Published in final edited form as:

Ear Hear. 2012 ; 33(3): 315–329. doi:10.1097/AUD.0b013e31823d7917.

Behavioral Hearing Thresholds Between 0.125 and 20 kHz Using Depth-Compensated Ear Simulator Calibration

Jungmee Lee, Ph.D.¹, Sumitrajit Dhar, Ph.D.^{1,3}, Rebekah Abel, M.S.¹, Renee Banakis, AuD.¹, Evan Grolley, AuD.¹, Jungwha Lee, Ph.D.², Steven Zecker, Ph.D.¹, and Jonathan Siegel, Ph.D.^{1,3}

¹The Roxelyn & Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208.

²Biostatistics Collaboration Center, Department of Preventive Medicine, Northwestern University, Chicago, IL 60611.

³The Hugh Knowles Center, Northwestern University, Chicago, IL 60208.

Abstract

Objectives—The purpose of this study was to obtain behavioral hearing thresholds for frequencies between 0.125 and 20 kHz from a large population between 10 and 65 years old using a clinically feasible calibration method expected to compensate well for variations in the distance between the eardrum and an insert-type sound source. Previous reports of hearing thresholds in the extended high frequencies (> 8 kHz) have either used calibration techniques known to be inaccurate or specialized equipment not suitable for clinical use.

Design—Hearing thresholds were measured from 352 human subjects between 10 to 65 years old having clinically normal hearing thresholds (< 20 dB HL) up to 4 kHz. An otoacoustic emission probe fitted with custom sound sources was used and the stimulus levels individually tailored based on an estimate of the insertion depth of the measurement probe. The calibrated stimulus levels were determined based on measurements made at various depths of insertion in a standard ear simulator. Threshold values were obtained for 21 frequencies between 0.125 and 20 kHz using a modified Békésy technique. Forty six of the subjects returned for a second measurement months later from the initial evaluation.

Results—In agreement with previous reports hearing thresholds at extended high frequencies were found to be sensitive to age related changes in auditory function. In contrast with previous reports, no gender differences were found in average hearing thresholds at most evaluated frequencies. Two aging processes, one faster than the other in time scale, appear to influence hearing thresholds in different frequency ranges. The standard deviation of test-retest threshold difference for all evaluated frequencies was 5–10 dB, comparable to that reported in the literature for similar measurement techniques, but smaller than that observed for data obtained using the standard clinical procedure.

Conclusions—The depth-compensated ear-simulator-based calibration method and the modified Békésy technique allow reliable measurement of hearing thresholds over the entire frequency

Corresponding author: Jungmee Lee, FAX: (847) 491-2523; Tel: (847) 491-2469; jmlee6@northwestern.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Portions of this study were presented in a poster at the 32nd MidWinter Meeting of the Association for Research in Otolaryngology.

range of human hearing. Hearing thresholds at the extended high frequencies are sensitive to aging and reveal subtle differences, which are not evident in the frequency range evaluated regularly (8 kHz and below). Previously-reported gender-related differences in hearing thresholds may be related to ear canal acoustics and the calibration procedure and not due to differences in hearing sensitivity.

Keywords

hearing threshold; auditory threshold; extended high frequency; aging

INTRODUCTION

The first commercially produced electronic audiometer, the Western Electric 1A, could be used to measure behavioral hearing thresholds between 32 and 16,384 Hz. The later, more portable and hence convenient, Western Electric 2A had a more limited frequency range between 64 and 8,192 Hz (see Vogel et al., 2007 for a recent review of the history of the audiogram). Convenience led to custom, custom led to practice, and practice led to the established standard of limiting the measurement of hearing thresholds up to 8 kHz. The importance of measuring hearing sensitivity at frequencies above 8 kHz is well known, especially in the case of detecting and monitoring specific pathologies. However, a convenient method for calibrating and delivering well-controlled stimuli at these extended high frequencies (EHFs) has remained elusive, and so has the adoption of routine hearing evaluation at EHF.

The sensitivity of hearing thresholds at EHF to anomalies of the auditory periphery has been evident from the very first systematic reports (Bunch, 1929) and has been reiterated in countless reports since (e.g., Harris, 1978; Sindhusake, et al. 2001). The importance of measuring hearing sensitivity at frequencies above 8 kHz is most evident in certain populations and pathologies. Perhaps the most recognized and well-studied use of hearing thresholds at EHF is in monitoring auditory function in patients undergoing chemotherapy using agents known to be ototoxic (Fletcher and Cairns, 1967; Jacobson et al., 1969; Dreschler et al., 1985; Tange et al., 1985). In these patients changes in hearing thresholds between 10 and 14 kHz occur prior to any measurable changes at frequencies below 8 kHz. Age-related changes in hearing thresholds also appear first at frequencies above 8 kHz (Bunch, 1929; Rosen et al., 1964; Fletcher, 1965; Northern et al., 1972, Ahmed et al., 2001).

The value of using hearing thresholds at EHF to monitor noise-induced changes to the auditory periphery is not as clear in the literature. Some groups have reported no change in hearing thresholds above 8 kHz concurrent with measurable shifts in thresholds between 2 and 8 kHz (Osterhammel, 1979; Laukli and Mair, 1985). In contrast, others have reported thresholds at EHF to be equally or more sensitive to noise damage compared to thresholds below 8 kHz (Satallof et al., 1967; Corliss et al., 1970). Fausti et al. (1979) found thresholds at EHF to be altered prior to any observable changes in thresholds below 8 kHz in cases of chronic exposure to hazardous levels of noise. The additive effects of noise-induced and age-related hearing loss have also been documented in the EHF range. Although not extensively studied, the consensus of published reports appears to be that age-related mechanisms dominate the changes at EHF, while noise-induced changes are more prominent at frequencies below 8 kHz (Morton & Reynolds, 1991; Ahmed et al., 2001). Hearing thresholds at EHF have also been used to assess middle ear surgical techniques (Laukli and Mair, 1985) and as predictors of speech-in-noise difficulty in the absence of clinical hearing loss (Shaw et al., 1996), albeit in isolated reports.

Several groups have published reports on EHF hearing sensitivity from relatively large numbers of subjects (Osterhammel and Osterhammel, 1979; Dreschler et al., 1985; Stelmachowicz et al., 1989; Frank, 1990; Buren et al., 1992; Hallmo et al., 1994; Sakamoto et al., 1998a; Ahmed et al., 2001). The results from these studies are unanimous in that: (1) EHF sensitivity worsens with age at a rate faster than frequencies below 8 kHz, (2) EHF sensitivity is better in women than men, (3) thresholds at EHF have higher inter-subject variability than at frequencies below 8 kHz, while intra-subject variability and test-retest reliability is similar across all frequencies (Henry and Fast, 1984; Henry et al., 1985; Green et al., 1987; Stelmachowicz et al., 1988; Frank and Dreisbach, 1991; Zhou and Green, 1995; Ahmed et al., 2001; Schmuziger et al., 2004).

While the data available in the literature on hearing sensitivity at EHF may seem adequate, the complications in obtaining such data sets are also plentiful. Almost every published report acknowledges the difficulty in delivering controlled stimuli at frequencies where standing waves in the ear canal create spatially non-uniform sound pressures. Most studies of age-related changes in threshold have used circumaural headphones calibrated using a standard coupler with a flat plate adapter (IEC 60318-1 ed2.0:2009) (Osterhammel and Osterhammel, 1979; Dreschler et al., 1985; Frank, 1990; Buren et al., 1992; Hallmo et al., 1994; Sakamoto et al., 1998a; Ahmed et al., 2001). The remaining studies used a prototype audiometer that estimates the pressure at the medial end of the ear canal and at the eardrum from pressure measurements in a coupling tube just outside the ear canal (Green et al., 1987; Stelmachowicz et al., 1988, 1989). Stimuli calibrated in a coupler do not account for individual variations in ear canal anatomy. The methods of Green et al. (1987) and Stelmachowicz et al. (1988, 1989) are similar in spirit to those employed here.

We have developed a procedure for estimating eardrum stimulus levels using a depth-compensated ear simulator method similar to that previously described by Gillman and Dirks (1966). While our results are generally consistent with those of previous studies, we will present test-retest data on a subset of our subjects to compare the reliability of the instrumentation and calibration used here with those used in previous studies. An important consideration in our study was to incorporate a wide range of diagnostic tests into a single hardware/software system. Using circumaural headphones in a separate method to measure EHF thresholds would have defeated that purpose. The results reported here were obtained using a standard otoacoustic emission probe fitted with custom sound sources, making the stimulus delivery and measurement system usable for other tests such as otoacoustic emissions, ear-canal reflectance, and other psychophysical measures including speech perception. The use of a standard otoacoustic emission probe allows this method to be implemented in equipment available currently. As explained later, the use of the simulator and a 50 ft long copper water supply tube is limited to the set up stage. That is, these procedures are performed initially during hardware installation and are not repeated during the test session.

Behavioral hearing thresholds between 0.125 and 20 kHz obtained using reliable calibration and measurement techniques are reported in this study. Hopefully our report will serve as another building block for the incorporation of reliable threshold measurements up to 20 kHz in clinical practice.

METHODS

Subjects

Three hundred fifty two (352) subjects, ranging in age from 10–65 years were evaluated. Of the total subject pool, 131 were male and 221 were female. When asked to provide racial information, 298 subjects self-identified as Caucasian, 51 as African-American, 41 as Asian,

1 as Native Hawaiian, and 12 declined to answer. One ear was chosen at random for hearing threshold evaluation. Immittance measures and clinical audiometry were performed on both ears with an Interacoustics AA220 Audiometer and Middle Ear Analyzer, though subjects were not excluded based on immittance results. All data were collected in a sound-treated booth. Subjects were excluded from the study if both ear canals were occluded with cerumen or debris. The threshold measurements were performed in one of two sound treated chambers that each met the current maximum allowable ambient noise standard (ANSI S3.1-1999, R2008). All subjects had hearing thresholds no worse than 20 dB HL through 4 kHz and had visible and healthy tympanic membranes, as evaluated by otoscopy. Subjects were grouped according to age and fell into one of five categories: 10–21 years (group A), 22–35 years (group B), 36–45 years (group C), 46–55 years (group D), or 56–65 years (group E). Forty six subjects returned for a repeat evaluation between 3 and 9 months after the initial test. Each subject signed a consent form, and all procedures were executed in compliance with the Northwestern University Institutional Review Board guidelines.

Instrumentation

Stimuli were generated using custom software on an Apple Macintosh computer. A MOTU 828 MkII FireWire device (44,100 Hz, 24-bits) was used for digital-to-analog conversion, and then buffered by a custom headphone amplifier. Custom-built MB Quart 13.01HX drivers were used to generate stimuli that were delivered to the ear via an Etymotic Research, ER10B+ otoacoustic emission probe, coupled using a 13mm foam tip. Subjects used a computer mouse button to control the level of the stimulus.

Eardrum Sound Pressure Level (SPL) Estimation using a Depth-compensated Ear Simulator Method

In every subject, the frequency of the first half-wave resonance of the ear canal was measured to estimate the distance between the emission probe and the eardrum. The objective was to be able to compensate for differences in insertion depth and ear canal length in individual subjects in each test session. We use the first half-wave frequency rather than the first quarter-wave pressure null because we have found the latter to be sensitive to factors such as the distance the source tubes are extended beyond the calibrated plane of the emission probe microphone (Siegel, 1994), crosstalk and early reflections from irregularities in the ear canal. The frequency response to a constant voltage pure-tone swept from 20–20,000 Hz was recorded in the ear canal (Figure 1, red curve). This pressure response was normalized (Figure 1, dashed red curve) by a similar measurement performed previously with the emission probe inserted into a 50 ft long copper plumbing supply tube with an internal diameter of 7.9mm (3/8 in od), approximately the same as that of the average adult human ear (Figure 1, green curve). The loss of acoustic energy over distance leaves no measurable reflections from the distal end of this tube. The pressure measured in the tube thus represents the system's response in an infinite transmission line with a characteristic impedance similar to the average adult ear canal which is the incident pressure of the probe system (Keefe, 1997; Goodman et al., 2009; Keefe and Schairer, 2011). Contributions of irregularities in the frequency responses of the sound sources and the emission probe microphone transfer function are largely removed in the normalized emission probe pressure response, making it much easier to identify accurately the first half-wave resonance.

The half-wave frequency of the subject's ear canal could then be matched to the pressure response in a standard ear simulator (IEC 60318-4; Bruel and Kjaer 4157) with the emission probe inserted to a comparable depth. A family of pressure frequency responses was measured with the emission probe inserted into the simulator to different depths in ~0.5 – 1 mm increments. For each depth, the pressure response was measured using both the emission probe microphone and the simulator's internal 0.5 in microphone (Figure 1, gray

curves), with each pair of measurements repeated for the second sound source of the emission probe. The first half-wave resonant frequency was measured for each insertion depth using the normalized pressure response measured by the emission probe, exactly as for the measurement in real ears. To facilitate making measurements using a range of insertion depths comparable to those we have encountered in real ears, a customized cylindrical extension piece was used in place of the flared “ear canal” extension. Although this deviates from the flared shape of real ears, the modification was necessary, as the fixed geometry of the simulator would not allow the foam ear-tip to seal adequately for the shallowest insertion depths (corresponding to the longest real ear canals we have encountered). The measurements in the long tube and the simulator were performed once during hardware setup. The protocol does not require these steps to be repeated for each test session or on a regular basis. These measurements would have to be repeated only if the frequency response of the sound drivers changed.

The eardrum pressure estimate for each test session was the response measured by the simulator’s microphone (“eardrum”) selected from the family of calibration measurements that best matched the half-wave resonance frequency in each subject. In the example shown in Figure 1, the pressure response measured with the emission probe reveals a first half-wave resonance at 8.0 kHz (red curves). The simulator calibration with the same half-wave resonance (blue curve) is selected from the entire set (gray curves) to represent the estimate of the subject’s eardrum pressure. The eardrum pressure near the first half-wave resonance is thus expected to have been well compensated by our procedure. All subsequent stimulus levels in the subject were adjusted by the software at each stimulus frequency, based on the selected simulator pressure response, to control the stimulus level at the eardrum.

Even though the ear simulator does not accurately represent the human ear above 10 kHz, the standard specifies that it may be used as a coupler that allows repeatable calibration of insert earphones up to 16 kHz. The pressure in either the ear simulator or the real ear is enhanced at frequencies near the second and, for long probe-to-eardrum distances, the third half-wave resonance. A reflection coefficient in adult human ears on the order of 0.6 at high frequencies (Farmer-Fedor and Rabbit, 2002) implies an eardrum impedance on the order of 340 cgs Ohms (roughly 4× the characteristic acoustic impedance of about 84 cgs Ohms). The measured Thévenin source impedance of our emission probe system is on the order of 150–400 cgs Ohms. These terminating impedances at both ends of the ear canal limit the height of the higher-order half-wave resonances. The simulator has a large resonance at ~13 kHz that greatly exceeds that of real ears. However, the fact that the higher-order resonances in our simulator calibration set above 10 kHz do not generally exceed 10 dB indicates that the source impedance of the probe may control the resonances in the simulator to a range similar to real ears. The more important source of systematic error is the fact that the frequencies of half-wave resonance of successive order in the simulator are spaced roughly by integer multiples of the first resonance, while the higher-order resonances in real ears are systematically lower in frequency than those in the simulator. In the example in Figure 1, the second half-wave resonances in the simulator and real ear are 16.39 and 14.36, respectively. Thus eardrum stimulus levels are consistently lower in the simulator near the second resonance than in the real ear and consistently higher in the simulator than in the real ear near the simulator’s second half-wave resonance. Note that the pressure developed by the system in the long tube lies at the minimum pressure in the simulator calibration set in the frequency range where half-wave resonances are evident. Using the pressure in the long tube (green curve) as the stimulus estimate would be a viable alternative for frequencies well above the first half-wave resonance in the simulator as the eardrum pressure would deviate in only one direction (higher) and only near the higher-order resonance frequencies. We chose to simplify our calibration protocol by relying on the simulator pressure for all frequencies. The pressure in the long tube, like the ear simulator, does not compensate for

differences in the diameters of individual ear canals, as the corresponding differences in characteristic acoustic impedance result in nearly frequency-independent changes in stimulus pressure in real ears.

Several recent studies have established the promise of using forward pressure, sound intensity or transmitted pressure or power to the middle ear to specify stimulus levels to the ear (Scheperle, et al., 2008; Withnell, et al, 2009; Keefe and Schairer, 2011). These procedures offer the ability, at least in principle, to compensate for differences between individual ears more accurately than our simulator method, particularly above 10 kHz with the systematic differences between higher-order half-wave resonances in the simulator and real ear noted above. Although we now calibrate stimulus levels routinely in forward pressure, this study was begun before we were confident that forward pressure estimation was practical and accurate for high frequencies and the procedures could be applied successfully on the large scale of this study. Based on our experience since then, we are now confident that forward pressure would have provided somewhat better stimulus level control to the ear than our simulator method (Souza et al., 2010), but the differences are not likely large enough to have affected any of the conclusions of this study.

Threshold Tracking

Custom software, built on Max/MSP (cycling74.com), was used to present stimuli and record data points. Threshold values were obtained for 21 standard audiometric frequencies between 0.125 and 20 kHz using a modified Békésy technique described below.

Stimuli were pulsed tones, 250 ms in duration with 25 ms rise and fall times, presented twice per second. The adaptive threshold task required the subject to press a computer mouse button to indicate that the pulsed tone was heard. Once the button was pressed, the presentation level decreased until the subject indicated that it had become inaudible by releasing the button (first reversal). The presentation level then increased until the button was pressed again (second reversal). The attenuation step size, initially 6 dB, was reduced to 2 dB after the first two reversals. Midpoints between reversals were calculated for each ascending run (i.e., as the tones crossed the subject's threshold from below to above audibility). Descriptive statistics of these individual threshold crossings were calculated after six ascending runs, excluding the first two reversals that constituted "training" runs. The tracking procedure was considered to have converged to the subject's threshold if the standard error of the mean was less than 1. If this convergence criterion was not reached for the first six runs, additional runs were accumulated until convergence was achieved. Threshold crossings that fell either below or above 1.5 times the interquartile range were considered outliers and were not included in subsequent convergence computations. Frequencies were presented in a fixed order, starting at 1 kHz, proceeding to the highest frequency, repeating the measurement at 1 kHz, then proceeding downward to the lowest test frequency. If a subject did not respond to the tone at the output limit of the equipment, the threshold was recorded as the maximum output for that frequency.

RESULTS

Figure 2 represents averaged hearing thresholds as a function of frequency for five different age groups. Thresholds are presented in standard format using a logarithmic frequency axis in Figure 2a. The same data are plotted on a linear frequency axis in Figure 2b to make the higher frequencies more discernible. The error bars represent 95% confident intervals, which are used in lieu of standard deviation or standard error, as they are more amenable to clinical comparisons across age groups. Note that thresholds for all age groups plateau at 105 dB SPL, essentially due to a ceiling effect imposed by the maximum output of the hardware. This is considered "no response" as it represents the output limit of our equipment and does

not represent true thresholds. The results of a three-way analysis of variance (frequency $\times$ age group $\times$ gender) indicated significant effects for frequency and age group ($F_{20,6604} = 1778.10$, $p < 0.0001$, and $F_{4,6604} = 651.11$, $p < 0.0001$, respectively), but surprisingly not for gender. A significant interaction between age group and frequency ($F_{80,6604} = 34.33$, $p < 0.0001$) was observed. In contrast, interactions between other factors were weak but still statistically significant (i.e., age group $\times$ gender and gender $\times$ frequency, $F_{4,6604} = 4.19$, $p = 0.0022$, and $F_{20,6604} = 1.61$, $p = 0.0426$, respectively). That is, when data were collapsed across all frequencies and age groups, average thresholds did not differ between male and female subjects. Gender-related variations in thresholds are presented in detail later. A post-hoc pair-wise multiple comparison using the Bonferroni method ($p = 0.05$) showed (1) no effect of age group on hearing thresholds between 0.125 and 4 kHz as well as for 19 and 20 kHz, (2) average hearing thresholds for age group A were significantly better than those for age groups D and E between 6 and 11.2 kHz, (3) for frequencies of 12.5 and 14 kHz, average thresholds of all age groups were significantly different with the exception of age groups A and B, (4) average thresholds for age group A were significantly better than the average thresholds of all other age groups for frequencies between 15 and 18 kHz, and (5) there were significant differences in average thresholds between age groups D and E between 8 and 14 kHz.

Comparisons of average hearing thresholds for male and female subjects within each age group are displayed in Figure 3. As in Figure 2, data are presented using logarithmic and linear frequency scales in Figures 3a and 3b, respectively. Data from each age group are presented in individual panels and the error bars represent 95% confidence intervals. As was reported above, statistically significant effects of gender were not detected in any age group. However, some trends, all statistically insignificant, are observable in Figure 3. Average hearing thresholds for male subjects were slightly worse than those for female subjects between 4 and 6 kHz in age groups B (22–35 yrs) and D (46–55 yrs) and between 4 and 14 kHz in age group E (56–65 yrs). In contrast female subjects in age group C (36–45 yrs) had slightly worse thresholds between 1 and 2 kHz.

Two alternate indices were extracted from the hearing thresholds measured in each subject. The frequency at which hearing thresholds crossed 60 dB SPL was determined for each subject after linearly interpolating between the thresholds immediately lower and higher than 60 dB SPL. This metric was termed the “60-dB intercept.” A “corner frequency” was also determined for each subject after visual examination of a display of his/her hearing thresholds as a function of frequency (logarithmic scale). A typical graphical display used to identify the corner frequency by trained research assistants is presented in Figure 4. The “corner frequency” marked a boundary in frequency below which thresholds were relatively independent of frequency and above which thresholds increased precipitously to the end of the measurement range. Average values for both the “60-dB intercept” and the “corner frequency” for each age group are displayed in Figure 5. The x and y error bars mark the 95% confidence interval in either frequency or threshold. Both the corner frequency and the 60-dB intercept decrease in frequency with increasing age. The average corner frequency decreases from ~12.5 kHz in age group A to 5.9 kHz in age group E. The average 60-dB intercept also shows a similar pattern, decreasing from ~16.7 kHz in age group A to ~10 kHz in age group E. The linear slope of a line joining the corner frequency and the 60-dB intercept was computed for each age group, and found to be ~ -0.009 dB/Hz for all age groups. This parallel shift in the corner frequency and 60-dB intercept represents an approximately invariant slope of the steeply rising segment of hearing thresholds in the highest frequencies.

Hearing thresholds measured at different frequencies are displayed as a function of age in individual panels of Figure 6. Each symbol in each panel represents the hearing threshold

measured from an individual subject. The threshold data were fitted using either simple linear regression (i.e., one-line regression) or piecewise linear regression (i.e., two-line regression) based on mean square error (MSE). A two-line regression function was chosen when it resulted in a smaller MSE than the one-line function at least by 5. While simple linear regression fits were found to be optimal between 0.125 and 4 kHz, piecewise linear regression fits were found to be optimal between 6 and 19 kHz. A simple linear regression was found to be the optimal fit at 20 kHz. These single and dual line fits are represented using the solid lines in Figure 6. Close examination of Figure 6 reveals that the cluster of thresholds around 105 dB SPL at 12.5 kHz and above significantly influences the fits at these frequencies. These thresholds represent no responses at the upper limit of the output range of our instrumentation and cannot be considered to be true thresholds. The prevalence of these non-responses increased from 25% at 12.5 kHz to 84% at 20 kHz. Once these non-responses were eliminated from the fitting pool, the data could be optimally characterized using a simple linear regression at and above 14 kHz (demonstrated using the dashed lines in Figure 6). In the frequency range between 6 and 12.5 kHz where the data were best described using the two-line regression, the age-junction between the two fitted lines was noted (“breakpoint”).

The same analysis of changes in hearing thresholds as a function of age was done separately for female and male subjects. The slope of the simple linear regression was not statistically different between female and male subjects except at 4 kHz ($p = 0.0052$). However, female and male subjects differed in whether a simple or piecewise regression best fit the data at any given frequency. While a piecewise linear regression was the optimal fit at most frequencies for female subjects, such was the case only at three frequencies for male subjects (4, 8 and 10 kHz). A comparison of hearing thresholds as a function of age between female and male subjects is shown in Figure 7. Data from female and male subjects are presented using open and filled circles, respectively. Either simple or piecewise linear regression lines are plotted in gray and black lines for female and male subjects, respectively.

When considering data for both genders, the two-line fits between 6 and 12.5 kHz seem to represent two aging processes, a relatively slow process operational at ages below the breakpoint and a process that causes much faster deterioration of hearing thresholds above the breakpoint (see Figure 6). This breakpoint in age between the two rates of decline in hearing threshold decreases with increasing frequency, as depicted in Figure 8 (filled circles). In the frequency range between 6 and 12.5 kHz, where the breakpoint can be reliably extracted, the age at the breakpoint changes from ~ 51 yrs for 6 kHz to ~ 30 yrs for 12.5 kHz. However, when female (filled triangles) and male (filled squares) data are considered separately, the breakpoints do not monotonically decrease with increasing frequency and the patterns in males and females are different. The breakpoint at any given frequency appears at a lower age for female subjects between 6 and 10 kHz compared to males. Comparisons above 10 kHz are not possible as data from male subjects at higher frequencies were best fit using a simple regression (one-line fit).

DISCUSSION

There has been long standing interest in hearing thresholds at EHF and the volume of reports on this topic in the archival literature reflects this interest. Groups that have published on this topic have typically been motivated by access to a special population, a new instrument, or a novel calibration technique. Indeed part of our motivation was to evaluate a new calibration technique as well. Here we have reported hearing thresholds from a sizeable population spread over a substantial portion of the typical human life span (10 – 65 yrs). Our use of a tracking technique for establishing thresholds that demanded focused use of a

computer mouse prevented us from including subjects younger than 10 years old. Our desire to examine hearing thresholds at EHF in a clinically normal-hearing population determined the upper limit of 65 years. It certainly is possible to find individuals older than 65 years with clinically normal hearing sensitivity. However, the proportion of such individuals in the population decreases dramatically above this age cut off.

The sample size and the age range of the subjects in this study were similar to that in previous reports. For example, the report by Osterhammel and Osterhammel (1979) contains hearing thresholds between 4 and 20 kHz from 286 normal-hearing subjects between 10–70 years old. Others have also reported hearing thresholds at EHF for expansive age ranges (Stelmachowicz et al., 1989; Hallmo et al., 1994; Sakamoto et al., 1998a). In contrast, a few groups have limited their examination to subjects belonging to specific age groups. Buren et al. (1992) reported hearing threshold between 0.125 and 20 kHz from 10, 14, and 18 year-old subjects. Matthews et al. (1997) and Lee et al. (2005) focused on older adults while Reuter et al. (1998) examined children between 4 and 7 years only. With the exception of the aforementioned reports, the majority of previous data sets on hearing thresholds at EHF were recorded from young adults (Harris and Myers, 1971; Dreschler et al., 1985; Green et al., 1987; Stelmachowicz et al., 1988; Frank, 1990; Frank and Dreisbach, 1991; Ahmed et al., 2001).

A. Age-related changes in thresholds at EHF

By only including subjects with hearing thresholds lower than 20 dB HL through 4 kHz, we hoped to rigorously test the sensitivity of thresholds at EHF to age-related changes in the auditory system. Figure 3 clearly demonstrates a separation of average thresholds for each age group at EHF. The frequency at which the average thresholds begin to separate from the adjacent groups decreases with increasing age. For example, the traces representing average thresholds for age groups A and B can be visually differentiated only above 12.5 kHz. However, the average thresholds for the older age groups can be visually differentiated from age groups A and B starting at lower frequencies. Clearly, age-related changes in hearing thresholds are evident at EHF significantly earlier than at regular audiometric frequencies, in agreement with previous studies on EHF thresholds. On this, we are in agreement with every other group that has asked the same question (Osterhammel and Osterhammel, 1979; Dreschler et al., 1985; Stelmachowicz et al., 1989; Hallmo et al., 1994; Sakamoto et al., 1998; Ahmed et al., 2001).

The similarity in hearing thresholds of all age groups at frequencies below 4 kHz provides support for the use of one set of normative values across the life span. However, the same cannot be said about thresholds at EHF. Others before us have suggested the use of age-specific normative ranges at EHF (Osterhammel & Osterhammel, 1979; Osterhammel, 1980). Our results support the need for such age-specific norms and the data provided in Table I could well serve this purpose. Normative values for hearing thresholds at EHF are critically dependent on the method of calibration. Issues related to calibration are discussed later.

In addition to documenting the effect of age on hearing thresholds at EHF, we extracted two additional parameters related to the frequency pattern of hearing thresholds for each subject. Both the corner frequency and the 60-dB intercept decreased in frequency as a function of age (see Figure 5). Extracting these two “landmarks” from the thresholds of individual subjects revealed some aspects lost in averaged group data. As can be seen in the example in Figure 4, thresholds are relatively constant until the corner frequency is reached, following which there is a rapid (and in most cases linear) rise in thresholds until the highest audible frequency is reached. This abrupt change in the trajectory of thresholds seen in almost all individual subjects is lost if the data are averaged across subjects due to variation in the

corner frequency across subjects. It is also remarkable that while the corner recedes in frequency with age there is relatively little change in its level, hovering around 20 dB SPL for all age groups. Finally, knowing the corner frequency and the 60-dB intercept allows us to estimate the slope of this high-frequency segment of hearing thresholds reliably. The 60-dB intercept was convenient for us as most of our subjects had measurable thresholds at the relevant frequencies, allowing accurate estimation of the intercept. Hearing thresholds typically deteriorated rapidly (> 100 dB/octave) between the corner frequency and the 60-dB intercept (see Figure 5). The steepness of this slope and its invariance with age suggests that the processes of aging that dominate at these frequencies are cochlear in origin.

B. Evidence for multiple aging processes

Plotting hearing thresholds at each frequency as a function of age (Figure 6) allowed us to critically examine the effect of age at each frequency. The data were fitted using regression techniques and we used the slopes of these fitted functions as metrics of aging at any given frequency. Very shallow slopes of the fitted functions (between 0.05 and 0.39) between the frequencies of 0.125 and 4 kHz suggested the lack of any significant age-related changes in hearing thresholds at these frequencies. The effects of age were more pronounced at and above 6 kHz. Interestingly, the data between 6 and 12.5 kHz were fit more optimally using two-fit functions revealing one aging process in the low frequencies (up to 4 kHz) and an additional faster process at higher frequencies (6–12.5 kHz). The junction (in age) between these two aging processes decreased with increasing frequency (51.35, 46.56, 46.27, 36.47 and 30.22 years for 6, 8, 10, 11.2 and 12.5 kHz, respectively; see Figure 7). It is remarkable that the rapid aging process is operational starting at 12.5 kHz in subjects as young as 30 years old. Perhaps a similar observation was reported by Sakamoto et al. (1998a) who fit independent regression lines to thresholds between 8 and 10 kHz and between 14 and 17 kHz. They showed that the intercept (in dB SPL) for the fitted line between 8 and 10 kHz increased significantly for subjects older than 30–39 yrs old, suggesting the onset of rapid aging of the auditory system. The data presented here provide the clearest indication yet available for the existence of two separate processes of aging operating on two different time scales. The origin of the slow process observed here is uncertain and could be peripheral or central. The rapid aging process is likely the result of progressive degeneration of hair cells in the cochlea as previously identified in histology (i.e., Gacek and Schuknecht, 1969; Schuknecht and Gacek, 1993).

C. The effect of gender on hearing thresholds

Hearing sensitivity in human females has been reported to be better than males in various reports. This difference has been demonstrated at frequencies up to 8 kHz in both longitudinal and cross sectional studies (see Gordon-Salant, 2005, for a review) and at EHF (Matthews et al., 1997; Lee et al., 2005; Silva et al., 2006; Stelmachowicz et al., 1989; Wiley et al., 1998). More specifically, age-related changes in hearing have been reported to start earlier in life in males compared to females (Pearson et al., 1995), and the rate of deterioration of hearing has also been reported to be faster in males (Moscicki et al., 1985; Cruickshanks et al., 1998). Surprisingly, no statistically significant differences in hearing thresholds were found between male and female subjects (between 10 to 65 yrs old) in this data set. To the best of our knowledge, Green et al. (1987) and Dunkley and Dreisbach (2004) provide the only other reports of a similar indifference in hearing thresholds between males and females (18 to 26 yrs old and 18 to 29 yrs old, respectively). It is notable that Stelmachowicz et al., (1989) observed a difference in male and female thresholds for subjects between 10 and 59 years old using the same calibration method as Green et al., (1987). Frequency regions where thresholds from female subjects were better on average than those from male subjects could be observed in age groups B and D (see Figure 3). In contrast, thresholds from male subjects were marginally better on average than those from

female subjects in age group C. It should be noted that none of these differences were statistically significant (i.e., $p < 0.05$).

Although there was no significant gender difference in hearing threshold, the breakpoints for different aging processes (slow vs. fast) were different between female and male subjects (see Figure 8); while two aging processes were observed for most frequencies (except 16 kHz) in females, the two aging processes were evident only for 4, 8, and 10 kHz in males. It should be noted that our sample is biased towards females with an overall female to male sample size ratio of 1.6:1. However, the results reported here held even when the statistical tests were repeated with the group sizes equated by randomly sampling female subjects in each age group (details later).

Comparing mean hearing thresholds (Figure 4) could overlook a possible difference in the distribution of hearing thresholds between female and male subjects. To address this possibility, hearing threshold distributions were compared between genders for each frequency using the Kolmogorov-Smirnov test (Conover, 1999). The results revealed a significant difference in the distribution only at 1.5 kHz ($p = 0.006$). Further, the equality of hearing threshold distribution was tested for each age group and some examples are plotted in Figure 9. The histogram of threshold distributions was fitted by a Kernel density function. The Kolmogorov-Smirnov test revealed a statistically significant gender effect on threshold distributions only for a few frequencies ($p < 0.05$) for each age group: 0.125 kHz for age group A; 2 and 4 kHz for age group B; 0.125, 4, and 6 kHz for age group D; 4 kHz for age group E. There are two interesting observations from Figure 9: (1) the frequencies (except 0.125 kHz) showing a gender effect seem to be in the range of frequencies where thresholds are known to be vulnerable to noise-induced hearing loss; (2) gender effects in hearing threshold distribution are not observed at any frequency for age group C. One distinctive difference between age group C and other age groups is a sample size ratio between female and male: 1:1 (group C) vs. 1.6:1 (other groups). To eliminate the possibility of a sampling bias on this analysis, the equality test of threshold distribution was done using a random sample of females to match the male sample size. The general trends reported before held for this analysis: a significant gender effect was observed at 1.5 kHz only for the entire data set, and a gender effect was not observed at any frequency for the age group C. The frequencies at which the gender effects were observable changed minimally (see Table II). Notably, the differences between male and female threshold distributions observed at 0.125 kHz were eliminated once the sample sizes were balanced.

We attribute this lack of a gender effect on thresholds to the calibration technique employed in this study. Recall that the stimuli delivered were individually compensated based on the depth of insertion of the probe into the subject's ear canal during the test session. This essentially neutralized effects of ear-canal-length (and partially compensated for differences in volume) between subjects. Indeed the average length and volume is smaller for female ear canals and could be responsible for the persistent finding of better hearing thresholds in female subjects. Hearing level is positively correlated with ear canal volume and length (Hellström, 1994; 1995). The fact that gender differences are insignificant in our study suggests that our depth-compensated ear simulator calibration method better controls the eardrum stimulus levels than calibration methods used previously.

D. Comparison with previous studies

The complexity of delivering well calibrated stimuli at frequencies as high as 20 kHz to the human eardrum is well recognized. The hardware (amplifier and sound source) necessary to deliver distortion-free stimuli at high enough sound pressure levels to elicit sensation at frequencies as high as 20 kHz is not trivial either. These complications have led to a variety

of hardware solutions and calibration approaches. Thus comparison of data across studies has to be done with caution, accounting for instrumentation and calibration differences.

The techniques proposed and used to calibrate stimuli in the EHF's can be grossly divided into two categories – flat-plate-coupler calibration and in-the-ear calibration. Stelmachowicz et al. (1988, 1989) performed in-the-ear calibration using a highly specialized acoustic assembly and a method to predict the eardrum pressure from the pressure measured in a coupling tube just outside the ear canal. The use of a probe microphone to measure the sound pressure at the tympanic membrane has also been proposed. However, this method is inaccurate at high frequencies, as the sound pressure varies considerably over the surface of the tympanic membrane (Stinson and Shaw, 1982; Stinson, 1985; Stevens et al., 1987), so a single placement of the probe tube does not uniquely define the eardrum pressure. The vast majority of the previous studies have been conducted using circumaural or over-the-ear headphones calibrated with a flat-plate coupler (e.g., Hallmo et al., 1994; Sakamoto et al., 1998a). The critical difference between calibration using a flat-plate coupler and any in-the-ear technique is that the flat-plate-coupler method does not account for individual differences in ear canal acoustics resulting from differences in geometry. In other words, calibrating using a flat-plate-coupler aims to standardize headphone output in the coupler, but it is accepted that individual ear canal acoustics will impose deviations from this standard. In contrast, in-the-ear calibration techniques attempt to deliver the stimulus at the desired level to each individual's ear. Below we compare our data to three previous studies, recognizing these differences in methodology.

A comparison of average hearing thresholds for different age groups reported by Stelmachowicz et al. (1989), Hallmo et al. (1994), Sakamoto et al. (1998a) with those reported here is presented in Figure 10. The filled circles with dotted lines represent the percentage of subjects in a given age group who did not have reliable or recordable responses at a particular frequency in the current data set (re: right y-axis). An arbitrary value of 105 dB SPL was assigned in these cases as a representative value greater than the maximum output of our instrumentation. Our data were reconstructed using age ranges reported in the previous studies. In general, average hearing thresholds for lower frequencies are in agreement across the four studies. However, noticeable differences exist at high frequencies, and the data from Sakamoto et al., (1998a) appear to be most different from the remaining sets. Average thresholds appear to reach a plateau around 16 kHz and above. Sakamoto et al. (1998a) attributed this plateau to be an artifact of an audible electrical noise floor. Reportedly, some of the subjects in the Sakamoto et al. study heard this intermittent noise during threshold measurement at the highest frequencies. These plateaus at the highest frequencies are present in the current data set as well, but without any complaints of audible noise or distortion at these frequencies from our subjects.

To verify that the data presented here were free of electrical or acoustic artifacts, the output spectrum of the hardware was measured in a standard ear simulator (IEC 60318-4; Bruel and Kjaer 4157) using the internal Bruel and Kjaer 4134 ½" microphone. Continuous tones generated using the MOTU digital audio interface at 0 dB attenuation at all test frequencies were produced from the test hardware and spectral averages were measured using a second computer that contained an RME Multiface II digital audio interface. The results at 19 and 20 kHz are presented in Figure 11 and exemplify the extreme of artifacts in the system. Hearing thresholds at each frequency representing the first and twenty-fifth percentiles of the data set are also reported in the same figure. As is evident in the figure, the system artifacts were considerably lower in amplitude than the lower limit of thresholds at high frequencies.

We attribute the plateaus at the highest frequencies to the preponderance of non-responses in this as well as previous data sets, in agreement with Stelmachowicz et al., (1989). The filled circles in each panel of Figure 10 support this observation as the percentage of non-responses increases steadily with increasing age and frequency reaching a high of 87% for all age groups at 19 and 20 kHz. Finally, the difference in average thresholds at the highest frequencies between the current study and two previous studies is related to the upper limit of the output of equipment used. Delivering 130 or 150 dB SPL at 20 kHz to the human eardrum at 20 kHz without distortion is a difficult, if not impossible task in our experience. However, whether those levels were truly produced and delivered may be somewhat irrelevant as the majority of the subjects in the Sakamoto et al., and the Stelmachowicz et al. studies did not respond to the stimuli at 19 and 20 kHz even at these elevated levels.

The surprising correspondence of the average hearing thresholds in age group across the four studies may be unexpected, given the differences in calibration, equipment, etc. However, it should be noted that these similarities are demonstrated for averaged data, where variability due to individual differences in ear canal geometry have been eliminated by construction.

E. Reliability of threshold measurement

Various methods, ranging from a yes-no one-interval procedure to an alternative forced choice procedure to the modified Hughson-Westlake procedure, have been employed for the measurement of hearing thresholds at EHF. The majority of the data sets available for comparison with the current data were obtained using a modified Hughson-Westlake procedure (American National Standards Institute, 1997; American Speech-Language-Hearing Association, 1978; Carhart and Jerger, 1959), which is a combination of ascending and descending methods of limits.

Using the modified Hughson-Westlake procedure and calibration in a standard cavity using a flat-plate coupler, Frank and Dreisbach (1991) reported four estimates of thresholds between 10–18 kHz, measured using a Beltone 2000 audiometer and Sennheiser HD 250 circumaural earphones, to be within ± 10 dB for at least 94% of ears. Schmuziger et al. (2004) reported test-retest variation for 0.5–16 kHz in 99% of ears tested to be within ~ 10 dB and additionally, no differences in variability between thresholds measured using Sennheiser HDA 200 and Etymotic ER-2 insert earphones. The standard error between trials was between 3 and 6 dB and increased with increasing frequency when the same measurement procedure but an improved calibration method indirectly estimating the sound pressure level at the medial end of the ear canal was used (Stevens et al., 1987; Stelmachowicz et al., 1988). Arguably the improvement in accuracy of measurement across trials in the second group of studies was a direct consequence of the compensation for differences in ear canal geometry.

The standard deviation of threshold estimates across test sessions has been reported to be within 5 dB when using a Békésy procedure to measure hearing thresholds for frequencies up to 8 kHz (Erlandsson et al., 1980; Laroche and Héту, 1997; Richards and Dunn, 1974). Erlandsson et al. (1980) reported that the standard deviation could be further reduced when earphone position (TDH-49P) was fixed across trials. Using a two-alternative forced choice (2AFC) procedure with a three-down one-up adaptive rule (thresholds correspond to 79.4 % percent correct on the psychometric function) and a calibration procedure similar to Stevens et al. (1987), Zhou and Green (1995) reported standard deviations of threshold estimates less than 1 dB for both low and high frequencies. However, standard deviations of hearing thresholds at frequencies higher than 7 kHz increased to ~ 4 dB when the position of the sound source was intentionally altered across trials.

Test-retest threshold differences from forty six subjects are plotted in Figure 12. The standard deviation of the differences in threshold was typically 5 dB or smaller between 0.5 and 10 kHz, with somewhat larger variability outside this frequency range, not exceeding 10 dB from 0.125 to 17 kHz. The larger test-retest variability at lower frequencies is attributed to leaks precipitated by a suboptimal fit of the test probe to the ear canal. These test-retest threshold differences are comparable to those reported in previous studies using a flat-plate-coupler calibration but not better. In fact, greater variability in real ears might be expected from an insert earphone that has generally higher source impedance than circumaural or supra-aural headphones (Voss, et al., 2000b; Voss and Herrmann, 2005). Compensating for differences in the distance between the eardrum and the probe appears to have dealt with this problem fairly well.

Another factor that likely influenced the test-retest variability observed here is the variability in a subject's threshold responses when measured months apart, rather than within the same test session (e.g., Zhou & Green, 1995; Schmuziger et al., 2004). Indeed Green et al. (1987) reported the test-retest variability to be greater for a one-month retest compared to a one-hour retest. Additionally, we allowed the two tests on a given subject to be conducted by different testers as may be the case in a realistic clinical situation. This could have resulted in variations in response criteria induced by differences in instructions. Our measure of reliability is realistic in that these factors are certainly in play in real hearing clinics.

The depth-compensated ear simulator calibration we used thus appears to be a practical approach for application in the clinical setting, however three complications must be considered. First, the sound source calibration requires an ear simulator that is not universally available for routine use. Measuring a set of pressure responses with finely-graded insertion depths is time consuming and impractical for most clinicians. Both of these practical issues would in most cases need to be performed by service personnel who regularly calibrate audiometers. Second, the probe assembly must contain a microphone to measure the pressure response in each ear to estimate insertion depth and thus to select the appropriate ear simulator calibration. Conventional earphones and tubephones obviously do not have this capability. Finally, coupler or simulator calibration can be inaccurate in children, persons who have had temporal bone surgery or have certain structural anomalies (Voss, et al, 2000a). The primary advantage of using a probe assembly that can both deliver stimuli and measure the sound pressure in the ear canal allows a wide-range of tests to be performed with a single device. Our research demonstrates that reliable measurement of hearing thresholds over the full frequency range of human hearing is one such test.

SUMMARY

1. In agreement with previous reports hearing thresholds at extended high frequencies were found to be sensitive to age related changes in auditory function.
2. The corner frequency and 60-dB intercept evident in nearly all subjects were age-dependent.
3. In contrast with previous reports, no significant gender differences were found in average hearing thresholds and hearing threshold distribution.
4. Two aging processes, one faster than the other, appear to influence hearing thresholds in different frequency ranges.
5. The depth-compensated ear-simulator calibration method and the modified Békésy technique allow reliable measurement of hearing thresholds over the entire frequency range of human hearing.

Acknowledgments

The authors wish to thank Gayla Poling and three anonymous reviewers for their valuable comments. The authors also wish to thank Erica Choe, Helen Han, Lauren Hardies, Kelly Waldvogel, Coryn Weissinger, Darrin Worthington and Wei Zhao for help in data collection. Kathleen Dunckley and Lauren Calandruccio participated in many stimulating discussions related to these data. Vickie Hellyer managed research subject recruitment and participation. This research was supported by NIDCD grant R01 DC008420 and Northwestern University.

REFERENCES

- Ahmed HO, Dennis JH, Badran O, et al. High-frequency (10–18 kHz) hearing thresholds: reliability, and effects of age and occupational noise exposure. *Occup Med.* 2001; 51:245–258.
- American National Standards Institute. Maximum permissible ambient noise levels for audiometric test rooms. ANSI S3.1-1999-R2008. New York: American National Standards Institute, Inc.; 2008.
- Bunch CC. Age variations in auditory acuity. *Arch Otolaryngol.* 1929; 9:625–636.
- Burén M, Solem BS, Laukli E. Threshold of hearing (0.125–20 kHz) in children and youngsters. *Br J Audiol.* 1992; 26:23–31. [PubMed: 1586846]
- Carhart R, Jerger J. Preferred method for clinical determination of pure tone thresholds. *J Speech Hear Dis.* 1959; 24:330–345.
- Corliss LM, Doster MD, Simonton J, et al. High frequency and regular audiometry among selected groups of high school students. *J Sch Hlth.* 1970; 40:400–405.
- Cruikshanks KJ, Wiley TL, Tweed TS, et al. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study. *A J Epidem.* 1998; 148:879–886.
- Dorn PA, Piskorski P, Gorga MP, et al. Predicting audiometric status from distortion product otoacoustic emissions using multivariate analyses. *Ear Hear.* 1999; 20:149–163. [PubMed: 10229516]
- Dreschler WA, Hulst RJAMvd, Tange RA, et al. The role of high-frequency audiometry in early detection of ototoxicity. *IJA.* 1985; 24:387–395.
- Dunckley KT, Dreisbach LE. Gender effects on high frequency distortion product otoacoustic emissions in human. *Ear. Hear.* 2004; 25:554–564. [PubMed: 15604916]
- Erlandsson B, Håkanson H, Ivarsson A, et al. The reliability of Békésy sweep audiometry recording and effects of the earphone position. *Acta Oto-laryngologica.* 1980; 89:99–112.
- Farmer-Fedor BL, Rabbitt RD. Acoustic intensity, impedance and reflection coefficient in the human ear canal. *J Acoust Soc Am.* 2002; 112:600–620. [PubMed: 12186041]
- Fausti SA, Frey RH, Erickson DA, et al. A system for evaluating auditory function from 8000–20,000 HZ. *J Acoust Soc Am.* 1979; 66:1713–1718. [PubMed: 521555]
- Fausti SA, Rappaport BZ, Schechter MA, et al. An investigation of the validity of high-frequency audition. *J Acoust Soc Am.* 1982; 71:646–649. [PubMed: 7085969]
- Fletcher JL. Reliability of high-frequency thresholds. *J Audiol Res.* 1965; 5:133–137.
- Fletcher JL, Cairns AB, Collins FG, et al. High frequency hearing following meningitis. *J. Aud Res.* 1967; 7:223–227.
- Frank T. High-frequency hearing thresholds in young adults using an available audiometer. *Ear Hear.* 1990; 11:450–454. [PubMed: 2073979]
- Frank T, Dreisbach LE. Repeatability of High-Frequency Thresholds. *Ear Hear.* 1991; 12:294–295. [PubMed: 1783231]
- Gacek RR, Schuknecht HF. Pathology of presbycusis. *Int Audiol.* 1969; 8:199–209.
- Gilman S, Dirks DD. Acoustics of ear canal measurement of eardrum SPL in simulators. *J Acoust Soc Am.* 1986; 80:783–793. [PubMed: 3760332]
- Goodman SS, Fitzpatrick DF, Ellison JC, Jesteadt W, Keefe DH. High-frequency click-evoked otoacoustic emissions and behavioral thresholds in humans. *J Acoust Soc Am.* 2009; 125:1014–1032. [PubMed: 19206876]
- Gordon-Salant S. Hearing loss and aging: New research findings and clinical implications. *J Rehabil Res Dev.* 2005; 42:9–24. [PubMed: 16470462]

- Green DM, Kidd G Jr, Stevens KN. High-frequency audiometric assessment of a young adult population. *J Acoust Soc Am*. 1987; 81:485–494. [PubMed: 3558966]
- Groh D, Pelanova J, Jilek M, et al. Changes in otoacoustic emissions and high-frequency hearing thresholds in children and adolescents. *Hear Res*. 2006; 212:90–98. [PubMed: 16364580]
- Hallmo P, Sundby A, Mair IWS. Extended high-frequency audiometry: Air- and bone conduction thresholds, age and gender variations. *Scand Audiol*. 1994; 23:165–170. [PubMed: 7997833]
- Harris JD, Myers CK. Tentative Audiometric Threshold-Level Standards from 8 through 18 kHz. *J Acoust Soc Am*. 1971; 49:1971–1972.
- Hellström, PA. Individual differences in peripheral sound transfer function: relationship to NIHL. Vth International Symposium. Effect of Noise on Hearing; 12–15 May, 1994; Gothenburg, Sweden. 1994.
- Hellström PA. The relationship between sound transfer functions and hearing levels. *Hear Res*. 1995; 88:54–60. [PubMed: 8576003]
- Henry KR, Fast GA. Ultrahigh-frequency auditory thresholds in young adults: Reliable responses up to 24 kHz with a quasi-free-field technique. *IJA*. 1984; 23:477–489.
- Henry KR, Fast GA, Nguyen HH, et al. Extra-high-frequency auditory thresholds: Fine structure, reliability, temporal integration and relation to ear canal resonance. *IJA*. 1985; 24:92–103.
- International Electrotechnical Commission. Electroacoustics - Simulators of human head and ear - Part 1: Ear simulator for the measurement of supra-aural and circumaural earphones. 60318-1 ed2.0:2009. Geneva, Switzerland: International Electrotechnical Commission; 2009.
- International Electrotechnical Commission. Electroacoustics - Simulators of human head and ear - Part 4: Occluded-ear simulator for the measurement of earphones coupled to the ear by means of ear inserts 60318-4:2010. Geneva, Switzerland: International Electrotechnical Commission; 2010.
- International Standards Organization. Acoustics: reference zero for the calibration of audiometric equipment. Part 5: Reference equivalent threshold sound pressure levels for pure tones in the frequency range 8 kHz to 16 kHz. ISO 389-5:1998. Geneva, Switzerland; International Standards Organization; 1998.
- Jacobson EJ, Downs MP, Fletcher JL. Clinical findings in high frequency thresholds during known ototoxic drug usage. *J. Aud Res*. 1969; 9:379–385.
- Keefe DH. Othereflectance of the cochlea and middle ear. *J Acoust Soc Am*. 1997; 102:2849–2859. [PubMed: 9373972]
- Keefe DH, Schairer KS. Specification of absorbed-sound power in the ear canal: application to suppression of stimulus frequency otoacoustic emissions. *J Acoust Soc Am*. 2011; 129:779–791. [PubMed: 21361437]
- Laroche C, Héту R. A study of the reliability of automated audiometry by the frequency scanning method (AUDIOSCAN). *Audiology*. 1997; 36:1–18. [PubMed: 9063557]
- Laukli E, Mair IWS. High-frequency audiometry normative studies and preliminary experiences. *Scand Audiol*. 1985; 14:151–158. [PubMed: 4059853]
- Lee F, Matthews LJ, Dubno JR, et al. Longitudinal Study of Pure-Tone Thresholds in Older Persons. *Ear Hear*. 2005; 26:1–11. [PubMed: 15692300]
- Matthews LJ, Lee F, Mills JH, et al. Extended High Frequency Thresholds in Older Adults. *J Speech Lang and Hear Res*. 1997; 40:208–214. [PubMed: 9113871]
- Morton LP, Reynolds L. High frequency thresholds: variation with age and industrial noise exposure. *S Afr F Common Discord*. 1991; 38:13–17.
- Moscicki EK, Elkins EF, Baum HM, et al. Hearing loss in the elderly: an epidemiologic study of the Framingham Heart Study Cohort. *Ear Hear*. 1985; 6:184–190. [PubMed: 4043571]
- Northern JL, Downs MP, Rudmose W, et al. Recommended High-Frequency Audiometric Threshold Levels (8000–18 000 Hz). *J Acoust Soc Am*. 1972; 52:585–595.
- Osterhammel D. High frequency audiometry clinical aspects. *Scand Audiol*. 1980; 9:249–256. [PubMed: 7466287]
- Osterhammel D, Osterhammel P. High-frequency audiometry: Age and sex variations. *Scand Audiol*. 1979; 8:73–80. [PubMed: 515692]

- Pearson JD, Morrell CH, Gordon-Salant S, et al. Gender differences in a longitudinal study of age-associated hearing loss. *J Acoust Soc Am*. 1995; 97:1196–1205. [PubMed: 7876442]
- Reuter W, Schonfeld U, Mansmann U, et al. Extended high frequency audiometry in pre-school children. *Audiology*. 1998; 37:285–294. [PubMed: 9776205]
- Richards AM, Dunn J. Threshold measurement procedures in brief-tone audiometry. *J Speech Lang and Hear Res*. 1974; 17:446–454.
- Rosen S, Plester D, El-Mofty A, et al. High-frequency audiometry in presbycusis. *Arch Otolaryngol (Chicago)*. 1964; 79:18.
- Sakamoto M, Sugawara M, Kaga K, et al. Average Thresholds in the 8 to 20 kHz Range as a Function of Age. *Scand Audiol*. 1998a; 27:190–192.
- Sakamoto M, Sugawara M, Kaga K, et al. Average Thresholds in the 8 to 20 kHz Range in Young Adults. *Scand Audiol*. 1998b; 27:169–172. [PubMed: 9728777]
- Sataloff J, Vassallo L, Menduke H. Occupational hearing loss and high frequency thresholds. *Archs Envir Health*. 1967; 14:832–836.
- Schechter MA, Fausti SA, Rappaport BZ, et al. Age categorization of high-frequency auditory threshold data. *J Acoust Soc Am*. 1986; 79:767–771. [PubMed: 3958317]
- Scheperle RA, Neely ST, Kopun JG, Gorga MP. Influence of in situ, sound-level calibration on distortion-product otoacoustic emission variability. *J Acoust Soc Am*. 2008; 124:288–300. [PubMed: 18646977]
- Schmuziger N, Probst R, Smurzynski J. Test-Retest Reliability of Pure-Tone Thresholds from 0.5 to 16 kHz using Sennheiser HDA 200 and Etymotic Research ER-2 Earphones. *Ear Hear*. 2004; 25:127–132. [PubMed: 15064657]
- Schuknecht HF, Gacek MR. Cochlear pathology in presbycusis. *Ann Otol Rhinol Laryngol*. 1993; 102:1–16. [PubMed: 8420477]
- Seta E, Bertoli GA, Filipo R. High-frequency audiometry above 8 kHz comparative results of normative thresholds obtained with a headphone system and a quasi-free-field system. *IJA*. 1985; 24:254–259.
- Shaw GM, Jardine CA, Fridjhon P. A pilot investigation of high-frequency audiometry in obscure auditory dysfunction (OAD) patients. *Br J Audiol*. 1996; 30:233–237. [PubMed: 8879688]
- Siegel JH. Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes. *J Acoust Soc Am*. 1994; 95:2589–2597.
- Siegel, JH. Calibrating otoacoustic emission probes. In: Robinette, MS.; Glatke, TJ., editors. *Otoacoustic Emissions: Clinical Application (3rd Edition)*. New York: Thieme Medical Publishers, Inc.; 2007. p. 403-427.
- Silva IM, Ângela M, Feitosa G. High-frequency audiometry in young and older adults when conventional audiometry is normal. *Rev Bras Otorrinolaringol*. 2006; 72:665–672.
- Souza NN, Dhar S, Siegel JH. A critical test of alternate stimulus level measures for the human ear. *ARO Abst*. 2010; 33:14.
- Stevens SN, Berkovitz R, Kidd G Jr, et al. Calibration of ear canals for audiometry at high frequencies. *J Acoust Soc Am*. 1987; 81:470–484. [PubMed: 3558965]
- Stinson MR. The spatial distribution of sound pressure within scaled replicas of the human ear canal. *J Acoust Soc Am*. 1985; 78:1596–1602. [PubMed: 4067075]
- Stinson MR, Shaw EAG. Wave effects and pressure distribution in the ear canal near the tympanic membrane. *J Acoust Soc Am Suppl*. 1982; 71:S88.
- Stelmachowicz PG, Beauchaine KA, Kalberer A, et al. Normative thresholds in the 8- to 20-kHz range as a function of age. *J Acoust Soc Am*. 1989; 86:1384–1391. [PubMed: 2808912]
- Stelmachowicz PG, Beauchaine KA, Kalberer A, et al. The reliability of auditory thresholds in the 8- to 20-kHz range using a prototype audiometer. *J Acoust Soc Am*. 1988; 83:1528–1535. [PubMed: 3372868]
- Stelmachowicz PG, Beauchaine KA, Kalberer A, et al. High-frequency audiometry: Test reliability and procedural considerations. *J Acoust Soc Am*. 1989; 85:879–887. [PubMed: 2926002]
- Tange RA, Dreschler WA, Hulst RJ. The importance of high-tone audiometry in monitoring for ototoxicity. *Arch Otorhinolaryngol*. 1985; 242:77–81. [PubMed: 3899068]

- Valete-Rosalino CM, Rozenfeld S. Auditory screening in the elderly: comparison between self-report and audiometry. *Rev Bras Otorrinolaringol.* 2005; 71:193–200.
- Voss SE, Herrmann BS. How does the sound pressure generated by circumaural, supra-aural, and insert earphones differ for adult and infant ears? *Ear Hear.* 2005; 26:636–650. [PubMed: 16377999]
- Voss SE, Rosowski JJ, Merchant SN, et al. Middle ear pathology can affect the ear-canal sound pressure generated by audiologic earphones. *Ear Hear.* 2000; 21:265–274. [PubMed: 10981602]
- Voss SE, Rosowski JJ, Shera CA, et al. Acoustic mechanisms that determine the ear-canal sound pressures generated by earphones. *J Acoust Soc Am.* 2000; 107:1548–1565. [PubMed: 10738809]
- Wiley TL, Cruickshanks KJ, Nondahl DM, et al. Aging and High-Frequency Sensitivity. *J Speech Lang and Hear Res.* 1998; 41:1061–1072. [PubMed: 9771629]
- Withnell RH, Jeng PS, Waldvogel K, Morgenstein K, Allen JB. An in situ calibration for hearing thresholds. *J Acoust Soc Am.* 2009; 125:1605–1611. [PubMed: 19275318]
- Zhou B, Green DM. Reliability of pure-tone thresholds at high frequencies. *J Acoust Soc Am.* 1995; 98:828–836. [PubMed: 7642821]

SHORT SUMMARY

Behavioral hearing thresholds were measured between the frequencies of 0.125 to 20 kHz for 352 clinically normal hearing subjects, ranging in age between 10–65 years. A depth-compensated ear simulator method was used to calibrate stimulus levels at the eardrum, based on an acoustic estimate of the insertion depth of the measurement probe for each subject. The results were in general agreement with previous reports of hearing thresholds measured over this frequency range, albeit using less practical or less accurate calibration techniques. These results confirmed the sensitivity of hearing thresholds above 8 kHz to aging processes. They also revealed possible multiple, frequency-dependent aging processes that have not been reported before. In contrast to previous reports, age-matched groups of male and female subjects did not show statistically significant differences in average thresholds. Test-retest variability was evaluated in forty-nine subjects and found to be comparable to that reported in the literature.

Figure 1.

Depth-compensated ear simulator calibration. The frequency of the first half-wave resonance is used to characterize each member of the family of pressure responses measured using the simulator's microphone with the emission probe inserted to a range of depths that encompass the range of comparable depths in human ears (gray curves). The pressure response measured using the emission probe's microphone (red curve), normalized by its response in the long copper tube (green curve) removes irregularities due to the sound sources and allows the half-wave resonance to be measured accurately in the subject's ear (dashed red curve). The simulator response with the closest half-wave resonance (blue curve) is selected to represent the subject's eardrum pressure.

Figure 2.

Average hearing thresholds for different age groups. The error bars represent the 95% confidence interval. Data are plotted using a logarithmic frequency axis in (a) and a linear frequency axis in (b). The plot using a linear frequency axis is provided for greater visual clarity of the high frequencies.

Figure 3.

Comparison of average hearing thresholds of male and female subjects in each age group. The error bars represent 95% confidence interval. Data are plotted using a logarithmic frequency axis in (a) and a linear frequency axis in (b). The plot using a linear frequency axis is provided for greater visual clarity of the high frequencies. Data from each age group are displayed in an independent panel.

Figure 4.

An example of hearing thresholds measured over the entire test frequency range from one subject. Note the abrupt change in trajectory of thresholds marked by the solid arrow at 12.5 kHz. These junctions were identified as “corner frequencies” in each subject. The frequency at which the threshold crossed an arbitrary 60 dB SPL boundary was identified as the “60-dB intercept” and is marked by the dashed arrow at 17 kHz.

Figure 5.

Average “corner frequencies” and “60-dB intercept” for different age groups. See text for description of method of estimating these parameters. The x error bars represent the 95% confidence interval in frequency, while the y error bars represent the 95% confidence interval in thresholds at the corner frequency.

Figure 6.

Hearing thresholds from individual subjects plotted as a function of age. Data recorded at different frequencies are displayed in individual panels. The solid lines represent a one- or two-line linear regression fit to the data. The dashed lines represent a similar fit after eliminating all non-responses.

Figure 7.

Female (open symbols) and male (filled symbols) hearing thresholds from individual subjects plotted as a function of age for 4, 6, 11.2 and 12.5 kHz. The solid lines represent a one or two-line linear regression fit to the data: gray lines for female and black lines for male.

Figure 8.

The “transition age” where the two segments of a two-line fit intersect in Figures 5 and 6, plotted as a function of frequency. The transition age is only available where the two-line fit was found to be the optimal solution even after eliminating the non-responses.

Figure 9. Hearing threshold distribution for Female (F) and male (M) subjects for each age group (A,B, C, D, and E). The distributions were fitted using a Kernel density function. Star symbols indicate the comparisons where gender effects were statistically significant.

Figure 10.

Comparison of average data from current data set with three previous studies represented by the lines and various symbols. The dotted lines with filled circles represent the percentage of non-responses at a given frequency for a particular age group.

Figure 11.

Spectra of system output at 0 dB attenuation for 19 and 20 kHz tones. Hearing thresholds representing the first and twenty-fifth percentiles at each frequency are also presented.

Figure 12.

Test-retest threshold differences in dB measured in 46 subjects. The black line connects the average test-retest variation across frequency. The error bars represent ± 1 standard deviation.

Table 1

Mean and 95% confidence intervals (in *italics*) for hearing thresholds (in dB SPL) of subjects from different age groups. The number on top is the mean and the values in the bottom of each cell represent the 95% confidence interval.

	10–21 yrs	22–35 yrs	36–45 yrs	46–55 yrs	56–65 yrs
125	34.83 31.57 38.08	39.19 37.24 41.15	36.46 33.80 39.13	38.76 35.40 42.11	38.16 34.90 41.42
250	25.40 23.57 27.23	26.55 25.21 27.88	24.68 22.87 26.49	26.56 24.08 29.05	26.44 24.52 28.36
500	15.17 13.77 16.58	15.66 14.43 16.90	15.76 14.13 17.40	16.42 14.97 17.86	17.54 15.74 19.34
750	11.67 9.66 13.69	12.13 10.63 13.64	13.79 9.22 18.36	12.99 10.84 15.13	14.09 11.97 16.20
1000	10.52 9.12 11.93	11.45 10.40 12.49	12.72 8.90 16.54	13.16 11.35 14.98	13.20 11.45 14.95
1500	14.96 13.07 16.85	15.11 13.80 16.41	15.89 11.21 20.57	15.36 13.45 17.26	17.39 14.89 19.89
2000	17.22 15.66 18.79	16.83 15.66 18.00	18.56 16.92 20.19	19.10 17.18 21.01	21.41 19.58 23.23
3000	15.24 13.83 16.66	15.65 14.56 16.74	17.75 15.88 19.61	20.94 18.92 22.96	25.02 22.48 27.56
4000	12.67 11.22 14.12	13.97 12.52 15.43	16.27 13.99 18.99	20.97 18.41 23.54	26.91 23.89 29.93
6000	14.25 12.58 15.91	15.93 14.23 17.63	20.90 18.68 23.11	26.19 23.76 28.62	36.52 31.97 41.07
8000	16.35 14.36 18.33	18.44 16.02 20.87	22.10 19.26 24.93	28.39 25.58 31.21	42.76 37.92 47.60
10000	20.44 18.16 22.72	20.69 18.29 23.09	27.87 24.66 31.09	37.01 32.40 41.61	54.92 48.80 61.04
11200	22.99 19.55 26.43	24.30 21.21 27.39	31.20 26.80 35.60	44.46 38.60 50.31	65.75 58.75 72.75
12500	27.16 24.16 30.17	29.56 26.43 32.70	43.99 38.85 49.13	57.33 51.61 63.05	85.33 79.74 90.92

	10–21 yrs	22–35 yrs	36–45 yrs	46–55 yrs	56–65 yrs
14000	33.43	39.44	67.12	80.23	99.00
	29.46	35.61	60.78	74.80	85.65
	40.69	51.14	82.57	95.90	104.50
15000	36.11	45.28	75.89	91.42	103.80
	48.00	63.03	92.14	100.68	104.73
16000	43.44	58.45	87.18	97.82	103.54
	64.70	80.12	98.32	103.77	105.00
17000	58.99	70.42	94.11	102.54	105.00
	83.45	94.95	100.35	103.49	104.25
18000	78.43	88.47	97.04	103.66	103.22
	93.81	99.71	102.41	104.08	104.38
19000	89.32	98.30	99.87	104.95	102.81
	93.07	95.19	101.93	102.19	103.61
20000	88.55	97.60	98.84	105.03	99.79
					104.59
					101.63
					105.58

Table II

Frequencies at which significant differences in the distribution of hearing thresholds between genders were detected in each age group ($p < 0.05$, marked by asterisks). Results for analyses conducted with both unmatched and matched sample sizes are presented.

Frequency(kHz)	Age group	Unmatched (p value)	Matched (p value)
0.125	A	0.039*	0.095
0.125	D	0.021*	0.074
0.25	E	0.222	0.035*
1.5	A	0.167	0.017*
2	B	0.007*	0.033*
4	B	0.004*	0.021*
4	E	0.005*	0.068
6	D	0.013*	0.035*
8	D	0.045*	0.560
8	B	0.115	0.033*

Table III

Average test-retest threshold differences from 46 subjects. Standard deviation (SD) values as well as the number of threshold comparisons available (n) at each frequency are also presented.

Frequency(kHz)	Average (dB)	SD	n
0.125	-0.91	8.32	46
0.25	0.53	5.89	45
0.5	1.24	5.34	45
0.75	0.61	3.94	45
1	0.12	3.62	45
1.5	0.51	4.79	43
2	0.42	3.36	43
3	0.17	4.53	44
4	0.08	4.34	46
6	1.48	4.28	45
8	1.08	4.90	41
10	3.72	6.68	36
11.2	3.45	7.36	34
12.5	4.22	5.27	30
14	4.15	6.50	16
15	1.51	3.77	9
16	0.77	1.84	6
17	-0.50	2.30	5
18	0.07	8.97	3
19	-3.58	11.84	3
20			0