de López-Poveda y Alves-Pinto (2008).

2.4.2 Muescas y mesetas de OEAPD

Muescas y mesetas son comunes en las curvas de E/S de OEAPD (Figs. 2.4–2.7, en particular en baja f_2, tabla 2.2). Esto contrasta con las conclusiones de Kummer *et al.* (1998) que han reportado que muescas y mesetas fueron menos comunes cuando las curvas de E/S de OEAPD eran medidas con su regla de primarios que con otras reglas. La regla de niveles primarios de Kummer *et al.* se basa en un promedio de grupo y, en algunos casos, podría desviarse considerablemente de la regla óptima

individual (la regla óptima sería la que evoca la máxima respuesta de OEAPD en todos los niveles). Por lo tanto, una posible explicación es, simplemente, que la regla de Kummer *et al.* (1998) no era óptima para los sujetos en el presente estudio. Esta explicación es apoyada por Neely *et al.* (2005), que reportaron que L_1 debería ser sistemáticamente más alto que el prescrito por Kummer *et al.* (1998) y que la relación $L_1 - L_2$ debería variar con f_2. Johnson *et al.* (2006) confirmaron esto y además sugirieron que la ratio f_2/f_1 debería variar ligeramente con f_2. Curiosamente, es precisamente en baja frecuencia y en niveles moderados-altos donde la regla de Kummer *et al.* se desvía más de la regla de Neely *et al.* (2005) y Johnson *et al.* (2006). Estas fueron también las condiciones donde las mesetas y muescas fueron más comunes en los presentes datos, lo cual sugiere que la regla de Kummer *et al.* no es óptima en baja frecuencia para los sujetos del presente estudio. Por otro lado, Kummer *et al.* (2000) verificaron su paradigma original [derivado de datos de Gaskill y Brown (1990)] con una muestra más grande y todavía resultó ser independiente de la frecuencia. Esta hipótesis se explorará en el Capítulo 5.

Entre las posibles razones para estas muescas y mesetas está que otra fuente de generación de PD (véase Mills, 1997), a altos niveles de estimulación, empieza a crear interferencia destructiva con la fuente de generación normal de nivel medio-bajo. Las muescas y mesetas también podrían ser debidas a la estructura fina, incluso con las medidas tomadas para minimizar su efecto. Sin embargo, parece una explicación menos probable dado que la estructura fina tiene importancia, sobretodo, a niveles bajos (Mauermann y Kollmeier, 2004) y las muescas y mesetas ocurrieron a niveles medios-altos. Se discutirán estas posibles explicaciones en más detalle en la sección 6.4 y 6.6.

2.4.3 El nivel donde ocurre el exponente mínimo de compresión

Se desconoce la causa para la baja correlación de este parámetro entre los dos métodos (OEAPD y CET).

2.4.4 Umbral de compresión

Hubo una correlación moderadamente alta entre los umbrales de compresión inferidos de los dos métodos [Fig. 2.9 (B)]. Las estimaciones basadas en OEAPD fueron, en promedio, 7 dB inferiores a las estimaciones a partir de CET. Se midieron un total de 22 curvas de E/S en 10 sujetos y se ha podido estimar el umbral de compresión en solo 14 casos, lo cual podría interpretarse en contra de la correspondencia entre los dos métodos. Sin embargo, hay buenas razones que explican la mayoría de los 8 casos sin umbral de compresión. Cuatro corresponden a datos de OEAPD de 0.5 kHz (Fig. 2.4), donde el rango de estímulo (L_2 > 45 dB SPL) era insuficiente para revelar el umbral de compresión. Otros 3 casos pertenecen a CET, donde no se alcanzó el criterio de pendiente de 0.4 dB/dB (véase sec. 2.3.4). Tres de los 14 casos con umbral de compresión fueron considerados valores atípicos, porque tenían mucha dificultad para realizar las CET y se han excluido de la comparación de umbral de compresión [círculos en Fig. 2.9 (B)]. Se desconoce la razón del porqué las estimaciones de umbral de compresión, basada en OEAPD, fueron inferiores a los basados en CET. Quizás la respuesta OEAPD estuvo influenciada por la supresión mutua entre los primarios, (véase también sección 6.1) lo cual podría haber hecho más lineal la curva de E/S (disminuye la ganancia coclear) y, por lo tanto, haber incrementado el umbral de compresión. Pero se descarta esta explicación porque la tendencia de los datos es la contraria. También es posible que la diferencia de las estimaciones de los dos métodos se deba al uso de parámetros sub-óptimos de OEAPD (véase capítulo 5 y sec. 6.5).

La correlación moderada-alta entre estimaciones de umbral de compresión, también a frecuencia baja [Figs. 2.9 (B), 2.10 (E), y 2.10 (F)], podría sorprender, dada la baja correlación entre estimaciones de exponente de compresión. Una posible explicación es que los parámetros de OEAPD fueron adecuados en niveles bajos L_2, donde está el umbral de compresión, pero no a niveles más altos de L_2, donde ocurren las muescas y mesetas.

El promedio de umbral de compresión a 4 kHz fueron 47 y 40 dB SPL para CET y OEAPD, respectivamente. Estos valores son comparables a los que obtuvieron Yasin y Plack (2003) mediante métodos psicofísicos (promedio de 37 dB SPL a 3–4 kHz). Se obtiene un valor de ~35 dB SPL aplicando nuestra definición de umbral de compresión a los datos de OEAPD de Neely *et al.* (2003). Los umbrales de compresión también fueron estimados a partir de los datos de OEAPD de Gorga *et al.* (2007) utilizando nuestro criterio de umbral. Sus resultados (30 y 45 dB SPL, en 0.5 y 4 kHz, respectivamente) concuerdan con los nuestros a 4 kHz pero no a 0.5 kHz, donde nuestros resultados indican umbrales de compresión independientes de la frecuencia [Figs. 2.10 (E) y 2.10 (F)].

2.4.5 Ganancia y el nivel de vuelta a la linealidad

Es todavía controvertido si la curva de E/S de la MB es más lineal a niveles muy altos en la cóclea sana (Robles y Ruggero, 2001). Suponiendo que es una característica fisiológica real, no se considera OEAPD un predictor fiable del umbral de vuelta a la linealidad. Primero, podría haber fuentes adicionales de generación a altos niveles (sec. 2.4.2). Segundo, el sitio de la MB, sintonizado a una frecuencia determinada, cambia a medida que se incrementa el nivel y como consecuencia, el sitio de máxima excitación por los dos primarios de las OEAPD, probablemente cambia. Esto implica que cambia con el nivel la zona de máximo solapamiento por los dos primarios, y con esto la respuesta OEAPD. Por esta razón, no se ha estimado ganancia y el umbral de vuelta a la linealidad a partir de OEAPD.

Las curvas de E/S basadas en CET sugieren que la ganancia coclear se incrementa con la frecuencia paralelamente al umbral de vuelta a la linealidad a niveles altos (Fig. 2.11). Esto concuerda con los datos fisiológicos (Robles y Ruggero, 2001). Las estimaciones de ganancia basadas en diferencias pico-base de supresión de curvas de sintonización (Gorga *et al.*, 2008) incrementaron también con la frecuencia.

2.4.6 Acerca de los méritos de OEAPD y CET para estimar curvas de E/S cocleares

La presencia de muescas y mesetas en las curvas de E/S de OEAPD en 0.5 y 1 kHz resulta en exponentes de compresión cero o negativos, lo cual no ocurre en las curvas de E/S basadas en CET. Se han encontrado muescas profundos en curvas de E/S apicales de la MB [por ejemplo, Fig. 7(a) de Rhode y Cooper (1996)], pero típicamente ocurren para frecuencias de estimulación más alta que la FC (Rhode y Cooper, 1996). Por eso resulta improbable que las muescas encontradas reflejen muescas reales en las respuestas de la MB (véase también sec. 2.4.2).

Sin embargo, sería equivocado concluir que es inapropiado utilizar OEAPD en baja frecuencia para inferir curvas de E/S cocleares. Como se mencionó anteriormente (sec. 2.4.2), las muescas podrían ser debidas al uso de niveles primarios sub-óptimos en baja frecuencia, y podría ser posible encontrar parámetros de OEAPD que consigan mejor correlación entre curvas de E/S de CET y OEAPD. También sería equivocado concluir que las curvas de E/S inferidas de CET son mas correctas (es decir, que reflejan mejor la respuesta real de la MB) que las de OEAPD, dado que el método CET también se basa en varias supuestos, algunos de los cuales han sido cuestionados recientemente (véase también sec. 6.1).

En resumen, la falta de correlación entre los dos métodos en baja frecuencia aporta poca información respecto a la precisión para inferir curvas de E/S cocleares.

2.5 Conclusiones

(1) La correlación entre estimaciones del exponente de compresión individual inferidas de CET y OEAPD es razonablemente alta en 4 kHz, pero no en 0.5 y 1 kHz. Ambos métodos sugieren que la máxima compresión es aproximadamente 10:1 y constante con la frecuencia (0.5 a 4 kHz).

(2) La baja correlación en baja frecuencia pone en duda los postulados e interpretaciones de las curvas de E/S de uno y/o ambos métodos. La causa más

probable para la falta de correlación en baja frecuencia (0.5–1 kHz) es la presencia de muescas y mesetas en las curvas de E/S de OEAPD. Esto sugiere que el paradigma de estimulación de Kummer *et al.* (1998) podría no ser óptimo en baja frecuencia.

(3) Hay una alta correlación entre las estimaciones de umbral de compresión inferidos de OEAPD y CET entre 1 y 4 kHz. OEAPD y CET indican umbrales de 40 y 47 dB SPL, respectivamente y constante con la frecuencia entre 0.5 a 4 kHz para CET y entre 1 y 4 kHz para OEAPD.

(4) La ganancia coclear y el umbral de vuelta a la linealidad fueron inferidas solo de CET. Ambos parámetros se incrementaron ~16 dB con frecuencia (0.5 a 4 kHz).

(5) Parece razonable usar curvas de E/S de CET y OEAPD de forma intercambiable a 4 kHz pero existe duda si lo mismo se aplica a baja frecuencia entre 0.5 y 1 kHz.

CAPÍTULO 3

SIMULACIÓN DE LA DEPENDENCIA DE LA CURVA DE ENTRADA/SALIDA DE PRODUCTOS DE DISTORSIÓN DE LA REGLA DE PRIMARIOS[3]

3.1 Introducción

En el presente estudio, como en muchos anteriores, (por ejemplo, Dorn *et al*., 2001; Williams y Bacon, 2005; Johannesen y López-Poveda, 2008; Neely *et al*., 2009), se supone, implícitamente, que la curva de E/S del componente $2f_1–f_2$ del OEAPD es una descripción razonable de la curva de E/S de la MB (medida con tono puro igual a la FC) de la región coclear con FC igual a f_2. Con otras palabras, sería posible inferir la curva de E/S en la frecuencia característica de la MB a partir del crecimiento del componente $2f_1–f_2$ del OEAPD con L_2 incrementando y con f_2 = FC. Este capítulo aporta evidencia que muestra que esta suposición es más razonable cuando se utilizan niveles L_1 óptimos. También se muestra evidencia que apoya que niveles L_1 sub-óptimos podrían producir curvas de E/S del PD $2f_1–f_2$ no-monótonas, incluso en ausencia de una segunda fuente de generación (estructura fina).

[3] Este capítulo se basa en el apéndice A del siguiente artículo: Johannesen, P. T., and López-Poveda, E. A. **2010**. "Correspondence between behavioural and individually "optimized" otoacoustic emission estimates of human cochlear input/output curvas," publicado en J. Acoust. Soc. Am. **127**, 3602-3613. Se reproduce aquí parte del artículo con el permiso de "the Acoustical Society of America".

Idealmente, se debería mostrar matemáticamente de forma analítica la suposición en cuestión para la no-linealidad subyacente de la respuesta de la MB. Sin embargo, esto no es posible porque los sistemas no-lineales no tienen una forma exacta de la función de transferencia, y porque tampoco se conoce con certeza la forma de la no-linealidad de la MB. La presente sección aporta solo una demostración, muy simple, para dos tipos de no-linealidad: una función no-lineal de ganancia de dos segmentos rectos (Meddis *et al*., 2001) y una función doble-Boltzmann (por ejemplo, Lukashkin y Russell, 1998, 2001). Estas funciones no-lineales se han usado con éxito para describir las respuestas de la MB y de las CCE en modelos computacionales. No se considera la filtración coclear en aras de la simplicidad.

3.2 Métodos

Denominamos $x(t)$ e $y(t)$ a las señales temporales de entrada y salida de la no-linealidad. La no-linealidad de dos segmentos rectos tiene la siguiente forma:

$$y(y) = \min[ax(t), bx^c(t)], \tag{3.1}$$

donde a y b son parámetros de ganancia y c es el exponente de compresión. Asimismo, la no-linealidad doble-Boltzmann tiene la forma:

$$y(t) = G_M \left\{1 + e\left(\frac{x_1 - x(t)}{s_1}\right)\left[1 + e\left(\frac{x_2 - x(t)}{s_2}\right)\right]\right\}^{-1}, \tag{3.2}$$

donde G_M, x_1, x_2, s_1 y s_2 son parámetros. Aquí, se ha supuesto que la señal de entrada es en unidades de μPa, y los parámetros de las dos funciones fueron arbitrariamente elegidos a los siguientes valores: $a = 1$; $b = 437$; $c = 0.2$; $x_1 = 41$; $x_2 = 24$; $s_1 = 62.5$; y $s_2 = 15.38$.

Las características de las respuestas de la curva de E/S se obtuvieron considerando señales digitales de uno o dos senos. Se denominaran curvas de E/S de tono puro (TP) y de producto de distorsión (PD) respectivamente. Las curvas de E/S se obtuvieron utilizando un solo seno, con frecuencia fija (f_2) y amplitud (A_2), que variaba en pasos de 2 dB (re 20 μPa). Una transformación Fourier rápida (TFR)

fue aplicada a las señales de salida, y la amplitud de la frecuencia f_2 fue dibujada en función de A_2, en escala log-log, para obtener la curva de E/S. Las curvas de E/S de PD se obtuvieron presentando dos senos simultáneos, con frecuencias f_2 y $f_1 = f_2/1.2$. La amplitud A_2 del seno f_2 fue igual a la condición con un seno. Para cada valor de A_2, la amplitud A_1 del seno f_1, variaba en pasos de 1 dB. Se aplicó un TFR a las señales de salida y la amplitud del componente PD de $2f_1–f_2$ fue utilizada para dibujar la curva de E/S en función de la amplitud de A_2.

3.3 Resultados

Los paneles superiores e inferiores de la Figura 3.1 ilustran los resultados para la no-linealidad de dos segmentos rectos y de Boltzmann, respectivamente. Los paneles izquierdo y derecho ilustran las curvas de E/S y las reglas de los primarios A_1-A_2 respectivamente. Las curvas de E/S de TP se muestran como líneas acentuadas continuas. Las curvas de E/S de PD se muestran para tres condiciones de A_1-A_2: amplitud igual de primarios (A_1=A_2) (cruces); A_1 óptima (círculos); y A_1 5 dB por debajo del valor óptimo (triángulos). Las amplitudes A_1 óptimas se definen como la amplitud A_1 que produce la máxima amplitud del PD para cada valor de A_2.

Una inspección visual de las curvas de E/S de la Figura 3.1 revela que las curvas de E/S de TP son más parecidas a las curvas de E/S de PD obtenidas con amplitudes A_1 óptimas. La Figura también revela que amplitudes sub-óptimas de A_1 pueden producir curvas de E/S de PD no-monótonas (por ejemplo, triángulos) (véase también Lukashkin y Russell, 2001). Además, la curva de E/S de PD en la condición de amplitudes A_1 sub-óptimas muestra muescas profundas. Estos resultados apoyan las afirmaciones realizadas al inicio de esta sección.

Varios aspectos de estos resultados son importantes. Primero, los presentes resultados no implican que la correspondencia máxima entre curvas de E/S de PD y TP ocurre para amplitudes óptimas de primarios; solo muestran que amplitudes

de primarios óptimas llevan a una correspondencia razonable entre curvas de E/S de PD y de TP. Este resultado es útil en la práctica, porque la combinación de A_1-A_2, que maximiza la correspondencia entre las dos curvas de E/S, normalmente se desconoce de entrada. Segundo, las conclusiones de este ejercicio son independientes de los valores de los parámetros de la no-linealidad. Tercero, el presente modelo es muy simplificado y no considera la filtración coclear. Por eso, la presente simulación no depende de la frecuencia actual de los senos o de los parámetros de la no-linealidad. Sin embargo, teniendo en cuenta la filtración, casi seguramente produciría distintas combinaciones de amplitudes óptimas a las ilustradas en la Figura 3.1(C)-(D).

3.4 Conclusiones

Aún siendo un modelo muy simplificado, este ejercicio muestra:

(1) Los niveles primarios influyen mucho en la forma de la curva de E/S de PD.

(2) Niveles sub-óptimos pueden producir curvas de E/S de PD no-monótonas con mesetas y muescas anchas y profundas.

(3) El uso de niveles óptimos podría mejorar el grado de semejanza entre las curvas de E/S de TP y de PD.

CAPÍTULO 4

LAS TEORÍAS DE GENERACIÓN DE OTOEMISIONES ACÚSTICAS Y LAS ESTIMACIONES CONDUCTUALES DEL MOVIMIENTO DE LA MEMBRANA BASILAR SON MUTUAMENTE CONSISTENTES[4]

4.1 Introducción

El nivel de las OEAPD es comúnmente usado como un indicador del estado fisiológico del mecanismo activo del oído interno. La sensibilidad de las OEAPD para este fin es mejor cuando los primarios tienen niveles que evocan máxima respuesta (Mills y Rubel 1994; Whitehead *et al*. 1995). Denominamos regla óptima de OEAPD a la combinación de primarios que evocan máxima respuesta. Hasta ahora, se han obtenido estas reglas óptimas de forma *empírica* (Kummer *et al*. 1998), pero la forma de la regla todavía es controvertida (Johnson *et al*. 2006; Kummer *et al*. 2000).

La controversia podría esclarecerse conociendo las condiciones cocleares que maximizan el nivel de OEAPD. La percepción común es que ocurre cuando los primarios excitan, al mismo grado, la región coclear sintonizada a f_2 (Kummer *et*

[4] Este capítulo se basa en el siguiente artículo: López-Poveda, E. A., and Johannesen, P. T. **2009**. "Otoacoustic emission theories and behavioral estimates of human basilar membrane motion are mutually consistent," publicado en J. Assoc. Res. Otolaryngol. **10**, 511–523. Se reproduce aquí parte del artículo con el permiso de "the Journal of the Association for Research in Otolaryngology".

al. 2000; Neely *et al*. 2005). La medición simultanea de OEAPD y de la MB ha revelado que esto es aproximadamente cierto en roedores (Rhode 2007), pero queda por confirmarse en humanos, dado que no es posible grabar el movimiento de la MB y, a la vez, medir OEAPD.

Por otro lado, se ha afirmado que es posible inferir los niveles de dos tonos puros iguales de eficiencia, en un sitio coclear determinado, a partir de curvas de enmascaramiento temporal (CET), y este último es el método más usado para inferir curvas de E/S cocleares en humanos (López-Poveda *et al*. 2003; Nelson *et al*. 2001). El método CET podría parecer una herramienta apropiada para verificar la conjetura sobre la generación de OEAPD de Kummer *et al*. (2000) en humanos. Desafortunadamente, sus supuestos (véase abajo) han sido validados solo indirectamente utilizando modelos computacionales, u otros métodos psicofísicos, pero sin evidencia fisiológica directa.

Una correlación alta entre reglas óptimas de OEAPD y sus correspondientes reglas conductuales, inferidas a partir del método CET, aportaría evidencia fuerte a favor de la conjetura de Kummer *et al*. (2000) sobre la máxima generación de OEAPD y también a los supuestos del método CET, utilizado para inferir respuestas de la MB humana. El objetivo del presente estudio es analizar dicha correlación. Será demostrado que hay una correspondencia alta para frecuencias de 1 y 4 kHz y para niveles inferiores a 65 dB SPL.

4.2 Métodos

4.2.1 Base lógica

Una gráfica CET muestra el nivel de un tono puro (máscara), que justo enmascara otro tono (sonda), en función del intervalo temporal que hay entre máscara y sonda. El nivel de la sonda es fijo y se sitúa justo encima del umbral de la sonda. El nivel de la máscara se incrementa al aumentar el intervalo temporal y, supuestamente, depende de dos variables (Nelson *et al*. 2001). Primero, depende del intervalo

temporal: el enmascaramiento disminuye al aumentar el intervalo temporal (Duifhuis 1973; Moore y Glasberg 1983; Nelson y Freyman 1987). Segundo, depende de la excitación relativa producida por la máscara y por la sonda en el sitio de la MB, sintonizado a la frecuencia de la sonda (Nelson *et al*. 2001; Oxenham *et al*. 1997; Oxenham y Moore 1995; Oxenham y Plack 1997). Dado que el nivel de la sonda es fijo, se supone que una CET representa el nivel de máscara necesario para generar un nivel fijo de excitación, después del decremento del efecto interno de la máscara, que se produce durante el intervalo entre máscara y sonda. Por eso, se denominan a las funciones resultantes como curvas de iso-respuesta de enmascaramiento temporal o CET (Nelson *et al*. 2001).

Hay fuerte evidencia que la tasa de recuperación del enmascaramiento temporal es, aproximadamente, la misma para distintas frecuencias de máscara y en un amplio rango de niveles (Wojtczak y Oxenham 2009). Por lo tanto, resulta razonable suponer que, para cualquier intervalo máscara-sonda, dos máscaras, de frecuencia ligeramente distinta (por ejemplo, f y f/1.2), con niveles al umbral de enmascaramiento, producen idénticos grados de excitación, en un sitio coclear sintonizado a la frecuencia de la sonda. Esta suposición es común para inferir curvas de E/S cocleares a partir de CET (López-Poveda *et al*. 2003, 2005; Nelson *et al*. 2001; Plack *et al*. 2004; Wojtczak y Oxenham 2009).

Basado en lo anterior, el enfoque consiste en medir dos CET, ambas con la frecuencia de sonda igual a la frecuencia de interés de las OEAPD (f_2), y con máscaras iguales a los tonos primarios de las OEAPD (f_1, f_2; con f_2/f_1=1.2). A continuación, se dibujan los niveles (L_1) de la máscara f_1 en función de los niveles (L_2) de la máscara f_2, pareados según el intervalo temporal entre máscara y sonda. Basado en la anterior interpretación de las CET, la gráfica resultante mostraría la combinación de niveles L_1-L_2 para que dos tonos puros, con frecuencias f_1 y f_2, excitaran al mismo grado el sitio coclear f_2. Si el vigente modelo de generación de OEAPD (como descrito por Kummer *et al*., 2000) y las suposiciones de las CET son correctos, la regla conductual debería coincidir con la regla óptima empírica de OEAPD.

4.2.2 Sujetos

Catorce sujetos normoyentes participaron en este estudio. Todos tenían umbrales auditivos inferiores a 20 dB HL en las frecuencias estudiadas (véase tabla 2.1 y 4.1).

4.2.3 Reglas conductuales

El procedimiento de medición de las CET fue prácticamente igual al usado en el Capítulo 2 (véase sec. 2.1.8). Se midieron CET para frecuencias (f_P) de sonda de 0.5, 1 y 4 kHz y para frecuencias de máscara igual a f_P y f_P/1.2. Estas frecuencias de máscara fueron iguales a las de los tonos primarios (f_1 y f_2, respectivamente) usados para medir OEAPD (véase abajo).

Las CET fueron ajustadas con la función (1) de López-Poveda *et al*. (2005). Se obtuvieron reglas de niveles individuales conductuales, dibujando el nivel ajustado de la máscara f_1 en función de los de la máscara f_2, pareados según el intervalo entre sonda y máscara.

4.2.4 Reglas óptimas de OEAPD

La magnitud (en dB SPL) del $2f_1-f_2$ del OEAPD fue medida para frecuencias f_2 de 1 y 4 kHz y para un ratio de frecuencia fijo de $f_2/f_1 = 1.2$. Se obtuvieron las reglas individuales óptimas de OEAPD variando sistemáticamente los niveles de los primarios (L_1 y L_2 para f_1 y f_2, respectivamente), para encontrar la combinación L_1-L_2 que produce la respuesta más alta de OEAPD. Se variaba L_2 en pasos de 5-dB en el rango de 35 a 75 dB SPL. Para cada L_2, se variaba L_1 en pasos de 3 dB, y el valor individual óptimo (el nivel de máxima respuesta) fue anotado.

La magnitud de las OEAPD puede variar rápidamente con un pequeño cambio de la frecuencia (Gaskill y Brown, 1990) lo que se denomina estructura fina de las OEAPD (véase sección 2.1 para más detalle). En un intento de evitar su influencia sobre las reglas óptimas de OEAPD, se midieron tres reglas óptimas para tres frecuencias ($0.99f$, f, y $1.01f$), alrededor de la frecuencia de interés, y el promedio se consideró la regla óptima resultante.

4.2.5 Curvas de E/S de OEAPD

Las curvas de E/S de OEAPD se midieron para frecuencias f_2 de 1 y 4 kHz, con regla óptima individual conductual, con regla óptima de OEAPD y también con la regla de Kummer *et al*. (1998) ($L_1 = 0.4L_2+39$). También se midieron curvas de E/S para una frecuencia f_2 de 500 Hz, pero solo con la regla individual conductual y la regla de Kummer *et al*.

Igual que en la investigación descrita en la sección 2.2.5, se midieron cinco curvas de E/S, para cinco frecuencias alrededor de la frecuencia de interés, y la curva de E/S final se calculó como el promedio de las cinco curvas de E/S. Se midieron y promediaron tres repeticiones de curvas de E/S para la regla de Kummer y para las reglas individuales conductuales, para cada frecuencia f_2 y sujeto, mientras que para la regla óptima individual solo se midió una única curva de E/S.

4.2.6 Procedimiento de medición de OEAPD

El procedimiento de medición de OEAPD fue prácticamente igual que en el Capítulo 2 (véase sec. 2.2.8).

4.3 Resultados

4.3.1 Curvas de enmascaramiento temporal

Las Figuras 4.1 a 4.3 ilustran las CET para frecuencias de sonda (f_P) de 0.5, 1 y 4 kHz, respectivamente. Cada panel ilustra las CET para un sujeto y para dos frecuencias de máscara $f_1 = f_P/1.2$ y $f_2 = f_P$.

Las características de las presentes CET son similares a otras de la literatura (López-Poveda *et al*. 2003; Nelson y Schroder 2004; Plack y Drga 2003). Las CET con frecuencia de sonda y máscara igual son generalmente más pendientes que las de la condición de frecuencia de sonda y máscara desigual, y esta diferencia de pendiente se interpreta como compresión de la respuesta de la MB.

4.3.2 La influencia de la estructura fina sobre las reglas de primarios óptimos de OEAPD

La Figura 4.4 aporta ejemplos de la influencia de la estructura fina de OEAPD, sobre las reglas óptimas individuales de OEAPD para las frecuencias 1 (paneles izquierdos) y 4 kHz (paneles derechos). La Figura muestra que la estructura fina solo podría haber afectado ligeramente las reglas óptimas de OEAPD, lo que justifica nuestro abordaje, que consiste en usar el promedio de tres curvas de tres frecuencias adyacentes como la regla óptima final.

4.3.3 Reglas conductuales de primarios óptimos *vs*. reglas de OEAPD

Las Figuras 4.5 y 4.6 ilustran el nivel de umbral de la máscara f_1 en función de los de la máscara f_2, pareados según intervalo máscara-sonda. Según la interpretación anterior (véase métodos), esta Figura ilustra combinaciones de niveles igual de eficientes en el sitio coclear sintonizado a la frecuencia de la sonda (f_2). Estas reglas conductuales se comparan con las reglas óptimas individuales de OEAPD, para frecuencias primarias iguales a las de las máscaras. La correspondencia entre las dos reglas para niveles inferiores a 65 dB SPL es casi perfecta a 1 kHz y ligeramente inferior a 4 kHz. La razón para la peor correspondencia por encima de 65 dB SPL se discutirá más adelante.

Las reglas conductuales individuales se basan en el promedio de tres repeticiones de L_1 y L_2, mientras las reglas óptimas de OEAPD se basan en un solo valor de L_1 para cada L_2 (este L_1 fue el promedio de tres frecuencias alrededor de la frecuencia de interés). Si esta única regla de OEAPD estuviera dentro del rango de las posibles combinaciones de L_1 y L_2, basadas en datos CET (puntos pequeños grises en Figs. 4.5 y 4.6), la diferencia entre las dos reglas individuales se considera estadísticamente no significativa.

La causa de la variabilidad de las combinaciones conductuales de L_1-L_2 es incierta. Refleja la variabilidad típica de las curvas CET y es normalmente mayor

en las partes más pendientes de las CET (por ejemplo, López-Poveda *et al*., 2003, 2005; Nelson *et al*., 2001; Plack *et al*., 2004).

La Figura 4.7 ilustra reglas de promedio de todos los sujetos. El promedio de la regla de OEAPD y de la regla conductual se solapa para 1 kHz en el rango 35-65 dB SPL. En 4 kHz no se solapan, pero están dentro de una desviación estándar y la diferencia no es estadísticamente significativa para $L_2 \leq 65$ dB SPL.

4.3.4 La dependencia de frecuencia de la regla conductual de primarios óptimos y de OEAPD

Las reglas conductuales de promedio (cuadrados grises en Fig. 4.7) alcanzaron niveles casi iguales ($L_1 \sim L_2$) a niveles altos de L_2 (~75 dB SPL). La diferencia entre L_1 y L_2 fue más grande a niveles bajos que a niveles altos de L_2 y se incrementa con la frecuencia. Fueron ajustadas a los datos líneas rectas para $L_2 \leq 65$ dB SPL. Las líneas tienen pendientes similar a 1 y a 4 kHz (tabla 4.2) siendo estas últimas más llanas que la pendiente de la regla conductual a 0.5 kHz. Las diferencias por frecuencia de las reglas conductuales no fueron estadísticamente significativas.

4.3.5 Curvas de E/S de OEAPD

Las curvas de E/S de OEAPD se midieron para cada sujeto utilizando su regla óptima de OEAPD y su regla conductual. La Figura 4.8 muestra el promedio de las curvas de E/S de todos los sujetos y, también, el promedio de curvas de E/S medido con la regla de Kummer *et al*. (1998). Las OEAPD se incrementaron considerablemente menos de 1 dB/dB. Los niveles de OEAPD medidos con regla óptima de OEAPD fueron, consistentemente, de 3 a 5 dB más altos que los medidos con reglas conductuales.

Los niveles de OEAPD evocados por la regla de Kummer *et al*. fueron idénticos o inferiores a aquellos de la regla conductual en 1 y 4 kHz excepto para $L_2 < 50$ dB SPL. La regla de Kummer *et al*. evocó niveles ligeramente más altos que la regla conductual a 0.5 kHz. En 1 y 4 kHz la diferencia entre los niveles de

OEAPD, medidos con la regla óptima y con la regla de Kummer, se incrementaron al incrementar L_2.

4.4 Discusión

El solapamiento general, entre reglas conductuales y reglas óptimas empíricas de OEAPD en 1 y 4 kHz (Figs. 4.5-4.7) para $L_2 \leq 65$ dB SPL, apoya la conjetura que los niveles de OEAPD son máximos cuando los primarios producen un grado similar de excitación en la región coclear de f_2 (Kummer *et al*., 2000).

Las presentes reglas conductuales se infirieron asumiendo que las CET de las dos máscaras (f_1 y f_2), reflejan solo la diferencia de excitación coclear, por los dos tonos, en el sitio sintonizado a la frecuencia de la sonda (f_2). Es decir, se asume que la interacción post-coclear entre máscara y sonda es lineal e idéntica, para las dos máscaras, para todos los intervalos máscara-sonda y todos los niveles de máscara. Esta suposición es comúnmente utilizada para inferir curvas de E/S cocleares a partir de las CET (López-Poveda *et al*. 2003; Nelson *et al*. 2001) y recibe apoyo de estudios de modelos experimentales (Oxenham y Moore 1994), por lo menos para niveles inferiores a 83 dB SPL (López-Poveda y Alves-Pinto 2008; Wojtczak y Oxenham 2009). La interacción post-coclear podría ser ligeramente distinta para máscaras a una frecuencia más lejana que un octavo y a altos niveles de máscara (López-Poveda y Alves-Pinto 2008; Wojtczak y Oxenham 2009), pero es poco probable que invalide el presente enfoque, dado que las máscaras eran cercanas en frecuencia y las conclusiones se afirman solo para $L_2 \leq$ ~65 dB SPL. Además, la correspondencia entre ambos métodos aporta apoyo adicional a las suposiciones del método CET, por lo menos por debajo de 65 dB SPL, dado que las reglas conductuales y las reglas óptimas de OEAPD se infirieron basándose en distintas suposiciones y métodos.

El nivel del $2f_1–f_2$ de la OEAPD es muy probablemente la suma de varias contribuciones. Entre ellas tenemos, la distorsión por la interacción no-lineal de los

primarios en la región coclear de f_2 (Martin *et al*. 1998), la reflexión de esta distorsión en el sitio coclear de $2f_1-f_2$ (Kalluri y Shera 2001), y la interacción no-lineal entre f_2 y la primera harmónica de f_1 ($2f_1$) en un sitio coclear más basal ($2f_1$) (Fahey *et al*. 2000). El peso de cada contribución al OEAPD medido es incierto. Las presentes reglas conductuales se obtuvieron de CET para frecuencias de sonda iguales a la frecuencia de la OEAPD (f_2). Basándose en las interpretaciones actuales de las CET (López-Poveda *et al*. 2003; López-Poveda y Alves-Pinto 2008; Nelson *et al*. 2001), las reglas conductuales reflejan la combinación de L_1-L_2, para la cual las dos máscaras (f_1, f_2) producen la misma excitación en el sitio coclear con FC ~f_2. Por lo tanto, la correspondencia entre reglas conductuales y óptimas de OEAPD, junto con la evidencia de que las OEAPD consisten en la suma de las contribuciones de varios sitios cocleares, sugiere que la contribución de la región f_2 es dominante y/o que las otras contribuciones son proporcionales a la generada en la región f_2.

Dado la razonablemente buena correspondencia entre reglas óptimas conductuales y reglas de OEAPD (Fig. 4.7B-C), no resulta claro porqué los niveles de las OEAPD, en promedio, fueron de 3 a 5 dB más altos para la regla óptima que para la regla conductual (Fig. 4.8B-C). Una posibilidad es que la supresión mutua entre los primarios haya afectado el resultado. En las OEAPD se presentan *simultáneamente* los dos tonos primarios y, por lo tanto, se tiene en cuenta implícitamente cualquier interacción no-lineal (por ejemplo supresión) entre los tonos primarios, lo cual es ignorado para las reglas conductuales ya que se infiere de respuestas de un tono solo. La grabación simultánea de OEAPD y respuestas de la MB en roedores sugieren que los niveles de OEAPD son sub-máximos para combinaciones de L_1-L_2 presentados simultáneamente y que producen la misma respuesta coclear y que esto posiblemente ocurre porque la respuesta del tono f_1 suprime la respuesta de la MB del tono f_2 [Fig. 1 de Rhode (2007)]. Dado que la regla conductual refleja un criterio de igual-respuesta para primarios no-simultáneos, esto podría explicar porque los niveles promedios de OEAPD de la regla conductual fueron consistentemente inferiores a los medidos con la regla

óptima de OEAPD. Por otro lado, el mismo mecanismo de supresión hubiera conducido a valores conductuales de L_1 inferiores a los valores de L_1 de OEAPD, lo cual no fue el caso (Figs. 4.5-4.7). Por lo tanto, resulta improbable que la supresión mutua pueda explicar la diferencia de niveles de OEAPD que hay entre las dos reglas. [véase también la discusión de (Kummer *et al*. 2000)].

Una sencilla explicación alternativa es que las reglas óptimas individuales de OEAPD fueron, por definición, las que evocaron los niveles máximos de OEAPD. Cualquier desviación de esta regla de la regla conductual (Figs. 4.5-4.6) siempre produciría niveles inferiores para cada sujeto y esto se reflejaría en el promedio de las curvas de E/S (Fig. 4.8).

La correspondencia entre las reglas óptimas conductuales y de OEAPD tendieron a ser peor en los datos individuales (Figs. 4.5-4.6) para L_2 por encima de 65 dB SPL. En varios casos (círculos abiertos en Figs. 4.5-4.6), el nivel óptimo de L_1 fue superior al límite de salida del sistema y la mayoría fueron superiores a los valores conductuales de L_1. La causa es incierta, pero podría reflejar un cambio de sitio de la generación de OEAPD, hacia la base de la cóclea con incremento en L_2. Es bien conocido, que el pico de la onda viajera cambia hacia la zona basal con incrementos en el nivel sonoro (Robles y Ruggero 2001). Por eso, es probable que la región de máxima interacción, entre las ondas viajeras de los dos tonos primarios, se desplaza desde la región con FC~f_2 hacia regiones más basales cuando aumenta L_2. La Figura 4.9 ilustra este cambio y la región de máxima interacción a niveles altos L_2 se denomina f'_2. La Figura 4.9 también ilustra como el incremento en L_1, para un incremento dado en L_2, sería mayor si los dos primarios tuvieran que evocar respuestas iguales en f'_2 que en la región coclear f_2.

Es improbable que las presentes reglas conductuales estuvieran afectadas por este cambio en la región de excitación, porque se basaron en CET para una sonda de nivel bajo y fijo y, por lo tanto, reflejan la respuesta coclear en un sitio *fijo* con FC~f_2 para todos L_2 (Nelson et al. 2001). Si el nivel máximo de OEAPD ocurriera para primarios igual-efectivos cerca del pico de la onda viajera de f_2, para cada nivel de L_2, entonces los valores L_1 siempre serían mayores que los valores conductuales

a altos niveles de cada L_2 (Fig. 4.9). Esto podría explicar porque los valores conductuales de L_1 a veces son inferiores a los L_1 óptimos de OEAPD a niveles altos de L_2. Además, si este cambio, que depende de nivel, fuera gradual, explicaría porque algunas de las reglas óptimas L_1-L_2 de OEAPD aparecieron generalmente más pendientes que sus contrapartidas conductuales en el rango donde se ha podido medir ambas (Figs. 4.5-4.6).

4.4.1 La controversia sobre la regla de primarios óptimos de OEAPD

Kummer *et al*. (2000) han argumentado que maximizando la respuesta de OEAPD en normoyentes también se maximiza la sensibilidad para discriminar entre una cóclea sana y una dañada y, por lo tanto, esto tiene mucho interés clínico. Hay consenso que el óptimo L_1 incrementa con L_2 siguiendo una relación lineal ($L_1 = aL_2 + b$), mientras unos estudios han concluido que a y b deberían ser constantes en el rango de frecuencias f_2 desde 1 a 8 kHz (Kummer *et al*. 2000), otros han concluido que deberían variar sistemáticamente con frecuencia f_2 (Johnson *et al*. 2006; Neely *et al*. 2005).

Los datos conductuales y de OEAPD producidos apoyan que la regla óptima es, aproximadamente, similar tanto en 1 como en 4 kHz (Fig. 4.7B-C y tabla 4.2). Los datos de 0.5 kHz sugieren que la regla conductual difiere significativamente de las otras reglas de frecuencia más alta (Fig. 4.7 y tabla 4.2). Desafortunadamente, esto no se ha podido apoyar con reglas empíricas de OEAPD en esta frecuencia. Dicho esto, las curvas de E/S de la Figura 4.8 sugieren que la regla de Kummer *et al*. (2000), que es constante con la frecuencia, evocó niveles de OEAPD en promedio indistinguibles de las evocadas por la presente regla conductual, aunque la regla es distinta en particular en 0.5 kHz.

4.5 Conclusiones

La coincidencia entre reglas óptimas conductuales y empíricas de OEAPD, apoya la conjetura que la máxima respuesta de OEAPD ocurre cuando los primarios excitan al mismo grado la región coclear de f_2. También apoya las suposiciones del método CET para inferir respuestas de la MB humana.

CAPÍTULO 5

CORRESPONDENCIA ENTRE ESTIMACIONES DE CURVAS COCLEARES HUMANAS CONDUCTUALES DE ENTRADA/SALIDA Y DE OTOEMISIONES ACÚSTICAS INDIVIDUALMENTE OPTIMIZADAS[5]

5.1 Introducción

En el Capítulo 2, se observó que, para frecuencias por encima de ~2 kHz, hay buena correspondencia a nivel individual entre curvas de E/S de la MB, inferidas conductualmente a partir de CET, y las curvas de E/S de OEAPD obtenidas con parámetros de promedio de grupo. La correspondencia fue peor para bajas frecuencias (0.5 y 1 kHz) porque las curvas de E/S de OEAPD tenían muescas y mesetas lo cual no ocurría en las correspondientes curvas conductuales. Se sugirieron dos explicaciones. Primero, podría ser debido a la estructura fina (Gaskill y Brown, 1990), a pesar del intento de reducir su influencia utilizando promediación espectral (Kalluri y Shera 2001). Segundo, podría ser debido al uso de parámetros sub-óptimos de estimulación de OEAPD.

El término estructura fina se refiere, como ya se señaló, a variaciones rápidas presentes en la representación del nivel del $2f_1–f_2$ del OEAPD en función

[5] Este capítulo se basa en el artículo: Johannesen, P. T., and López-Poveda, E. A. **2010**. "Correspondence between behavioural and individually "optimized" otoacoustic emission estimates of human cochlear input/output curvas," publicado en J. Acoust. Soc. Am. **127**, 3602-3613. Se reproduce aquí parte del artículo con permiso de "the Acoustical Society of America".

de la frecuencia (f_2) (el PD-grama). Estas variaciones tienen su origen en la suma vectorial de contribuciones de PD procedentes de varios sitios de la cóclea (véase sec. 2.1 y Shera y Guinan, 2008). El mismo mecanismo de interferencia influye en las curvas de E/S de OEAPD (He y Schmiedt, 1993, 1997; Mauermann y Kollmeier, 2004).

La estructura fina está igualmente presente en todas las frecuencias (Fig. 2 de Mauermann *et al.*, 1999; Fig. 3 de Dhar y Schaffer, 2004) al contrario de las muescas y mesetas de nuestras curvas de E/S que solo ocurrieron en baja frecuencia. Johnson *et al.* (2006b) han mostrado una incidencia más alta de muescas en 2 que en 4 kHz, lo cual sugiere mayor incidencia de mesetas y muescas en baja frecuencia, lo que a su vez es consistente con los resultados del Capítulo 2. Sin embargo, la evidencia de Johnson *et al*. (2006b), se basa en niveles bajos de primarios (L_2 = 30 dB SPL), donde la estructura fina es más pronunciada (He y Schmiedt, 1993, 1997; Mauermann y Kollmeier, 2004; Johnson *et al*., 2006b). Por lo tanto, se desconoce si los resultados de Johnson *et al.* (2006b) se pueden generalizar a 50-60 dB SPL, que es el rango donde se observaron las mesetas y muescas en los resultados del Capítulo 2. Además, también se desconoce si los resultados de Johnson *et al*. (2006b) se pueden generalizar a 0.5 y 1 kHz, donde las mesetas y muescas fueron más frecuentes en nuestros datos. Adicionalmente, las mesetas y la mayoría de las muescas, en nuestras curvas de E/S del Capítulo 2, se extendieron a un rango de niveles más amplio que el de las muescas causadas por interferencia de varias fuentes de PD [en torno a 10 dB, según estimación de la curva de E/S de Figuras 6-8 de He y Schmiedt (1993)]. Finalmente, las curvas de E/S de OEAPD del Capítulo 2, fueron un promedio de cinco (típicamente) frecuencias f_2 adyacentes en un intento de reducir la influencia de la estructura fina (Kalluri and Shera, 2001). Este número fue suficiente para controlar la influencia de la estructura fina en frecuencias por encima de 2 kHz y parecía razonable asumir, que también sería el caso para frecuencias bajas. En total, esto sugiere que, la explicación de que la falta de correspondencia entre curvas de E/S conductuales y de OEAPD en baja frecuencia se debe a la estructura fina, es posible pero podría ser insuficiente.

En efecto, una explicación complementaria podría ser que las curvas de E/S de OEAPD del Capítulo 2 se midieron con parámetros sub-óptimos. Las OEAPD tenían frecuencias primarias (f_1 y f_2) con un ratio de $f_2/f_1 = 1.2$ (Gaskill y Brown 1990) y sus niveles según la regla de Kummer *et al.* (1998): $L_1 = 39 + 0.4L_2$. Estos son parámetros óptimos de promedio de grupo y son constantes con la frecuencia; por lo tanto, es improbable que sean óptimos a nivel individual. Existe evidencia de que el ratio f_2/f_1 de los primarios tiene poco efecto sobre la forma y la pendiente de las curvas de E/S (Johnson *et al.*, 2006a). Por el contrario, hay evidencia significativa de que los niveles primarios tienen gran influencia sobre la forma y la pendiente de las curvas de E/S. Primero, las reglas de niveles individuales varían bastante entre sujetos y con la frecuencia (Neely *et al.*, 2005). Segundo, solo el 10% de los sujetos tienen curvas de E/S no-monótonas en el rango de frecuencias entre 1 a 8 kHz, utilizando niveles primarios de OEAPD optimizados individualmente (Kummer *et al.,* 2000). Tercero, las simulaciones numéricas del Capítulo 3 y de Lukashkin y Russell (2001) sugieren, en primer lugar, niveles primarios sub-óptimos que podrían conducir a curvas de E/S de PD no-monótonas, incluso en la ausencia de una fuente secundaria de PD; y, en segundo lugar, que el uso de niveles óptimos mejora la correspondencia entre curvas de E/S de PD y curvas de E/S de tono puro. Finalmente, hemos confirmado, en el Capítulo 4 y en López-Poveda y Johannesen (2010), que las reglas óptimas individuales difieren de la regla de Kummer *et al*. En total, esto sugiere que se podría mejorar la correspondencia entre curvas de E/S de la MB conductuales y curvas de OEAPD utilizando niveles individuales óptimos.

El uso de niveles óptimos individuales de OEAPD también podría reducir la estructura fina. La contribución relativa de varias fuentes cocleares de PD depende de los niveles primarios. (He y Schmiedt, 1993). Supuestamente se maximiza la contribución de la fuente "distorsión" (f_2) utilizando parámetros óptimos, y esto disminuye, comparativamente, la contribución de la fuente de "reflexión" ($2f_1$-f_2), reduciendo de esta forma la magnitud de la estructura fina. Si este fuera el caso, las curvas de E/S de OEAPD, medidas con parámetros óptimos

individuales, podrían reflejar mejor la curva de E/S de la MB y por lo tanto, mejorar la correspondencia con la curva de E/S inferida a partir de pruebas conductuales.

En resumen, no se puede descartar la estructura fina como explicación para la falta de correlación en baja frecuencia entre curvas de E/S de la MB, basadas en OEAPD, e inferidas de pruebas conductuales, pero el uso de niveles óptimos individuales probablemente podría mejorar la correspondencia entre los resultados de los dos métodos.

Este estudio es un re-análisis de datos anteriormente mostrados en los Capítulos 2 y 4 con el objetivo de comprobar esta posibilidad para un ratio fijo de frecuencias primarias de f_2/f_1 = 1.2. Se intenta minimizar la influencia de la estructura fina promediando curvas de E/S de OEAPD para un número de frecuencias alrededor de la frecuencia de prueba.

5.2 Métodos

5.2.1 Enfoque

El planteamiento consiste en comparar, a nivel individual, curvas de E/S de OEAPD para estímulos óptimos, con curvas de E/S cocleares, inferidas a partir de curvas de enmascaramiento temporal (CET). El Capítulo 2 describe, en detalle, cómo se infieren curvas de E/S cocleares y el Capítulo 4 contiene una descripción de la obtención de estímulos óptimos individuales de OEAPD.

El abordaje consiste en reanalizar los datos de los capítulos anteriores con el objetivo de comprobar si la correspondencia a nivel individual, entre estimaciones de curvas de E/S cocleares conductuales y las basadas en OEAPD, mejora cuando las OEAPD se miden utilizando niveles óptimos individuales. Con el fin de ser exhaustivo se incluyen curvas de E/S de OEAPD, medidas con reglas individualizadas y basadas en CET de Capítulo 4, y la regla de promedio de grupo de Kummer *et al.* (1998).

Los detalles metodológicos se han proporcionado en los Capítulos 2 y 4.

5.2.2 Sujetos

Quince sujetos normoyentes participaron en este estudio. Todos tenían umbrales auditivos inferiores a 20 dB HL en las frecuencias estudiadas.

5.2.3 Estimulación de CET

Los parámetros de la estimulación de CET, la referencia lineal usada y la forma de inferir curvas de E/S fueron idénticos a los del Capítulo 2 (véase 2.2.2).

5.2.4 Procedimiento CET

El procedimiento de la medición de CET fue idéntico al usado en el Capítulo 2 (véase sec. 2.2.3).

5.2.5 Inferencia de curvas de E/S de la MB a partir de CET

Las curvas de E/S cocleares fueron inferidas de la misma manera que en el Capítulo 2 (véase sec. 2.2.4).

5.2.6 Inferencia de reglas de primarios de OEAPD a partir de CET

Las reglas de primarios óptimos individuales conductuales (basadas en curvas CET) fueron las mismas obtenidas en el Capítulo 4.

5.2.7 Estimulación de OEAPD

Las curvas de E/S de OEAPD se obtuvieron dibujando la magnitud (en dB SPL) del $2f_1-f_2$ OEAPD en función del nivel (L_2) del tono primario f_2. Las curvas de E/S se obtuvieron para un ratio fijo entre los primarios de $f_2/f_1 = 1.2$ y para las siguientes tres reglas de niveles primarios.

- Se obtuvieron niveles óptimos individuales buscando el valor L_1 para cada valor de L_2 que maximiza la respuesta.

- Los niveles basados en CET fueron derivados de niveles de máscaras f_1 y f_2 como se explicó en la sección 4.2.1.
- Se utilizo la regla de Kummer *et al.* (1998): $L_1 = 0.4L_2+39$, con L_1 y L_2 en dB SPL.

Se midieron curvas de E/S de OEAPD para frecuencias de 0.5, 1 y 4 kHz para las tres reglas de primarios.

Se utilizó un enfoque similar al de la sección 2.2.5, de promediación de cinco frecuencias adyacentes (por ejemplo, Kalluri y Shera, 2001; Mauermann y Kollmeier, 2004), en un intento de reducir la influencia de la estructura fina sobre las curvas de E/S de OEAPD. Para la regla óptima se promediaron solo tres frecuencias por razones de tiempo.

5.2.8 Calibración de estimulación de OEAPD y artefactos de sistema

La calibración del instrumento de OEAPD y el control de respuestas falsas evocadas por el instrumento se hizo como en las secciones 2.2.6 y 2.2.7.

5.2.9 Procedimiento de OEAPD

El procedimiento de medición de OEAPD fue igual que en la sección 2.2.8.

5.2.10 Análisis de correspondencia entre curvas de E/S conductuales y de OEAPD

El grado de correspondencia a nivel individual entre curvas de E/S, inferidas de CET y de OEAPD, fue estimado a partir de polinomios de tercer orden ajustados a todas las curvas de E/S (por ejemplo, Capítulo 2). La primera derivada del polinomio fue calculada analíticamente y evaluada en el rango de niveles de entrada disponible. La semejanza entre curvas de E/S conductuales y de OEAPD fue evaluada calculando el RMS (Root Means Square, por sus siglas en ingles) de las

diferencias de las pendientes. El uso de la primera derivada tiene la ventaja de reflejar la forma de la curva de E/S, pero ignora cualquier desplazamiento vertical.

Dos métricas adicionales fueron empleados para evaluar la semejanza entre el grado de compresión de la MB, estimado a partir de la curva de E/S de la prueba conductual, y de OEAPD. Primero, se aplicó una regresión lineal a la pendiente mínima del polinomio de tercer orden, ajustado a las curvas de E/S. Segundo, también se aplicó una regresión lineal a la pendiente de líneas rectas ajustadas al segmento compresivo de las curvas de E/S.

5.3 Resultados

5.3.1 CET

Excepto las referencias lineales de los sujetos S11 a S15, todas las curvas de CET se han presentado en anteriores capítulos.

5.3.2 Reglas de primarios de OEAPD

Las reglas individuales basadas en CET fueron derivadas de las máscaras f_1 y f_2 como se describió en la sección 4.2.1. Los niveles óptimos individuales también fueron obtenidos como se describió en la sección 4.2.4.

5.3.3 Curvas de E/S de OEAPD

Las Figuras 5.1-5.3 ilustran curvas de E/S de OEAPD para las tres reglas de estimulación y para todas las frecuencias. Los niveles de OEAPD para la regla óptima casi siempre fueron superiores a los de la regla basada en CET y los de la regla de Kummer.

Las curvas de E/S de OEAPD típicamente se pueden describir con un segmento pendiente (casi linealidad) a niveles bajos de L_2 seguido de otro segmento con pendiente llana a niveles medios. De los segmentos llanos muchos tienen mesetas (pendiente de ~0 dB/dB) o muescas (regiones con pendiente negativo) en

baja frecuencia de 0.5 y 1 kHz (Figs. 5.1 y 5.2) pero casi ninguno en 4 kHz. La tabla 5.1 resume el número de curvas de E/S con mesetas o muescas por frecuencia y para cada una de las tres reglas de niveles. Se aprecia que las muescas y mesetas ocurrieron a 0.5 y 1 kHz no solo para la regla de Kummer *et al.* (1998), sino también para la regla de niveles óptimos individuales.

Una inspección visual (Figs. 5.2 y 5.4), revela que las curvas de E/S de las tres reglas tienen, a menudo, una forma similar. Cuando hay excepciones, las características de la regla óptima y la regla basada en CET son similares, pero difieren de las características de la regla de Kummer.

5.3.4 Correspondencia entre curvas de E/S conductuales y de OEAPD

Las curvas de E/S conductuales, en general, tienen las mismas tendencias que las de OEAPD, es decir, tienen un segmento pendiente a niveles bajos de entrada, seguido de otro segmento más compresivo a niveles medios, terminando, a veces, con un segmento más pendiente a niveles muy altos. Sin embargo, las curvas de E/S conductuales casi nunca muestran mesetas y muescas como es habitual para las de OEAPD.

Una inspección visual revela una correspondencia, razonablemente buena, entre curvas de E/S conductuales y de OEAPD de todas las tres reglas a 4 kHz (Fig. 5.3). Esto era esperable para la regla de Kummer, dado que ya se ha mostrado este resultado en el Capítulo 2. De las reglas óptimas individuales, la basada en CET no mejora la correspondencia, mientras que la regla óptima algunas veces mejora la correspondencia (por ejemplo, S6 y S7) y otras veces la empeora (por ejemplo, S8 y S9).

La correspondencia entre curvas de E/S conductuales y de OEAPD en 0.5 y 1 kHz (Figs. 5.1 y 5.2) fue escasa, incluso para ambas reglas individuales. Normalmente, las curvas de E/S conductuales fueron claramente distintas de las de OEAPD para todas las reglas. La diferencia se debe, en general, a las mesetas y muescas que solo ocurrieron en las curvas de E/S de OEAPD.

La Figura 5.4 ilustra las diferencias en RMS, de las pendientes de las curvas de E/S conductuales y de OEAPD, por frecuencia y por regla de primarios. Las diferencias medias fueron estadísticamente iguales por frecuencia y regla de primarios. La diferencia media tendía a ser menor en 4 kHz que en 0.5 kHz (prueba-*t* no pareado, $0.05 < p < 0.10$) para la regla de Kummer y la regla óptima. En efecto, la diferencia media para las tres reglas tendía a disminuir al incrementar la frecuencia.

5.3.5 Correspondencia de estimaciones de compresión

Algunos autores consideran la pendiente del segmento compresivo de la curva de E/S como un indicador del estado fisiológico de la cóclea (revisión por Robles y Ruggero 2001). Otros, sugieren que la pendiente no cambia al incrementarse la pérdida auditiva. (Plack *et al.,* 2004). A menudo se calcula la pendiente como el valor promedio de todo el segmento compresivo (por ejemplo, López-Poveda *et al.*, 2003; Plack *et al.*, 2004) y esta variable podría ser menos sensible frente a las muescas de las curvas de E/S de OEAPD y, por lo tanto, ser una variable valiosa para evaluar la correspondencia entre curvas de E/S conductuales y de OEAPD. Sin embargo, en otras ocasiones la pendiente puede cambiar gradualmente a lo largo de la región compresiva y, en este caso, la pendiente mínima seria la variable adecuada. Aquí se contemplan ambas variables, para evaluar la correspondencia entre curvas de E/S conductuales y de OEAPD, en el rango compresivo.

Las Figuras 5.5 y 5.6 muestran la pendiente mínima y media de las curvas de E/S de OEAPD en función de la pendiente de la curva conductual. Las líneas gruesas, en la figura, muestran una regresión lineal (RL). Las tablas 5.2 y 5.3 resumen los resultados de las regresiones lineales de estas estimaciones de la compresión. La correspondencia se considera buena cuando la pendiente de la RL (a) y el intercepto (b) son cercanos a uno y cero, respectivamente, y cuando la estadística F permite rechazar la hipótesis cero (que no hay relación estadística entre los dos variables).

La correspondencia entre OEAPD y estimaciones conductuales fue baja a 0.5 y 1 kHz, independientemente del descriptor de compresión (mínimo o promedio de la pendiente), y de la regla utilizada para medir OEAPD (tablas 5.2 y 5.3).

En 4 kHz, se observó una correspondencia alta para la estimación de la pendiente mínima (Fig. 5.5 y tabla 5.2) para OEAPD, medidas con la regla de Kummer. La correspondencia fue inferior y no significativa para reglas individuales (tabla 5.2). Para el promedio de la pendiente hubo una moderada correspondencia para OEAPD medidas con la regla de Kummer y la regla óptima, pero no alcanzaron el nivel significativo (ambos $p = 0.11$).

5.4 Discusión

En el Capítulo 2 se mostró una buena correspondencia entre estimaciones de curvas de E/S conductuales y de OEAPD medidas con la regla de Kummer, solo para frecuencias encima de ~2 kHz. La falta de correspondencia en frecuencias bajas parece deberse a la presencia de muescas y mesetas en las curvas de E/S de OEAPD, que no ocurrieron en las curvas conductuales. El presente estudio pretende investigar si la correspondencia mejora utilizando dos reglas individuales para medir las OEAPD. Una regla fue optimizado individualmente para maximizar la respuesta y la otra pretende evocar el mismo grado de excitación en el sitio coclear f_2. La correspondencia, en baja frecuencia, no mejoró con respecto a la regla de Kummer *et al.* (1998). En efecto, ha sido mostrado que mesetas y muescas son igualmente comunes en las curvas de E/S de OEAPD, para las reglas individuales y para la regla de grupo promedio, a 0.5 y 1 kHz (tabla 5.1).

La morfología similar de las curvas de E/S de OEAPD, para los distintos niveles primarios (Figs. 5.1-5.3), sugiere menor dependencia de niveles individuales de lo esperado (véase capítulo 3). Por ejemplo, todas las curvas de E/S de OEAPD sugieren umbrales similares de compresión y, también pendientes similares, tanto por debajo como por encima del umbral de compresión (Fig. 5.1-

5.3). La ligera dependencia de la pendiente, respecto de la regla de niveles, podría explicar el sorprendente resultado de que la correspondencia entre curvas de E/S conductuales y de OEAPD fue mejor para la regla de Kummer, aunque esta es una regla de promedio de grupo, que desestima la idiosincrasia individual.

5.4.1 Las causas de las muescas y mesetas en baja frecuencia

Se ha mostrado que muescas y mesetas son igualmente comunes en las curvas de E/S de OEAPD, para reglas de primarios individuales y de grupo promedio, en 0.5 y 1 kHz (tabla 5.1). Las mesetas y muescas podrían ser causadas por la estructura fina y también por estímulos sub-óptimos de OEAPD (véase la introducción). Las dos explicaciones no son necesariamente independientes, dado que la estructura fina depende de los niveles primarios (He and Schmiedt, 1993). La hipótesis de la introducción era que, maximizando individualmente la respuesta de OEAPD, podría aumentar la contribución de la fuente denominada "distorsión" (f_2), relativa a la de la fuente "reflexión" ($2f_1-f_2$) y, por lo tanto, contribuir a reducir la magnitud de la estructura fina. Los presentes resultados no apoyan dicha hipótesis, dado que la frecuencia y la magnitud de muescas y mesetas no se redujo utilizando reglas óptimas individuales. Si acaso, los resultados sugieren, que la magnitud de la contribución de la fuente "reflexión" ($2f_1-f_2$) es proporcional a la de la fuente "distorsión" (f_2).

5.5 Conclusión

(1) Para un ratio fijo entre las frecuencias primarias de $f_2/f_1 = 1.2$, la respuesta de OEAPD depende, moderadamente, de la regla de primarios, pero la morfología fundamental de las curvas de E/S apenas cambia.

(2) La frecuencia de mesetas y muescas, en curvas de E/S de OEAPD, fue comparable para las reglas de primarios individuales y para la regla de promedio de grupo de Kummer *et al*. Estas características fueron más frecuentes en

frecuencias bajas (0.5 y 1 kHz) que en altas (4 kHz), y sigue siendo la causa más probable de la baja correspondencia en frecuencias bajas, entre curvas de E/S conductuales y de OEAPD.

(3) La correspondencia fue razonablemente alta entre curvas de E/S conductuales y de OEAPD en 4 kHz. La regla promedio de grupo de Kummer *et al.* (1998) es suficiente para estimar las curvas cocleares de E/S individuales en frecuencias altas.

(4) Es discutible cual es el método más adecuado, el conductual o el de OEAPD, para inferir curvas de E/S cocleares individuales en 0.5 y 1 kHz, pero los métodos no son intercambiables, ni siguiera cuando se utilizan reglas de primarios óptimos individuales para medir las OEAPD.

CAPÍTULO 6

DISCUSIÓN GENERAL

La hipótesis principal que motivó esta tesis era que los métodos CET y OEAPD se pueden usar indistintamente para inferir curvas de E/S. El abordaje consistió en comparar curvas de E/S medidas con OEAPD y curvas inferidas a partir de CET en sujetos normoyentes. Los resultados sugieren que las curvas de E/S cocleares en normoyentes estimadas con ambos métodos son consistentes entre sí para frecuencias mayores que ~2 kHz pero no para bajas frecuencias. Los resultados también sugieren que la correspondencia entre curvas de E/S de los dos métodos apenas mejora en bajas frecuencias cuando se optimizan individualmente los niveles primarios de las OEAPD. Sigue siendo incierto cuál de los dos métodos refleja mejor la curva de E/S coclear en bajas frecuencias, si es que alguno lo hace. Varios factores pueden influir en las curvas de E/S estimadas de ambos métodos y en la baja correspondencia entre ellas en bajas frecuencias. Estos factores se discuten en las siguientes secciones.

6.1 Los supuestos del método CET

El método psicoacústico CET para inferir curvas de E/S de la MB se basa en varios supuestos (véase sec. 5.2 y 2.4). La mayoría de ellos tienen apoyo experimental para un amplio rango de condiciones (véase el Capítulo 4; López-Poveda y Alves-Pinto 2008; Wojtczak y Oxenham 2009). Aun así, sería erróneo considerar las curvas psicoacústicas como un “golden standard” (Stainsby y Moore, 2006; Wojtczak y Oxenham, 2009), en particular en baja frecuencia donde la

correspondencia entre las curvas de E/S conductuales y las de OEAPD es menor (Capítulo 2).

El método CET es un método psicofísico e indirecto y, por lo tanto, sus resultados podrían estar influidos por mecanismos retro-cocleares desconocidos. De hecho, hay diferencias intra-sujeto entre curvas de E/S inferidas con distintos métodos psicofísicos (Rosengard *et al.*, 2005). Uno de los supuestos más importantes es que la tasa de recuperación del efecto interno (post-coclear o post-compresión) de la máscara es constante e independiente de la frecuencia y del nivel sonoro (Nelson *et al.*, 2001; López-Poveda *et al.*, 2003). La validez de estos supuestos todavía es controvertida. Stainsby y Moore (2006) han argumentado que la tasa de recuperación es mayor para sondas de baja frecuencia, o lo que es igual para FC bajas, por lo menos para sujetos hipoacúsicos. Por el contrario, López-Poveda y Alves-Pinto (2008) han aportado evidencia indirecta a favor de que la tasa de recuperación es independiente de frecuencia, por lo menos para normoyentes. Además, hay evidencia de que para cualquier frecuencia, la tasa de recuperación es menor a niveles altos (López-Poveda y Alves-Pinto, 2008; Wojtczak y Oxenham, 2007). Estas cuestiones complican la elección de la CET de referencia lineal, y por lo tanto, ponen en duda las correspondientes curvas de E/S, en particular en baja frecuencia (Stainsby y Moore, 2006; López-Poveda y Alves-Pinto, 2008). Si ambos supuestos fueran ciertos, el exponente de compresión sería mayor que el sugerido por los presentes datos (Fig. 2.8), en particular para baja frecuencia, lo cual incrementaría la diferencia entre los exponentes de compresión basados en CET y OEAPD (Fig. 2.8, 5.5 y 5.6).

6.2 Los efectos del sistema eferente

La activación del reflejo eferente ipsilateral podría influir tanto en las curvas de E/S inferidas de las CET como en las curvas de E/S de OEAPD. El sistema eferente se puede activar por sonidos ipsilaterales y/o contralaterales con magnitudes superiores a 40 dB SL (Hood *et al.* 1996) y tiene una latencia de

activación en el rango de 40 a 80 ms (Backus and Guinan, 2006; Roverud and Strickland, 2010). La magnitud de los efectos eferentes es mayor en bajas frecuencias (Lilaonitkul and Guinan, 2009; Lopez-Poveda *et al.*, 2013; Aguilar *et al.* 2013), que es a su vez el rango donde hay mayor discrepancia entre las curvas de E/S de CET y OEAPD. Los estímulos empleados por ambos métodos son suficientemente altos para evocar los reflejos eferentes. La activación refleja del sistema eferente reduce la ganancia coclear y linealiza las curvas de E/S (Cooper and Guinan, 2006; Lopez-Poveda *et al.*, 2013). Si el reflejo eferente estuviera activado durante las mediciones CET, implicaría que la curva "verdadera" de E/S inferida de CET sería mas compresiva que lo sugerido por los presentes datos (Fig. 2.8, 5.5 y 5.6) y la diferencia entre la compresión estimada a partir de CET y OEAPD sería menor.

Si el reflejo eferente estuviera activado durante las OEAPD de niveles altos de estimulación, su efecto sería atenuar sólo las OEAPD de niveles altos. Por lo tanto, la curva real de E/S de OEAPD sería menos compresiva que lo sugerido por los presentes datos de OEAPD y la diferencia entre el exponente de compresión de CET y OEAPD sería menor (Fig. 2.8, 5.5 y 5.6). Dado que el efecto sobre OEAPD es que la respuesta disminuye ~2 dB, parece que el efecto sobre OEAPD es insuficiente para explicar la discrepancia entre los resultados de CET y OEAPD.

6.3 Mejoras en la ratio de frecuencias de los primarios de OEAPD

En el Capítulo 2, mostramos una baja correspondencia, en bajas frecuencias, entre curvas de E/S inferidas de CET y de OEAPD y sugerimos que podría deberse al uso de primarios con niveles sonoros mejorables (Sec. 2.4.2 y Sec. 5.1), lo cual también apoyaban las simulaciones del Capítulo 3, aunque esta explicación fue rechazada en el Capítulo 5. Otro aspecto de esta explicación de parámetros mejorables es que aquí se han usado primarios con una ratio de frecuencias fija $f_2/f_1 = 1.2$ sobre la base

de que esta ratio maximiza los niveles de OEAPD (Gaskill y Brown 1990). Sin embargo, según un estudio más reciente, la ratio óptima de f_2/f_1 se incrementa ligeramente al disminuir la frecuencia f_2 y al incrementar el nivel L_2 (Johnson *et al.*, 2006a), en particular para frecuencias bajas. La ratio óptima para f_2 = 0.5 kHz está todavía por determinar. Sin embargo, hay evidencia de que el procesamiento coclear en regiones apicales difiere significativamente del de la zona basal (por ejemplo, López-Poveda *et al.*, 2003; Plack *et al.*, 2004), lo cual sugiere que es probable que la ratio óptima a 0.5 kHz difiera de 1.2. Por lo tanto, la correspondencia entre curvas de E/S conductuales y de OEAPD quizás hubiera mejorado si se hubieran considerado no sólo primarios con niveles óptimos, sino también con frecuencias óptimas. Dicho esto, sin embargo, la ratio de f_2/f_1 parece tener poco efecto sobre la pendiente del promedio de las curvas de E/S, por debajo de 65 dB SPL (por ejemplo, Fig. 3 de Johnson *et al.*, 2006a), por lo menos en el rango de frecuencias entre 1 y 8 kHz. Esto pone en duda si la correspondencia entre curvas de E/S conductuales y de OEAPD mejoraría al optimizar la ratio de frecuencias, aunque el beneficio a nivel individual todavía está por investigar.

6.4 Supresión mutua de los estímulos de OEAPD

Hay una diferencia fundamental entre los paradigmas de estimulación conductual y de OEAPD: los estímulos del método conductual de CET no son simultáneos, mientras que medir OEAPD implica presentar *simultáneamente* los dos primarios f_1 y f_2. Una posible causa de las diferencias entre estimaciones de curvas de E/S de CET y de OEAPD, sería quizás que el primario f_1 podría suprimir la respuesta de la MB evocada por el primario f_2. Rhode (2007) midió la excitación de la MB y las OEAPD simultáneamente en roedores y demostró que, cuando la amplitud L_1 del primario f_1 excede 60 dB SPL, éste suprimió la respuesta de la MB del primario f_2, que tenía una amplitud L_2 fija de 60 dB SPL (véase su Fig. 1). Como resultado, las curvas de E/S de OEAPD podrían no corresponder a la curva de E/S de la MB medida con tonos puros, como se supone a menudo.

Al contrario de las curvas de E/S de OEAPD, las curvas de E/S inferidas a partir de las CET no estarían afectadas por supresión, porque la máscara y la sonda no se presentan simultáneamente. De hecho, esta es una de las razones por las que el método CET se usa comúnmente para inferir curvas de E/S de la MB. Por eso, se podría pensar que la supresión mutua afecta a las OEAPD, pero no a las curvas de E/S basadas en CET, y que el efecto es mayor en bajas frecuencias porque en el ápex coclear los efectos no lineales se extienden a un rango de frecuencias más amplio que en la base (Rhode y Cooper, 1996; López-Poveda *et al.*, 2003). Es improbable, sin embargo, que esta sea la explicación para la baja correlación entre estimaciones de compresión de los dos métodos en baja frecuencia. Hay evidencia fisiológica y psicofísica de que la supresión produce curvas de E/S más pendientes que las curvas de E/S de tonos puros (Nuttall y Dolan, 1993; Rhode, 2007; Yasin y Plack, 2007). Esto sucede, en particular, con combinaciones de supresor/suprimido similares a las usadas aquí. Si la supresión hubiera afectado a las curvas de E/S de OEAPD, estas indicarían una compresión inferior que la de las curvas de E/S basadas en CET, lo cual *no* fue el caso (Fig. 2.8). Por lo tanto, una explicación más probable para la baja correlación entre estimaciones de compresión conductuales y de OEAPD sigue siendo la presencia de muescas y mesetas en las curvas de E/S de OEAPD, que ocurren más frecuentemente en bajas frecuencias (Figs. 2.4–2.7, Tabla 2.2).

6.5 Fuentes secundarias de OEAPD a niveles de estimulación altos

Una primera posible explicación para las muescas y mesetas podría ser que hay otro mecanismo de generación de distorsión que contribuye en niveles de estimulación altos y que las muescas reflejan la interferencia destructiva entre esta nueva fuente de alto nivel y la fuente habitual (véase Mills, 1997). Liberman *et al.* (2004) mostraron que ratones genéticamente modificados, sin la proteína prestina que es necesaria para alimentar la electro-motilidad de las CCE, todavía generaron una

respuesta atenuada de PD, lo que apoyaría la existencia de un mecanismo secundario de generación de PD en niveles altos. Algunas de sus curvas de E/S de OEAPD mostraron muescas a niveles similares a los del presente estudio. Por otro lado, Avan *et al.* (2003) atribuyeron OEAPD de nivel alto y bajo al mismo mecanismo de generación. También, Lukashkin *et al.* (2002) y Lukashkin y Russell (2002) han mostrado que una función no linealidad con saturación es suficiente para explicar una muesca en una función de E/S de PD (véase también la Fig. 3.1).

6.6 La estructura fina de OEAPD

Otra posible explicación, para las mesetas y muescas en baja frecuencia, es que algunas de las curvas de E/S de OEAPD todavía están afectadas por la estructura fina en baja frecuencia, a pesar de las medidas adoptadas para minimizar su influencia. Las muescas ocurrieron en niveles en torno a 45–50 dB, donde la estructura fina seguramente puede tener influencia. Pero es discutible si esta explicación también vale para las mesetas, porque éstas siempre ocurrieron a niveles moderados-altos (60–70 dB SPL) y la estructura fina tiene mayor influencia en niveles bajos (Mauermann y Kollmeier, 2004).

En la introducción al Capítulo 5, se ha argumentado que la explicación basada en la estructura fina es posible pero podría ser insuficiente. Los argumentos propuestos se basaron en evidencia recogida para frecuencias iguales o mayores a 2 kHz. La importancia de la estructura fina por debajo de 1 kHz es todavía incierta, pero los presentes resultados sugieren que su importancia es mayor de lo inicialmente pensado. Si este fuera el caso, podría explicar porqué el método de promediación espectral, aquí utilizado para minimizar la estructura fina, parecía más eficiente para frecuencia alta (4 kHz) que para baja (0.5 y 1 kHz). Si la interferencia entre contribuciones de la distorsión generada en diferentes regiones cocleares (Shera y Guinan, 2008) fuera más pronunciada y menos sensible a los niveles y las frecuencias de los primarios para frecuencias bajas que altas, esto podría explicar la mayor incidencia relativa de mesetas y muescas en frecuencia

baja que en alta y, para todos los primarios contemplados. Por extensión, esto podría también explicar las discrepancias entre las curvas de E/S conductuales y de OEAPD. Se desconoce porqué las interferencias destructivas habrían de ser más pronunciadas, o menos sensibles, a los niveles primarios en frecuencias bajas que en altas. Sin embargo, la desorganización de las CCE (por ejemplo, Lonsbury-Martin *et al.,* 1988) y el rango de frecuencias con compresión, parece comparativamente mayor para las regiones cocleares apicales que para las basales (Rhode y Cooper 1996; López-Poveda *et al.,* 2003; Plack y Drga 2003). Como resultado, la región de generación de OEAPD por el mecanismo de "reflexión" y también por el de "distorsión" es probablemente más amplia en el ápex que en la base, y las interacciones potenciales entre contribuciones desde regiones adyacentes podrían ser más significativas para frecuencia baja. Sin embargo, esto es solo una conjetura que habría que demostrar.

6.7 Ideas para futuros estudios

Una línea importante de investigación futura debería ser comprobar si la curva de E/S de OEAPD realmente refleja (con precisión) la curva de E/S de la MB. La evidencia proporcionada en el Capítulo 3 es sólo circunstancia. Procedería realizar un estudio similar pero con un modelo más realista de la MB o incluso estudios fisiológicos similares a los de Rhode (2007) para medir simultáneamente curvas de E/S de la MB y de OEAPD.

La estructura fina de OEAPD resulta ser la causa más probable para la falta de correlación entre las curvas de E/S conductuales y las de OEAPD en baja frecuencia. En los estudios descritos en los Capítulos 2 a 5, se aplicó un método sencillo para reducir la influencia de la estructura fina, causada por la interferencia entre la fuente "distorsión" de OEAPD y la de "reflexión". Sin embargo, existen mejores métodos para evitar la contribución de la fuente de "reflexión" que no se emplearon aquí, en parte por motivos prácticos, pero también por otras razones que se explican más abajo. Un buen método para evitar la contribución de la fuente de

"reflexión" es probablemente el paradigma de la transformación rápida de Fourier (TRF) inversa (Stover *et al*., 1996). Su mayor desventaja es que necesita un tiempo de medición extremadamente largo, porque requiere datos con muchísima resolución de frecuencia (~20 Hz) y habría que repetirlo para cada nivel L_2. Por ello, su uso pareció inviable para el presente objetivo.

Otro método consiste en la aplicación de un tercer tono cerca de la frecuencia de la fuente de "reflexión" ($2f_1$-f_2) que suprima la contribución de esta fuente. Es igual de rápido que el procedimiento normal de OEAPD. Su mayor problema es que es difícil elegir el nivel correcto del supresor porque varía mucho de un individuo a otro (Johnson *et al*., 2006). Si se elige un supresor demasiado bajo, la fuente de "reflexión" no es suprimida y si es demasiado alto, se suprime también la fuente "distorsión" (f_2), lo cual afectaría la curva de E/S. A pesar de este problema, podría ser una buena alternativa porque se podría emplear un supresor con un nivel sonoro fijo y relativamente bajo (50 dB SPL), que todavía podría suprimir algo la fuente de "reflexión" sin apenas afectar la fuente de "distorsión".

Finalmente, Long *et al*. (2008) han propuesto un método que consiste en un barrido en frecuencia de los primarios, lo cual separa las contribuciones de la fuente "distorsión" y "reflexión" por un procedimiento de análisis temporal que, supuestamente, no requiere mucho tiempo de medición. Sería sumamente interesante investigar si la correspondencia entre las curvas de E/S basadas en CET y OEAPD mejoraría en baja frecuencia, utilizando el método de supresión o el de barrido de frecuencia de primarios.

También valdría la pena investigar la influencia de parámetros óptimos en baja frecuencia (f_2 < 1 kHz). De hecho, la ratio de frecuencias, f_2/f_1, y la regla de niveles sonoros, L_1-L_2, óptimos están todavía por determinar en baja frecuencia y definirlas podría tener interés más allá del objetivo de este estudio, por ejemplo, en el ámbito clínico.

Existen otros métodos psicofísicos para inferior curvas de E/S cocleares como el método de crecimiento de enmascaramiento (en inglés: '*growth of masking*') (Oxenham and Plack 1997; Rosengard *et al*. 2005) así como una variante del

método CET (Lopez-Poveda and Alves-Pinto, 2008). Las curvas de E/S inferidas con estos métodos podrían corresponderse mejor con las curvas de E/S de OEAPD cuando la excitación de la MB cambia hacia la zona basal al aumentar el nivel de estimulación, como se ha discutido en Sección 4.4. Además, la variante del método CET no se basa en los supuestos habituales del método CET (discutidos en Sec. 6.1) sino en otros supuestos. Por tanto, sería interesante comparar las curvas de E/S de OEAPD con las inferidas usando estos otros métodos psicoacústicos.

También el presente método CET podría mejorarse en dos aspectos: 1) minimizando el efecto del reflejo eferente sobre la curva de E/S inferida (Yasin *et al.*, 2013); 2) simplificando a la vez el método para que sea más viable en un contexto clínico.

Una limitación del presente trabajo es que sólo se ha centrado en personas normoyentes. Evidentemente, se debería extender el estudio a sujetos hipoacúsicos y, de hecho, ya estamos trabajando en esta línea. Es incierto si las curvas de E/S estimadas con CET y OEAPD serían más o menos consistentes entre sí en hipoacúsicos. Dorn *et al.* (2001) mostraron curvas de E/S de OEAPD promedio para distintos grados de pérdida auditiva (supuestamente por disfunción de las CCE) y la forma de estas curvas es incompatible con el modelo de curvas de E/S de la MB para hipoacúsicos de Plack *et al.* (2004) que sugiere que el umbral de compresión aumenta con la pérdida auditiva, pero que el exponente de compresión no cambia con el grado de pérdida auditiva. La similar morfología de las curvas de E/S de CET y OEAPD para frecuencias mayores que ~2 kHz observado en normoyentes quizás no es esperable en hipoacúsicos. Por otro lado, esto no excluye que pueda haber buena correlación entre, por ejemplo, la pérdida de ganancia coclear y alguna métrica basada en curvas de E/S de OEAPD. Probablemente el problema de la estructura fina sigue siendo relevante (He y Schmiedt, 1996). Cuando la fuente de "reflexión" ($2f_1-f_2$) coincide con una frecuencia de una región de la MB dañada, la estructura fina es reducida (Mauermann *et al.*, 1999), mientras una estructura fina aumentada puede darse cuando la fuente de "distorsión" (f_2)

coincide con una región dañada, por lo tanto, la influencia de la estructura fina dependerá de la configuración de la pérdida auditiva (Konrad-Martin *et al*., 2002).

El objetivo a largo plazo de este trabajo es proporcionar un método clínicamente útil para estimar curvas de E/S cocleares. La idea es que, usando información acerca de la curva de E/S de la MB, la programación de un audífono podría ser más individualizado y así reducir el número de sesiones necesarias para ajustarlo, lo cual mejoraría el rendimiento del audífono.

CONCLUSIONES

El objetivo general del presente estudio era adaptar los procedimientos de medida de OEAPD con miras a usar las curvas de E/S de OEAPD como alternativa rápida y fiable al método psicoacústico CET para inferir curvas de E/S cocleares. Las principales conclusiones son:

1. Las curvas cocleares de E/S basadas en el CET y las basadas en OEAPD son consistentes mutuamente para frecuencias iguales o mayores que 2 kHz. Dado que los dos métodos están basados en supuestos diferentes, esto sugiere que ambos conjuntos de curvas de E/S dependen de los mismos mecanismos subyacentes y que probablemente reflejan las “verdaderas” curvas de E/S de la MB.
2. La peor correspondencia entre curvas de E/S basadas en CET y OEAPD a frecuencias más bajas puede deberse a defectos en los supuestos de ambos métodos así como al uso de estímulos no óptimos.
3. La peor correspondencia a frecuencias más bajas está, sin embargo, asociada a la presencia de muescas y mesetas en las curvas de E/S de OEAPD. Aunque las simulaciones sugieren lo contrario, es improbable que esos rasgos se deban al uso de primarios con niveles sonoros promedio en lugar de optimizados individualmente. La peor correspondencia podría deberse, por tanto, a la estructura fina de OEAPD.
4. Los niveles máximos de OEAPD ocurren cuando los dos tonos primarios producen idéntica excitación de la MB en el sitio coclear sintonizado a la frecuencia del primario f_2.