A comparison of pure tone and distortion product audiometric thresholds

J. B. Allen
Acoustics Research Dept.
AT&T Bell Labs
M. Connington
H. Levitt
Dept. of Speech and Hearing
City Univ. of N.Y.

12/11/91

Abstract

In this paper we compare objective measurements of ear canal acoustic distortion products (DP's) in the human ear canal to subjective measures of hearing level (HL) thresholds. To date we have measured 22 subjects (39 ears). These subjects are normal hearing as defined by their hearing level being less than 25 dB. The correlation between the distortion products and the hearing levels, averaged over subjects and frequency, is -0.51, and the regression slope (Δ_{DP}/Δ_{HL}) has been found to be -0.525 dB/dB. Working under the hypothesis that the two measures should be more highly correlated, we discuss possible sources of the remaining variance. When measuring the hearing level for the subjective thresholds, the transducer was calibrated in a DB-100 coupler. When making the DP measurements, an ear canal calibration was used. Preliminary analysis suggests that the variance is partially due to the nonlinear quantization of the HL to 5 dB levels and to the subjective bias in measuring the psychophysical hearing thresholds. Above 4 kHz the variance is mainly due to the differences in coupler calibration, ear canal standing waves, and subject criterion bias. Further experiments are needed to confirm this initial analysis.

1 Introduction

Otoacoustic emission distortion product measurements are now known to be an important objective measure of the state of the cochlea[4, 9, 15, 13, 16, 6, 17]. The acoustic distortion product, or DP, is defined as the ear canal sound pressure $P_d(f_d)$, at frequency $f_d = 2f_1 - f_2$, which results from imposed ear canal primary signals $P_1(f_1)$ and $P_2(f_2)$. This distortion product is generated in the cochlea by the nonlinear basilar membrane motion which results from outer hair cell (OHC) movement.

The more we learn about acoustic distortion products, the more it appears that they are fundamentally related to the hearing process. Two important questions must be asked. *First,* what is the nature of the relationship between the DP pressure and the psychophysical hearing threshold, or *Hearing Level* (HL). The hearing level is the pure tone hearing threshold relative to that of the average "normal" hearing listener. *Second,* assuming that these measures are related, how well can HL be predicted given DP.

The most direct demonstration of an HL-DP relationship would be to show a tight correlation between hearing thresholds and ear canal DP levels. If DP's were highly correlated to hearing thresholds, then an understanding of their properties could lead to a much deeper understanding of cochlear function. A method of objectively predicting the HL from the DP could be very important in the clinic. DP's are objective, and depend only on the state of the cochlea and the middle ear, whereas HL, which is subjectively determined, depends on the state of not only the middle ear and cochlea, but also on the state of the auditory nerve, the CNS, and the psychological state of the subject.

Not only are the DP's objective, but they are the result of physical mechanisms in the cochlea. The conclusions of most research into the source of DP's is that they result from the electrical motility of the outer hair cells. As the basilar membrane moves, the outer hair cells depolarize in response to the shear displacement between the tectorial membrane and the reticular lamina. In response to the resulting outer hair cell membrane voltage, the OHC's change their length. Because they move, the outer hair cells are said to be motile. This change in OHC length modifies the mechanical impedance of the basilar membrane (BM) thereby introducing non-linearities in the mechanical behavior of the BM. We call this impedance change the motility induced change. Because of the hair cell membrane capacitance and fluid viscosity, the OHC motility is limited in its frequency response to some cutoff frequency f_m. These few basic ideas lead to many important insights and consequences. ref

The most commonly studied DP is at frequency $f_d = 2f_1 - f_2$, where $f_2 > f_1$ are the frequencies of the primary tones. We define the pressure level of the primaries to be P_1 and P_2, and that of the DP to be P_d. The $2f_1 - f_2$ distortion product results from the injection of energy in that region of the basilar membrane at which the two primaries interact to produce OHC motion. This non-linear behavior is believed to occur near, or at, the f_2 place x_2[7, 14, 5]. Apical to x_2, the BM does not respond to the higher frequency primary because the traveling wave rapidly attenuates.

Important possible applications of the DPOAEs include the following:

1. The amplitude of the DP might be used as an objective measurement of hearing level. It has been observed that correlations exist between DP Input-Output functions (defined as the DP pressure, in the ear canal, as a function of the canal primary pressure) and hearing level (see references given above). This relationship will be explored further in this paper.

2. By simultaneously recording the DP pressure in the ear canal and with a neuron associated with basilar membrane activity at a particular place, one can study the power gain for waves traveling along the basilar membrane[3].

3. DP's can be used[10, 2, 6] to probe the frequency response of cochlear micromechanics. This can give insight into how the BM signals are transformed during signal transduction.

4. The DP phase can be used to measure the BM traveling wave latency as a function of place and frequency[4].

5. DP's can give information about the mechanical saturation of the OHC response.

Our working hypothesis in this paper is that hearing level (obtained by subjective measurement) and the magnitude of the distortion product (obtained by objective measurement) are highly correlated.

The purpose of this study is *first* to explore the relation between DP's and hearing thresholds, with an emphasis on understanding the sources of error affecting the two measures, and *second,* to study the prediction of HL from the DP.

This study differs from previous studies in that *(a)* hearing thresholds and DP's have been measured using the same transducer, and at about the same time (i.e., the two measurements are obtained within minutes of each other), *(b)* the response of the transducer in the ear canal has been measured, thereby controlling the pressure in the ear canal, and *(c)* DP's have been measured as a function of frequency for one level rather than as a function of level for one frequency.

Several earlier studies have measured DP's as a function of stimulus level, and have attempted to define a threshold based on the noise floor[15, 13, 9, 17]. These studies have found the relationship between DP and stimulus level to be non-monotonic. It has also been assumed that this non-monotonicity is the largest source of variability in defining the DP threshold. In this paper we look for other sources of variability, such as the acoustic variability in making the hearing threshold measurements, ear drum impedance variation across subject, and the variability due to subjective measurement factors.

2 The Experiment

Even if the DP and hearing level were perfectly correlated, the observed correlation between the two variables must be less than one because of variability in the measurements. The sources of variability may be subdivided into acoustic variability, cochlear variability, and psychophysical variability. Included in the acoustic category are middle ear variations and pathology, ear drum impedance variations, microphone calibration errors due to standing waves in the ear canal, and impedance loading of the DP generator by the plugged ear canal. Included in the cochlear category are the variations due to differences in the cochlear micromechanics, standing waves in the cochlea due to reflections at the stapes, and effects of cancellation due to the distribution of the generator site along the basilar membrane[5]. Included in the psychophysical category are test-retest variability, and subject response bias (e.g., a change in subject's criterin as to whether the stimulus is audible or not). The variability may be further categorized by frequency and subject. For example, acoustic calibration errors due to standing waves is a problem at high frequencies. Acoustic impedance variations can occur at any frequency, but are easily measured and controlled. Subject response bias should be independent of frequency. Test-retest variability must be measured directly since it could depend on the measurement method, the subject, the experimenter, as well as the trial-to-trial variations in the subject's sensitivity.

2.1 Methods

Subject hearing levels (HLs) for 29 subjects (more than 50 ears) ranged from –10 dB to + 40 dB HL. As will be described below, this group was pruned to 22 subjects (39 ears) having 'normal' hearing. *Normal hearing* is defined here as a loss of less than 25 dB HL between 500 Hz and 8 kHz.

2.1.1 Distortion Product measurements

The experiments were carried out in a single walled booth. The distortion products were measured using CUBDISP© (Etymōtic Research, Elk Grove Village, IL). The level of each of the two primary tones was 65 dB SPL, as determined by a reference microphone in the ear canal. Measurements were obtained at discrete frequencies, varied in steps of 1/10 octave over the range of f_2 from 500Hz to 10,000 Hz. The ratio of f_2/f_1 was held at 1.2.

CUBDISP uses an Ariel DSP-16 signal processing board in an IBM compatible PC, a pair of Etymōtic ER-2 earphones, and an ER-10B ear canal microphone. After an initial calibration of the two ER-2 earphone frequency responses in the ear canal, pairs of tones are presented to the ear canal. The resulting ear canal DP is measured with the Etymōtic ER-10B reference microphone. CUBDISP tests at a rate of up to one frequency point every 4 seconds. This is done by time averaging a periodic stimulus for 4 seconds and then using an FFT to find the energy in 0.25 Hz frequency bands. The acoustic noise floor is simultaneously measured along with the DP. This allows for the signal-to-noise ratio (SNR) to be determined simultaneously with the DP. leave for discussion

The distortion product is plotted as a function of f_2, since we believe that it is generated at the f_2 place[7, 8, 11, 12].

2.1.2 Calibration

The calibration was done according to commonly accepted methods. The pressure sensitivity of the ER-2 earphone in a DB-100 Zwislocki-type coupler was measured in Pascals/volt. The conversion from Pascals/volt as measured in the coupler to HEARING-LEVEL/volt was made at each frequency using the pressure to hearing level conversion tables for the ER-2. ref) This conversion is shown graphically in Figure 1. The sound pressure level in the coupler corresponding to the threshold of hearing for normal subjects is referred to as the *Reference Equivalent Threshold Sound Pressure Levels* (RETSPLs). (

2.1.3 Audiometric threshold measurements

Immediately after measuring the DP pressure, while the transducer package was still in place, audiometric thresholds were measured. The hearing level (HL) for each subject, expressed as a function of frequency, was obtained by subtracting the normal threshold curve from the measured threshold of hearing curve. The standard audiometric procedure for measuring hearing levels was used since we were intrested in potential audiological applications of DP measurements and wished to relate these measurements to the traditional method of measuring hearing level[1]. leave footnote for discussion

Each threshold was found by decreasing the level of a tone in 5 dB steps, until the subject could no longer detect the tone. This sequence was repeated until a reliable estimate of the threshold level was obtained, at frequencies 0.25, 0.5, 0.75, 1, 1.5, 3, 4, 6, and 8 kHz, which are the standard audiometric frequencies.

The data have been processed in several ways. First we removed data points having a poor signal-to-noise ratio. We could do this because we measured the noise floor along with the DP signal, and therefore knew the signal-to-noise ratio. Then we removed all data points where the hearing level was outside of the normal range. Means ($\overline{x} = mean(x)$) and standard deviations (σ) were computed across subjects as a function of frequency, and scatter plots of (DP,HL) were made across subjects and frequency, and across subjects for a given frequency.

Does fig 1 show RETSPL vs freq for 0 dB HL?

RETL should be sufficient, since ...

3 RESULTS

The first issue to be addressed is the nature of the relationship between the distortion product (DP) and hearing level (HL). A measure of this relationship is the correlation coefficient ρ. Unfortunately, both DP and HL are subject to errors of measurement. As a consequence, the observed correlation between DP and HL will be lower than the true correlation assuming no measurement error. In general,

$$r_{HL,DP} = \rho_{HL,DP}(1 - \frac{\sigma^2_{eHL}}{\sigma^2_{HL}})(1 - \frac{\sigma^2_{eDP}}{\sigma^2_{DP}}), \quad (1)$$

where

$r_{HL,DP}$	=	observed correlation between HL and DP
$\rho_{HL,DP}$	=	true correlation between HL and DP
$\sigma^2 HL$	=	variance of HL
$\sigma^2 eHL$	=	error variance of HL measurements
$\sigma^2 DP$	=	variance of DP
$\sigma^2 eDP$	=	error variance of DP measurements.

[1] In retrospect, this may have been a poor choice because of subjective bias that results from the clinical procedure. The thresholds should have been measured by a two-interval forced choice, or PEST type, procedure.

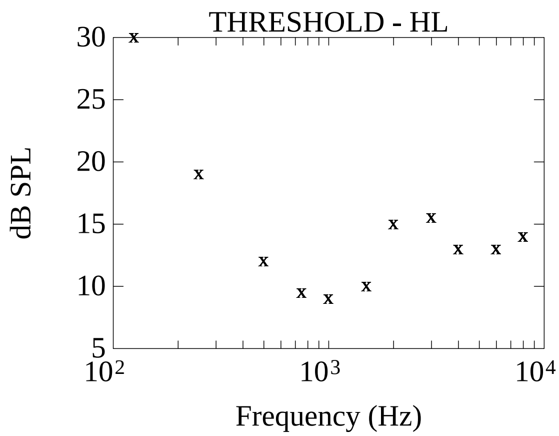

Figure 1: We plot here the Reference-Equivalent-Threshold-Sound-Pressure-Levels (RETSPLs) in dB. The audiometer is first calibrated in a DB-100 coupler in terms of Pascals/volt. Given the voltage on the transducer at the subject's threshold of hearing, the corresponding sound pressure level in the coupler, referred to as the Reference-Equivalent-Theshold-Sound-Pressure-Level (RETSPL), is computed. The hearing level is then calculated from the RETSPL. For example, threshold measurements on a given subject show that at 8 kHz the ER-2 transducer requires an electrical signal level of x volts for the tone to be just detectable. This voltage corresponds to a sound pressure level of 10 dB SPL in a DB-100 coupler, as determined from the manufacturer's tables; i.e., the RETSPL is equal to 10 dB SPL. According to the definition of normal audiometric zero at 8 kHz (Weissler, 19xx) for an equivalent transducer (these data also provided by the manufacturer) 0 dB HL = 14 dB RETSPL. The HL for this subject at 8 kHz is therefore 10-14 = -4 dB HL.

In order to reduce the effect of measurement error on the measured correlation coefficient, data suspected of having a high error variance were omitted. These included,

i) data for hearing levels outside the normal range; i.e., hearing levels in excees of 25 dB HL were not considered, and

ii) DP measurements with a low signal-to-noise ratio. The method for determining the noise floor for the DP measurements is described below.

The sound pressure level due to brownian motion is given by $P_T = \sqrt{4kTBR}$, where P_T is the thermal pressure, k is Boltzman's constant, T is the temperature in degrees Kelvin, B is the bandwidth, and R is the acoustic resistance. Given the 0.25 Hz bandwidth of our measurements, these values give a noise floor of -37 dB SPL, assuming plane waves at room temperature,

Figure 2 shows the observed and predicted noise levels as a function of frequency. The solid line shows the mean noise level, $\overline{P_n}$, averaged over subjects at each frequency. The two dashed lines show $\overline{P_n(f)} \pm 1.5\sigma(f)$, where $\sigma(f)$ is the standard of the noise levels at each frequency. The horizontal dotted line at -37 dB SPL is the predicted noise level due to brownian motion. Note that the measured noise levels approach the theoretical limit in the range of 2 to 8 kHz; i.e., the measured noise levels in this frequency range were on the order of -30 dB SPL, which is only 7 dB above the predicted noise floow.

The upper dashed line, corresponding to $\overline{P_n(f)} + 1.5\sigma(f)$, was used as a threshold for determing if the background noise was sufficiently low for a reliable DP measurement. The DP and noise level measurements were obtained concurrently, only those DP measurements with an observed noise level less than $\overline{P_n(f)} + 1.5\sigma(f)$ were considered in the analysis.

A scatterplot of illustrating the relationship between DP and HL is shown in Figure 3a. The data have been collapsed across frequency and subject. The measured correlation coefficient, as shown in the lefthand panel, is -0.50. While the correlation is not high, it is statistically significant at the 0.001 level.

Although steps were taken to reduce extraneous sources of variance, it was not possible to remove all sources of measurement error. As a consequence, it is likely that the true correlation is higher than 0.50. Estimates of the test-retest standard deviation for the DP and HL measurements were $\sigma_{eDP} = 1$ dB and $\sigma_{eHL} = 5$ dB, respectively. (these estimates to be checkec). The standard deviation of the measured DP and HL values were $sigma_{DP} = x$ dB and $\sigma_{HC} = y$ dB, respectively. Substituting these estimates in Eq. 1 yields a slightly higher estiamte of the combination coefficient, i.e., ck eq no

$$\rho_{DP,HL} = 0.50(1 - \frac{a}{b}^2)(1 - \frac{c}{d}^2) \quad (2)$$

$$= XX. \quad (3)$$

Another way of examining the relationship between DP and HL is to assume a linear relationship between DP and HL, fit an appropriate regression line, and then to test if the assumption of linearity provides an adequate description of the relationship between DP and HL. If not, a more complex regression line, such as a quadratic or higher order polynomial, would be fitted to the data.

Three linear regression lines have been fitted to the data, as shown in Figure 3a. The solid line corresponds to the regression of DP on HL; i.e., $DP(HL) = a_h + b_h HL$. The fitted line minimizes the mean square deviation along the DP axis. The chain line corresponds to the regression of HL on DP; i.e., $HL(DP) = a_h + b_h DP$. The fitted line minimizes the mean square deviation along the HL axis. The dotted line is obtained from the singular value decomposition (SVD) of the matrix of $[DP - \overline{DP}, HL - \overline{HL}]$. This third regression line minimizes the squared perpendicular distances between the data points and the fitted line. Note that although the three regression lines have very different slopes, all three regression lines pass through a common point, $(\overline{DP}, \overline{HL})$, the bi-variate mean of the data.

The mean square deviation of the data points from the HL(DP) regression provides a measure of the error variance when estimating HL from DP, assuming DP is known without error. The issue of prediction is discussed shortly. The mean square deviation for the DP(HL) regression provides a measure of the error variance when estimating DP from HL, assuming HL known without error. The SVD regression provides measures of the variance on two dimensions, along the principle and minor directions, respectively. These variances are obtained from the singular values of the matrix, λ_i, according to the formula $S_i^2 = \lambda_i^2/(N-1)$, where N is the length of the vector DP[18]. The variance along the major axis provides a measure of the precision with which the data points are fitted by a straight line, the variability along the minor axis provides a measure of the variance not accounted for by a linear relationship.

(Jont, can we cite values for the different variances, or std deviations, here in an additional paragraph.)

The best fit to the data, in terms of minimizing the error variance, is obtained for the regression HL(DP). Figure 3b shows this regression line and the mean DP values for each value of HL. The standard deviation of each mean is also shown by the error bars. Note that the measured HL values occur at discrete 5 dB intervals. This is because the standard audiometric method of measuring hearing level was used in which the data are quantized in steps of 5 dB.

Since the noise level is determined during the measurement process, it is possible to remove noisy data points. This was done by setting a threshold on the noise floor, as shown in Fig. 2. All data points were removed if *(a)* the signal to noise ratio was less than 6 dB, or *(b)* if the noise floor was greater than the pooled mean plus K times the standard deviation, or $\overline{P_n(f)} + K\sigma_{P_n}(f)$, where $K = 1.5$. Harry stopped here

3.1 Linear Dependence

After pruning the noisy data, and subject measurements with HL > 25, we formed a scatter plot between DP and HL across frequency and subject, as seen in Fig. 3. The correlation coefficient, as shown in the left-hand panel, is -0.5, and is significant at the 0.001 level. While the correlation is not high, it is very significant.

Another important question is the slope of the relation between DP and HL. The definition of the best slope depends on the question being asked. For example, if one wishes to predict HL given DP, then the slope of the linear regression of DP given HL is not of interest. The regression line DP given HL would be of interest if we wished to predict DP from HL.

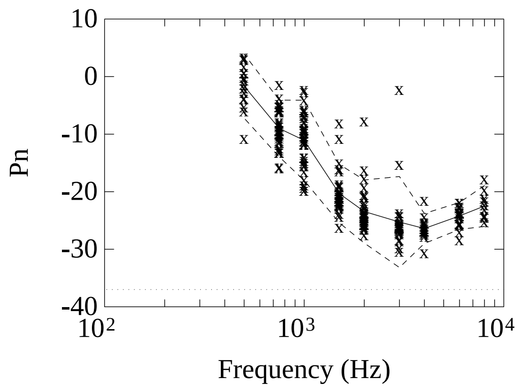

Figure 2: This plot shows the noise floor pressure P_n, in db SPL, as a function of frequency. The raw data was first processed to remove all points having a signal-to-noise ratio less than 6 dB. We also removed data points having large noise, regardless of the signal. We set the rejection threshold of the noise P_n to be $\overline{P_n} + 1.5\sigma(f)$, where σ is computed over all the measurements at a given frequency. The solid line shows the mean noise floor, while the upper dashed line defines $\overline{P_n(f)} + 1.5\sigma(f)$. Removing noisy data using this threshold improved the correlation between P_{HL} and P_{DP} while only removing a small number of data points. The dotted line at the bottom of the plot at -37 shows the theoretical limit of the noise floor in air at the 0.25 Hz bandwidth used by our measurement system.

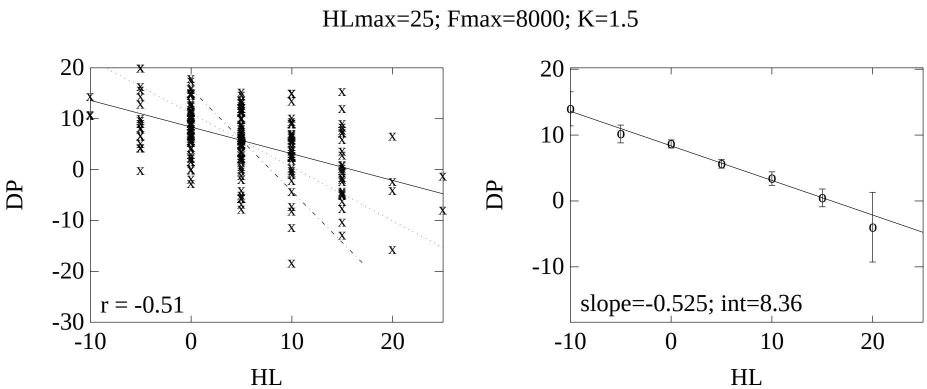

Figure 3: This figure (left panel) shows a scatter plot of DP verses HL averaged across both frequency and subjects. Three lines are superimposed on the data points. The solid line corresponds to the regression line $DP = a_p HL + b_p$. The dotted line is a 'regression' line that has been determined from a singular value decomposition of the (DP,HL) vector pairs. The dash-dot line is a regression line of for $HL = a_h DP + b_h$. The correlation coefficient is given by r. In the right panel we see the mean values along with their standard error. The solid line is the DP(HL) regression line (the solid line from the left panel).

In the left panel we see a scatter plot of DP(HL) and three different regression lines, as described in the figure caption. It is clear from the figure that the slopes of the three regression lines differ significantly. Initially it is not clear which slope is correct. The SVD regression line minimizes the perpendicular distance, while the linear regressions minimize the vertical distance in the case of DP(HL), and the horizontal distance for the case of HL(DP).

In the right panel we show the linear least squares regression line for HL(DP) (the solid line from the left panel), the HL group means (the circles), and the standard error of the means (error bars). The group means for each HL group have a standard error that is much less than the differences in means at any two HL values. This, in an intuitive way, answers the question of the significance of the correlation since, by inspection, a statistical t test between any two HL groups would clearly be significant (the difference between any two means is greater than the average standard error of the mean.)

It is striking that the means of the DP's at each HL value fit a straight line within the standard errors of the means. This argues very strongly for a *linear* relation between DP and HL for these normal hearing subjects. *Because the DP(HL) regression line so nicely represents the HL group means, we take this line as the best fit to the data.* The HL(DP) regression line and the SVD regression lines do not fit the group means, and are therefore rejected.

In attempting to account for the remaining variance, one would like to know if the slope is a function of frequency and/or subject.

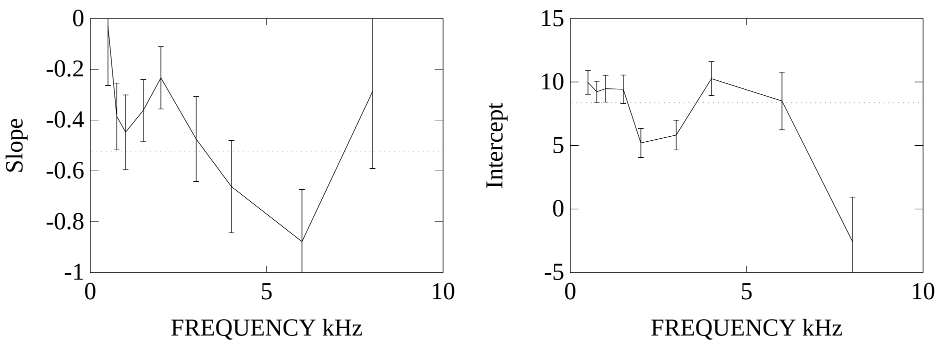

Figure 4: The left panel shows the slope $a(f_i)$ determined from the linear regression of HL as a function of DP. The right panel gives the intercept $b(f_i)$. The error bars confidence intervals. The dotted line gives the slope and intercept when the regression is formed over both frequency and subject.

3.1.1 Frequency Dependence

We next look at the question of the frequency dependence of the correlation and regression slope. The results of linear regressions at each frequency of the form

$$DP(f_i) = a(f_i)HL(f_i) + b(f_i). \tag{4}$$

which defined $a(f_i)$ and $b(f_i)$, are shown in Fig. 4.

The regression confidence interval is computed using a standard method from regression analysis. If the regression matrix equation is $Y = X\beta$, where β is a 2x1 vector containing the slope and intercept, X is the abscissa data matrix, and Y is an Nx1 vector containing the observations to be predicted, then $\beta = (X'X)^{-1}X'Y$, and the confidence interval matrix is given by $\sigma^2 = (X'X)^{-1}Y'(Y - X\beta)$. The dotted line in each panel shows regression slope and intercept for a regression computed over frequency as well as subject data.

Strictly speaking the slopes and intercept are dependent on frequency. However, taking the confidence intervals into account, the regression is approximately independent of frequency, with the possible exception of the slopes at 0.5, 2, and 6 kHz, and the intercept at 8 kHz. Between 1 and 4 kHz, the slope varies between -0.45 and -0.25, with a standard error of less than ± 0.2. The intercept has smaller relative standard error, and therefore seems to be somewhat more significant. These results tell us that we should be careful when pooling the data across frequency. Ideally we should not pool the data across frequency. However it is premature to assign any significance to the frequency dependence given the complexity of the experiment and the small N at each frequency.

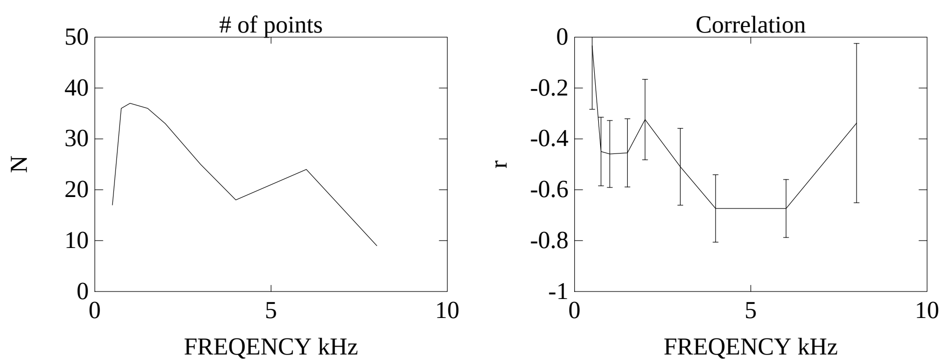

Figure 5: On the left we give the number of points that were used at each frequency for computing the means and correlations as a function of frequency. In the right panel is the correlation r as a function of frequency.

3.1.2 Other effects over frequency

In Fig. 5 the number of data points N and the correlation r are shown as functions of frequency. The error bars for the correlation were computed using the approximate formula $\sigma_r = (1 - r^2)/\sqrt{N-1})$. The number of points N drops off above 2 kHz because of the limiting effects of standing waves in the ear canal for many subjects. At frequencies where there are standing wave nulls, a pressure of 65 dB SPL could not be reached, and thus the DP could not be measured. The standing wave null frequencies depend directly on the depth of insertion of the ER-2 1 mm diameter transducer probe tubes into the ear canal. Depending on this depth, the frequency of the standing wave null ranged from 4 kHz to 10 kHz.

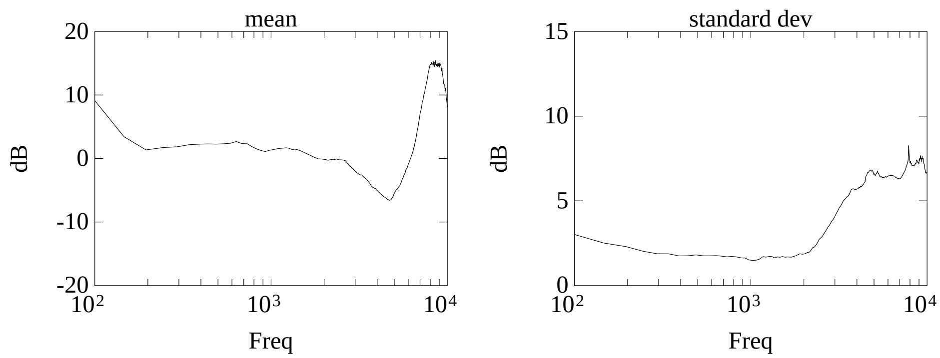

Figure 6: In the left panel we show the average ER-2 pressure response in the ear canal, normalized by the response of the ER-2 earphone in a DB-100 calibration coupler, averaged over all subjects. In the right panel we see the standard deviation of the pressure. At frequencies above 2 kHz σ increases because of deep nulls in the ear canal pressure due to standing waves. If these standing waves were reduced by placing the ER-2 probe tubes closer to the ear canal, then the resulting variation should drop.

In Fig. 6 we show the mean and standard deviation of the ear canal pressure, normalized by the DB-100 calibration response, averaged over all the subjects. At frequencies below 4 kHz these plots characterize the difference in acoustic impedance between an average subject and the DB-100, since it is the eardrum impedance variations that lead to earcanal pressure differences at low frequency. Above 4 kHz, standing waves complicate this interpretation because the probe microphone pressure and the ear drum pressure can differ, depending on the ER-2 probe depth. From the right panel it appears that these standing waves resulted in an increase in the standard deviation to about 7.5 dB above 4 kHz.

Since we have the ear canal pressure for the subjects for frequencies below 4 kHz, we could correct the HL thresholds by the difference between the known pressure and the DB-100 pressure that was used as the reference. If this correction were made, then the standard deviation of HL for frequencies below 4 kHz shown in Fig. 4 should decrease because the subject variations of the acoustic impedance of the ear canal would be compensated. From Fig. 6 this would be about 2 dB. We have not attempted to make this (small) correction because of the programing effort involved.

Based on Fig. 4 to Fig. 6, we conclude that we should remove the 6 and 8 kHz data points from the data base, and average over the remaining frequency points to avoid the standing wave artifact which introduces a difference between the ER-10B pressure and the true ear canal pressure.

3.2 Predicting DP from HL

We have argued that because of the linear relation between the HL group means, and the excellent agreement of the DP(HL) regression lines with these means, that the best representation of the data is the DP(HL) linear regression line. We then concluded that there are no significant frequency effects below 4.0 kHz, and that we may therefore work with averages across frequency.

Recomputing the least squares linear regression, excluding frequencies above 4 kHz, gave

$$\hat{DP}_{LSQ} = -0.398HL + 8.24. \tag{5}$$

Since this regression is based on normal hearing subjects, having a hearing level of better than 25 dB HL, it is impossible to predict the range of validity extrapolated into the nonnormal hearing range. If the above formula remains valid for impaired ears, and the noise floor were -20 dB SPL, then it might be possible to measure hearing losses at levels to -50 dB SPL. This extrapolation ignores the variability about the mean levels. If the variability could be controlled, the noise floor further reduced, and the levels optimized, then, assuming the extrapolation were correct, the method might be extended to a 60 dB hearing level.

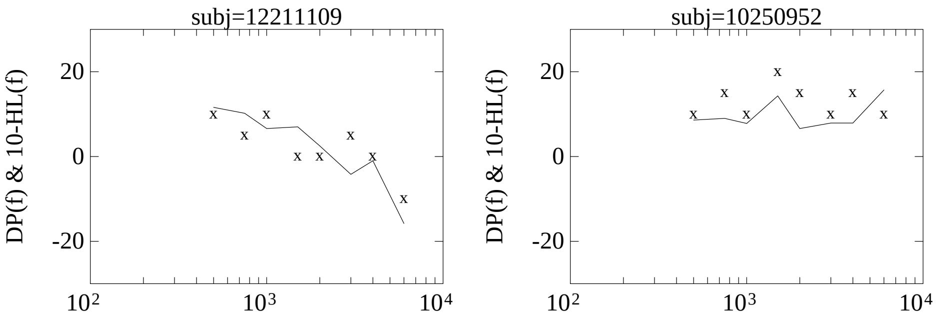

Figure 7: Example of subjects without (right), and with (left), a criterion bias. The solid lines show the DP and the x's show 10-HL. The DP and HL data on the left match fairly closely, whereas the data on the right are grossly out of alignment. We say that the subject on the right has a *subject bias* relative to the subject on the left.

3.2.1 Subject Bias

Next we investigate the variability of subject bias. When measuring the hearing level, each subject responds with a threshold criterion that is assumed to be independent of frequency, but is of course subject dependent. We assume that each subject responds when the sound is some number of dB above their absolute internal threshold. One subject might respond when they heard the tone faintly, and the next only when they clearly heard the tone. An example of the variability we are attempting to explain is shown in Fig. 7. In the left panel the two measures are similar, while in the left panel, the two are separated by about 8 dB. In this section we shall assume that this average difference is due to subject criterion bias. We *estimate* this bias as the average difference between HL and DP, averaged over frequency for each subject's ear, since we assume that the criterion is not a function of the tone frequency (we did not average over both ears).

In trying to estimate the subject effect we initially set up two regression models of the form

$$DP = aHL + S_i + F_j, \tag{6}$$

and,

$$DP = aHL + S_i, \tag{7}$$

where a gives the slope between DP and HL, and S_i, and F_j are subject and frequency dependent constants, respectively. Solving Eq. 7 for a and S_i would, in principle, be the proper way of simultaneously estimating the slope and the subject bias. The weakness of this model is that there are 39 subject constants S_i and only one slope regression constant a. The individual subjects do not have sufficient variation in hearing loss over frequency, and as a result, the 39 subject constants S_i account for most of the variability. Perhaps if the subjects had a large variation in hearing loss, then this type of a regression model could work. It was necessary, therefore, to apply the subject regression *after* the slope regression. After finding a for the linear regression, a second regression was done for the subject constants.

The data of Fig. 8, right panel, is the same as that of Fig. 3, left panel, except that the estimated subject bias, shown as the histogram to the left, has been removed. This raised the correlation from 0.5 to 0.64. The standard deviation s_2 was found to be 4.3 dB in this case, which is the error not accounted for by the regression line.

The left panel is a histogram of the computed bias across subjects. For example, from this histogram, 14 subjects had a subjective bias between ± 1.0 and 4 subjects were between 1 and 3 dB.

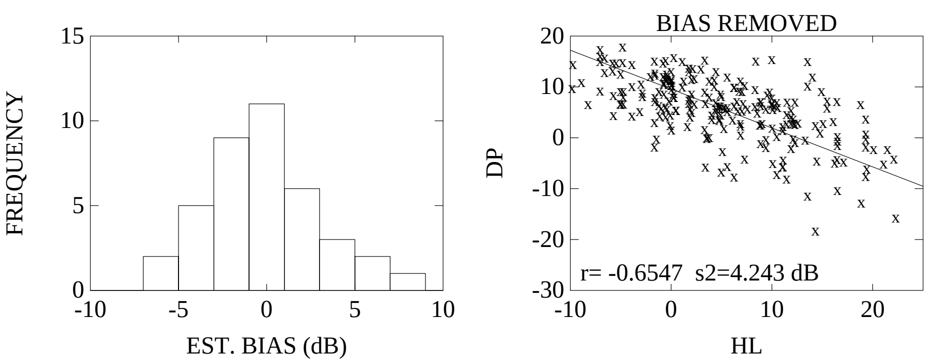

Figure 8: The effect of subject criterion bias may contribute to the standard deviation between HL and DP. To test this hypothesis we compute the average difference between HL(f) and DP(f) for each subject. We interpret this average difference as a subject bias. In the right panel we see a scatter diagram of the data after removing the subject bias. In the left panel we see a histogram of the bias values for the subjects.

3.2.2 A data model

We can summarize our error analysis, and slightly extend it by defining an error model.

Besides the subject threshold variations, which we are trying to quantify, there are at least two categories of measurement error – acoustic and psychophysical. Our statistical model is:

$$HL(f,S) + \epsilon_{TR}(f,S) + \epsilon_B(S) = \beta DP(f,S) + \epsilon_{CAL}(f) + \epsilon_{|Z|}(f).$$

The left side of the equation correspond to psychophysical effects, while the right side corresponds to physical (acoustical) effects. The acoustical sources of error include calibration artifact errors $\epsilon_{CAL}(f)$, such as those due to standing waves, and ear drum impedance errors $\epsilon_{|Z|}(f)$ which are known to be functions of frequency f. The calibration artifacts are important at higher frequencies, such as above 4 kHz. We have measured the impedance variations during our experiment. Therefore, in principle, this source of variation can be removed.

Psychophysical sources of error include test-retest variability, $\epsilon_{TR}(f,S)$, and subject criterion bias $\epsilon_B(S)$. We have assumed that subject criterion bias is not a function of frequency but depends only on the subject S.

The DP measure was calibrated in the ear canal for each subject. Thus the variability in the DP due to subject impedance has been removed. A calibration error remains at high frequencies however because of the ear canal standing wave problem. If the ER-2 earphone probe tubes were placed closer to the ear drum, then $\epsilon_{CAL}(f)$ should depend less on frequency.

4 Conclusions

We summarize some of our conclusions as follows:

- We found that the use of the noise floor to reject noisy data improved the quality of the data. Since it is easy to estimate the noise floor along with the DP, the use of the noise floor is a convenient way to initially screen the data.

- We found that the best way to fit the data was to form a linear regression of the form $DP = a * HL + b$. The slope was -.525, the intercept 8.36, and the correlation coefficient was -0.51 and was highly significant at the 0.001 level. The group-means fell on the regression line and regression line remained within the standard error of the mean. We interpreted this to mean that the relation between DP and HL was linear.

- We believe that the quantization of the HL levels to the nearest 5.0 dB, along with the subject bias, appear to be the largest source of error in the experiment. In retrospect, it seems to have been a mistake to quantize the HL levels. However, since this is the procedure used in the clinical setting, it initially seemed like a reasonable thing to do, since one of the goals of the experiment was to predict HL from DP in the clinical setting.

- We found the slope and the intercept to be functions of frequency above 4 kHz, but we were not able to place any significance on the frequency dependence.

- Standing waves were a serious problem with the ER-10B for frequencies above about 4 kHz. One possible solution to this problem is to position the sound probe tubes closer to the eardrum, which may be easily solved by the use of a foam tip rather than the hard rubber tip of the ER-10B.

- Based on an extrapolation of our normal-hearing data into the impaired range, DPOAE's could screen for mild hearing levels of up to about 50 dB HL, assuming a slope for the HL and DP regression line of -0.5 and a noise floor of -30.0 dB.

- We have found that DPOAE's are correlated with HL to at least 8 kHz, and that DPOAE is approximately independent of frequency up to 4 kHz. In the cat we have measured DPOAE's for $f_2 \leq 20$ kHz, using a different transducer[5]. This means that DPOAE's could be used for early screening of presbyacusis, and could be used for monitoring cochlear function during ototoxic drug administration.

4.1 Acknowledgements

I would like to thank the following people for their generous help in discussing the data: David Berkley, John Durrant, Pat Jeng, Mead Killion, Joe Kruskal, Jan van Santen, and Mohan Sondhi. I would also like to thank Tony Agnello and Duck on Kim for their help in the development of the CUBDIS software and hardware.

References

[1] 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.

[2] Allen, J.B. (1990). "Modeling the noise damaged ear," *The Mechanics and Biophysics of Hearing,*, Edited by Dallos *et al.*, , Springer-Verlag, New York, pp. 324-332.

[3] Allen, J.B. and Fahey, P.F. (1991). "Using acoustic distortion products to measure the cochlear amplifier gain on the basilar membrane," In preperation.

[4] Fahey, P.F. and Allen, J.B. (1985a). "Characterization of cubic intermodulation distortion products in the cat external auditory meatus," in *Peripheral Auditory Mechanisms* Edited by Allen, Hall, Hubbard, Neely, and Tubis. Springer Verlag New York, pp 314-321.

[5] Fahey, P.F. and Allen, J.B. (1985b). "Nonlinear phenomena as observed in the ear canal and at the auditory nerve," J. Acoustical Soc. Am. 77(2) 599-612.

[6] Gaskill, S.A. and Brown, A.M. (1990). "The behavior of the acoustic distortion product, $2f_1 - f_2$, from the human ear and its relation to auditory sensitivity," J. Acoust. Soc. Am. 88:2, 821-839.

[7] Hall, J.L. (1974). "Two-tone distortion products in a nonlinear model of the basilar membrane," J. Acoust. Soc. Am. 56, 1818-1828.

[8] Hall, J.L. (1981). "Observations on a Nonlinear Model for Motion of the Bassilar Membrane," in *Hearing Research and Theory* edited by Tobias, J.V. and Schubert, E.D., Academic Press New York, NY, pp 1-61.

[9] Harris, F. P., Lonsbury-Martin, B. L., (et al.) (1989). "Acoustic distortion products in humans: Systematic changes in amplitude as a function of f2/f1 ratio." J. Acoust. Soc. Am. 85 (1), pp 220-229.

[10] Kemp, D.T. (1979). "Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea," Arch. Otorhinolaryngol. 224, 37-45.

[11] Kim, D.O. and Molnar, C.E. (1975). "Cochlear Mechanics: Measurements and Models," in *The Nervous System* Edited by D.B. Tower, Raven Press, New York NY, pp 57-68.

[12] Kim, D.O., Molnar, C.E., and Matthews, J.W. (1980). "Cochlear mechanics: Nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure," J. Acoust. Soc. Am. 67, 1704-1721.

[13] Leonard, G., Smurzynski, J., Jung, M., Kim, D.O. (1990). "Evaluation of distortion product otoacoustic emissions as a basis for the objective clinical assessment of cochlear function," *Advances in Audiology* Vol 7, chapter 13, S. Karger, Switzerland.

[14] Matthews, J.W., and Molnar, C.E. (1986). "Modeling of intracochlear and ear canal distortion product (2f1-f2)," in *Peripheral Auditory Mechanisms*, edited by J.B. Allen *et al.* (Springer-Verlag, New York), pp. 258-265.

[15] Nelson, D. A. and Kimberley, B. P. , (1992). "Distortion product emissions (DPEs) and auditory sensitivity in human ears with normal hearing and cochlear hearing loss." JSHR **35** 1142-1159.

[16] Probst R., Lonsbury-Martin, B.L., Martin, G.K.(1991). "A review of otoacoustic emissions," J. Acoust. Soc. Am. 89:5, 2027-2067.

[17] Smurzynski , J., Leonard, G., (et al.) (1990), "Distortion Product Otoacoustic Emissions in Normal and Impaired Adult Ears." Arch. Otolaryn. Head Neck Surg., Nov. 116, pp 1309-1316.

[18] Sokal, Rohlf *BIOMETRY* 2nd Ed. (1981).

5 Design of a new experiment

- Measure thresholds using method that does not quantize the results, and that does not have a subject bias, such as a PEST type method.

- DPs should be measured with $A_2 = A_1 - 15$ dB, to reduce the problem of multiple sources and nothches.

- Measure the DP's at 65 dB, 50 dB, and only at the frequencies where the thresholds have been measured.

- Think about the question of frequency resolution of the measurements. We know that the microstructure of the hearing threshold is detailed. Would we expect to mirror this in the DP measurements? If so, then perhaps this would be the experiment to go for. There is a problem, that the DP is measured at 65 dB, while the HL is measured at 10 dB. Thus the micro-structure could be gone at the higher levels. This might be checked by looking at the impedance as a function of level.

- Looking at the impedance as a function of level would be intersting because of the correlation of the micro-structure of HL with the spontaneous emissions. One might expect the impedance to be correlated also.

6 Questions that need further investigation:

- The biggest problem is the fitting the transducer in the ear canal with out leaks, and inserting the probe tubes deep into the canal. Dips between 3-10 kHz are due either to standing waves in the canal or to the spreading inheritance caused by the non-planer wave propagation. The resulting dip in the reference microphone pressure means that the required primary pressure can not be attained. Furthermore, when standing waves are present, the pressure at the drum is not the pressure that is seen by the microphone. This problem caused a paucity of data in the present study between 2-8 kHz, as shown in Fig. 5, left.
- We need to remeasure with $P_2 = P_1 - 15$ dB.
- What might account for the remaining variance in our results?
  - Does the subject bias for the threshold account for the variance? We need to remove the subject bias directly by using a threshold measurement technique that is insensitive to this bias, such as a two-interval forced choice method.
  - How does middle ear effect the results? We need air-bone gap data on these subjects.
- What happens at lower frequencies? Below 500 Hz the subjective threshold sound pressure increases to about 30 dB SPL. The noise in the CUBDISP system is also increasing. Will the signal to noise ratio remain useable? Answer: the noise floor will not allow us to measure the DP if it stays constant, as it seems to be doing. This question needs further study.

7 Basic Questions about DP generation

- Does the existence of DP prove that the cochlea is active? Answer: No. Their existence shows that the cochlea is nonlinear.
- Is the cochlea Active? Answer: Our experiments with DP's in cats seem to show that the cochlear amplifier gain is unit! ([3])
- P_{DP} is maximum for f_2/f_1 near 1.2. Does this maximum represent a proximity condition ($x_2 - x_1$ on the basilar membrane) or does it represent a filtering action in the micromechanics? Answer: by observing that the frequency response of $3f_2 - 2f_1$ is nearly the same as that of $2f_1 - f_2$, we conclude that the maximum is the result of an internal filter in the micromechanics, and not some form of nonlinear self-suppression, or the proximity of the f_2 and f_1 excitation patterns on the BM!
- How can we account for the peak in the $2f_1 - f_2$ Answer: The series mechanical impedance of the resonant tectorial membrane has a minimum 1/2 octave below f_{CF}. This translates to an f_2/f_1 of about 1.2 ([2]).
- What are the 'NULLS' due to? Answer: When $P_2 >> P_1$, we see many more nulls. We may conclude that the OHC's and the DP's have saturated. This gives many nonlinear DP generates at different sites along the basilar membrane with about the same generator amplitude. These have varying phase relationships between them because of the delay of the traveling wave. (ARO Figure).
- Where is the DP being generated? Answer1: by looking at the correlations between PTA microstructure and that of the DP's, we should be able to answer this question. Answer2: Our original argument was in terms of latency (phase slope) as a function of frequency along the BM. (Killion Figure)