Relation of distortion-product otoacoustic emission input-output functions to loudness

Daniel M. Rasetshwane,^{a)} Stephen T. Neely, Judy G. Kopun, and Michael P. Gorga
Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131

(Received 27 July 2012; revised 21 March 2013; accepted 26 April 2013)

The aim of this study is to further explore the relationship between distortion-product otoacoustic emission (DPOAE) measurements and categorical loudness scaling (CLS) measurements using multiple linear regression (MLR) analysis. Recently, Thorson *et al.* [J. Acoust. Soc. Am. **131**, 1282–1295 (2012)] obtained predictions of CLS loudness ratings from DPOAE input/output (I/O) functions using MLR analysis. The present study extends that work by (1) considering two different (and potentially improved) MLR models, one for predicting loudness rating at specified input level and the other for predicting the input level for each loudness category and (2) validating the new models' predictions using an independent set of data. Strong correlations were obtained between predicted and measured data during the validation process with overall root-mean-square errors in the range 10.43–16.78 dB for the prediction of CLS input level, supporting the view that DPOAE I/O measurements can predict CLS loudness ratings and input levels, and thus may be useful for fitting hearing aids.

© 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4807560]

PACS number(s): 43.64.Jb, 43.66.Cb [ELP]

Pages: 369–383

I. INTRODUCTION

Otoacoustic emission (OAE) input/output (I/O) functions and categorical loudness scaling (CLS) functions both typically become steeper when hearing loss is present (e.g., Neely *et al.*, 2003; Thorson *et al.*, 2012). The common mechanism for the steepening of both functions presumably is outer hair cell (OHC) dysfunction, despite the fact that more central factors may influence loudness judgments. Because loudness is a subjective measure that may be difficult to obtain in certain populations, there has been interest in estimating loudness from OAEs, an objective measurement that does not require a voluntary response from the patient and may require less time. In the current study, we extended the results of a previous study by further exploring the efficacy of estimating loudness functions from distortion-product OAE (DPOAE) I/O functions.

OAEs are acoustic signals that originate within the cochlea as by-products of its normal signal-processing function and are apparently produced by the OHCs (e.g., Brownell, 1990). DPOAEs are evoked by simultaneous presentation of a pair of primary tones having slightly different frequencies. The level of the DPOAEs depends on the status of the cochlea. Hearing impairment, especially when it is associated with OHC damage or dysfunction, results in a reduction or absence of DPOAEs, and OAEs in general. This reduction in DPOAE level is usually accompanied by elevation of audiometric thresholds (e.g., Boege and Janssen, 2002; Gorga *et al.*, 2003) and loss of compression. In a healthy cochlea, DPOAEs grow approximately linearly with level at low levels, but growth becomes more compressive as levels increase (e.g., Neely *et al.*, 2003). This normal compressive characteristic manifests as a DPOAE I/O function with slope that is steeper at

low levels compared to high levels, when the primary levels are selected using the paradigm of Kummer *et al.* (1998). With hearing loss, the DPOAE I/O function becomes steeper, less compressive (more linear) and has a reduced range over which compression takes place (e.g., Dorn *et al.*, 2001). The relation between degree of hearing loss, as measured using audiometric thresholds, and DPOAE level has been demonstrated in several studies (e.g., Boege and Janssen, 2002; Gorga *et al.*, 2003; Oswald and Janssen, 2003; Rogers *et al.*, 2010). As a consequence of the relationship between DPOAEs and hearing loss, DPOAEs have been used as objective non-invasive probes of cochlear status (e.g., Gorga *et al.*, 1997), and are currently used in many newborn-hearing-screening programs.

Loudness is defined as a listener's subjective response to the intensity of a sound (Scharf, 1978). Several psychoacoustic procedures have been developed to measure the loudness of a sound. In one of these procedures, called categorical loudness scaling (CLS), signals of different intensities are presented to subjects who are asked to rate the loudness of the signals using meaningful labels, such as "very soft," "medium," and "very loud" (e.g., Allen *et al.*, 1990; Brand and Hohmann, 2002; Al-Salim *et al.*, 2010). The CLS procedure is attractive compared to other measures of loudness because (1) it relates more to a listener's experience and informal descriptions of their loudness percepts, (2) it requires minimal subject training, and (3) it is fast (Allen *et al.*, 1990; Brand and Hohmann, 2002; Al-Salim *et al.*, 2010).

Loudness is mediated by the peripheral and central auditory systems. Peripheral measures of cochlear compression based on OAE have been shown to be correlated with growth of loudness (Neely *et al.*, 2003; Muller and Janssen, 2004; Epstein and Florentine, 2005; Epstein and Silva, 2009; Thorson *et al.*, 2012). Neely *et al.* (2003) showed that the Fletcher and Munson (1933) log-loudness function and DPOAE I/O can both be described by similar log functions.

^{a)}Author to whom correspondence should be addressed. Electronic mail: daniel.rasetshwane@boystown.org

Muller and Janssen (2004) demonstrated the similarity of estimates of hearing-aid gain based on DPOAE I/O functions to gain based on CLS functions. Epstein and Silva (2009) obtained high correlations between estimates of loudness from toneburst OAE (TBOAE) and direct loudness measurements at 1 kHz. Thorson *et al.* (2012) obtained promising results using a multiple linear regression (MLR) model to predict CLS functions from DPOAE I/O functions.

For purposes of simplicity, we will consider the central auditory system (CNS) to include those portions the auditory pathway above the level of the sensory cells. In the CNS, increases in sound intensity and loudness results in increases in firing rates of auditory-nerve fibers (e.g., Heinz *et al.*, 2005). For band-limited stimuli such as tones, increases in loudness may also result from an increase in the total number of excited neurons, i.e., there may be spread of excitation from neurons with primary sensitivity to the particular frequency to adjacent neurons (e.g., Epstein, 2011). While most hearing loss in humans is initiated with damage at the sensory-cell level (particularly the OHCs), impairments to the CNS, such as impaired connection between neurons and inner hair cells and dysfunction of the auditory neurons, can result in elevation of hearing thresholds (Smeds and Leijon, 2011). Although loudness is presumably affected by CNS impairments, DPOAEs are insensitive to these impairments. Therefore, the use of DPOAEs to predict loudness implicitly assumes that CNS impairments are negligible.

Although some studies have suggested that loudness is mainly mediated by central auditory processes (e.g., Relkin and Doucet, 1997; Heinz *et al.*, 2005), the correlation between OAEs and growth of loudness suggest that OAEs and loudness growth may be determined by the same underlying mechanism. Additionally, loudness has been shown to be correlated with audiometric thresholds (e.g., Launer *et al.*, 1996; Al-Salim *et al.*, 2010). Specifically, elevated thresholds are associated with loudness growth functions that have steeper slopes, compared to the slopes observed in ears with normal hearing. These relationships between (1) loudness and OAEs, (2) OAEs and hearing thresholds, and (3) loudness and hearing thresholds suggests that it may be possible to predict loudness from OAEs, especially in the case where hearing loss can be attributed mostly to dysfunction of the outer hair cells, as is the case with most hearing loss in humans, because OHCs amplify both DPOAEs and the signal delivered to the auditory nerve by a similar amount. Loudness predicted from OAEs may be used instead of CLS for fitting and adjustment of hearing aids that attempts to compensate for loss of compression (as has been proposed in several studies, e.g., Cox, 1995; Allen, 1997; Al-Salim *et al.*, 2010). Using loudness may be especially useful when fitting hearing aids to individuals who are unable to describe their percepts of loudness such as infants, young children and patients with developmental delays.

The current study expanded on the study of Thorson *et al.* (2012) by further exploring the efficacy of predicting loudness functions from DPOAE I/O functions. In particular, we extend the study of Thorson *et al.* (1) by designing a more robust MLR model, (2) by designing two models, one for prediction of the CLS function itself, and the other for

prediction of the input level required for a particular loudness perception or category, and (3) by including a validation that uses independent data to evaluate whether the prediction of loudness can generalize to other data. In the present study, the data from Thorson *et al.* were used to train the new MLR models, and then an independent set of data was used to validate the predictions of the models. Validation using an independent set of data is important and clinically relevant because if the models are to be applied in the clinic, the models must be able to provide predictions of a CLS function for a particular patient with knowledge of only their DPOAE I/O function. A more robust model (compared to the one described in Thorson *et al.*, 2012) was developed by including a constant term in the MLR coefficients. Two different predictions were made using the new MLR models. The first analysis evaluated the predictability of a CLS function by predicting loudness rating as function of input level specified in sound pressure level (SPL). In the second analysis, a different MLR model was applied to predict the input level required for a particular loudness perception or category. The second analysis may be especially useful when determining the input level required for a particular loudness perception (e.g., “very soft,” “very loud”), which has relevance during a hearing-aid fitting task. In addition to our desire to improve upon the model used by Thorson *et al.*, we have elected to use MLR analyses because previous research has shown that they improve both the accuracy of determining the correlation between OAEs and audiometric thresholds and of determining auditory status (e.g., Dorn *et al.*, 1999; Gorga *et al.*, 1999; 2005; Kirby *et al.*, 2011; Hussain *et al.*, 1998; Vinck *et al.*, 1998).

II. METHODS

A. Subjects

A total of 120 subjects, aged 11–76 yr, participated in this study. A subset of 74 subjects constituted the training set for this study. Sixteen of these subjects had normal hearing and the remaining 58 had hearing loss. CLS data for these subjects were previously analyzed and described by Al-Salim *et al.* (2010) and Thorson *et al.* (2012). Their DPOAE data were previously analyzed and described by Thorson *et al.* (2012). An additional 46 subjects constituted the validation set for this study. Sixteen of these subjects had normal hearing and 30 had hearing loss. Data from these subjects were not used during the development of the MLR models. For the training set, data were collected during two sessions separated by as little as one week to as much as six months. Data were collected during a single session for the validation set.

All measurements were made monaurally. Pure-tone air-conduction thresholds were measured at octave and inter-octave frequencies from 0.25 to 8 kHz, while pure-tone bone-conduction thresholds were measured at octave frequencies from 0.25 to 4 kHz. Subjects with air-conduction thresholds ≤ 15 dB hearing loss (HL) at all frequencies were considered normal hearing. Subjects were categorized as hearing impaired when they had thresholds >15 dB HL at one of the three test frequencies (1, 2, or 4 kHz) at which CLS and DPOAE data were collected. Hearing loss did not

exceed 75 dB HL at the test frequency for any subject. Subjects were excluded from the study if air-bone gaps were greater than 10 dB at any frequency. Middle-ear function was assessed using 226-Hz tympanometry prior to each session in which DPOAE measurements were made. Subjects were included if static compliance was between 0.3 to 2.5 cm³ and middle-ear pressure ranged from +25 to −75 daPa. Subjects not meeting these criteria were excluded from the study. The study described in this article was conducted under an approved Institutional Review Board protocol.

B. DPOAE measurements

DPOAE measurements were made using custom-designed software (EMAV version 2.89, Neely and Liu, 1994). The measurement system included a 24-bit soundcard (CardDeluxe; Digital Audio Labs, Chanhassen, MN) and a DPOAE probe-microphone system (ER-10 C, Etymotic Research, Elk Grove Village, IL). Separate channels of the soundcard were used to generate two primary tones (f_1 and f_2), which were sent to separate sound sources that were housed in the probe-microphone system. The probe-microphone system was coupled to the ear using a foam tip. The ER-10 C had been modified to remove the built-in 20 dB of attenuation in order to present stimuli up to 80 dB SPL during DPOAE measurements.

DPOAE measurements were made in both normal-hearing (NH) and hearing-impaired (HI) subjects at three f_2 frequencies of 1, 2, and 4 kHz, and L_2 levels ranging from −20 to 80 dB SPL in 5 dB steps. The lowest L_2 used for a particular subject was equal to 15 dB below that subject's audiometric threshold at the f_2 frequency. L_1 for a given f_2 and L_2 was set according to the following equation (Johnson *et al.*, 2006a):

$$L_1 = 80 + 0.137 \log_2(18/f_2)(L_2 - 80). \quad (1)$$

The f_2/f_1 ratio was set using according to Eq. (2) (Johnson *et al.*, 2006a),

$$f_2/f_1 = 1.22 + \log_2(9.6/f_2)(L_2/415)^2. \quad (2)$$

Equations (1) and (2) result in the largest DPOAE levels (based on group averages) in normal-hearing humans (Neely *et al.*, 2005; Johnson *et al.*, 2006a).

DPOAE data were collected into two separate buffers and the level of the DPOAE, L_d , was obtained by summing the contents of the two buffers in the $2f_1$ - f_2 frequency bin. The level of the noise, L_n , was estimated by subtracting the contents of the two buffers and then averaging the level in the $2f_1$ - f_2 frequency bin along with the level in the five bins on either side of the $2f_1$ - f_2 frequency bin. The buffer length was 8192 and the sampling rate was 32 kHz, giving a frequency resolution of 3.9 Hz.

Measurement-based stopping rules were used during data collection. Data collection was stopped if one of the following conditions were met: (1) the noise floor was less than −25 dB SPL (i.e., $L_n \leq -25$ dB SPL), (2) artifact-free averaging time exceeded 64 s, or (3) the signal-to-noise ratio (SNR) was greater than 60 dB. The primary criterion was

noise-floor and was chosen as a conservative estimate of the level at which system distortion occurred. For DPOAE I/O functions to be considered for further analysis, they had to have three or more consecutive points for which the L_d level was greater than −25 dB SPL and had an $\text{SNR} \geq 3$ dB. For further details on DPOAE measurements, see Thorson *et al.* (2012).

C. CLS measurements

CLS measurements were made in the same group of subjects. The CLS procedure consisted of a response scale with 11 loudness categories. A subset of these loudness categories were assigned labels such as “cannot hear,” “very soft,” “soft,” “medium,” “loud,” “very loud,” and “too loud.” The procedure determined the input level of pure tones that correspond to the loudness categories. The loudness categories were presented on a computer touch screen as horizontal bars with increasing width from the softest level to the loudest level. Subjects were asked to rate the loudness of pure tones using one of the 11 loudness categories by pressing a button that best described the loudness perception. The loudness categories were assigned categorical units (CUs) of 0 to 50 in steps of 5, for numerical representation of loudness categories.

The CLS measurements were made at 1, 2, and 4 kHz (the same frequencies used for the DPOAE measurements) and repeated three times for each frequency during a test session. As a result, for a given frequency, subjects who participated in the training stage of the study had a total of six CLS measurements (three per session $\times$ two sessions), while subjects who participated in the validation stage of the study had a total of three CLS measurements (three measurements in a single session). The stimuli were 1 s in duration with rise/fall times of 20 ms.

Stimulus presentations were separated by an interval of 2 s. The pure tones were presented at varying intensity levels using the same sound card used for the DPOAE measurements, but this time stimulus delivery was controlled using different custom-designed software (BEHAVIORAL AUDITORY RESEARCH TESTS, version 2.3.32). The stimuli were routed via a headphone buffer (HB7; Tucker-Davis Technologies, Alachua, FL) to ER-3 A sound sources (Etymotic Research, Elk Grove Village, IL) connected to the sound ports of an ER-10B+ probe-microphone system (Etymotic Research, Elk Grove Village, IL). The combination of the headphone buffer and ER-3 A sound sources allowed for production of high stimulus levels up to a maximum of 105 dB SPL. The probe-microphone system was coupled to the ear using a foam tip. In-the-ear calibrations, using the same ER-10B+ probe-microphone system used for the data collection, were performed at each frequency before and after each measurement. For further details on the measurement procedure, see Al-Salim *et al.* (2010) and Thorson *et al.* (2012).

CLS data from the three trials (and the two visits for subjects who participated in the training study) were analyzed to remove outliers and to obtain a mean input level that corresponded to each CU. Two definitions of outliers at different stages of the analysis were used. In the first stage,

data from the three trials were considered together and then for each CU, data that deviated by more than 12 dB from the median input level for the three trials were considered outliers and removed from further analysis. The remaining data points were averaged to obtain a mean input level for a given CU. The second definition of outliers was intended to make the CLS function monotonic because the perceptual correlate of an increase in stimulus intensity is always expected to be an increase in loudness. To achieve this, the change in input level, when going from low to high CUs, between two successive CUs was required to be a minimum of 1 dB. Any data points that did not meet this requirement were considered outliers and excluded. The first definition of outliers removed 14% of the total data and the second definition removed 34% of the total data.

CLS data for some subjects can have missing CUs (and hence input levels) because either the subject did not use a particular CU in his/her loudness judgments or the data were considered outliers and removed. To have CLS data that includes all CUs and that is also smooth, CLS data were interpolated prior to use in the determination of MLR coefficients. Different interpolation rules were used depending on whether the analysis was the prediction of CLS-CU or the prediction of CLS-SPL, and they are each described with the corresponding prediction. Note, however, that the performance of the model predictions were based on comparison of predicted data to non-interpolated data.

D. Prediction of CLS-CU

DPOAE I/O functions were used to predict CLS-CU for each input level. Separate predictions were performed at each of the three frequencies –1, 2, and 4 kHz. DPOAE I/O functions can have missing L_d values at low L_2 levels for some subjects, especially hearing-impaired subjects when data-collection was stopped because one of the three measurement-based stopping rules was satisfied. However, the MLR model used for prediction required DPOAE I/O data to be specified at all L_2 for all subjects. To satisfy this requirement, missing L_d were set to the noise floor, i.e., –25 dB SPL. Although the DPOAE measurements were made at L_2 levels of –20–80 dB SPL, only responses to levels from –5 to 80 dB SPL were used for the predictions. Responses to lower levels are useful for determining whether there is a biological response in the measurement, and for determining the noise floor. CLS data were interpolated prior to the prediction so that the data would include all input levels. To interpolate the CLS function, it was fit with two linear functions, one for the $0 < \text{CU} < 25$ segment and the other for the $25 \leq \text{CU} < 50$ segment of the function. We will refer to the segments $0 < \text{CU} < 25$ and $25 \leq \text{CU} < 50$ as the *soft* CU and *loud* CU segments, respectively (Al-Salim *et al.*, 2010). We selected a transition from soft to loud CU of 25 CUs because this CU corresponds to “medium” loudness and divides the loudness scale used in this study into two equal segments. The lines fit to the CLS function were dependent on the input level with the input level ranging from 0 to 105 dB SPL in 5 dB steps. This implies that the interpolated CLS function, which we will refer to as the *target CLS*

function or simply *target data*, can have CUs that are not multiples of 5 (the measured data only has CU that are multiples of 5). The slope of the line fit to the loud CU segment was limited to not exceed 3 CU/dB in order to avoid having target data with infinite (or near infinite) slope. If it exceeded this limit, a line with a slope of 3 CU/dB and a CU offset based on the data available in the loud CU segment of the function was fit to the data points. This occurred for 11% of the total data. In the case where a subject did not have data in the loud CU segment of the CLS function, a line with a slope of 3 CU/dB that passes through $\text{SPL}_{\text{max}} = 105$ dB SPL was used. This occurred for 30% of the total data. Limiting the slope reduces the variability in the loud segment of the target CLS function used as input to the MLR model. However, the performance of the MLR model is based on comparison of predicted data to actual data, i.e., data before interpolation and application of the slope limit. Additionally, the limit of 3 CU/dB for the loud CU slope is a reasonable limit since it is greater than the slope of a typical CLS function in the loud or high CU segment (e.g., Brand and Hohmann, 2002 obtained a median slope of 1.7 dB/CU for the loud CU segment). In the transition region between CUs of 20 and 25, the larger (greater CU value) of the two line segments fit to the data was used as the target. Target CUs with values less than zero were set equal to zero.

The CLS target data, together with the DPOAE I/O data, were used to obtain MLR coefficients that characterize the relationship between CLS-CU and DPOAE I/O for the training set. Let A be a $n \times (p + 1)$ matrix of DPOAE L_d ’s at a particular frequency and B be a $n \times q$ matrix of CLS CUs at the same frequency; that is,

$$A = \begin{bmatrix} 1 & a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ 1 & a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{bmatrix}, \quad (3)$$

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,q} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,q} \end{bmatrix}, \quad (4)$$

where $[a_{i,1}, a_{i,2}, \dots, a_{i,p}]$ is the DPOAE I/O function for the i th subject and $[b_{i,1}, b_{i,2}, \dots, b_{i,q}]$ is the CLS functions for the i th subject. n is the number of subjects, with $n = 74$ for the training stage and $n = 46$ for the validation stage. $p = 19$ is the number of DPOAE L_2 ’s and $q = 22$ is the number of input levels for CLS. The MLR model relating B to A is

$$B = AX, \quad (5)$$

where X is a $(p + 1) \times q$ matrix of MLR coefficients. Equation (5) expresses the CLS-CUs (responses) as linear combinations of DPOAE L_d (observations). The column of ones in the matrix A results in a matrix of coefficients with intercept parameters (or constant terms), which provides a better fit to the data. Having an intercept term to better fit the data is one of the improvements made over the MLR model used by Thorson *et al.* (2012). Separate sets of MLR

coefficients were obtained at the three test frequencies. The MLR coefficients X in Eq. (5) were calculated by Gaussian elimination with partial pivoting.¹ Using the MLR coefficients, predicted CUs $\hat{B}$ were obtained from measurement of DPOAE I/O functions, A , as

$$\hat{B} = (AX)^T, \quad (6)$$

where T denotes matrix transposition. Separate sets of MLR coefficients were obtained for each test frequency. The same coefficients developed during the training stage were used during the validation stage without alteration.

E. Prediction of CLS-SPL

DPOAE I/O functions were also used to predict the input level required for a particular CU, i.e., CLS-SPL for each of the nine loudness categories from *very soft* (CU = 5) to *very loud* (CU = 45). CLS data were interpolated prior to the prediction so that the data would include all CUs. To interpolate the CLS function, it was fit with two lines, one for the soft CU segment and the other for the loud CU segment. For this analysis, however, the interpolated CLS function was such that the CUs were always multiples of 5 from 0 to 50, but the input level was a continuous value. The slope of the line fit to the loud CU segment was again limited to 3 CU/dB. If the slope exceeded this limit, a line with a slope of 3 CU/dB and a CU offset based on the data available in the loud CU segment was fit to the data points. In the case where a subject did not have data in the loud CU segment of the CLS function, a line with a slope of 3 CU/dB that passes through $SPL_{\max} = 115$ dB SPL was used. Predicted input levels greater than 120 dB SPL were set to 120 dB SPL. This is a reasonable limit since an input level of 120 dB SPL is higher than the upper limit of our sound-delivery system (105 dB SPL) and is also higher than the level that would be judged as too loud by both normal-hearing and hearing-impaired listeners (see, e.g., Al-Salim *et al.*, 2010; Brand and Hohmann, 2002). The target CLS function created for the prediction of CLS-SPL is consistent with the target CLS function created for the prediction of CLS-CU. When the differences between the two functions were compared for all subjects and test frequencies in terms of rms difference in input level, 82% of the functions had an rms differences less than 2 dB.

The CLS target data and the DPOAE I/O data were used to obtain MLR coefficients that characterize the relationship between CLS-SPL and DPOAE I/O for all subjects in a manner similar to Sec. IID. In this case however, the matrix of CLS CUs B was $n \times 10$, resulting in a 19×10 matrix of MLR coefficients X . These coefficients were subsequently used to obtain predictions of CLS-SPL, again using the procedures described in Sec. IID. Separate sets of MLR coefficients were obtained for each test frequency.²

III. RESULTS

Examples of actual, target and predicted CLS functions for individual subjects for the training stage of the analysis are presented first. The examples are separated by hearing

status; normal hearing versus hearing impaired. These examples are followed by summary results for all the subjects. Separate results are presented for prediction of CLS-CU and for prediction of CLS-SPL. Recall that separate MLR models were created for prediction of CLS-CU and for prediction of CLS-SPL. These results are intended to give a sense of the performance of the training stage of the study. Finally, CLS predictions for the validation set are presented to validate the two MLR models. These results give a sense of the efficacy of using DPOAEs for prediction of loudness.

A. Prediction of CLS-CU

Examples of individual predictions of CLS-CU from DPOAE I/O for several subjects are shown in Fig. 1, in which measured (open circles), target (solid line), and predicted (dashed line) CLS functions are shown. The predictions were obtained by using the DPOAE I/O data for each subject as input to the MLR model. Examples for four normal-hearing (first two columns) and four hearing-impaired (last two columns) subjects are shown at each frequency (as indicated on the right Y axis for each cluster of four panels). From each hearing category (normal hearing versus hearing impaired), examples for subjects whose root mean square (rms) error was around the median rms error for the respective group and frequency are presented. Because the subjects were ranked by rms error separately for each frequency, the subjects whose data are shown at 1 kHz, for example, are not necessarily the same subjects whose data are shown at 2 or 4 kHz. Hearing thresholds and rms errors (signified by θ in the figure) are included as insets in each panel. Normal-hearing subjects have a lower CLS threshold (stimulus level at which CU become greater than zero³) compared to hearing-impaired subjects. CLS functions for both groups are characterized by two segments, a soft CU segment with a shallow slope and a loud CU segment with a steep slope. The soft CU slopes for normal-hearing subjects are shallower than the soft CU slopes for hearing-impaired subjects with mean and standard deviation (SD) of 0.24 ± 0.04 , 0.25 ± 0.05 , and 0.24 ± 0.05 CU/dB at 1, 2, and 4 kHz for normal-hearing subjects and mean $\pm$ SD = 0.34 ± 0.13 , 0.40 ± 0.14 , and 0.50 ± 0.21 CU/dB at 1, 2, and 4 kHz for hearing-impaired subjects. The target data (solid lines) provide excellent fits to the actual data (open circles) for all subjects. In general, the predicted CLS functions (dashed lines) are similar to the target CLS functions, and, consequently, provide a good prediction of the actual CLS function. In the examples of Fig. 1, individual rms errors are, in most cases, less than 5 CU (a single step size on the CLS function).

DPOAE I/O functions for subjects whose CLS functions are shown in Fig. 1 are shown in Fig. 2. Each panel in Fig. 2 presents data from the same subject whose data were provided in the corresponding panels of Fig. 1. The DPOAE I/O functions for HI subjects have L_d 's of -25 dB SPL at low L_2 's for reason explained earlier in Sec. II. In general, DPOAE I/O functions for normal-hearing subjects have a lower DPOAE threshold (defined as the L_2 at which $L_d > -22$ dB SPL, equivalent to $SNR > 3$ dB used for

FIG. 1. (Color online) Examples of prediction of CLS-CU using the training data set. Open circles are actual, solid lines are target, and dashed lines are predicted CLS data. The first two columns provide examples for normal hearing subjects and the last two columns provide examples for hearing impaired subjects. The two rows at the top show examples for 1 kHz, the two rows in the middle show examples for 2 kHz, and the two rows at the bottom show examples for 4 kHz. Audiometric thresholds in dB HL and estimates of rms error in CU are provided as insets in each panel.

accepting the DPOAE response as valid), a wider range of L_2 's over which L_d can be measured, and a higher L_d , compared to DPOAE I/O functions of hearing-impaired subjects. The expected steep slopes of the DPOAE I/O functions at low levels that then become shallower as level increases can be observed in normal-hearing subjects, especially at 1 and 2 kHz.

To summarize the results of the prediction of CLS-CU using the training set, Fig. 3 shows the predicted CLS-CU plotted against the measured CU, with data for a different frequency shown in each column. Data for all the subjects (top row) are presented, together with separate plots of data for normal-hearing subjects (second row) and for hearing-impaired subjects (bottom row). Simple linear regression (SLR) lines that characterize the relationship between the predicted and measured CUs for each condition are plotted as dashed lines. SLR analysis is used here to evaluate the success of the predictions, and should not be confused with the MLR model used for making the predictions. The correlation coefficients r and slopes for the SLR lines are also shown in each panel of Fig. 3 and in Table I. For the training set, the predicted CUs are correlated with the measured CUs, with correlation coefficients of about 0.90 and slopes ranging from 0.89 to 0.99, and with all the correlations having statistical significance ($p < 0.001$). Average rms errors between the predicted CUs and measured CUs are reported in Table II. The average rms errors are less than 5 CU, except at

1 kHz for normal-hearing subjects, which indicates that overall, the error made by the MLR-based prediction was not more than one loudness category.

B. Prediction of CLS-SPL

Examples of individual predictions of CLS-SPL from DPOAE I/O for several subjects are shown in Fig. 4. The symbols and line-styles, as well as the criterion for subject selection, are the same as those used in Fig. 1 (prediction of CU). Although the input to the prediction model was CUs and the output was predicted input level, the data in Fig. 4 have been plotted with input level on the x axis and CUs on the y axis, to maintain consistency with Fig. 1 and with a standard CLS function (e.g., Brand and Hohmann, 2002). The rms error (θ) here is in units of dB. Different rules were used to obtain the target CLS function from the actual data, as described in Secs. IID and IIE. As in Fig. 1, (1) normal-hearing subjects have a lower CLS threshold compared to hearing-impaired subjects, (2) CLS functions for both groups are characterized by two segments, a soft CU segment with a shallow slope and a loud CU segment with a steep slope, and (3) soft CU slopes for normal-hearing subjects (mean $\pm$ SD: 0.25 ± 0.04 , 0.26 ± 0.06 , and 0.25 ± 0.05 CU/dB at 1, 2, and 4 kHz) are shallower than soft CU slopes for hearing-impaired subjects (mean $\pm$ SD: 0.35 ± 0.13 , 0.42 ± 0.15 , and 0.53 ± 0.21 CU/dB). The target CLS functions (solid

FIG. 2. (Color online) DPOAE I/O functions for subjects whose CLS functions are shown in Fig. 1. A panel in Fig. 2 corresponds to the same panel in Fig. 1.

lines) provide an excellent fit to the actual CLS functions (open circles) and the predicted CLS functions (dashed lines) are similar to the target CLS functions, and, consequently, provide a good prediction of the actual CLS function.

As a summary of the results of the prediction of CLS-SPL using the training set, Fig. 5 shows the predicted input level for the CLS function plotted against the measured input level. Data for all the subjects (top row) are presented, together with separate plots of data for normal-hearing subjects (second row) and for hearing-impaired subjects (bottom row). SLR lines that characterize the relationship between the predicted and measured input level for each condition are plotted as dashed lines. The correlation coefficient and slope for each SLR are also shown in each panel of the figure and in Table III. For the training set, the predicted input levels are highly correlated with the measured input levels, with correlation coefficients ranging from 0.92 to 0.96 (all of which were statistically significant, $p < 0.001$) and slopes of the SLR line ranging from 0.87 to 0.96. Average rms errors between the predicted input level and measured input level are reported in Table IV. The average rms error is less than 10 dB in all cases, which is slightly more than the 7 or 8 dB standard deviation (SD) reported by Al-Salim *et al.* (2010) for repeated measures. Figure 6 shows the rms error between predicted and measured input level as a function of CU for the training set. At all three frequencies, there is a general trend of higher average rms errors at low

CUs compared to high CUs. The average rms errors for normal-hearing subjects are higher than those for hearing-impaired subjects.

C. Validation of predictions

The results described above were based on the training set. To provide validation for the predictions, MLR coefficients obtained using the training set were applied to an independent validation data set without any alterations. This analysis will not only provide validation but also determine the robustness of the MLR coefficients and the efficacy of predicting loudness from DPOAE data. The first of two validation studies applies MLR coefficients that were created for the purpose of predicting CUs for the training set to the validation data to predict CUs. In the second validation study, MLR coefficients that were created for predicting CLS-SPL for the training set are applied to the validation set for prediction of CLS-SPL.

Predicted CU plotted against the measured CU for the validation set are shown in Fig. 7. The top row shows data for all subjects, the second row shows data for normal-hearing subjects, and the bottom row shows data for hearing-impaired subjects. The correlation coefficients (presented as insets in the figure and in Table V) are lower and the average rms errors (Table II) between predicted and actual data are higher, compared to the training set (see Fig. 2). However,

FIG. 3. (Color online) Summary of prediction of CLS-CU from DPOAE I/O using the training data set. Different columns present data at different frequencies, as indicated by labels at the top. The open circles are the individual data and the dashed lines are the simple regression line relating measured and predicted CUs. Correlation coefficients (r) and slopes of simple regression lines are included as insets in the figure panels. p -values for the correlation coefficients (not shown) were all less than 0.001.

the correlations are 0.72 or higher and all are statistically significant ($p < 0.001$). Average rms errors for the validation set are in the range 6.85 to 9.10 CU. Compared to the training data, slopes of the SLR lines are lower at 1 kHz, but similar at 2 and 4 kHz.

Predicted CLS-SPL plotted against the measured CLS-SPL for the validation set is shown in Fig. 8, following the convention used in Figs. 5 and 7. Like the predicted CU results, the correlation coefficients (presented in the figure and Table VI) are lower and the average rms errors (see Table IV) between predicted and measured data are higher compared to the training set (see Fig. 4). However, the correlations are 0.80 or higher and all are statistically significant ($p < 0.001$). The average rms errors for the validation set are in the range 10.41 to 16.78 dB. The slopes of the SLR lines

did not show a systematic change between training and validation, with some slopes increasing and others decreasing. Figure 9 shows the rms error between predicted and measured input levels as a function of CU for the validation set. The trends observed during training (see Fig. 6) cannot be generalized to this figure; that is, there is no clear trend in the variation of the error with CU. The rms error at 4 kHz is greater than at 1 and 2 kHz. The rms errors in Fig. 9 are higher than the errors in Fig. 6, which is expected because the errors of Fig. 9 are based on validation data that is independent of the MLR coefficients.

An interesting matter is how well the prediction of loudness from DPOAE I/O data using MLR models would perform at an individual level. This would be useful, for example, if our MLR models were to be used as part of the

TABLE I. Correlation coefficients (r) and slopes of simple regression lines for prediction of CLS-CU using training data. Correlation coefficients and slopes are shown for all subjects, normal-hearing subjects and hearing-impaired subjects. The p -values for the correlation coefficients (not shown) were all less than 0.001.

	1 kHz		2 kHz		4 kHz	
	r	slope	r	slope	r	slope
All subjects	0.90	0.91	0.88	0.94	0.90	0.89
Normal hearing	0.89	0.99	0.90	0.93	0.91	0.89
Hearing impaired	0.91	0.89	0.88	0.95	0.90	0.89

TABLE II. Average rms error in units of CU between predicted and actual CLS-CU at 1, 2, and 4 kHz for the training and validation data set. Average rms errors are shown for all subjects, normal-hearing subjects and hearing-impaired subjects.

	1 kHz		2 kHz		4 kHz	
	Training	Validation	Training	Validation	Training	Validation
All subjects	4.84	8.30	5.06	7.37	4.75	8.51
Normal hearing	5.27	6.95	4.53	6.76	4.50	9.18
Hearing impaired	4.71	8.86	5.22	7.65	4.82	8.14

FIG. 4. (Color online) Examples of prediction of CLS-SPL using the training data set. Line styles and data organization are the same as in Fig. 1, however, the subjects are different and the rms error is now in units of dB.

hearing-aid fitting process. To evaluate the performance at an individual level, Fig. 10 provides distributions of individual correlation coefficients for the validation data set. Specifically, Fig. 10 provides standard and cumulative distributions of correlation coefficients between measured and predicted CLS functions for the prediction of CLS-CU (left column) and prediction of CLS-SPL (right column). Correlation coefficients are provided separately for the three test frequencies. The individual correlations are generally high for both types of predictions, with 90% of subjects having correlation coefficients greater than 0.9.

IV. DISCUSSION

The aim of this study was to extend the study of Thorson *et al.* (2012) by further exploring the efficacy of predicting loudness functions from DPOAE I/O functions using MLR analyses. Three improvements were made over the previous study: (1) a more robust MLR model was designed by including an intercept term to better fit the data, (2) two models were designed, one for prediction of the CLS function itself, and the other for prediction of the input level required for a particular loudness perception or category, and (3) a validation stage that uses independent data to evaluate whether the prediction of loudness can generalize to other data was included.

In the smoothing of the measured CLS functions to obtain target CLS functions, a slope limit of 3 CU/dB for the loud CU segment of the CLS function was imposed to avoid having CLS functions with infinite or near infinite slope in the loud CU segment and to improve the fit to the data. The slope limit does not limit the analysis because it is greater than the slope of a typical CLS function (e.g., Brand and Hohmann, 2002 obtained a median slope of 1.7 CU/dB for loud CUs). Additionally, the loud CU segment does not vary much with hearing threshold, (i.e., there are small differences between NH and HI subjects at high input levels). Most of the variability occurs in the soft CU segment, as evidenced by the different mean slopes. Our procedures for removing outliers and interpolating the CLS function are intended to provide a smooth and monotonic CLS function, similar to the fitting procedure of Brand and Hohmann (2002). The smoothed CLS function, which we referred to as the target CLS function was used as input to the MLR models, but the performance of the MLR model was based on comparison of predicted CLS functions to the CLS function before interpolation. CLS functions of normal-hearing subjects had lower CLS thresholds, soft CU slopes that were shallower, and a wider range of input levels, compared to CLS functions for hearing-impaired subjects, consistent with previous observations. Similar trends (lower thresholds, shallower I/O slopes and wider dynamic range for normal-hearing subjects) were observed when

FIG. 5. (Color online) Summary of prediction of CLS-SPL from DPOAE I/O using the training data set. The open circles are the individual data and the dashed lines are the simple regression line relating measured and predicted CUs. Correlation coefficients (r) and slopes of simple regression lines are included as insets in the figure panels. p -values for the correlation coefficients (not shown) were all less than 0.001.

comparing the DPOAE I/O functions of the two groups of subjects.

Our DPOAE measurements include contributions from the distortion component, the initial component generated near f_2 , and the reflection component, a second component generated at the characteristic site of the distortion frequency $2f_1-f_2$. These two sources interact constructively and destructively, resulting in DPOAE fine structure-quasi-periodic variation in DPOAE level across frequency. This DPOAE fine structure is responsible for the variability (notches and peaks) observed in DPOAE I/O functions, and this variability might have been reduced by removing the reflection component from the total DPOAE measurement (e.g., Mauermann and Kollmeier, 2004). In the current study, the total DPOAE measurement was used for prediction of CLS

functions without removal of the reflection component. Therefore, DPOAE fine structure may be one of the reasons for some of the variability observed in our measurements. This variability affects the subsequent predictions of CLS functions from DPOAE I/O. This implies that our prediction errors may potentially be reduced by accounting for DPOAE fine structure and removing the reflection component using methods for DPOAE source separation such as time windowing (e.g., Kalluri and Shera, 2001) or selective suppression (e.g., Dhar and Shaffer, 2004; Johnson *et al.*, 2006b). However, it is difficult to select a suppressor tone that will sufficiently suppress the DPOAE reflection component without having adverse effects on the distortion component (Dhar and Shaffer, 2004; Johnson *et al.*, 2007), and

TABLE III. Correlation coefficients (r) and slopes of simple regression lines for prediction of CLS-SPL using training data. Correlation coefficients and slopes are shown for all subjects, normal-hearing subjects, and hearing-impaired subjects. The p -values for the correlation coefficients (not shown) were all less than 0.001.

	1 kHz		2 kHz		4 kHz	
	r	slope	r	slope	r	slope
All subjects	0.95	0.93	0.94	0.90	0.94	0.90
Normal hearing	0.94	0.87	0.95	0.90	0.95	0.87
Hearing impaired	0.96	0.96	0.94	0.90	0.92	0.94

TABLE IV. Average rms error in units of dB between predicted and actual CLS-SPL at 1, 2, and 4 kHz for the training and validation data set. Average rms errors are shown for all subjects, normal-hearing subjects, and hearing-impaired subjects.

	1 kHz		2 kHz		4 kHz	
	Training	Validation	Training	Validation	Training	Validation
All subjects	7.55	12.39	7.03	11.01	7.31	13.71
Normal hearing	9.71	10.41	7.76	12.26	9.02	16.78
Hearing impaired	6.86	13.22	6.80	10.36	6.73	11.77

FIG. 6. (Color online) Average rms errors as function of CU for the prediction of CLS-SPL using the training set at 1, 2, and 4 kHz. Separate rms errors are shown for all subjects, normal hearing subjects, and for hearing impaired subjects.

suppression of the reflection component does not greatly improve clinical test performance (Johnson *et al.*, 2007).

Statistically significant correlations between measured and predicted CLS functions and simple linear regressions with slopes close to one were obtained during the training stage of the study. Correlation coefficients were equal or greater than $r=0.89$ (see Tables I and III). Average rms errors were, at most, 5 CU or one loudness category for predictions of CLS-CU (see Table II) and less than 10 dB for prediction of CLS-SPL (see Table IV). When the MLR coefficients obtained from the training set were applied to the validation set, the correlations between the measured and predicted CLS functions were reduced, as expected. However, statistically significant ($p < 0.001$) correlations were still observed with correlation coefficients in the range $r=0.73$ to $r=0.93$ (see Tables V and VI). There were no

systematic changes in the slopes of the simple regression lines between training and validation; the slope increased for some conditions and decreased for other conditions. Average rms errors for the validation stage of the study were larger than those for the training stage. This was expected since the data for the validation stage represented an independent data set from those used to develop MLR models used for prediction. However, the error rates, coupled with the correlation and the slopes of the simple regression lines, support the view that both DPOAE I/O functions and loudness functions depend on the same underlying compressive cochlear nonlinearity and that loudness may be predicted from DPOAEs. The ability to predict loudness from measurements of DPOAEs may provide objective assessment of recruitment and potentially provide information that might be useful when selecting compression parameters when fitting hearing

FIG. 7. (Color online) Validation of the prediction of CLS-CU from DPOAE I/O using the validation data set. The open circles are the individual data and the dashed lines are the simple regression line relating measured and predicted CUs. Correlation coefficients (r) and slopes of simple regression lines are included as insets in the figure panels. p -values for the correlation coefficients (not shown) were all less than 0.001.

TABLE V. Correlation coefficients (r) and slopes of simple regression lines for prediction of CLS-CU using validation data. Correlation coefficients and slopes are shown for all subjects, normal-hearing subjects, and hearing-impaired subjects. The p -values for the correlation coefficients (not shown) were all less than 0.001.

	1 kHz		2 kHz		4 kHz	
	r	slope	r	slope	r	slope
All subjects	0.73	0.68	0.80	0.92	0.74	0.88
Normal hearing	0.78	0.76	0.81	0.84	0.76	0.93
Hearing impaired	0.73	0.66	0.81	0.97	0.74	0.84

aids to individuals who are not capable of providing loudness percepts, such as infants, young children and those with developmental delays. The high correlation observed in the distribution of individual correlations (see Fig. 10) supports the idea that loudness can be predicted reliably at an individual level. It is important to note, however, that this approach may be applicable only in cases of mild-to-moderate hearing loss, as DPOAEs are typically not measureable for greater degrees of loss. Additionally, this approach assumes that the hearing loss is due to dysfunction of the OHC only and not due to dysfunction of the IHC or other central auditory processes.

One of the improvements we made over the previous study of Thorson *et al.* (2012) was the use of a more robust MLR model that included an intercept term to provide a

better fit to the data. Unfortunately we cannot evaluate the importance of the intercept term by comparing our performance with that of Thorson *et al.* (2012) because their errors were based on comparison of predicted CLS function to what we have defined presently as “target CLS function” (which is less variable) and not comparisons to actual measured CLS data as was done in this study. Therefore, direct comparison of errors between the two studies may not be meaningful. However we can evaluate the value of an intercept term by comparing the performance of our current model with and without this term. When an intercept was not used, the rms errors reported in Table II increased by as much as 1.8 CU for the training set and 1.0 CU for the validation set. Similarly, the rms errors in Table IV increased by as much as 3.0 dB for the training set and by as much as 4.8 dB for the validation set. Thus, it appears that inclusion of an intercept term improved the predictions.

Two separate data sets, one for training and one for validation, were used in this study, and the validation set was independent of the training set. Although the accuracy of the model may be improved by using all the data, the usefulness of the model is best evaluated by applying it to an independent set of data that was not used to obtain this model. If the model were to be applied in the clinic, the model would need to provide predictions of a CLS function for a particular patient with knowledge of only their DPOAE I/O function. This is the reason why an

FIG. 8. (Color online) Validation of the prediction of CLS-SPL from DPOAE I/O using the validation data set. The open circles are the individual data and the dashed lines are the simple regression line relating measured and predicted CUs. Correlation coefficients (r) and slopes of simple regression lines are included as insets in the figure panels. p -values for the correlation coefficients (not shown) were all less than 0.001.

TABLE VI. Correlation coefficients (r) and slopes of simple regression lines for prediction of CLS-SPL using validation data. Correlation coefficients and slopes are shown for all subjects, normal-hearing subjects, and hearing-impaired subjects. The p -values for the correlation coefficients (not shown) were all less than 0.001.

	1 kHz		2 kHz		4 kHz	
	r	slope	r	slope	r	slope
All subjects	0.90	0.97	0.89	0.90	0.83	0.94
Normal hearing	0.93	0.92	0.90	0.93	0.85	0.98
Hearing impaired	0.88	0.99	0.89	0.80	0.80	0.81

independent data set is needed and was used in this study to evaluate the potential of the model. The use of an independent set of data to validate model predictions represents an important distinction between the current study and the study of Thorson *et al.* (2012), who did not provide validation for their model.

Our MLR models were able to predict the loud CU segment of the CLS function from DPOAE I/O function despite the fact that the input levels for the loud segments was as much as 20 dB above the highest DPOAE stimulus level. We had no expectation that we would be able to make this prediction with any reliability. It is worth noting that slopes of the loud segment correlate poorly with hearing thresholds (Thorson *et al.*, 2012), so prediction of the loud segment based on threshold would not be reliable.

A bootstrapping procedure was used to evaluate the stability of our MLR models (Fox, 2008). In this procedure, regressors for the model (in our case the L_d) were treated as fixed, the model coefficients were used to calculate fitted values and residuals, and then bootstrap samples of the residuals were used to calculate bootstrapped samples of the model outputs. Following this, the bootstrapped model outputs were regressed on the fixed L_d values to obtain bootstrap samples of regression coefficients, which were then used to obtain confidence intervals for the regression coefficients. This bootstrapping procedure was performed only for the analysis that predicts CLS-SPL from DPOAE I/O data, and only for the condition in which CU = 25. Five thousand bootstrap samples were used. The confidence intervals were large, indicating that the model coefficients may not generalize to other data sets. Therefore we do not recommend the use of our model coefficients as a standard for relating OAEs to loudness.

Previous predictions of hearing thresholds from our laboratories using DPOAEs have produced residual errors of 10 dB or greater (Neely *et al.*, 2009, Gorga *et al.*, 2003, Rogers *et al.*, 2010). These studies, however, did not include a validation stage. Thus, it is expected that the errors they reported would have increased if the prediction models were applied to an independent set of data. Different strategies for creating regression coefficients relating DPOAEs to thresholds were adopted in the previous studies. Given the success of the strategy of applying the entire DPOAE I/O function for the creation of the prediction models used here for predicting loudness, we evaluated how our strategy would perform in the prediction of hearing thresholds. Specifically, the predictions of threshold were made following procedures similar to those described in the Sec. II but with matrix B of CLS data in Eqs. (4) and (6) replaced by audiometric thresholds at the DPOAE f_2 frequency. Figure 11 shows average rms errors between measured and predicted hearing threshold obtained for training (open symbols) and for validation (closed symbols) data sets. Average rms errors obtained during training are less than 10 dB HL, which is less than the errors obtained in earlier studies. Average rms errors obtained for validation data (8.31–16.19 dB) are similar to the errors observed in the studies cited above. These results indicate that our strategy of using the entire DPOAE I/O function in the creation of MLR analyses produces robust models that apparently represent improvements over the previous models used to make threshold predictions. It is interesting to note that the rms error at 4 kHz is usually higher than the error at 1 and 2 kHz. The source of the error may be related to standing-wave errors which have been shown to occur near this frequency (e.g., Scheperle *et al.*, 2008; Richmond *et al.*, 2011).

A limitation in the predictability of loudness from DPOAEs and, therefore, the use of DPOAEs for hearing-aid fitting, are the assumptions that the hearing loss is due to OHC damage only and that the nonlinearity observed in loudness growth is influenced mostly by OHCs and not inner hair cells (IHC) or other more central processes beyond the periphery (e.g., Relkin and Doucet, 1997; Heinz *et al.*, 2005). If the notion that hearing loss less than about 60 dB HL is due to OHC damage or dysfunction is true (e.g., Moore and Glasberg, 2003), then DPOAE-based loudness estimation (and hearing-aid fitting) will only be useful if losses are less than 60 dB HL. When hearing loss is due to damage to both OHC and IHC (or other more central

FIG. 9. (Color online) Average rms errors as function of CU for the prediction of CLS-SPL using the validation set at frequencies of 1, 2, and 4 kHz. Separate rms errors are shown for all subjects, normal hearing subjects and for hearing impaired subjects.

FIG. 10. (Color online) Standard and cumulative distributions of correlation coefficients between measured and predicted CLS functions for the validation set. The left column presents distributions for prediction of CLS-CU and the right column presents distributions for prediction of CLS-SPL.

auditory structures), the portion of the hearing loss that is not due to OHC damage will not manifest in measurements of DPOAEs, potentially weakening correlation between DPOAE and loudness I/O functions. However, if the proportion of hearing loss that is due to OHC damage can be determined accurately (e.g., Moore *et al.*, 2000; Lopez-Poveda and Johannesen, 2012), a DPOAE-based fitting strategy may be helpful in remedying the part of the hearing loss that is due to OHC damage.

FIG. 11. (Color online) Average rms error for prediction of audiometric thresholds from DPOAE I/O functions at 1, 2, and 4 kHz. Separate rms errors are shown for all subjects, normal hearing subjects and for hearing impaired subjects. Separate rms errors are also shown for training and validation data sets.

V. CONCLUSION

Statistically significant correlations, slopes of simple regression lines close to one, and low residual errors between measured loudness and loudness predicted from DPOAE I/O functions using MLR models were observed when a training data set was analyzed. The correlations were reduced, and the error increased when the MLR models developed using the training data were applied to a new and independent set of data. However, statistically significant correlations were still observed that support the views that both DPOAE functions and loudness functions depend on the same underlying compressive cochlear nonlinearity and that loudness may be predicted from DPOAEs. Our strategy of using the entire DPOAE I/O function for prediction may also be useful for prediction of hearing thresholds.

ACKNOWLEDGMENTS

This research was supported by grants R01 DC8318 (STN), R01 DC2251 (MPG) and P30 DC4662, from the NIH-NIDCD. We thank Sarah Al-Salim and her co-authors, and Megan Thorson and her co-authors for sharing their data. We also thank Enrique Lopez-Poveda for constructive criticisms of an earlier version of this manuscript.

¹In MATLAB software version 7.11 (MathWorks, Natick, MA), this can be done using the command “mldivide” or the backslash operator.

²See supplementary material at <http://dx.doi.org/10.1121/1.4807560> for a complete set of multiple linear regression coefficients for predicting CLS-CU and CLS-SPL.

³CU = 0 corresponds to the loudness category labeled “cannot hear.”

- Allen J. B. (1997). “Derecruitment by multiband compression in hearing aids,” in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Lawrence Erlbaum Associates, NJ), pp. 99–112.
- Allen, J. B., Hall, J. L., and Jeng, P. S. (1990). “Loudness growth in 1/2-octave bands (LGOB)—A procedure for the assessment of loudness,” *J. Acoust. Soc. Am.* **88**, 745–753.
- Al-Salim, S. C., Kopun, J. G., Neely, S. T., Jesteadt, W., Stiegemann, B., and Gorga, M. P. (2010). “Reliability of categorical loudness scaling and its relation to threshold,” *Ear Hear.* **31**, 567–578.
- Boege, P., and Janssen, T. (2002). “Pure-tone threshold estimation from extrapolated distortion product otoacoustic emissions I/O-functions in normal and cochlear hearing loss ears,” *J. Acoust. Soc. Am.* **111**, 1810–1818.
- Brand, T., and Hohmann, V. (2002). “An adaptive procedure for categorical loudness scaling,” *J. Acoust. Soc. Am.* **112**, 1597–1604.
- Brownell, W. E. (1990). “Outer hair cell electromotility and otoacoustic emissions,” *Ear Hear.* **11**, 82–92.
- Cox, R. M. (1995). “Using loudness data for hearing aid selection: The IHAFF approach,” *Hear. J.* **48**(10), 39–44.
- Dhar, S., and Shaffer, L. A. (2004). “Effects of a suppressor tone on distortion product otoacoustic emissions fine structure: Why a universal suppressor level is not a practical solution to obtaining single-generator DP-grams,” *Ear Hear.* **25**, 573–585.
- Dorn, P. A., Konrad-Martin, D., Neely, S. T., Keefe, D. H., Cyr, E., and Gorga, M. P. (2001). “Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears,” *J. Acoust. Soc. Am.* **110**, 3119–3131.
- Dorn, P. A., Piskorski, P., Gorga, M. P., Neely, S. T., and Keefe, D. H. (1999). “Predicting audiometric status from distortion product otoacoustic emissions using multivariate analyses,” *Ear Hear.* **20**, 149–163.
- Epstein, M., and Florentine, M. (2005). “Inferring basilar-membrane motion from tone-burst otoacoustic emissions and psychoacoustic measurements,” *J. Acoust. Soc. Am.* **117**, 263–274.
- Epstein, M., and Silva, I. (2009). “Analysis of parameters for the estimation of loudness from tone-burst otoacoustic emissions,” *J. Acoust. Soc. Am.* **125**, 3855–3864.

- Epstein, M. J. (2011). "Correlates of loudness," in *Springer Handbook of Auditory Research: Loudness*, edited by M. Florentine, A. N. Popper, and R. R. Fay (Springer, New York), pp. 89–107.
- Fletcher, H., and Munson, W. (1933). "Loudness: Its definition, measurement, and calculations," *J. Acoust. Soc. Am.* **9**, 1–10.
- Fox, J. (2008). *Applied Regression Analysis and Generalized Linear Models*, 2nd ed. (Sage, Thousand Oaks, CA), Chap. 21.
- Gorga, M. P., Dierking, D. M., Johnson, T. A., Beauchaine, K. L., Garner, C. A., and Neely, S. T. (2005). "A validation and potential clinical application of multivariate analyses of distortion-product otoacoustic emission data," *Ear Hear.* **26**, 593–607.
- Gorga, M. P., Neely, S. T., and Dorn, P. A. (1999). "Distortion product otoacoustic emission test performance for a priori criteria and for multifrequency audiometric standards," *Ear Hear.* **20**, 345–362.
- Gorga, M. P., Neely, S. T., Dorn, P. A., and Hoover, B. M. (2003). "Further efforts to predict pure-tone thresholds from distortion product otoacoustic emission input/output functions," *J. Acoust. Soc. Am.* **113**, 3275–3284.
- Gorga, M. P., Neely, S. T., Ohlrich, B., Hoover, B., Redner, J., and Peters, J. (1997). "From laboratory to clinic: A large scale study of distortion product otoacoustic emissions in ears with normal hearing and ears with hearing loss," *Ear Hear.* **18**, 440–455.
- Heinz, M. G., Issa, J. B., and Young, E. D. (2005). "Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment," *J. Assoc. Res. Otolaryngol.* **6**, 91–105.
- Hussain, D. M., Gorga, M. P., Neely, S. T., Keefe, D. H., and Peters, J. (1998). "Transient evoked otoacoustic emissions in patients with normal hearing and in patients with hearing loss," *Ear Hear.* **19**, 434–449.
- Johnson, T. A., Neely, S. T., Garner, C. A., and Gorga, M. P. (2006a). "Influence of primary-level and primary-frequency ratios on human distortion product otoacoustic emissions," *J. Acoust. Soc. Am.* **119**, 418–428.
- Johnson, T. A., Neely, S. T., Kopun, J. G., Dierking, D. M., Tan, H., Converse, C., Kennedy, E., and Gorga, M. P. (2007). "Distortion product otoacoustic emissions: Cochlear-source contributions and clinical test performance," *J. Acoust. Soc. Am.* **122**, 3539–3553.
- Johnson, T. A., Neely, S. T., Kopun, J. G., and Gorga, M. P. (2006b). "Reducing reflected contributions to ear-canal distortion product otoacoustic emissions in humans," *J. Acoust. Soc. Am.* **119**, 3896–3907.
- Kalluri, R., and Shera, C. A. (2001). "Distortion-product source unmixing: a test of the two-mechanism model for DPOAE generation," *J. Acoust. Soc. Am.* **109**, 622–637.
- Kirby, B. J., Kopun, J. G., Hongyang, T., Neely, S. T., and Gorga, M. P. (2011). "Do 'optimal' conditions improve distortion product otoacoustic emission test performance?," *Ear Hear.* **32**, 230–237.
- Kummer, P., Janssen, T., and Arnold, W. (1998). "The level and growth behavior of the $2f_1-f_2$ distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss," *J. Acoust. Soc. Am.* **103**, 3431–3444.
- Launer, S., Holube, I., Hohmann, V., and Kollmeier, B. (1996). "Categorical loudness scaling in hearing-impaired listeners—Can loudness growth be predicted from the audiogram?," *Audiol. Akust.* **4**, 156–163.
- Lopez-Poveda, E. A., and Johannesen, P. T. (2012). "Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss," *J. Assoc. Res. Otolaryngol.* **13**, 485–504.
- Mauermann, M., and Kollmeier, B. (2004). "Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source," *J. Acoust. Soc. Am.* **116**, 2199–2212.
- Moore, B. C. J., and Glasberg, B. R. (2003). "A revised model of loudness perception applied to cochlear hearing loss," *Hear. Res.* **188**, 70–88.
- Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcantara, J. I. (2000). "A test for the diagnosis of dead regions in the cochlea," *Br. J. Aud.* **34**, 205–224.
- 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. (2003). "Cochlear compression estimates from measurements of distortion-product otoacoustic emissions," *J. Acoust. Soc. Am.* **114**, 1499–1507.
- Neely, S. T., Johnson, T. A., and Gorga, M. P. (2005). "Distortion product otoacoustic emission measured with continuously varying stimulus level," *J. Acoust. Soc. Am.* **117**, 1248–1259.
- Neely, S. T., Johnson, T. A., Kopun, J., Dierking, D. M., and Gorga, M. P. (2009). "Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears," *J. Acoust. Soc. Am.* **126**, 728–738.
- Neely, S. T., and Liu, Z. (1994). "EMAV: Otoacoustic emission averager," Technical Memo No. 17, Boys Town National Research Hospital, Omaha, NE.
- Oswald, J. A., and Janssen, T. (2003). "Weighted DPOAE input/output functions: A tool for automatic assessment of hearing loss in clinical application," *Z. Med. Phys.* **13**, 93–98.
- Relkin, E. M., and Doucet, J. R. (1997). "Is loudness simply proportional to the auditory nerve spike count?" *J. Acoust. Soc. Am.* **101**, 2735–2740.
- Richmond, S. A., Kopun, J. G., Neely, S. T., Tan, H., and Gorga, M. P. (2011). "Distribution of standing-wave errors in real-ear sound-level measurements," *J. Acoust. Soc. Am.* **129**, 3134–3140.
- Rogers, A. R., Burke, S., Tan, H., Neely, S. T., and Gorga, M. P. (2010). "Influence of calibration method on distortion-product otoacoustic emission measurements: II. Threshold predication," *Ear Hear.* **31**, 546–554.
- Scharf, B. (1978). "Loudness," in *Handbook of Perception: Vol. IV. Hearing*, edited by E. C. Carterette and M. P. Friedman (Academic Press, New York), pp. 187–242.
- 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**, 288–300.
- Smeds, K., and Leijon, A. (2011). "Loudness and hearing loss," in *Springer Handbook of Auditory Research: Loudness*, edited by M. Florentine, A. N. Popper, and R. R. Fay (Springer, New York), pp. 223–259.
- Thorson, M. J., Kopun, J. G., Neely, S. T., Tan, H., and Gorga, M. P. (2012). "Reliability of distortion-product otoacoustic emissions and their relation to loudness," *J. Acoust. Soc. Am.* **131**, 1282–1295.
- Vinck, B. M., Cauwenberge, P. B., Corthals, P., and Vel, E. D. (1998). "Multi-variant analysis of otoacoustic emissions and estimation of hearing threshold: Transient evoked otoacoustic emissions," *Audiology* **37**, 315–334.