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*

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.

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

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

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 common-place 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 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 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
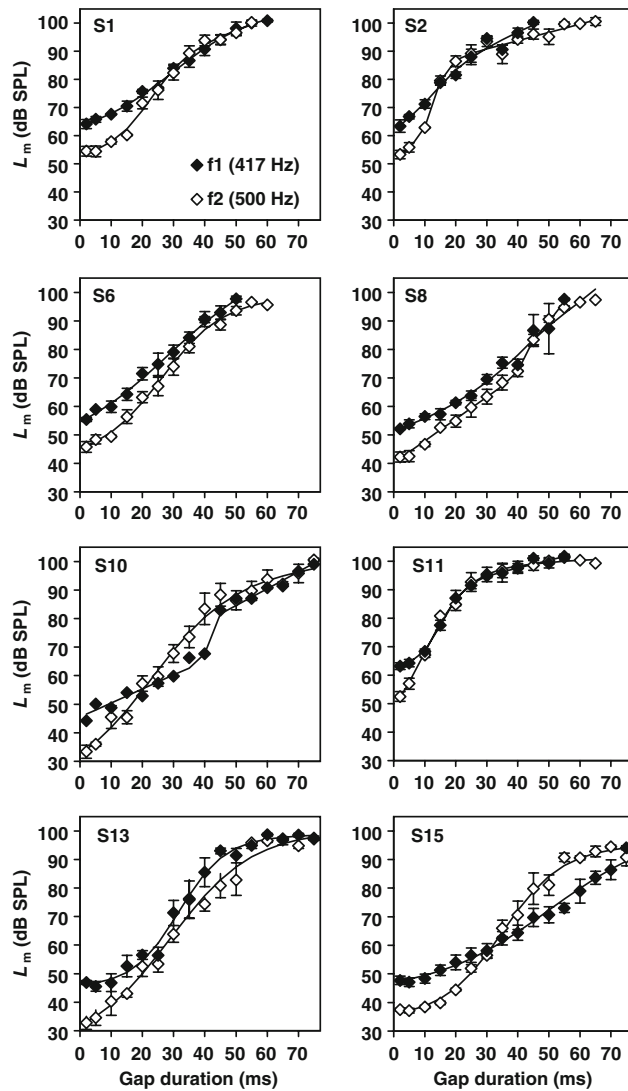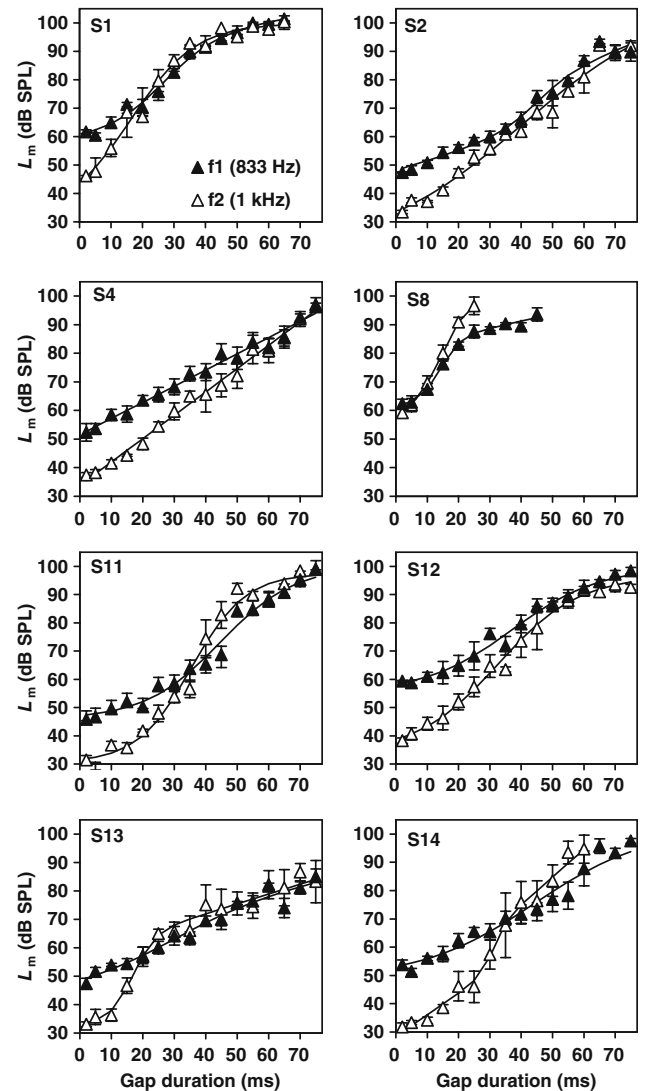
FIG. 1. TMCs for all listeners for probe frequencies (f_p) of 500 Hz. *Each panel* shows data for one subject. *Open symbols* illustrate TMCs for a masker frequency equal to the probe frequency ($f_2 = f_P$); *filled symbols* illustrate TMCs for a masker frequency below the probe frequency ($f_1 = f_P/1.2$). Each data point is the mean of at least three measurements. *Error bars* illustrate 1 SE of the mean ($n \geq 3$). *Lines* illustrate function fits to the data points. L_m masker level.

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

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

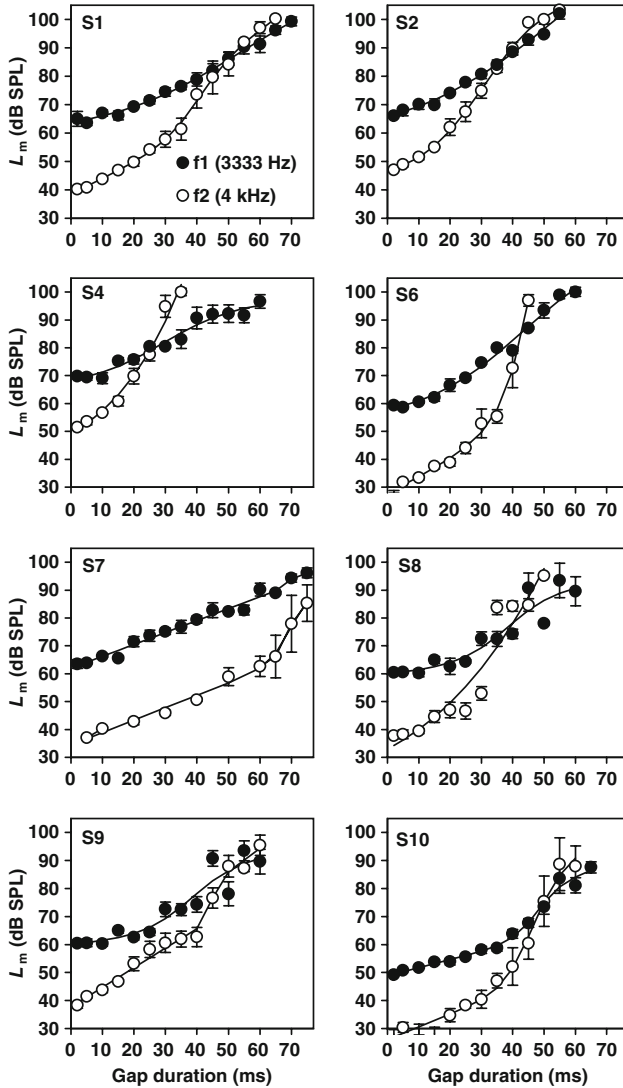**FIG. 2.** As Figure 1 but for a probe frequency of 1 kHz.

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
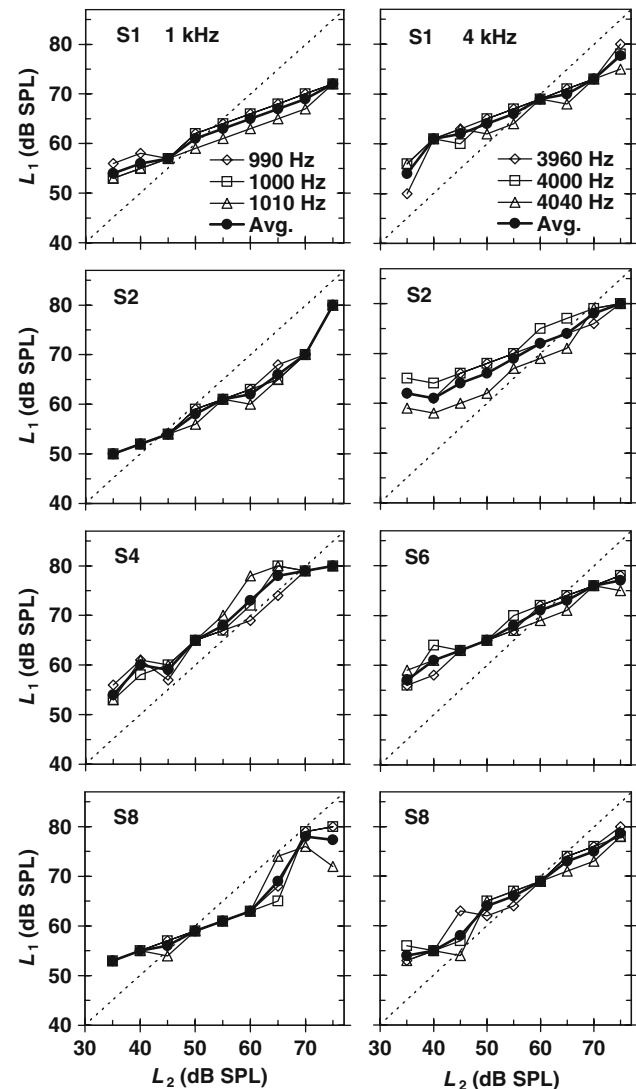

FIG. 4. Examples of the influence of the fine structure on individual DPOAE optimal rules at 1 and 4 kHz (*left and right panels*, respectively). The listener identifier is shown on the *top-left corner* of each panel. *Open symbols* illustrate DPOAE optimal rules for a different test frequency at or 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-
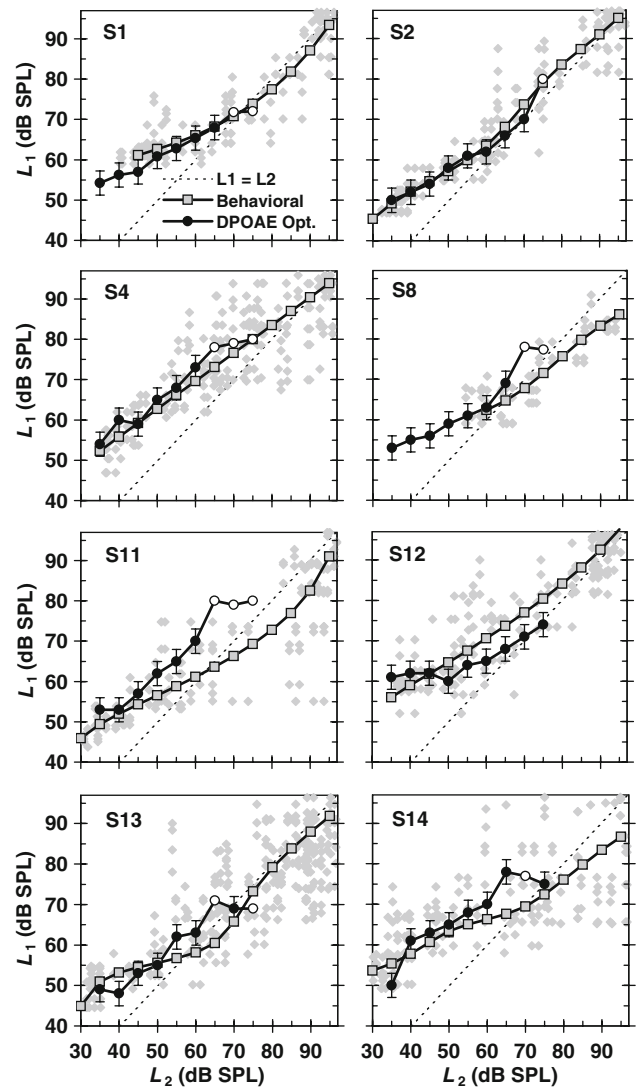

FIG. 5. Comparison of behavioral and DPOAE optimal level rules for a test frequency of 1 kHz. *Each panel* illustrates results for a single subject. *Gray squares* illustrate behavioral rules based on *mean* TMCs. *Small gray dots* illustrate *all* possible $L_1–L_2$ combinations based on the available 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 ($L_1=L_2$).

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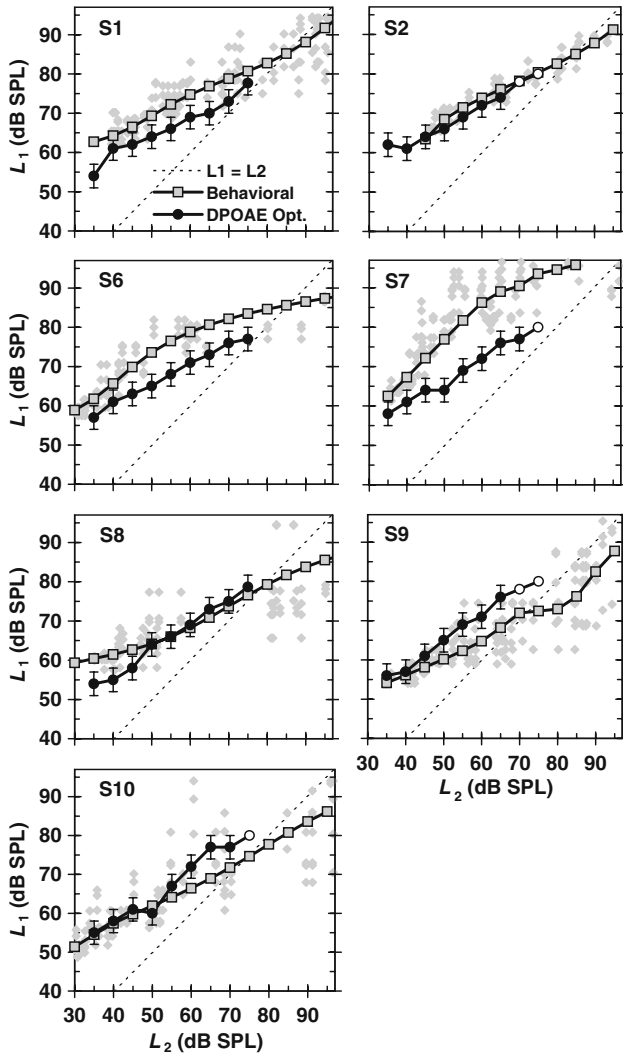
FIG. 6. As Figure 5 but for a test frequency of 4 kHz.

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
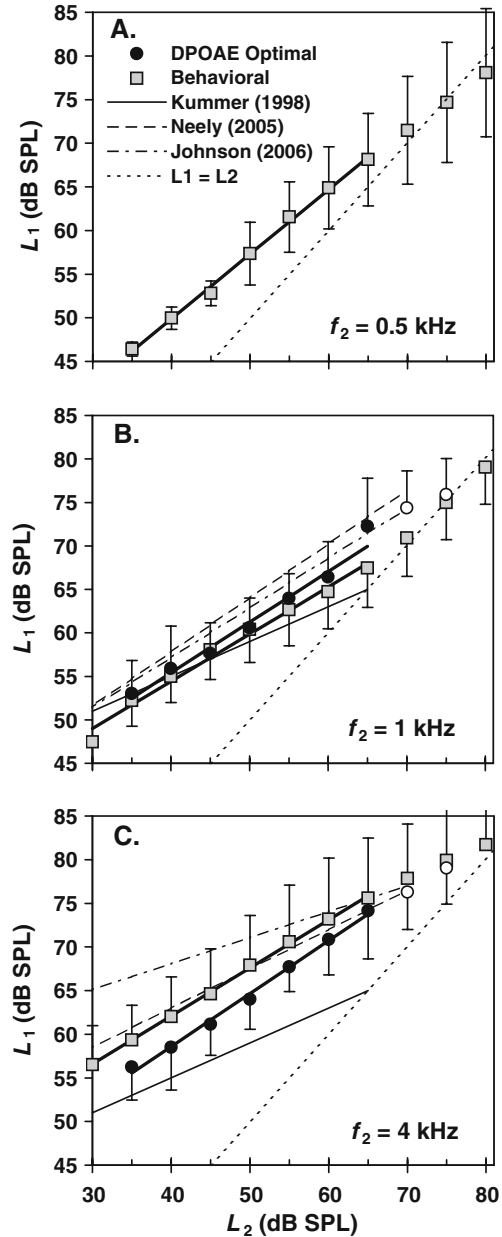

FIG. 7. Average L_1–L_2 rules at different frequencies and as proposed by the present and earlier studies. **A** For $f_2 = 0.5$ kHz. **B** For $f_2 = 1$ kHz. **C** For $f_2 = 4$ kHz. *Circles and squares* illustrate average DPOAE optimal and behavioral rules obtained in the present study, respectively. 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
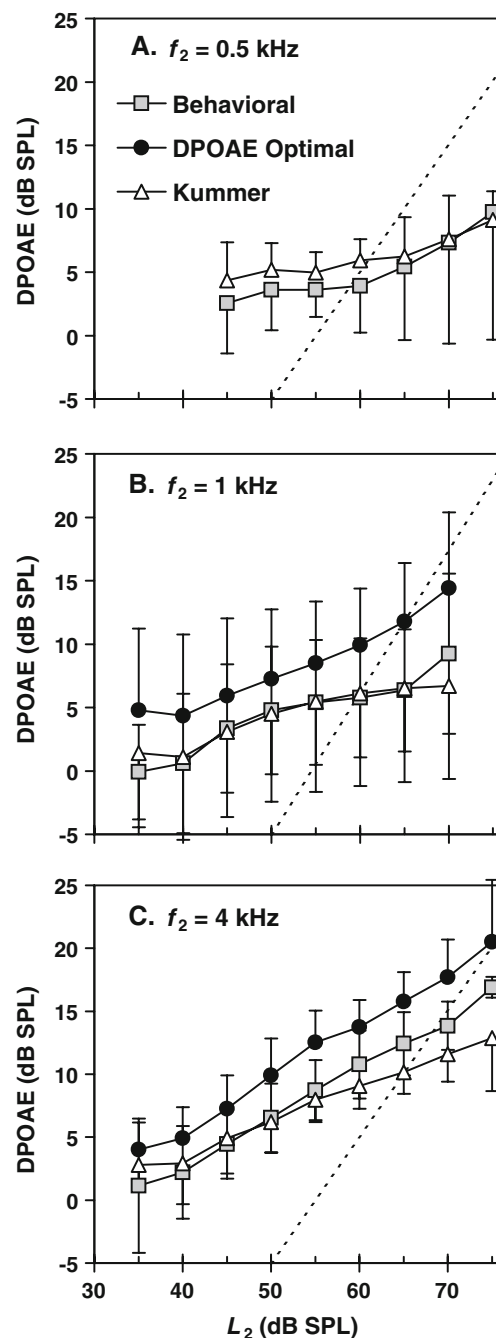

FIG. 8. Average I/O curves for the $2f_1 - f_2$ DPOAE for different primary level rules (as indicated in the *inset* of **A**). *Different panels* illustrate results for a different f_2 frequency. **A** $f_2 = 0.5$ kHz. **B** $f_2 = 1$ kHz. **C** $f_2 = 4$ kHz. Every point is the mean of data for between 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 $\sim f_2$ 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 $\sim f_2$ 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).
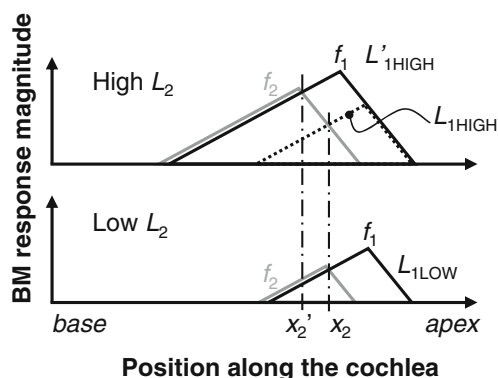
FIG. 9. Schematic cochlear excitation patterns of the two DPOAE primary tones, f_1 and f_2, for low and high L_2 levels (*lower and higher figures*, respectively). For low L_2 levels, the maximum interaction between the two excitation patterns occurs at the x_2 cochlear region whose CF is approximately 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

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.