Analysis of parameters for the estimation of loudness from tone-burst otoacoustic emissions

Michael Epstein[a)]
Department of Speech-Language Pathology and Audiology, Auditory Modeling and Processing Laboratory, Communications and Digital Signal Processing Center, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

Ikaro Silva
Department of Electrical and Computer Engineering, Auditory Modeling and Processing Laboratory, Communications and Digital Signal Processing Center, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

(Received 4 November 2008; revised 28 February 2009; accepted 4 March 2009)

There is evidence that tone-burst otoacoustic emissions (TBOAEs) might be useful for estimating loudness. However, within-listener comparisons between loudness and TBOAE measurements are an essential prerequisite to determine appropriate analysis parameters for loudness estimation from TBOAE measurements. The purpose of the present work was to collect TBOAE measurements and loudness estimates across a wide range of levels in the same listeners. Therefore, TBOAEs were recorded for 1- and 4-kHz stimuli and then analyzed using a wide range of parameters to determine which parameter set yielded the lowest mean-square-error estimation of loudness with respect to a psychoacoustical, cross-modality-matching procedure and the inflected exponential (INEX) loudness model. The present results show strong agreement between 1-kHz loudness estimates derived from TBOAEs and loudness estimated using cross-modality matching (CMM), with TBOAE estimation accounting for almost 90% of the CMM variance. Additionally, the results indicate that analysis parameters may vary within a reasonable range without compromising the results (i.e., the estimates exhibit some parametric robustness). The lack of adequate parametric optimization for TBOAEs at 4 kHz suggests that measurements at this frequency are strongly contaminated by ear-canal resonances, meaning that deriving loudness estimates from TBOAEs at this frequency is significantly more challenging than at 1 kHz.
[DOI: 10.1121/1.3106531]

PACS number(s): 43.66.Cb, 43.64.Jb [BLM] Pages: 3855–3864

I. INTRODUCTION

Typical procedures for the estimation of loudness growth require active participation by listeners giving subjective responses to stimuli. An alternative methodology for estimating loudness objectively would be very valuable in clinical populations that are unable to provide subjective responses. Schlauch *et al.* (1998) and Buus and Florentine (2001) showed compelling evidence that, for pure tone stimuli, loudness growth is proportional to the square of basilar-membrane velocity. This indicates that measures of cochlear activity may be useful for estimating loudness, at least for pure tones. A number of investigators have examined the general relationship between otoacoustic emissions and cochlear gain control, intensity, or loudness. Many of these studies have utilized distortion-product otoacoustic emissions (DPOAEs) and found at least some relationship between loudness and DPOAE input/output functions (Muller and Janssen, 2004; Janssen *et al.*, 2006; Neely *et al.*, 1997; Neely *et al.*, 2003; Buus *et al.*, 2001; Johannesen and Lopez-Poveda, 2008), but there has typically been difficulty selecting parameters that are suitable for estimating loudness over a wide range of levels in a diverse set of listeners. Despite evidence that tone-burst otoacoustic emissions (TBOAEs) might be useful for loudness estimation (Epstein *et al.*, 2004; Epstein and Florentine, 2005a), there is a dearth of measurements directly examining loudness estimates coupled with TBOAEs over a wide range of levels in a single group of listeners. In particular, there is some evidence that TBOAEs might be useful for estimating loudness at regional sites, as the response for a narrow-band stimulus is likely generated at around a characteristic location along the cochlea (Norton and Neely, 1987; Shera *et al.*, 2002) and OAEs are useful for estimating the peak of the traveling wave (Zweig and Shera, 1995). With the limited data in the literature, however, it is difficult to determine appropriate analysis parameters to be used for such an estimation procedure. The present work seeks to determine parameters that minimize the error in the loudness estimation from TBOAEs. Additionally, these parameters should be robust to small individual variations in optimality. For the present study, TBOAEs were recorded in response to 1- and 4-kHz tone bursts and then analyzed using a wide range of parameters to determine which parameter set yielded the average least-squares error estimation of loudness in reference to loudness functions derived using a cross-modality matching (CMM) procedure and the inflected exponential (INEX) loudness model (Florentine and Epstein, 2006). If TBOAEs are to be used for loudness estimation, it is important and desired that the esti-

---

[a)] Author to whom correspondence should be addressed. Electronic mail: m.epstein@neu.edu mation procedure is insensitive to small variations in parametric choices in order to allow a robust procedure that accounts for anatomical and physiological variability across listeners. In addition, great care must be taken to avoid the linear, non-cochlear portion of the response that results from ear-canal resonances and other trivial acoustic reflections within the auditory system (Ravazzani *et al.*, 1996), which are particularly prevalent around 4 kHz.

II. METHODS

A. Listeners

Six listeners with normal hearing (four females, two males), ages 19–31, participated in both TBOAE and loudness measurements. No listener had a history of hearing difficulties, and their audiometric thresholds did not exceed 15 dB hearing level (HL) at octave frequencies from 250 Hz to 8 kHz (ANSI, 1996). Additionally, all listeners had their middle-ear function evaluated via a clinical exam. Measurements were also made in a coupler to further verify when the measurement parameters resulted in stimulus artifact contamination.

B. Stimuli

The tone bursts used in all parts of the experiment were 2-cycle-up-2-cycle-down pure tones multiplied by Gaussian windows. Two frequencies were tested: 1 and 4 kHz. The 1-kHz tone had a 4-ms duration and the 4-kHz tone had 1-ms duration. This 2-cycle-up-2-cycle-down tone duration ensured energy consistency between the two different frequencies (Hall, 2007). The windowed tone was then end-padded with silence to generate a stimulus length of 41.7 ms. The stimulus levels varied from 25 to 100 dB sound pressure level (SPL) in steps of 5 dB. The measurements at 25 and 30 dB SPL were omitted from analysis and presentation due to limited success getting both TBOAE and CMM measurements at these levels. Levels matched the specifications of the voltage-to-level conversion provided by Etymotic Research for the ER-10c apparatus.

For the CMM procedure, each stimulus presentation consisted of 12 concatenated 41.7-ms intervals in order to generate a train of tone bursts that lasted approximately 0.5 s. This train was presented in place of a single tone-burst in order to minimize any potential temporal integration effects (Buus *et al.*, 1997; Florentine *et al.*, 1996; Zwicker and Fastl, 1999). Levels were determined by a pressure-proportional voltage-to-level conversion based on calibration levels measured in a coupler.

C. Apparatus

The stimuli were generated in MATLAB (2007b running on Windows 2000 for CMM and Ubuntu for TBOAEs) and converted from digital (48-kHz sampling frequency) to analog using a 32-bit Lynx Two Soundcard. The analog signal was then passed through either a Tucker-Davis Technologies (TDT) HB6 (CMM) or a TDT HB7 (TBOAEs) headphone buffer and presented monaurally via Sony MDR-V6 headphones (CMM) or the two transducers of the Etymotic ER-10C (TBOAEs) to a listener inside a double-walled sound-attenuating booth. During the TBOAE measurements, the recordings from the ER-10C were converted from analog to digital (48-kHz sampling frequency) via a Lynx Two soundcard. Routine calibration for each system was performed at the beginning of each session to test for proper wiring and ER-10C output in a plastic syringe coupler provided by Etymotic. In order to perform calibration, the signal was sent through the system and both the electrical signal and the acoustic signal from the coupled ER-10C were measured to be within 1 dB of the expected level. For the TBOAEs, all levels were determined using the rms of the windowed signal relative to the specifications which were provided by Etymotic and verified by doing an actual in ear measurement for a single listener using a Fonix 6500-CX real-ear system.

D. TBOAE recordings

Stimuli were presented in blocks of 1000 trials (about 41.7 ms per trial, at a presentation rate of about 24 Hz). Each block of trials was repeated eight times for each level yielding 8000 recordings per level. For each level, two averages of TBOAE recordings were made. The first average consisted of a weighted mean (Elberling and Wahlgreen, 1985) of all the trials in the first half of each of the eight blocks (total of 4000 trials), and the second average consisted of a weighted mean of all the trials in the second half of each of the eight blocks (total of 4000 trials). These averages were the basis for the loudness estimation procedure described in Sec. II G. Additionally, a separate analysis was done on the results recorded in a plastic coupler with the approximate size of a normal ear canal in order to yield a stable reference and an artificial approximation of the acoustic response of the ear canal.

E. CMM

Listeners were presented with six repetitions of each level in random order and asked to cut a string to be "as long as the sound is loud." After the listener cut each string, they taped it into a notebook and turned the page. Two blocks of trials were run separately, one for each of the two test frequencies. If a particular stimulus was not heard, no string was cut. Levels were omitted if fewer than four out of the six repetitions were heard. All listeners provided at least four strings for levels at 35 dB SPL and higher. The loudness estimate for each level was the transformed geometric mean of the string lengths produced for that level. The transformation was performed in response to the finding that CMM, although it provides access to the details of the shape of the loudness function for individual listeners, yields functions with shallower slopes than other procedures (Epstein and Florentine, 2005b). As such, a string-length multiplicative correction factor was determined by using a least-squares fit to match the average group data to a power function with an exponent equal to 0.3, widely used as a simple first approximation of the general form of the loudness function (Hellman and Zwislocki, 1963; Stevens and Guirao, 1964; Stevens, 1955; Stevens, 1957; Stevens, 1961). This correction factor was then applied to the individual data. While it is also known that the CMM string-length procedure may have edge effects at low levels, as listeners have a difficult time cutting very small string sizes, the string-length CMM procedure has been shown to yield reliable individual data for a wide range of levels (Epstein and Florentine, 2005b; Epstein and Florentine, 2006). A number of other similar cross-modal line-length methods have also been used to estimate loudness functions with consistent results (Hellman, 1999; Teghtsoonian and Teghtsoonian, 1983; Thalmann, 1965; Serpanos and Gravel, 2000, 2004; Serpanos *et al.*, 1998). The final loudness-growth curve was subtracted by an offset in order to yield a zero-mean loudness curve for comparison with loudness curves obtained through other modalities. The comparative parameters of interest and meaning are the slopes of the functions. Thus, the offsetting was performed on individual data sets. Because the scales differ, the vertical offsets between TBOAEs and CMM measurements are arbitrary.

F. INEX loudness model

The INEX loudness model was used only for comparisons with group average data and not for any individual fitting or parameter optimization. The INEX model is described in full in Florentine and Epstein (2006). It is a simple set of modifications to the classical power function that includes the subtle variations in slope as a function of level observed in a number of studies by a variety of investigators (Buus and Florentine, 2001; Buus *et al.*, 1998, 1997; Hellman and Zwislocki, 1961; Robinson, 1957; Stevens, 1972; Zwislocki, 1965). The plotted INEX used here is computed by the following equation:

$$\begin{aligned}\log_{10}(N) = {} & 1.7058 \times 10^{-9}L^5 - 6.587 \times 10^{-7}L^4 + 9.7515 \\ & \times 10^{-5}L^3 - 6.6964 \times 10^{-3}L^{2+}0.2367L \\ & - 3.4831,\end{aligned}$$

where N is the loudness in sones (Stevens, 1936) and L is the level in dB SPL. Because this model is designed for longer-duration sounds on the order of around 200–500 ms, the function was then adjusted by 28.75 dB to account for the difference in sensation level between the very brief tone bursts used in the present study and longer-duration sounds used in most psychoacoustical experiments (Epstein and Florentine, 2005b; Epstein and Florentine, 2006; Florentine *et al.*, 1996, 2001). This is the equivalent of replacing L in the equation with $(L+28.75)$. This value was determined by calculating a psychometric function from the number of responses to the CMM stimuli and determining the level at which, on average, 50% of trials were detected.

G. Estimation of loudness from TBOAEs

The procedure used to estimate a loudness function from the two averaged TBOAE waveforms at each level is similar to that of Epstein and Florentine (2005a). For the present study, four parameters were allowed to vary: window delay from the stimulus onset (from 0 to 39 ms in 1 ms steps), window size (2.5, 10, 20, or 30 ms), window type (rectangular or Hanning), and frequency bandwidth of the analysis region (F-ratio=1, 1.5, 2, or 3). Combinations in which the sum of window delay and window size exceeded the duration of the recording (41.7 ms) were omitted from analysis. F-ratio is defined here as the ratio by which the lower and upper bounds of the frequency bandwidth of the analysis region are related to the center frequency. An F-ratio of 1 indicates that only the single fast-Fourier-transform (FFT) bin nearest to the center frequency is used. An F-ratio of 2 indicates that the lower bound of analysis is one-half of the center frequency and the upper bound of analysis is two times the center frequency. For instance, with a 1-kHz tone-burst, an F-ratio of 2 corresponds to an analysis band from 500 Hz to 2 kHz (i.e., an octave-wide analysis region on each side of the center frequency).

While varying these parameters, loudness for each level was estimated using three steps. First, each of the two averaged waveforms (weighted, point-by-point means of 4000 trials) was windowed using the window delay, window type, and window size selected. In the second step, FFTs of the two individual averages were calculated and the real components of the cross-spectrum between the two averages determined. Instead of using the FFT absolute magnitude, the real components of cross-spectra were used in order to minimize noise artifacts by including only the portions of the wave that are synchronized in the two averages. In the third and final step, loudness is estimated by taking the logarithm of the sum of the positive, real components of the cross-spectrum within the frequency region specified by the F-ratio. The final loudness-growth curve was subtracted by an offset in order to yield a zero-mean loudness curve for comparison with loudness curves obtained through the other methods utilized.

H. Peak TBOAE latency estimation

The peak latencies of the recorded responses were also estimated as a function of level and stimulus frequency. The recordings at each level were filtered using an octave-wide filter with a center frequency equal to the stimulus frequency. The filter was generated in MATLAB as an eighth-order Butterworth infinite impulse response filter. The filtering was performed off-line in conjunction with the function FILTFILT in order to ensure that the filtered response had no phase shifts. The envelopes of these filtered responses were then estimated by passing the magnitude of the Hilbert transform of the filtered signals through a low-pass eighth-order Butterworth filter with cut-off frequency set to 200 Hz (again in conjunction with the FILTFILT function). The latency of the peak OAE in each condition was the maximum value of the filtered Hilbert transform observed between 2 ms after stimulus offset and 25 ms after the stimulus onset. The initial delay from the stimulus offset was chosen to help reduce the likelihood that the latency observed actually resulted from trivial acoustic artifacts in the ear canal rather than OAEs.

III. RESULTS AND DISCUSSION

A. Cross-modality measurement correction factor

The CMM measurements for 1 and 4 kHz were independently scaled by a fixed correction factor to ensure that the

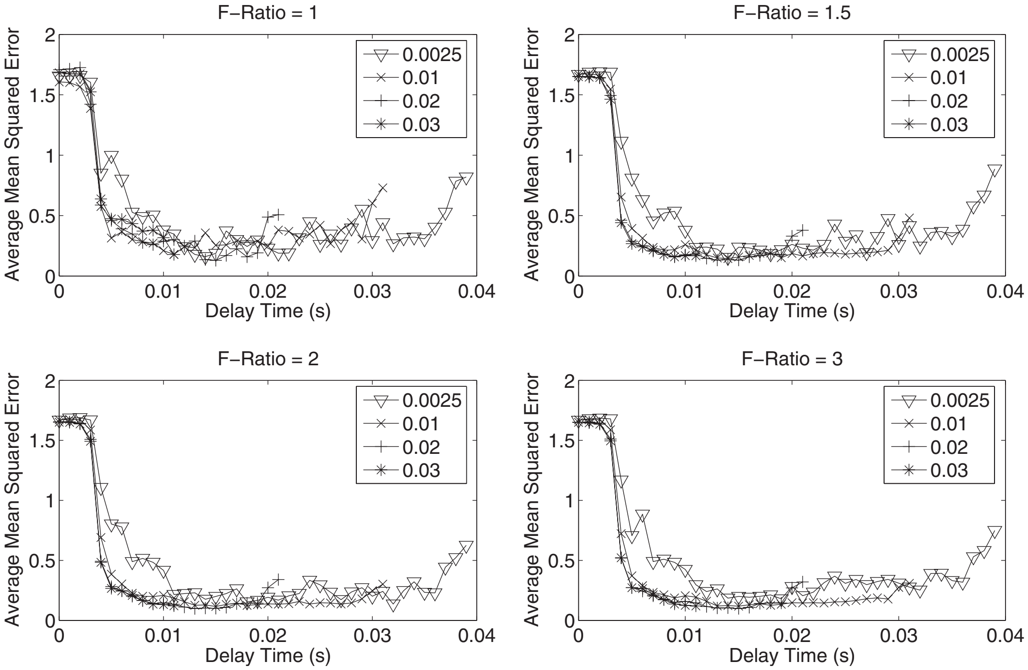

FIG. 1. AMSE between loudness measured using CMM and loudness estimated from TBOAEs in response to 1-kHz tone bursts using a rectangular analysis window. The plots show AMSE as a function of window delay, with each curve representing different window sizes and each plot representing a different *F*-ratio.

resulting functions were similar to the expected form of the loudness function (i.e., power function exponent equal to 0.3). The same scaling factor, 1.67, was used for all subjects based on the average data. The 1- and 4-kHz scale-factors were calculated separately, but agreed exactly to three decimal places, so the same value was used for both. This is consistent with the expectation that loudness functions for tones at different frequencies are parallel on a log scale within a relatively wide range of frequencies (Hellman, 1976).

B. Loudness Estimation from TBOAEs at 1 kHz

Figure 1 shows the averaged mean squared error (AMSE) across subjects between loudness measured using CMM (after scaling) and loudness estimated from TBOAEs in response to 1-kHz tone bursts using a rectangular analysis window. The measures are compared directly on an arbitrary normalized logarithmic scale with both functions vertically shifted to zero mean. An increase of 1 unit on this scale corresponds to the scaled string length multiplied by 10 and a change of 10 dB in TBOAE level. The plots show AMSE as a function of window delay, with each curve representing a different window size and each panel representing a different *F*-ratio. Figure 2 is identical to Fig. 1 except that the TBOAE loudness estimation was instead performed using a Hanning analysis window. All of the loudness estimations shown in both figures exhibit a similar pattern. Regardless of other parameters, when window delay is very short, the window captures stimulus artifact or the linear reflections within the auditory system. As the delay gets longer, the AMSE improves, indicating that the effects of the stimulus artifact diminish and the non-linear cochlear response becomes more pronounced. At very long delays, the window fails to capture the TBOAE at a level strong enough to overcome the background noise.

The Hanning window (Fig. 2) appears more robust than the rectangular window (Fig. 1) and results in near-optimal AMSE for a larger range of delays. In addition, the Hanning window reaches optimality sooner because it de-emphasizes the early portion of the signal more. It is thus important to note that the Hanning window has a total area equal to half of the rectangular window. As such, the Hanning window is more localized in time, resulting in an increased side-lobe attenuation in the spectral domain, which can improve spectral estimation and reduce any leakage artifacts (Harris, 1978). The AMSE curve is approximately flat for delays of 13–20 ms, similar to the findings of Epstein *et al.* (2004) in which no significant effect of window delay was found in that range. It is noteworthy that the apparatus and the duration of the stimulus differ somewhat between the two experiments and the results are still in agreement.

Each of the curves in Figs. 1 and 2 are for different window sizes. The shortest window size (0.0025 ms) tends to yield high variance estimates as a function of delay. This short window size is very sensitive to local variations in background noise because of the small number of points being collected. In addition, because the window is so short, it has a smaller range of near-optimal delays. At short delays, in particular, this window will place a lot of emphasis on the stimulus itself resulting in a poor loudness estimate. The loudness estimate becomes more robust as the window size increases (i.e., a classical trade-off between recording time and susceptibility to noise in measurements). The 20-ms and

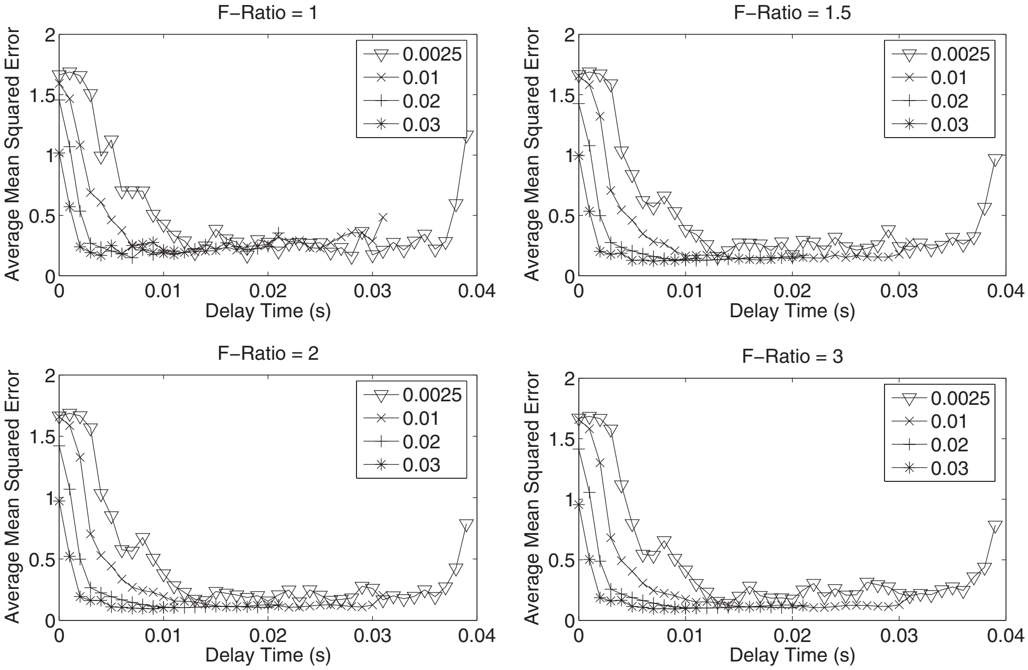

FIG. 2. AMSE between loudness measured using CMM and loudness estimated from TBOAEs in response to 1-kHz tone bursts using a Hanning analysis window. The plots show AMSE as a function of window delay, with each curve representing different window sizes and each plot representing a different *F*-ratio.

30-ms windows perform approximately equally well. As a result, a 20-ms window was arbitrarily chosen for subsequent loudness estimation.

The *F*-ratios of 2 and 3 result in lower optimal AMSEs (AMSE=0.10 for both) than *F*-ratios of 1 or 1.5 (AMSE =0.15 and 0.13, respectively). As the AMSEs for *F*-ratios of 2 and 3 are approximately equal, an *F*-ratio of 2 will be used for subsequent loudness estimation. The combination of parameters that provides the lowest AMSE is a Hanning window with a window delay of 10 ms, a window size of 20 ms, and an *F*-ratio of 2.

C. Loudness estimation from TBOAEs at 4 kHz

It is known that it is difficult to measure TBOAEs at 4 kHz due to ear-canal resonance and signal artifacts (Ravaz-

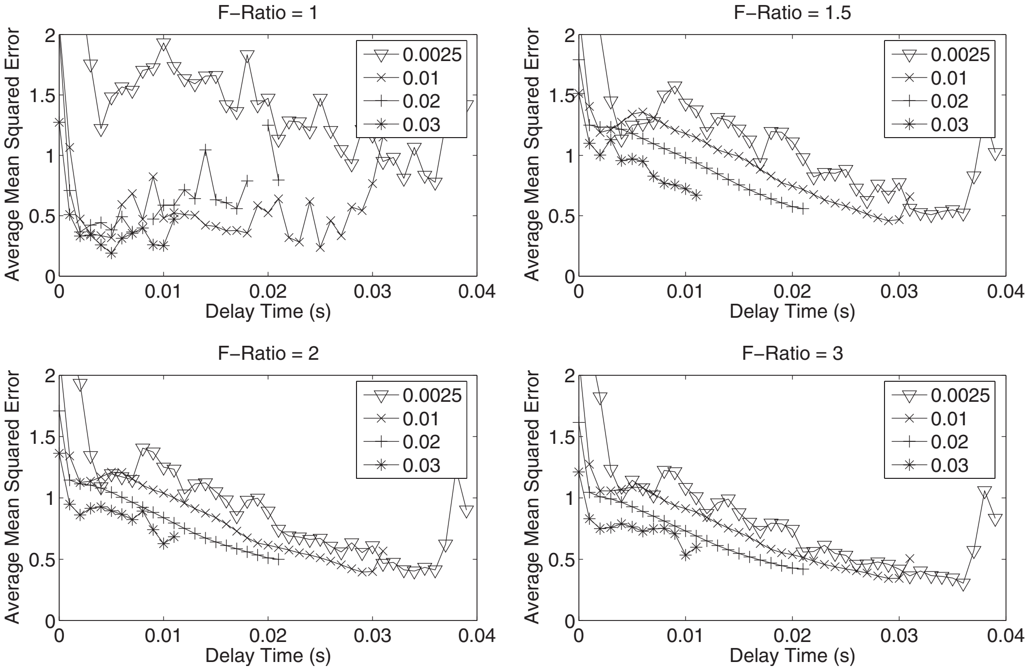

FIG. 3. AMSE between loudness measured using CMM and loudness estimated from TBOAEs in response to 4-kHz tone bursts using a Hanning analysis window. The plots show AMSE as a function of window delay, with each curve representing different window sizes and each plot representing a different *F*-ratio.

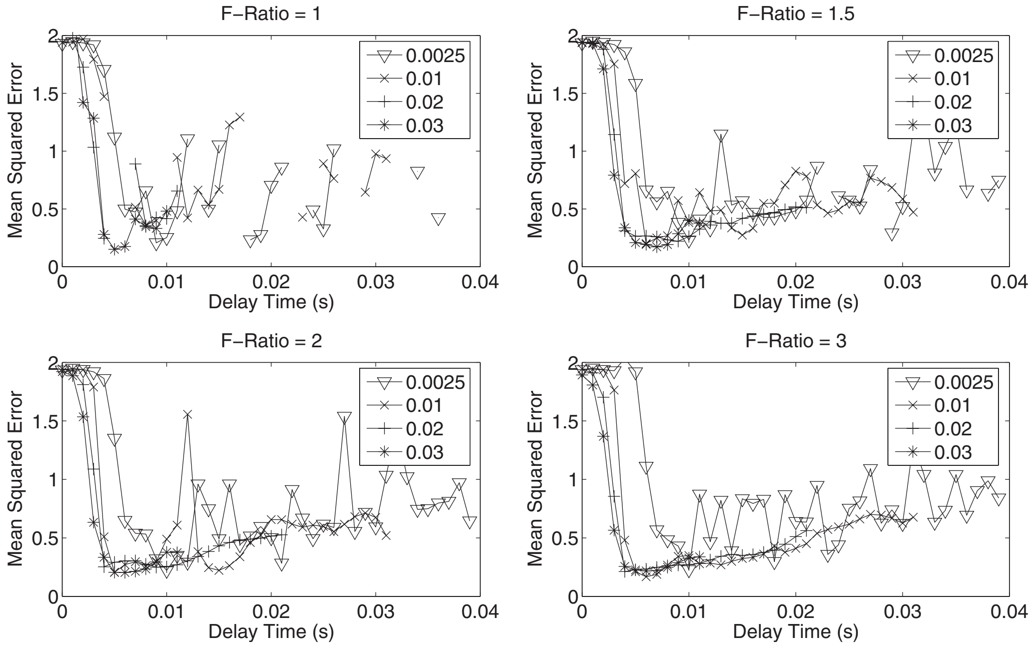

FIG. 4. MSE between loudness measured using CMM and loudness estimated from the response of the coupler to 1-kHz tone bursts using a Hanning analysis window. The plots show AMSE as a function of window delay, with each curve representing different window sizes and each plot representing a different *F*-ratio.

zani and Grandori, 1993; Whitehead *et al.*, 1994). Identical sets of analyses were performed for the responses to 4-kHz tone bursts to determine whether the current procedure could be used to estimate loudness of 4-kHz tone bursts using any possible combination of parameters. Although neither window produced particularly good results, both the Hanning and rectangular windows yielded roughly the same results, with the Hanning window providing a slightly more robust response. Therefore, only the Hanning-window results are shown in Fig. 3.

None of the parametric combinations resulted in AMSE approaching the values measured at 1 kHz. As the window delay increases and less of the stimulus resonance is captured, the error decreases. However, it seems like the linear response of the ear-canal resonance overpowers any possible non-linear cochlear response. The results support reports of difficulty making measurements at 4 kHz and support the claim that most of the 1-kHz responses are true non-linear responses and not simple linear reflections. The small number of missing data points resulted from noisy data that did not contain any positive real components in the cross-spectrum (see Sec. II G).

D. Coupler measurements

The goal of all of these analyses was not only to determine the parameters that minimize the error in the loudness estimation but also to identify a possible spurious relationship with stimulus artifacts. To this end, Fig. 4 shows the MSE comparing the results from the human listeners' average CMM and the loudness-estimation procedure in response to 1-kHz tone bursts compared to measurements in a coupler using a Hanning window as a function of window delay. These results exhibit a relatively clear MSE minimum despite the fact that the coupler generated no TBOAEs. However, this minimum still has a higher absolute estimation error than any of the minima measured using 1-kHz tones in real ears. Additionally, the lowest values of the MSE result immediately after the stimulus offset. As the CMM function has been shifted to have a zero mean, the MSE is then essentially just a sum of the squares of the logarithmic values of string length after the shift. This pattern is in great contrast with the results measured in the real ears, which show a relatively wide range of stable delays and lower MSE (see Fig. 2).

Identical measurements at 4 kHz in the coupler, shown in Fig. 5, exhibit a long, but asymptotic, decrease in MSE as the ringing diminishes (clearly, a much longer ringing decay time than for the 1-kHz stimulus). The optimal case results in a MSE much higher than the 1-kHz real-ear response (Fig. 1), but results that are not much worse than the 4-kHz real-ear responses (see Fig. 3). Again, this supports the idea that the 4-kHz measurements fail to capture the non-linear response, but that the 1-kHz non-linear response cannot be explained by simple resonance or linear reflections from the ear canal.

E. Comparison between TBOAE loudness estimates and CMM loudness measurements

In Figs. 6–8, functions were allowed to vary by a single parameter in order to adjust the location of the curve on the plot to have to zero-mean. Figure 6 shows estimates of loudness in response to 1-kHz tone bursts as a function of level derived from TBOAEs and CMM. Several loudness estimates are plotted, each for a different window delay to demonstrate the robustness of the estimation procedure and to show the contrast of the zero-delay stimulus response. There is excellent agreement between the TBOAE loudness estimates and the direct measure of loudness, except at the low-

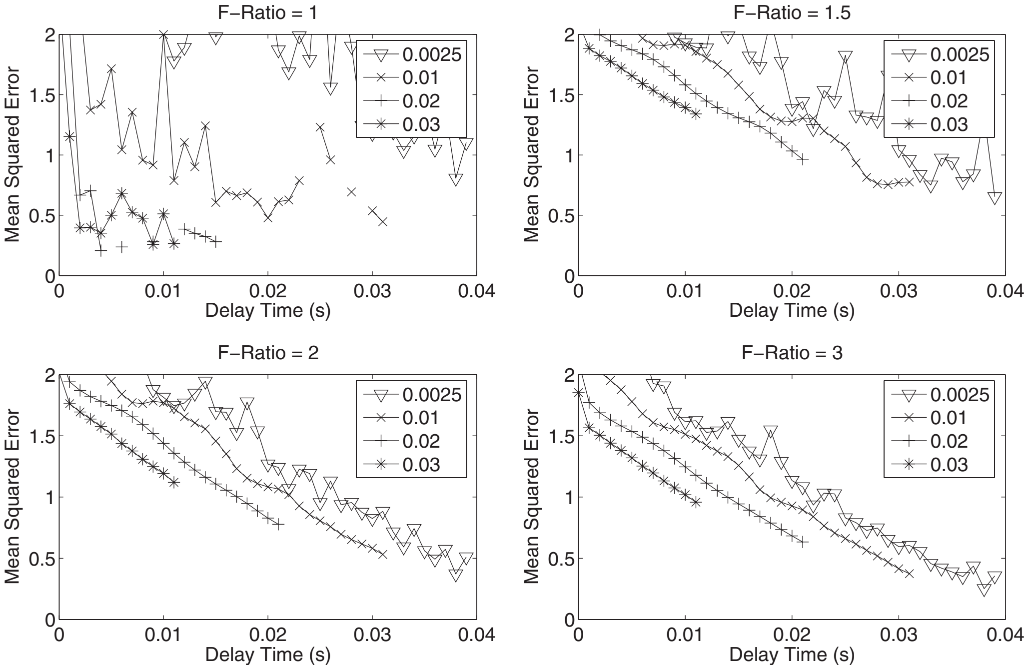

FIG. 5. MSE between loudness measured using CMM and loudness estimated from the response of the coupler to 4-kHz tone bursts using a Hanning analysis window. The plots show AMSE as a function of window delay, with each curve representing different window sizes and each plot representing a different *F*-ratio.

est levels. This is consistent with loss of steepness at low levels resulting from edge effects in the CMM procedure (Hellman and Meiselman, 1988). It is noteworthy that the analysis results in a nearly linear response when the window delay is set to zero and the stimulus itself is included in the analysis.

Figure 7 is identical to Fig. 6 for the 4-kHz tone bursts. The estimates derived from TBOAEs at 4 kHz are all far too steep and serve as very poor fits to the CMM function. In fact, they are more likely to match the linear response obtained when the window delay is set to zero. This is due to the difficulty in avoiding ear-canal resonances. The resulting OAEs contained too great a linear component to remove by simple windowing and averaging. As these results match those seen in the coupler, they are unlikely to arise from actual cochlear processing. In contrast, the OAEs in response to the 1-kHz tone bursts differed dramatically from the coupler recordings, indicating that these measurements did indeed result from cochlear processes.

Figure 8 shows a summary of all of the average functions in optimal parametric conditions. It is particularly noteworthy that the CMM functions at the two frequencies correspond quite closely, indicating that the psychoacoustical loudness functions are consistent across frequency. The INEX and the loudness function derived from TBOAEs at 1 kHz show excellent correspondence across the entire range.

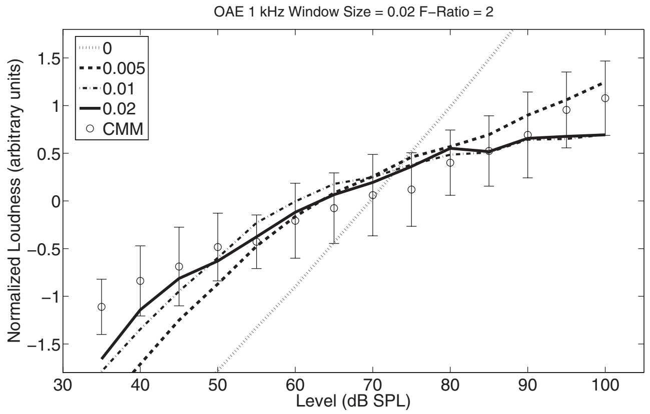

FIG. 6. Loudness resulting from 1-kHz tone bursts as a function of level for estimates derived from TBOAEs and for direct loudness measurements using CMM. Several loudness estimates for the TBOAEs are plotted, each for a different window delay. This analysis was performed with a 20-ms Hanning window and an *F*-ratio of 2. Error bars show ±1 standard deviation.

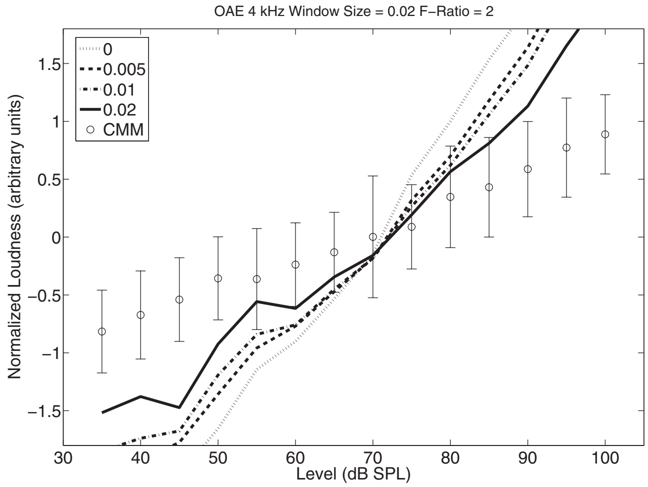

FIG. 7. Loudness resulting from 4-kHz tone bursts as a function of level for estimates derived from TBOAEs and for direct loudness measurements using CMM. Several loudness estimates for the TBOAEs are plotted, each for a different window delay. This analysis was performed with a 20-ms Hanning window and an *F*-ratio of 2. Error bars show ±1 standard deviation.

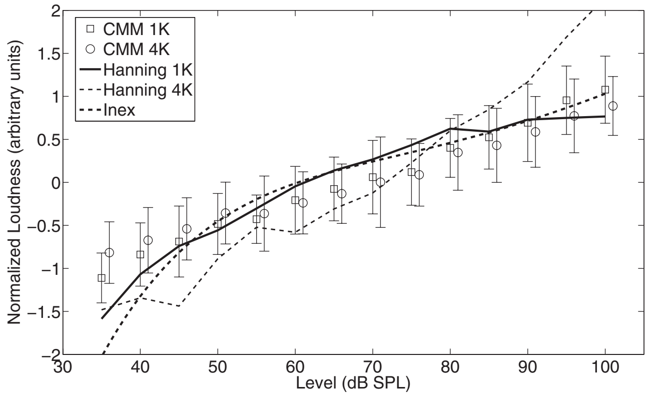

FIG. 8. Loudness as a function of level derived from CMM and Hanning-windowed TBOAEs in response to 1- and 4-kHz stimuli (The 4-kHz CMM data points are offset by 1 dB so that the error bars are visible). In addition, the INEX model loudness function is plotted for comparison with loudness measurements in the literature. Error bars show ±1 standard deviation.

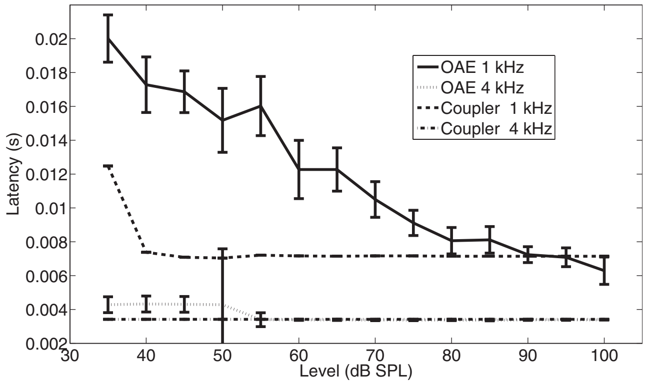

FIG. 9. TBOAE latency estimates as a function of level for both the real ears and coupler in response to 1- and 4-kHz stimuli. Error bars show ±1 standard deviation.

This indicates that TBOAE functions may be closely related to loudness measured using longer tones for an average of a group of listeners.

F. Latency/level relationship

Figure 9 shows the estimate of latency as a function of level for both the real ears and coupler at 1 and 4 kHz. For OAE measurements in real ears at 1 kHz, the results show a relatively smooth decrease in latency as a function of level, ranging from about 20 ms at 35 dB SPL to 6.3 ms at 100 dB SPL. Although this result differs somewhat from other reports of TBOAE latency (Hoth and Weber, 2001; Brass and Kemp, 1991; Schairer *et al.*, 2006; Sisto and Moleti, 2007), it is not surprising, since the methodologies for estimating OAE latency, the window sizes and shapes, and the definition of latency differ substantially among studies.

The present real-ear results are in agreement with the fact that the latency measured in the coupler at 1 kHz remains constant as a function of level. This indicates that the coupler recordings show a peak response at the biggest peak of the trivially reflected waveform, observed here to be close to 7 ms. In contrast, the acoustic ringing at 4 kHz was significant enough to render the procedure invariant to the initial delay chosen (i.e., the maximum would always occur at the beginning of the selected time region in both the real ears and the coupler). In this case, the latency was estimated at approximately 3.4 ms for both the real ears and the coupler. Even when the start of the observation window was delayed by several more milliseconds, the real ears and coupler produced identical, flat latency curves with latencies adjusted to match the start time of the window.

G. Analysis

Additional numerical analyses were performed by first assuming a "correct" loudness function and then determining the error (Model−Data) between the model and each of the other functions. Two analyses were performed: in one the correct model was the INEX function, and in the other the correct model was the CMM data at 1 kHz. The variability accounted for by the model function is then quantified by

$$1-[\mathrm{Variance}(\mathrm{Model}-\mathrm{Data})/\{\mathrm{Variance}(\mathrm{Model}-\mathrm{Data})$$
$$+\mathrm{Variance}(\mathrm{Model})\}].$$

This method is more useful than a simple correlation analysis because it does not heavily weight trends that are simply in the same linear direction, but rather looks at whether the functions deviate from each other. Table I shows the variability accounted for by each data set given a specific model. It is noteworthy that the INEX resulted from a relatively large quantity of data, but none of the data or listeners in the present study.

When the CMM $1k$ function was used as the ideal function, the CMM $4k$ function was found to correspond closely. This indicates that the average loudness functions at these different frequencies are consistent. Other measurements of loudness have shown similar results (e.g., Hellman, 1976). Additionally, the Hanning-windowed 1 kHz OAE input/output function also closely matched the CMM $1k$ function. This supports the idea that TBOAE measurements at 1 kHz are related to loudness. This idea is further supported by the second comparison, performed with the INEX model as the "ideal" function. If the INEX is assumed to be a good average loudness function, CMM is then an acceptable, but not ideal measure of loudness based on the variance accounted for by the CMM functions when using the INEX as the

TABLE I. Variance accounted for by using the CMM 1-kHz data and INEX function as models of the results of CMM and TBOAEs.

Model		CMM $1k$	CMM $4k$	INEX	OAE Hann. $1k$	OAE Hann. $4k$	OAE Rect. $1k$	OAE Rect. $4k$
CMM $1k$	Variance accounted	⋯	0.96	0.82	0.89	0.63	0.80	0.45
INEX	Variance accounted	0.88	0.76	⋯	0.96	0.72	0.95	0.60

model. In fact, the 1-kHz TBOAE loudness models matched the INEX function better than CMM, despite the fact that the analysis parameters were optimized to fit the CMM function. This indicates that TBOAEs may actually be a better measure of loudness than CMM. As expected, the 4-kHz TBOAE functions did not serve as very good models of either ideal function.

In order to examine potential viability for use as an individual metric for the assessment of loudness growth, the same analysis was performed comparing individual CMM at 1 kHz and loudness estimated using Hanning-window analysis on 1-kHz TBOAEs. The group average variance accounted for in this condition was 0.89 in the earlier analysis. The mean of the individual accounted variances was 0.80 with a standard deviation of 0.11. Of the six listeners, two estimates matched relatively poorly with variances accounted of 0.62 and 0.68. In both cases, the individuals for whom the variance accounted was lower had particularly shallow CMM slopes indicating that much of the problem may have arisen as a result of unreliable CMM data rather than unreliable OAE data.

The remaining listeners had variances accounted of 0.83, 0.86, 0.89, and 0.90. This indicates that, at least for two-thirds of the listeners, the two methods resulted in roughly the same estimates of loudness.

IV. CONCLUSIONS AND FUTURE WORK

These results indicate that there is significant potential for using tone-burst otoacoustic emissions for estimating loudness in response to 1-kHz stimuli across a wide range of levels in normal-hearing listeners. However, at frequencies near to ear-canal resonance, a simple reflection analysis procedure may not be sufficient to eliminate acoustic artifacts for loudness estimation. Tone-burst OAEs seem to be free from the complex interactions that occur with high-level DPOAEs, but the results of TBOAE measurements may not be as robust to ear-canal resonance as DPOAEs. If loudness functions could be estimated rapidly from TBOAEs at all audiometric frequencies in hearing-impaired listeners, this would provide an easy means for performing more advanced hearing-aid fitting using not only threshold but also loudness-growth functions. However, it remains to be seen if this measurement can be made across a wide enough range of frequencies and a separate strategy would likely need to be used for frequencies near the ear-canal resonances. More work still needs to be done to examine the frequency-specific applicability of TBOAEs for measuring loudness growth within the full range of audiometric frequencies. If the type of response measured at 1 kHz can be extended throughout the full frequency range, TBOAEs could be very useful clinically for performing improved hearing-aid fitting.

ACKNOWLEDGMENTS

The authors wish to thank Mary Florentine for helpful comments on an earlier version of this manuscript. The editor and two anonymous reviewers also provided helpful comments. This research was supported by The Capita Foundation.

ANSI (**1996**). "American National Standard Specification for Audiometers," ANSI S3.6-1996.

Brass, D., and Kemp, D. T. (**1991**). "Time-domain observation of otoacoustic emissions during constant tone stimulation," J. Acoust. Soc. Am. **90**, 2415–2427.

Buus, S., and Florentine, M. (**2001**). "Modifications to the power function for loudness," in *Fechner Day 2001*, edited by E. Sommerfeld, R. Kompass, and T. Lachmann (Pabst, Berlin).

Buus, S., Florentine, M., and Müsch, H. (**1998**). "Loudness function for tones at low levels derived from loudness summation," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Buus, S., Florentine, M., and Poulsen, T. (**1997**). "Temporal integration of loudness, loudness discrimination, and the form of the loudness function," J. Acoust. Soc. Am. **101**, 669–680.

Buus, S., Obeling, L., and Florentine, M. (**2001**). "Can basilar-membrane compression characteristics be determined from distortion-product otoacoustic-emission input-output functions in humans?," in *Physiological and Psychophysical Bases of Auditory Function*, edited by D. J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht, Netherlands).

Elberling, C., and Wahlgreen, O. (**1985**). "Estimation of auditory brainstem response, ABR, by means of Bayesian inference," Scand. Audiol. **14**, 89–96.

Epstein, M., Buus, S., and Florentine, M. (**2004**). "The effects of window delay, delinearization, and frequency on tone-burst otoacoustic emission input/output measurements," J. Acoust. Soc. Am. **116**, 1160–1167.

Epstein, M., and Florentine, M. (**2005a**). "Inferring basilar-membrane motion from tone-burst otoacoustic emissions and psychoacoustic measurements," J. Acoust. Soc. Am. **117**, 263–274.

Epstein, M., and Florentine, M. (**2005b**). "A test of the equal-loudness-ratio hypothesis using cross-modality matching functions," J. Acoust. Soc. Am. **118**, 907–913.

Epstein, M., and Florentine, M. (**2006**). "Loudness of brief tones measured by magnitude estimation and loudness matching," J. Acoust. Soc. Am. **119**, 1943–1945.

Florentine, M., Buus, S., and Poulsen, T. (**1996**). "Temporal integration of loudness as a function of level," J. Acoust. Soc. Am. **99**, 1633–1644.

Florentine, M., and Epstein, M. (**2006**). "To honor Stevens and repeal his law (for the auditory system)," in *Fechner Day*, edited by D. E. Kornbrot, R. M. Msetfi, and A. W. MacRae (International Society for Psychophysics, St. Albans, England).

Florentine, M., Epstein, M., and Buus, S. (**2001**). "Loudness functions for long and short tones," in *Fechner Day 2001*, edited by E. Sommerfeld, R. Kompass, and T. Lachmann (Pabst, Berlin).

Hall, J. W. (**2007**). *Handbook of Auditory Evoked Responses* (Allyn & Bacon, Boston, MA).

Harris, F. J. (**1978**). "On the use of windows for harmonic analysis with the discrete Fourier transform," Proc. IEEE **66**, 51–84.

Hellman, R. P. (**1976**). "Growth of loudness at 1000 and 3000 Hz," J. Acoust. Soc. Am. **60**, 672–679.

Hellman, R. P. (**1999**). "Cross-modality matching: A tool for measuring loudness in sensorineural impairment," Ear Hear. **20**, 193–213.

Hellman, R. P., and Meiselman, C. H. (**1988**). "Prediction of individual loudness exponents from cross-modality matching," J. Speech Hear. Res. **31**, 605–615.

Hellman, R. P., and Zwislocki, J. J. (**1961**). "Some factors affecting the estimation of loudness," J. Acoust. Soc. Am. **33**, 687–694.

Hellman, R. P., and Zwislocki, J. J. (**1963**). "Monaural loudness function at 1000 cps and interaural summation," J. Acoust. Soc. Am. **35**, 856–865.

Hoth, S., and Weber, F. N. (**2001**). "The latency of evoked otoacoustic emissions: Its relation to hearing loss and auditory evoked potentials," Scand. Audiol. **30**, 173–183.

Janssen, T., Niedermeyer, H. P., and Arnold, W. (**2006**). "Diagnostics of the cochlear amplifier by means of distortion product otoacoustic emissions," ORL **68**, 334–339.

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.

Muller, 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. (**1997**). "Distortion product and loudness growth in an active, nonlinear model of cochlear mechanics," in Symposium on Recent Developments in Auditory Mechanics, Sendai, Japan.

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.

Norton, S. J., and Neely, S. T. (**1987**). "Tone-burst-evoked otoacoustic emissions from normal-hearing subjects," J. Acoust. Soc. Am. **81**, 1860–1872.

Ravazzani, P., and Grandori, F. (**1993**). "Evoked otoacoustic emissions: Nonlinearities and response interpretation," IEEE Trans. Biomed. Eng. **40**, 500–504.

Ravazzani, P., Tognola, G., and Grandori, F. (**1996**). "'Derived nonlinear' versus 'linear' click-evoked otoacoustic emissions," Audiology **35**, 73–86.

Robinson, D. W. (**1957**). "The subjective loudness scale," Acustica **7**, 217–233.

Schairer, K. S., Ellison, J. C., Fitzpatrick, D., and Keefe, D. H. (**2006**). "Use of stimulus-frequency otoacoustic emission latency and level to investigate cochlear mechanics in human ears," J. Acoust. Soc. Am. **120**, 901–914.

Schlauch, R. S., DiGiovanni, J. J., and Ries, D. T. (**1998**). "Basilar membrane nonlinearity and loudness," J. Acoust. Soc. Am. **103**, 2010–2020.

Serpanos, Y. C., and Gravel, J. S. (**2000**). "Assessing growth of loudness in children by cross-modality matching," J. Am. Acad. Audiol. **11**, 190–202.

Serpanos, Y. C., and Gravel, J. S. (**2004**). "Revisiting loudness measures in children using a computer method of cross-modality matching (CMM)," J. Am. Acad. Audiol. **15**, 486–497.

Serpanos, Y. C., O'Malley, H., and Gravel, J. S. (**1998**). "Cross-modality matching and the loudness growth function for click stimuli," J. Acoust. Soc. Am. **103**, 1022–1032.

Shera, C. A., Guinan, J. J., Jr., and Oxenham, A. J. (**2002**). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. U.S.A. **99**, 3318–3323.

Sisto, R., and Moleti, A. (**2007**). "Transient evoked otoacoustic emission latency and cochlear tuning at different stimulus levels," J. Acoust. Soc. Am. **122**, 2183–2190.

Stevens, J. C., and Guirao, M. (**1964**). "Individual loudness functions," J. Acoust. Soc. Am. **36**, 2210–2213.

Stevens, S. S. (**1936**). "A scale for the measurement of a psychological magnitude: Loudness," Psychol. Rev. **43**, 405–416.

Stevens, S. S. (**1955**). "The measurement of loudness," J. Acoust. Soc. Am. **27**, 815–827.

Stevens, S. S. (**1957**). "Concerning the form of the loudness function," J. Acoust. Soc. Am. **29**, 603–606.

Stevens, S. S. (**1961**). "To honor Fechner and repeal his law: A power function, not a log function, describes the operating characteristic of a sensory system," Science **133**, 80–86.

Stevens, S. S. (**1972**). "Perceived level of noise by Mark VII and decibels (E)," J. Acoust. Soc. Am. **51**, 575–601.

Teghtsoonian, M., and Teghtsoonian, R. (**1983**). "Consistency of individual exponents in cross-modal matching," Percept. Psychophys. **33**, 203–214.

Thalmann, R. (**1965**). "Cross-modality matching in the study of abnormal loudness functions," Laryngoscope **75**, 1708–1726.

Whitehead, M. L., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (**1994**). "Measurement of otoacoustic emissions for hearing assessment," IEEE Eng. Med. Biol. Mag. **13**, 210–226.

Zweig, G., and Shera, C. A. (**1995**). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am. **98**, 2018–2047.

Zwicker, E., and Fastl, H. (**1999**). *Psychoacoustics: Facts and Models*, Springer Series in Information Sciences (Springer, New York), Vol. **22**.

Zwislocki, J. J. (**1965**). "Analysis of some auditory characteristics," in *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush, and E. Galanter (Wiley, New York).