Distribution of standing-wave errors in real-ear sound-level measurements

Susan A. Richmond
Department of Speech, Language, and Hearing Sciences, The University of Arizona, Tucson, Arizona 85721

Judy G. Kopun, Stephen T. Neely, Hongyang Tan, and Michael P. Gorga[a]
Boys Town National Research Hospital, Omaha, Nebraska 68131

(Received 17 September 2010; revised 22 February 2011; accepted 26 February 2011)

Standing waves can cause measurement errors when sound-pressure level (SPL) measurements are performed in a closed ear canal, e.g., during probe-microphone system calibration for distortion-product otoacoustic emission (DPOAE) testing. Alternative calibration methods, such as forward-pressure level (FPL), minimize the influence of standing waves by calculating the forward-going sound waves separate from the reflections that cause errors. Previous research compared test performance (Burke *et al.*, 2010) and threshold prediction (Rogers *et al.*, 2010) using SPL and multiple FPL calibration conditions, and surprisingly found no significant improvements when using FPL relative to SPL, except at 8 kHz. The present study examined the calibration data collected by Burke *et al.* and Rogers *et al.* from 155 human subjects in order to describe the frequency location and magnitude of standing-wave pressure minima to see if these errors might explain trends in test performance. Results indicate that while individual results varied widely, pressure variability was larger around 4 kHz and smaller at 8 kHz, consistent with the dimensions of the adult ear canal. The present data suggest that standing-wave errors are not responsible for the historically poor (8 kHz) or good (4 kHz) performance of DPOAE measures at specific test frequencies.
[DOI: 10.1121/1.3569726]

PACS number(s): 43.64.Jb, 43.58.Vb [BLM] Pages: 3134–3140

I. INTRODUCTION

Distortion-product otoacoustic emissions (DPOAEs) are evoked with two simultaneously presented tones. DPOAE measurements have traditionally been calibrated using sound-pressure level (SPL) measurements in individual ears. The low-level nature of the otoacoustic emissions (OAE) signals being recorded necessitates the sealing of probe-microphone system in the ear canal in order to minimize interference from external noise sources. This leads to the possibility that the reflective properties of the tympanic membrane (TM) may cause standing waves within the ear canal (Gilman and Dirks, 1986). These standing waves can cause calibration errors of up to 20 dB (Siegel, 1994), and therefore lead to measurement errors. Due to the physical distance from the plane of the probe to the eardrum in adult humans, these errors typically occur around 4 kHz and above. This potentially influences the reliability of DPOAE measurements at high frequencies, which may limit the utility of DPOAEs in monitoring the progression of high-frequency hearing loss associated with ototoxicity or exposure to noise. Although the problem is less pronounced in infants and young children because of the shorter length of their ear canals relative to adults (shifting the standing-wave errors toward higher frequencies), errors still occur during *in situ* calibration for very young patients. Given the importance of high-frequency sounds during the development of speech and language (Stelmachowicz *et al.*, 2004), the ability to accurately assess auditory status at higher frequencies is of value in these cases as well.

Previous studies have explored alternative methods of calibration that are insensitive to standing-wave errors, which might therefore provide more consistent DPOAE measurements. Scheperle *et al.* (2008) investigated DPOAE calibration variability using three different calibration methods: SPL, sound intensity level (SIL), and forward-pressure level (FPL). SPL represents the total pressure measured by the microphone. SIL measurements have been shown previously to be less susceptible than SPL to standing-wave errors in closed ear canals (Neely and Gorga, 1998); however, the study by Scheperle *et al.* was the first to apply SIL calibration to the collection of DPOAEs. In FPL calibration, prior measurements are made in cavities of known acoustic load, followed by *in situ* measurement; this allows the separation of the forward-going (or incident) sound wave (i.e., the pressure wave that propagates toward the TM) from the wave reflected off the TM that travels back to the plane of the probe. Scheperle *et al.* assessed the reliability of calibration procedure by varying the insertion depth of the probe, which changes the frequency at which standing-wave minima occur. The within-subject design reasonably assumed that there were no changes in cochlear or middle-ear status within a test session; thus, any variability in DPOAE measurements following changes in insertion depth was taken as an indication of calibration errors caused by standing waves. This study confirmed that *in situ* SPL calibration was susceptible to variations (errors) from standing waves, which were

---

[a] Author to whom Correspondence should be addressed. Electronic mail: Michael.Gorga@boystown.org

evident when insertion depth was changed. The results reported by Scheperle *et al.* also indicated that both SIL and FPL provided more reliable calibration for DPOAE measurement, compared to SPL. However, SIL required higher stimulus levels during the *in situ* phase of the calibration procedure, with a greater potential for subject discomfort.

In a follow-up to the Scheperle *et al.* (2008) study, Burke *et al.* (2010) and Rogers *et al.* (2010) investigated the impact of different calibration methods on DPOAE test performance and behavioral-threshold prediction, respectively. Data were collected from 155 normal-hearing and hearing-impaired subjects. These two studies explored four different FPL calibration conditions, each of which was less susceptible to standing-wave problems than SPL, and compared the results to those obtained following SPL calibration. Surprisingly, none of the four alternative approaches to calibration showed a significant effect on test performance or threshold prediction at any measurement frequency, with the exception of 8 kHz. However, DPOAEs were only evoked with f_2 frequencies ranging from 2 to 8 kHz at half-octave intervals. These frequencies were chosen because they correspond to frequencies commonly measured in clinical applications and because they cover the range of frequencies for which standing waves are expected in adult human ears. However, it is reasonable to assume that the frequencies at which standing-wave minima occur will vary among individuals, and that it may be the case that these standing-wave minima did not occur at the widely spaced test frequencies evaluated by both Burke *et al.* and Rogers *et al.* Furthermore, if standing-wave minima did occur at one of these frequencies but only rarely, the overall impact on both test performance and threshold prediction would have been reduced by the large sample size in their studies.

In a recent study by Kirby *et al.* (2011), an effect of calibration method was observed, but only when previously determined optimal stimulus conditions (including primary-level and primary-frequency ratios) were used (Neely *et al.*, 2005; Johnson *et al.*, 2006). This effect was accentuated when the alternative calibration method (daily FPL calibrations at body temperature) and optimal stimulus conditions were combined with multivariate analyses of DPOAE data. These results are in slight contrast with the results reviewed above because, unlike that of the previous work, they showed an effect of stimulus calibration on test performance. One explanation for the different outcomes is that the combined use of FPL calibrations, optimal stimulus conditions, and multivariate analyses produced the improved test performance. Burke *et al.* (2010), for example, used univariate analyses and non-optimized stimuli. It may be the case that standing-wave errors existed in the Burke *et al.* study when SPL calibration was used, but the impact on test performance was not evident because non-optimum stimuli and univariate analyses did not exploit the differences resulting from calibration errors.

The purpose of the present study was to analyze the calibration data from the 155 subjects who participated in the studies by Burke *et al.* (2010) and Rogers *et al.* (2010) in order to describe the frequencies at which standing-wave minima were observed and to determine the magnitude of their effects in pressure measurements. Stated another way, the present study is intended to describe the extent to which SPL, measured strictly *in situ*, deviates from FPL, derived from a combination of *in situ* and standardized measurements. The minimal effects seen in the Burke *et al.* study may not be due to the absence of standing waves, but rather due to the fact that the standing-wave pressure minima were not close enough to the measurement frequencies in a sufficient number of subjects in that study to cause an effect. Alternative calibration methods, such as FPL, have been demonstrated to lead to more reliable DPOAE measurements than SPL calibration; it is therefore important to quantify the frequencies and magnitudes of standing-wave effects in order to explore the potential value of alternative DPOAE calibration procedures in clinical settings.

II. METHODS

A. Subjects

The data used in the present study came from the calibration files that were collected as a part of recently published studies on the influence of calibration on DPOAE test performance and threshold prediction (Burke *et al*,. 2010; Rogers *et al.*, 2010). These calibration files were obtained from 155 subjects who ranged in age from 11 to 75 years. Because middle-ear status might affect calibration estimates by altering sound absorption at the plane of the TM, otoscopy, 226-Hz tympanometry (static acoustic admittance between 0.3 and 1.8 mmhos and tympanometric peak pressure between −100 and +50 daPa), and the absence of an air-bone gap >10 dB were used to establish that all subjects had normal middle-ear status. Most subjects ($n = 103$) had a sensorineural hearing loss (threshold >20 dB HL for at least one audiometric frequency). However, no distinction will be made between normal-hearing and hearing-impaired subjects in the present study as there is no reason to expect that calibration results depend on cochlear status.

B. Calibration procedures

Equipment specifications and procedures will be summarized only briefly here because comprehensive descriptions are available in the paper by Burke *et al.* (2010), from which the calibration data were drawn. All measurements were performed using a probe-microphone system (ER-10C, Etymōtic Research, Elk Grove Village, IL) with a PC using custom-designed software (EMAV version 3.01, Neely and Liu, 1994) and a 24-bit sound card (CardDeluxe, Digital Audio Labs, Chanhassen, MN). The calibration procedures used by Burke *et al.* were modeled after those described in Scheperle *et al.* (2008). Specifically, standard calibrations were completed, in which SPL at the plane of the probe was measured *in situ*. In order to estimate FPL, however, it was necessary to estimate Thévenin-equivalent source characteristics of the two ER-10C loudspeakers, which are housed in the ER-10C probe. To accomplish this task, measurements were completed in five brass tubes ($^{11}/_{32}$ inch outside diameter, 8 mm inside diameter), sealed at one end, of varying length (18.5–83 mm), with known acoustic impedances. The lengths of these calibration cavities were selected so that standing-wave peaks would occur at approximately 2, 3, 4, 6, and 8 kHz. The calibration stimulus for both SPL and FPL measurements was a wideband chirp with a sampling rate of 32 kHz.

Burke *et al.* (2010) completed the measurements in the calibration cavities under four different conditions. Calibrations were performed once each day (referred to as "daily"), meaning that the Thévenin-equivalent source characteristics were measured in the morning, just prior to data collection on each day's scheduled subjects. In addition, the same calibration was performed 25 separate times prior to any data collection, and the calibration file with the least error from these 25 files (referred to as the "reference") was also used with all subjects. Both the daily and reference calibrations were performed during two different temperature conditions: room temperature and an approximation of body temperature. The four calibration methods were termed daily body, daily room, reference body, and reference room. All four calibration methods were used with each subject. In order to calculate FPL, *in situ* calibration was performed in each subject with the same probe and stimulus used for source calibration. Burke *et al.* and Rogers *et al.* measured input/output functions at all test frequencies (with test order counter-balanced) in each subject, with *in situ* calibration repeated before each test run, resulting in multiple calibration files per calibration condition for each subject. Only the first chronological file for each calibration condition in each subject was selected for analysis in the present study. Although no significant differences in calibration pressure magnitude or frequency location of minima were expected among the four different calibration conditions, the four conditions were compared and found to yield essentially identical results. Therefore, only one calibration type (daily body) will be presented as representative of all conditions. The daily body calibration files were selected because Burke *et al.* (2010) demonstrated a larger effect in test performance (relative to standard SPL calibration) at 8 kHz with this condition.

C. Data analysis

Each calibration file was examined in the form of a pressure ratio of the total pressure to the forward pressure. This corresponds to the difference in magnitude between SPL (i.e., the total pressure measurement in the ear canal at the plane of the probe), and FPL. In an ideal cylindrical cavity, assuming long wavelengths and no absorption at the TM, SPL would be 6 dB greater than FPL because SPL represents the sum of the forward and reflected pressure. Thus, the ratio of the incident to total pressure would be –6 dB (independent of frequency) in this case. Deviations from the ideal case are expected in human ear canals. The extent to which the difference between SPL and FPL exceeds this amount (i.e., –6 dB) describes the amount by which reflected waves influence SPL measurements. In order to evaluate standing-wave effects as a function of frequency, calibration files were analyzed from 0.5 to 9.5 kHz in 10-Hz steps. At each frequency, cumulative distributions were constructed (using custom code) depending on the magnitude of the effect (cumulative percent as a function of the difference between SPL and FPL).

III. RESULTS

A. Pressure ratio and phase in individual subjects

When calibration files are examined in the form of differences between FPL and SPL, standing-wave minima appear as positive peaks rather than negative notches. Figure 1 is an example of such a calibration file, with FPL-SPL pressure ratio (top panel) and phase (bottom panel) plotted for one of the two ER-10C receivers. Because results were typically similar between the two receivers, the measurement from only one receiver is shown here and in all subsequent figures. While the frequency location and magnitude of the pressure-ratio peak varied among subjects, this calibration file is typical of the results observed in the majority of subjects, in that it has a single amplitude peak and a corresponding change in phase that occurs at the same frequency.

Some calibration files were not characterized by the typical single peak in FPL-SPL differences. Figure 2 demonstrates two cases (from two different subjects) in which the FPL-SPL differences were more complex. Approximately 17% exhibited two distinct peaks, as in Fig. 2(a), within the measurement-frequency range. As in Fig. 1, a change in phase was noted at the frequency location of each peak. An additional 26% showed an upward trend in magnitude approaching 10 kHz, which suggested a second peak outside of the measurement range. One explanation for a second peak would be the presence of a $^3/_4$-wavelength resonance.

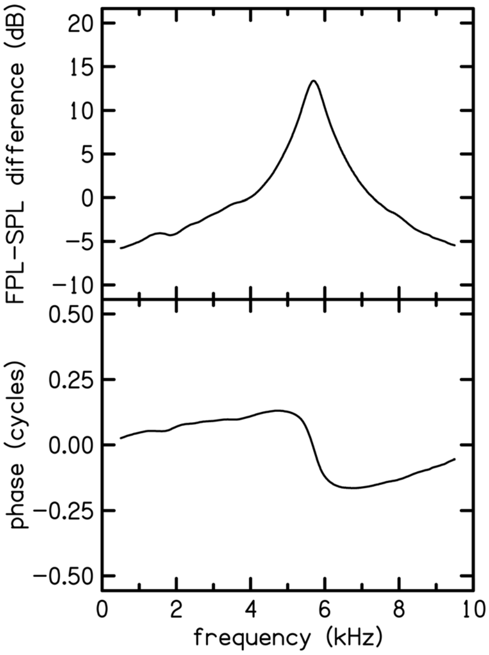

FIG. 1. Example of FPL-SPL pressure ratio (top panel) and phase (bottom panel) computed from measurements from one of the two ER-10C loudspeakers in a single subject. Because measurements were typically very similar between channels, results from only one channel will be displayed in this paper. Presentation of a single peak in magnitude with a change in phase at the corresponding frequency is typical. However, both frequency and magnitude of standing-wave minima vary across subjects.

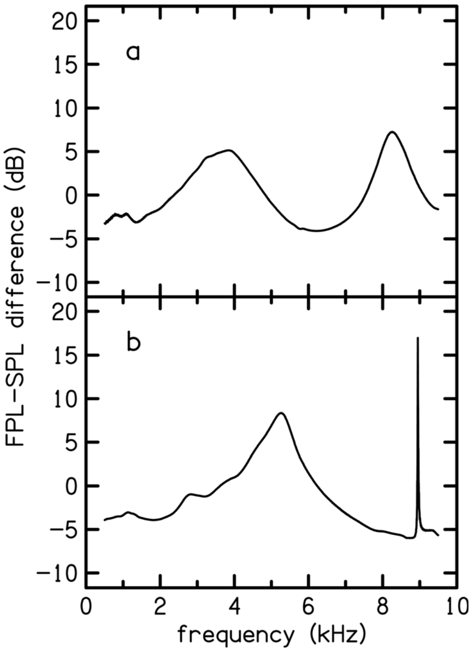

FIG. 2. Examples of FPL-SPL pressure ratios that differ from the expected typical single magnitude peak. Panel (a) is an example of two peaks in the magnitude, possibly due to $\frac{3}{4}$-wavelength resonance. Panel (b) is an example of artifact arising from errors in Thévenin-equivalent source parameter calculation.

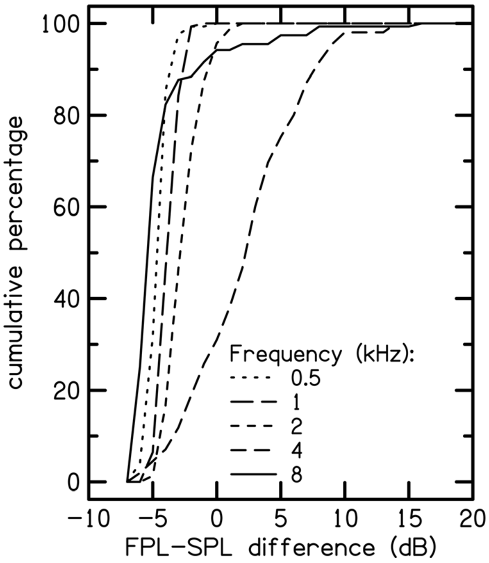

FIG. 3. Cumulative distribution of FPL-SPL differences at octave frequencies (0.5, 1, 2, 4, and 8 kHz). Variability is greater at 4 kHz than at other frequencies, where the range is approximately between –7 and –4 dB.

However, these second peaks sometimes occurred at frequencies lower than expected for $\frac{3}{4}$-wavelength resonances in a uniform tube. Because the human ear canal has a nonuniform diameter and a curved axis, it is not surprising that resonant peaks were not always at frequencies expected for a uniform tube. Human ear canals exhibit changes in diameter and angle that can be unique to each subject, resulting in SPL changes that are not adequately predicted by a uniform cylinder model (e.g., Hudde, 1983).

Measurement artifacts, due to numerical errors in the calculation of Thévenin-equivalent source parameters, appeared as random “spikes” in the pressure magnitude at frequencies above 8.5 kHz, and were present in approximately 17% of the sample. A similar spike also appeared in the corresponding phase at the same frequency. An example of an artifact in magnitude is shown in Fig. 2(b). Artifacts varied in presentation in the following ways: (1) were similar for the two receivers, (2) occurred at the same frequency but differed in magnitude for the two receivers, or (3) appeared on only one receiver. These artifacts are not due to physiological properties in human ear canals. Therefore, in order to avoid the influence of these artifacts (which only occurred at frequencies above the range of typical DPOAE measurements), frequencies above 8 kHz were excluded from estimates of distance from the plane of the probe to the TM (described in Fig. 6 below).

B. Distribution of magnitude differences

Cumulative distributions of FPL-SPL pressure-ratio magnitude differences were analyzed with 10-Hz frequency resolution. The results from this analysis are provided in Fig. 3 at five octave frequencies (0.5, 1, 2, 4, and 8 kHz) for one receiver of the ER-10C probe-microphone system. The distributions at 0.5, 1, 2, and 8 kHz cluster between about –7 and –4 dB and are not highly variable. Recall that the expected low-frequency difference between FPL and SPL in an ideal cylinder is –6 dB. Thus, the data at these frequencies approximate the expected result if all of the energy was reflected from the TM. The variability of magnitude differences around –6 dB may primarily be due to magnitude and phase properties of ear-canal reflectance. The extent to which the human ear canal deviates from an ideal cylinder as well as imperfect measurement due to, for example, an incomplete seal of the probe-microphone system in the ear canal may also contribute to FPL-SPL differences. There is greater variability in the magnitude of the difference between FPL and SPL at 4 kHz, a frequency for which, on average, the quarter wavelength approximates the distance from the plane of the probe to the TM in human adults. Thus, the results at 4 kHz indicate more variability and, therefore, more frequent occurrence of standing-wave effects at this frequency (compared to the other four octave frequencies). Despite the fact that variability is greater at 4 kHz, DPOAE test performance (identifying normal and impaired ears) is best at this frequency (e.g., Gorga *et al.*, 1993, 1997; Kim *et al.*, 1997; Johnson *et al*,. 2010; Burke *et al.*, 2010). Surprisingly, the distribution of magnitude differences at 8 kHz were only slightly greater than those observed at 0.5, 1, and 2 kHz, yet this is a frequency for which test performance is poor and for which calibration procedure had its largest influence (e.g., Gorga *et al.*, 1993, 1997; Burke *et al.*, 2010). The results at 4 and 8 kHz, taken together, indicate that whatever the influence of calibration procedure on test performance, the poorer performance at 8 kHz cannot be attributed to greater variability in calibration.

The results summarized in Fig. 3 provide information about magnitude differences only for octave frequencies from 0.5 to 8 kHz. However, it is reasonable to expect that among a sample of humans, there will be variability in the

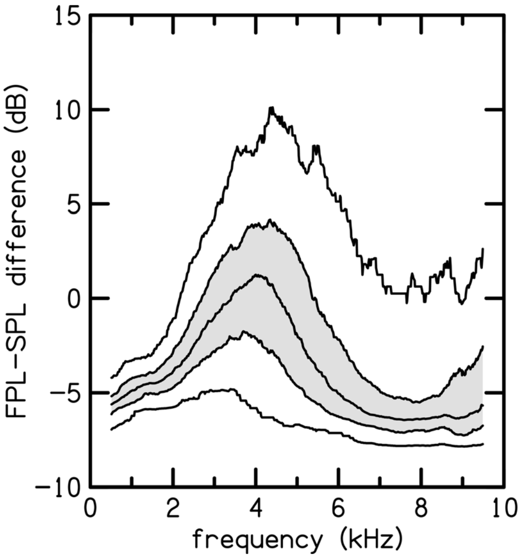

FIG. 4. Curves representing (top to bottom) 95th, 75th, 50th, 25th, and 5th percentile of cumulative distributions of FPL-SPL magnitude differences calculated in 10-Hz steps. Inter-quartile range is shaded. Variability is the greatest for frequencies around 4 kHz.

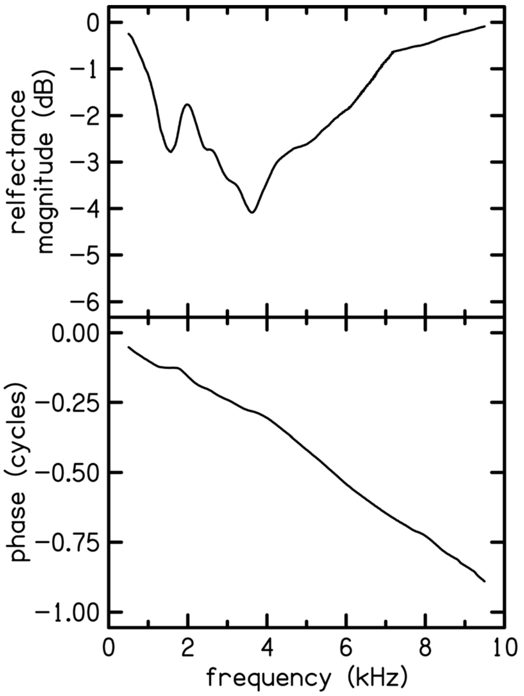

FIG. 5. Example of reflectance magnitude plotted on a decibel scale (top panel) and phase (bottom panel) for the same single-subject calibration file used in Fig. 1 The slope of the phase corresponds to time delay (τ).

frequencies at which standing-wave pressure minima are observed in ear canals. In order to describe the distribution of magnitude differences as a function of frequency, cumulative distributions like the ones shown in Fig. 3 were constructed from 0.5 to 9.5 kHz in 10-Hz steps, based on calibration files from the same 155 subjects who contributed data to the results plotted in Fig. 3. Again, only data from the daily body calibration condition were included in this analysis. Figure 4 plots FPL-SPL magnitude difference (expressed in decibels) as a function of frequency for a restricted set of percentiles. Specifically, data are shown for 95th, 75th, 50th, 25th, and 5th percentiles, with the inter-quartile range indicated with shading. Artifacts may inflate the magnitude in these curves for frequencies above 8 kHz, particularly for the 95th percentile. However, high frequencies were included in this analysis in order to provide a relatively complete description of the magnitude differences between FPL and SPL. There is a tendency for low-frequency data to cluster around –6 dB for percentiles at and below the 75th percentile; high-frequency data clusters in the same manner for the 50th percentile and below. As might be expected from Fig. 3, the greatest variability in magnitude differences (and, therefore, standing-wave effects) occurred around 4 kHz, but large effects were observed for a wider range of frequencies extending from about 3 kHz to about 5 kHz. This variability may be related to the influence of differences in ear-canal dimensions among subjects. There is variability in the FPL-SPL difference at all frequencies, but the variability is less for frequencies at and below 2 kHz, compared to higher frequencies.

C. Estimation of distance to reflection point

The phase information recorded during calibration measurements was used to estimate round-trip time of the signal (i.e., the time from when the forward-going wave was produced by the loudspeaker and the reflected wave was recorded at the plane of the probe microphone). These data were then used to estimate the distance from the plane of the probe microphone to the reflection point (the TM). To illustrate, Fig. 5 is an example of reflectance (R) calculated from the same individual subject's calibration data shown in Fig. 1. The top plot represents the ratio of the reflected pressure (P_r) to the forward pressure (P_f), presented in decibels. Reflectance can be related to pressure with the following formula (Withnell *et al.*, 2009), where total pressure $P_t = P_r + P_f$:

$$\frac{P_f}{P_t} = \frac{1}{1+R} \qquad (1)$$

P_f/P_t can be greater or less than one. If P_f and P_r differ in phase by more than 0.25 cycles due to propagation delay, the real part of R may be negative, which would allow the ratio to be greater than one.

The bottom panel of Fig. 5 displays the phase of the reflectance in cycles, the slope which corresponds to time delay (τ) and is expected to be a nearly linear function of frequency. The slope was estimated using linear regression, and subsequently was used to estimate distance from the plane of the probe to the TM (d) with the following formula incorporating the speed of sound (c = 352 m/s) at near-body temperature of 35°C:

$$d = \frac{\tau \cdot c}{2} \qquad (2)$$

The numerator ($\tau \cdot c$) estimates the round-trip distance. Thus, it was divided by two in order to convert it into a one-way estimate of distance from the probe to the TM. Artifacts at

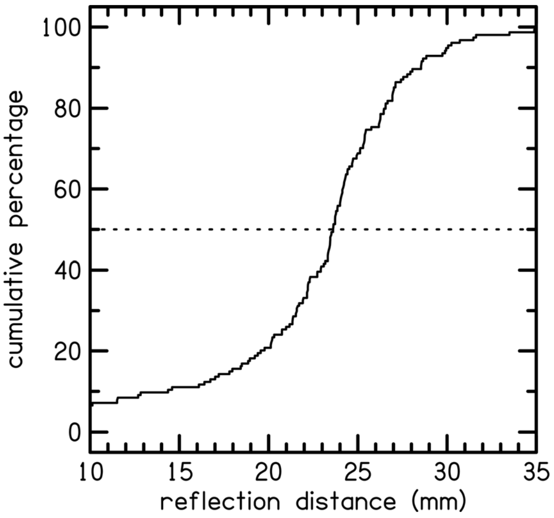

FIG. 6. Cumulative distribution of estimates of distance from sound source (probe-microphone) to reflection point (TM). Distances were derived by using linear regression to estimate the values of time delay (τ) across the group. The 50th percentile is indicated by a dotted line and corresponded to a distance of 23.5 mm, approximately equivalent to the 1/4-wavelength value of 22 mm for 4 kHz.

high frequencies would have had a significant effect on the estimates of the reflectance phase slope. Therefore, only calibration data up to 8 kHz were included in these calculations because artifacts were not observed at or below 8 kHz. The estimated distances from the plane of the probe to the TM are consistent with the frequencies of the peaks in the FPL-SPL difference curves. A cumulative distribution of these distances is plotted in Fig. 6. Again, the two receivers had similar distributions, so only the estimates of the distance for one receiver are included here. The expected wavelength of 4 kHz at this temperature is 88 mm. As shown in Fig. 6, the 50th percentile of the cumulative distribution occurs at approximately 23.5 mm, near the 1/4-wavelength of 4 kHz (22 mm). This explains why the largest FPL-SPL differences in Fig. 4 occurred near 4 kHz.

IV. DISCUSSION

The purpose of this study was to describe the frequency location and magnitude of standing-wave effects that occur during calibration of a probe-microphone system in human ear canals. Previous research has shown that standing-wave minima encountered during "traditional" SPL measurements can be as large as 20 dB (Siegel, 1994). FPL calibration provides an approach that attempts to avoid these errors by separating the forward-going pressure from reflections that cause standing waves. However, Burke *et al.* (2010) and Rogers *et al.* (2010) did not find improvements in DPOAE test performance or threshold prediction when using four FPL alternatives to SPL calibration, except at 8 kHz. Theoretically, one would predict that standing-wave effects would be greatest for frequencies around 4 kHz, given the dimensions of the ear canal in adult humans and the wavelength at this frequency. The data reported by Burke *et al.* and Rogers *et al.* provided no evidence that standing-wave effects during calibration led to poorer test performance or threshold prediction at 4 kHz.

One explanation for the minimal effects of calibration method seen in previous research is that standing-wave pressure minima do not occur frequently enough at or near the specific test frequencies used in previous studies to cause a significant effect. However, the present data do not support this hypothesis. Figure 3 demonstrates that standing-wave effects are most evident at 4 kHz, while Fig. 4 shows that the largest effects cluster around 4 kHz, although they extend to frequencies above and below 4 kHz. While there is also variability at higher frequencies, the magnitude of these effects is not nearly as large as those around 4 kHz. Even at 4 kHz, 95% of these magnitude effects were less than 16 dB. While maximum values fell outside of this range, these large values were not as common as might have been expected based on previous research (e.g., Siegel, 1994).

Reflectance is used clinically as a means of assessing middle-ear status. However, it is also useful in the present context for separating forward and reflected components of pressure in the ear canal. Reflectance minima and maxima vary in frequency location across subjects. The reflectance plot in Fig. 5 demonstrates a minimum magnitude at around 4 kHz and a maximum above 8 kHz, consistent with the results of previous studies (e.g., Keefe *et al.*, 1993; Lewis *et al.*, 2009). In other words, sound transmission through the TM is greatest around 4 kHz and is least around 8 kHz. These are frequencies where DPOAE test performance has historically been found to be good (4 kHz) and poor (8 kHz) (e.g., Gorga *et al.*, 1993, 1997; Burke *et al.*, 2010). Lower reflectance implies greater sound transmission at the TM and vice versa. The extent of transmission of sounds through the TM at certain frequencies could therefore be more predictive of test performance than standing-wave effects. However, this hypothesis was not supported by Kirby *et al.* (2011), who found no correlation between reflectance and DPOAE test performance.

The issue of measurement artifacts was left largely unaddressed in the present study. For example, Fig. 4 displayed variation in magnitudes above 8.5 kHz, possibly due to these artifacts. In future studies, it may be prudent to remove these artifacts in order to improve the accuracy of FPL measurements. Additionally, even at locations without standing waves or artifacts, the FPL-SPL difference still deviates from –6 dB, which is the expected value in an ideal cylinder at low frequencies: For example, note the slight variations at 1 kHz and below in Fig. 4, where we do not expect to see an effect of standing waves. The deviations from ideal could be due to imperfect sealing of the ear canal leading to low-frequency sound leakage or individual differences in human ear-canal shape.

Based on the results of the present study, the distribution of FPL-SPL differences provides no explanation for why DPOAE performance and threshold prediction in Burke *et al.* (2010) and Rogers *et al.* (2010) are worse at 8 kHz compared to other test frequencies. One concern that has yet to be resolved relates to the influence of system distortion on DPOAE measurements. All probe-microphone systems produce distortion at some level. The system used to collect the present data (Etymōtic ER-10C) is no exception. The levels of system distortion are higher at 8 kHz, compared to lower frequencies. System distortion could influence measures of test performance because it might result in the detection of energy at the distortion-product frequency that was generated by the hardware, and not by the biological system (cochlea). Under these circumstances it would be possible to commit a false-negative error because an ear with hearing loss might be misdiagnosed as normal hearing.

In any event, the results of the present study demonstrate that standing-wave minima exist in human ear canals and are most prevalent around 4 kHz. This observation is consistent with the expectation of standing-wave effects, given the wavelength at 4 kHz and the dimensions of the human ear canal. These standing waves cause errors in calibration, but their influence on test performance is minimal, given the observation of excellent results at 4 kHz in many other studies. Although calibration errors were evident at other frequencies, the distribution of error magnitudes does not provide an explanation for the poorer test performance at 8 kHz in the Burke *et al.* study. However, the results of Scheperle *et al.* (2008) and Kirby *et al.* (2011) suggest value in continuing to explore FPL as an alternative method of calibration for DPOAEs.

ACKNOWLEDGMENTS

This work was supported by the NIH (Grant Nos NIDCD R01 DC2251, R01 DC8318, P30 DC4662, and T35 DC8757). Interest in this project was stimulated by a discussion between Jont Allen and one of the authors (MPG).

Burke, S. R., Rogers, A. R., Neely, S. T., Kopun, J. G., Tan, H., and Gorga, M. P. (**2010**). "Influence of calibration method on distortion product otoacoustic emission measurements. I. Test performance," Ear Hear. **31**(7), 533–545.

Gilman, S., and Dirks, D. D. (**1986**). "Acoustics of ear canal measurement of eardrum SPL in simulators," J. Acoust. Soc. Am. **80**(3), 783–793.

Gorga, M. P., Neely, S. T., Bergman, B., Beauchaine, K. L., Kaminski, J. R., Peters, J., and Jesteadt, W. (**1993**). "Otoacoustic emissions from normal-hearing and hearing-impaired subjects: Distortion product responses," J. Acoust. Soc. Am. **93**(4), 2050–2060.

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**(6), 440–455.

Hudde, H. (**1983**). "Estimation of the area function of human ear canals by sound pressure measurements," J. Acoust. Soc. Am. **73**(1), 24–31.

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**(1), 418–428.

Johnson, T. A., Neely, S. T., Kopun, J. G., Dierking, D. M., Tan, H., and Gorga, M. P. (**2010**). "Clinical test performance of distortion-product otoacoustic emissions using new stimulus conditions," Ear Hear. **31**(1), 74–83.

Keefe, D. H., Bulen, J. C., Arehart, K. H., and Burns, E. M. (**1993**). "Ear-canal impedance and reflection coefficient in human infants and adults," J. Acoust. Soc. Am. **94**(5), 2617–2638.

Kim, D. O., Paparello, J., Jung, M. D., Smrzynski, J., and Sun, X. (**1997**). "Distortion product otoacoustic emission test of sensorineural hearing loss: Performance regarding sensitivity, specificity, and receiver operating characteristics," Acta Oto-Laryngol. **116**, 3–11.

Kirby, B. J., Kopun, J. G., Tan, H., Neely, S. T., and Gorga, M. P. (**2011**). "Do 'optimal' conditions improve distortion product otoacoustic emission test performance?" Ear Hear. **32**(2), 230-237.

Lewis, J. D., McCreery, R. W., Neely, S. T., and Stelmachowicz, P. G. (**2009**). "Comparison of *in-situ* calibration methods for quantifying input to the middle ear," J. Acoust. Soc. Am. **126**(6), 3114–3124.

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

Neely, S. T., and Gorga, M. P. (**1998**). "Comparison between intensity and pressure as measures of sound level in the ear canal," J. Acoust. Soc. Am. **104**, 2925–2934.

Neely, S. T., and Liu, Z. (**1994**). "EMAV: Otoacoustic emission averager," Technical Memo No. 17, Boys Town National Research Hospital, Omaha, NE.

Rogers, A. R., Burke, S. R., Kopun, J. G., Tan, H., Neely, S. T., and Gorga, M. P. (**2010**). "Influence of calibration method on distortion product otoacoustic emission measurements. II. Threshold prediction," Ear Hear. **31**(4), 546–554.

Scheperle, R. A., Neely, S. T., Kopun, J. G., and Gorga, M. P. (**2008**). "Influence of *in situ* sound-level calibration on distortion-product otoacoustic emission variability," J. Acoust. Soc. Am. **124**(1), 288–300.

Siegel, J. H. (**1994**). "Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes," J. Acoust. Soc. Am. **95**(5), 288–300.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., Lewis, D. E., and Moeller, M. P. (**2004**). "The importance of high-frequency audibility in the speech and language development of children with hearing loss," Arch. Otolaryngol. Head Neck Surg. **130**, 556–562.

Withnell, R. H., Jeng, P. S., Waldvogel, K., Morgenstein, K., and Allen, J. B. (**2009**). "An *in situ* calibration for hearing thresholds," J. Acoust. Soc. Am. **125**(3), 1605–1611.