

Published in final edited form as:

Ear Hear. 2008 December ; 29(6): 907–922. doi:10.1097/AUD.0b013e31818246f6.

Spectro-Temporal Characteristics of Speech at High Frequencies, and the Potential for Restoration of Audibility to People with Mild-to-Moderate Hearing Loss

Brian C. J. Moore¹, Michael A. Stone¹, Christian Füllgrabe¹, Brian R. Glasberg¹, and Sunil Puria^{2,3}

¹Department of Experimental Psychology, University of Cambridge, Cambridge, England

²EarLens Corporation, 101 Saginaw Dr., Redwood City, CA 94063, U.S.A ³Department of Mechanical Engineering and Otolaryngology - Head and Neck Surgery, Stanford University, Stanford, CA.

Abstract

Objectives—It is possible for auditory prostheses to provide amplification for frequencies above 6 kHz. However, most current hearing-aid fitting procedures do not give recommended gains for such high frequencies. This study was intended to provide information that could be useful in quantifying appropriate high-frequency gains, and in establishing the population of hearing-impaired people who might benefit from such amplification.

Design—The study had two parts. In the first part, wide-bandwidth recordings of normal conversational speech were obtained from a sample of male and female talkers. The recordings were used to determine the mean spectral shape over a wide frequency range, and to determine the distribution of levels (the speech dynamic range) as a function of center frequency. In the second part, audiometric thresholds were measured for frequencies of 0.125, 0.25, 0.5, 1, 2, 3, 4, 6, 8, 10, and 12.5 kHz for both ears of 31 people selected to have mild or moderate cochlear hearing loss. The hearing loss was never greater than 70 dB for any frequency up to 4 kHz.

Results—The mean spectrum level of the speech fell progressively with increasing center frequency above about 0.5 kHz. For speech with an overall level of 65 dB SPL, the mean 1/3-octave level was 49 and 37 dB SPL for center frequencies of 1 and 10 kHz, respectively. The dynamic range of the speech was similar for center frequencies of 1 and 10 kHz. The part of the dynamic range below the root-mean-square level was larger than reported in previous studies. The mean audiometric thresholds at high frequencies (10 and 12.5 kHz) were relatively high (69 and 77 dB HL, respectively), even though the mean thresholds for frequencies below 4 kHz were 41 dB HL or better.

Conclusions—To partially restore audibility for a hearing loss of 65 dB at 10 kHz would require an effective insertion gain of about 36 dB at 10 kHz. With this gain, audibility could be (partly) restored for 25 of the 62 ears assessed.

Introduction

Although the feasibility of wide-bandwidth hearing aids was demonstrated many years ago (Killion & Tillman, 1982), until recently most hearing aids provided gain only for frequencies up to 5 to 6 kHz (Dillon, 2001; Moore, et al., 2001). Because of this, most methods for fitting hearing aids, such as the Desired Sensation Level multistage input/output algorithm [DSL (i/o)] (Cornelisse, et al., 1995; Scollie, et al., 2005), NAL-NL1 (Byrne, et al., 2001), CAMEQ (Moore, 2005; Moore, et al., 1999), and CAMREST (Moore, 2000; 2005) give recommended gains only for frequencies up to 6 kHz. In recent years, it has become more common for hearing aids to provide gain at frequencies above 6 kHz, for example by use of a receiver in the ear canal, sometimes in combination with digital feedback cancellation. Another possibility for providing gain for frequencies above 6 kHz is provided by the “EarLens” system (Perkins, 1996). In the latter, a small magnet is embedded in a custom-made soft silicone “lens,” resembling a contact lens but being smaller, which is placed on the eardrum; the magnet is then “driven” by passing electrical current through a small coil in the ear canal, which causes the eardrum to vibrate. The maximum amount of vibration is equivalent to sound levels at the eardrum above 80 dB SPL for frequencies up to at least 10 kHz. In principle, this makes it possible to restore audibility of sounds with frequencies above 6 kHz, at least for people with mild-to-moderate hearing loss.

The main purpose of hearing aids, and of the EarLens system, is to improve speech intelligibility by increasing the audibility of speech sounds. However, the gains that are required at high frequencies to restore audibility of speech to a hearing-impaired person have not been determined. One reason for this is that the typical spectral shape of speech at high frequencies is not well established. Most studies of the spectral shape of speech have only considered frequencies up to about 8 kHz (Byrne, 1977; Cox & Moore, 1988; Olsen, et al., 1987; Pearsons, et al., 1976). Dunn and White (1940) measured the level of speech at the output of a highpass filter with a cutoff frequency of 8000 Hz. They reported that “we have given this band a practical upper limit of 12,000 cycles, due to the very small energy remaining in speech above that point” (page 278). Very few studies have presented information about the spectral shape of speech at higher frequencies. The distribution of levels at high frequencies is even less well established. One of the most comprehensive studies of the spectro-temporal characteristics of speech (Byrne, et al., 1994) did include measurements for frequencies above 6 kHz, but the data were obtained from cassette recordings, and the measurements at high frequencies were almost certainly affected by recording noise and by uncertainties in calibration. The authors stated that the results for center frequencies of 12 and 16 kHz should “be treated with some caution” (page 2111). One goal of the present study was to survey existing data, to make new wide-bandwidth low-noise recordings of speech from a representative sample of talkers, and to analyze those recordings so as to obtain a description of the typical spectro-temporal characteristics of speech at high frequencies.

Knowledge is also limited about the characteristics of audiometric thresholds at high frequencies. Most surveys of hearing loss have obtained audiometric thresholds for frequencies up to 8 kHz (Davis, 1995). Although there have been studies of absolute thresholds at high frequencies in populations of young people with normal hearing (Green, et al., 1987; Kurakata, et al., 2005; Northern, et al., 1972; Schechter, et al., 1986), or as a function of age in a “normal” population (Stelmachowicz, et al., 1989), relatively little is known about the typical degree of hearing loss at frequencies above 8 kHz in people with mild-to-moderate hearing loss at frequencies below 8 kHz. A second goal of the present study was to provide such information. The information from the two parts of the study was used to calculate insertion gains and output levels which will be required at high frequencies

to restore audibility of a reasonable proportion of the dynamic range of speech spoken at a normal conversational level.

Spectro-Temporal Characteristics of Speech at High Frequencies

Recordings of Speech

Recordings were made from nine male and eight female talkers. Their ages ranged from 21 to 53 years (mean = 36, SD = 13) for the males and from 17 to 44 years (mean = 24, SD = 8) for the females. All except one were native speakers of British English. One man had an Australian accent. The talkers were seated in a double-walled sound-attenuating booth with dimensions 2.7 m wide by 3.4 m long by 2.35 m high. All surfaces were lined with 10-cm thick sound-absorbing foam (there was a suspended mesh floor that was covered with thick underlay and carpet). There was a 1.5 cm air gap behind the foam, which was fixed to battens on the brick material of the walls. The foam was covered with an additional sound-absorbing quilt and decorative fabric. The room was designed to have minimal reflections and low background noise. There was some noise at very low frequencies caused by building vibration and the ventilation system, but this was largely filtered out during the analyses described below.

The microphone was a high-quality studio model made by Sennheiser (MKH 40 P 48 U3). Its on-axis frequency response was within ± 2 dB from 40 to 20,000 Hz. Its equivalent input noise level was 12 dB SPL (A-weighted broadband). The microphone was approximately 30 cm from the mouth of the talker in the horizontal direction and 15 cm below the mouth of the talker, and faced the talker's mouth. For all the female talkers and for six of the male talkers, the output of the microphone was fed to a Mackie 1202 VLZ Pro mixing desk, which acted as a buffer and preamplifier, and then to the analog-to-digital converter of an M-Audio 24-bit sound card, which was external to the personal computer to which it was connected. The resulting digitized signal was recorded on the hard drive of the personal computer. For three of the male talkers, recordings were made at an earlier time by recording directly to a digital-audio-tape recorder with high-quality 16-bit digital-to-analog converters (Sony DTC1000 ES, modified by HHB Professional Audio). All recordings were made at the standard sampling rate of 44.1 kHz, giving a bandwidth of about 20 kHz. The talkers were asked to read from a prose passage in a normal conversational manner. At least 55 sec (usually 114 sec or more) of speech was recorded for each talker.

We also analyzed high-quality recordings of the single male and female talkers from the compact disc produced by Bang and Olufsen called "Music for Archimedes" [connected discourse (CD)B&O 101]. The recordings in this case were taken under "moderately reverberant" conditions (track 6 for the female and track 7 for the male).

Analysis of Long-Term Average Spectra

The recordings were edited "by hand" using "CoolEdit" to remove pauses for breath and extraneous sounds, but natural sounding gaps between sentences of between 100 and 300 msec were left. The spectra were corrected to allow for the small deviations of the response of the measurement microphone from a "flat" response. The spectrum level (the level in a 1-Hz wide band re 20 μ Pa) at the center frequency of each 1/3-octave band was determined when the overall speech level was scaled to have a level of 65 dB SPL. This level is typical for normal conversational speech when the talker and listener are separated by about 1 m (Pearsons, et al., 1976), although Killion (1997) reported a range of 55 to 95 dB SPL "in normal group conversations, between offhand remarks and forceful arguments." The spectra were averaged separately for the male and female talkers, and the results are shown in Table 1. This table also shows long-term average spectra reported by Dunn and White (1940), Pearsons, et al. (1976), Byrne, et al. (1994), and taken from the Archimedes CD.

For male speech, the spectrum level was reasonably constant at low frequencies for all sources of data, being about 31 to 38 dB for frequencies up to 500 Hz. Above 500 Hz, the spectrum level declined progressively, reaching 0 dB between 6300 and 12500 Hz, depending on the source of the data. The SD of the spectrum level across male talkers recorded at Cambridge was typically 2 to 3 dB for center frequencies from 125 to 2000 Hz, and was somewhat larger (3–5 dB) for center frequencies outside this range. For female speech, the spectrum level was typically lower than for male speech at very low frequencies (100 and 125 Hz). For center frequencies from 200 to 500 Hz, the spectrum level for female speech was similar to that for male speech, being between 30 and 41 dB across sources. The spectrum level again declined with increasing frequency, but reached 0 dB at a somewhat higher frequency than for male speech. The SD of the spectrum level across female talkers recorded at Cambridge was typically 2 to 3 dB for center frequencies from 250 to 5000 Hz, but was 3 to 5 dB for higher center frequencies, and 4 to 6 dB for lower center frequencies.

Overall, the data were reasonably similar across the different sources, although the spectrum for the female talker from the Archimedes CD showed an unusually high spectrum level at 100 Hz and an atypical peak around 8000 Hz. The mean spectrum for male and female voices obtained in the present study was similar to that obtained by Byrne, et al. (1994) for center frequencies from 200 to 6300 Hz. For the male voices, the difference over this frequency range was never greater than 2.6 dB. For the female voices, the difference was never greater than 4.7 dB. The mean spectrum is also similar to that specified in the American National Standards Institute (ANSI, 1997) standard for typical speech, which is characterized as being “flat” up to 500 Hz, and to decline at 9 dB/octave above that. This confirms that the spectra are consistent with earlier measurements over the frequency range for which the earlier measurements were reliable.

To obtain representative spectra for the current study, the spectrum level at each frequency was averaged across all of the talkers recorded at Cambridge. The result is shown in the second column of Table 2. The upper panel of Figure 1 compares the composite male/female spectrum determined here with that determined by Byrne, et al. (1994). Over the frequency range where the results of Byrne, et al. were probably relatively unaffected by noise introduced by the microphone or recording, namely 100 to 8000 Hz, the spectrum levels determined in the present study are similar to those determined by Byrne, et al. except that the present values are about 3 dB higher in the range 2000 to 3200 Hz. At 10,000, 12,500, and 16,000 Hz, the spectrum levels determined here lie below those determined by Byrne, et al. by 1.1, 3.3, and 6.5 dB, respectively. The discrepancy at high frequencies probably reflects the influence of microphone and recording noise on the data of Byrne, et al. The microphone used in the study of Byrne, et al. had an equivalent input noise level of 25 dB SPL, about 13 dB higher than for the microphone used in the present study. Also, the cassette recorder used by Byrne, et al. would have led to a higher background noise level in the recordings than for the present study.

The third column of Table 2 shows the spectrum levels converted to sound pressure levels in each 1/3-octave band. The fourth column shows the diffuse-field to eardrum transfer function as estimated by Moore, et al. (1997) and as specified in ANSI (2007). The fifth column shows the estimated 1/3-octave band levels at the eardrum, obtained by adding the numbers in columns three and four. It is interesting that the 1/3-octave band levels at the eardrum vary over a relatively narrow range, 52 to 61 dB SPL, for center frequencies from 200 to 5000 Hz, the range that is most important for speech intelligibility (ANSI, 1969; 1997). The sixth column shows the estimated normal absolute threshold, specified as the minimum audible pressure (MAP) at the eardrum for monaural listening. This was estimated using the model described by Moore and Glasberg (2007), but with the middle-ear transfer function specified in Glasberg and Moore (2006). The equivalent rectangular bandwidths of

the auditory filters of normally hearing listeners at moderate sound levels, ERB_N , are somewhat smaller than 1/3-octave at medium and high frequencies (Glasberg & Moore, 1990), but the difference is relatively small, equivalent to a difference in level of 2 to 3 dB. To a first approximation, the normal auditory system integrates energy over 1- ERB_N -wide bands for the purpose of signal detection (Moore, et al., 1997). Hence, the audibility of the speech in each band, for normal-hearing listeners, can be estimated by comparing the numbers in columns 5 and 6. For center frequencies from 160 to 10000 Hz, the estimated 1/3-octave root-mean-square (RMS) level of the speech at the eardrum is >29 dB above the MAP, which means that the majority of the speech spectrum would be audible. Only at the highest center frequency, for which the MAP rises markedly, is the mean speech level below the MAP.

Another way of estimating audibility as a function of frequency, for normal-hearing listeners, is to calculate the excitation pattern of the average speech spectrum. Again, this was estimated using the model described by Moore and Glasberg (2007), but with the middle-ear transfer function specified in Glasberg and Moore (2006). The results for the speech spectra of Byrne, et al. and from the present study are shown in the lower panel of Figure 1. The figure also shows the excitation level at the absolute threshold for monaural listening (dashed line). Again, this figure leads to the conclusion that the average excitation evoked by speech with an overall level of 65 dB SPL is well above the threshold excitation over a wide frequency range.

Distribution of Speech Levels

In this section we consider the dynamic range of the speech, that is the distribution of levels, and whether it varies with center frequency. Before estimating this, it was necessary to choose appropriate analysis bandwidths and durations within which the distribution of levels would be measured. Hearing-impaired people usually have broader auditory filters than people with normal hearing (Glasberg & Moore, 1986; Moore, 2007; Pick, et al., 1977). Typically, for people with mild-to-moderate hearing loss, the equivalent rectangular bandwidth of their auditory filters is about 2 ERB_N (Moore, 2007). This means that stimulus energy is integrated, for the purpose of detection, over a frequency range of about 2 ERB_N . We decided to use analysis bands of this width to determine the distribution of speech levels. However, the distribution does not depend strongly on the bandwidths chosen for analysis, and very similar results would be obtained using 1- ERB_N -wide or 1/3-octave-wide bands. The 2- ERB_N -wide bands used here are similar to the commonly used 1/3-octave-wide bands for medium and high center frequencies. For example, for a center frequency of 2000 Hz, 2 ERB_N corresponds to 481 Hz, whereas 1/3 octave corresponds to 460 Hz.

The speech was filtered into 2- ERB_N -wide bands using an array of finite-impulse response filters. The filters had a variable number of taps, chosen so that the transition bands of each filter had similar slopes when plotted on a logarithmic frequency scale. Each filter was designed to have a response of -6dB at the frequencies at which its response intersected with the responses of the two adjacent filters. The crossover with the next-but-one filter was below the -50-dB point. The stopband response was below -70 dB. Initially, the distribution of levels in each band was measured using 125-msec rectangular windows, which were moved in time in steps of 125/3 msec. Windows of this duration have often been used in the past to characterize the distribution of levels in speech (Dubno, et al., 1989; Dunn & White, 1940; French & Steinberg, 1947; Kryter, 1962; Pavlovic, 1987; Pavlovic & Studebaker, 1984). However, it is not obvious that this window duration is appropriate. Shorter windows result in a wider measured range of levels (Cox, et al., 1988; Levitt, 1982; Pavlovic, 1993; Rhebergen, 2006). Also, shorter windows are often used in hearing-aid signal-processing algorithms. We consider later the effect of using a shorter window.

Figures 2 and 3 show examples of the distributions for a female talker (Fig. 2) and a male talker (Fig. 3). Figure 4 shows average results for all of the eight female talkers recorded at Cambridge and Figure 5 shows average results for the six male talkers whose voices were recorded using the 24-bit system. The abscissa shows the center frequency of the 2-ERB_N-wide band. The ordinate shows relative level in dB. The dashed line shows the relative RMS level in each band, scaled so that the level of the band with the highest level is set to 0 dB. This line indicates the spectral shape of the speech for that specific talker or group of talkers. We use the word “exceedance” to refer to the percentage of time that the level in a given band exceeds a certain level, relative to the RMS level in that band. The solid lines with circles show “equal-exceedance contours” (EECs). For example, the line labeled 20 is the 20% EEC; it shows the level that was exceeded for 20% of the windows. In Figures 2 to 5, this EEC is approximately horizontal, and it lies close to 0 dB on the ordinate, although for some talkers there was a small (5–6 dB) dip around 0.7 to 1.6 kHz and around 6 to 10 kHz. Overall, the relatively flat EEC indicates that the RMS level was exceeded for roughly 20% of the 125-msec windows, independent of the center frequency. Similar patterns were reported by Fletcher (1953) and by Cox, et al. (1988) for speech analyzed using 125-msec or 120-msec windows and by Pavlovic (1993) for speech analyzed using 13, 80, or 200-msec windows.

The top EEC shows the level that was exceeded for only 1% of the windows. It tended to lie <10 dB above the RMS level for low center frequencies, but lay slightly >10 dB above the RMS level for high center frequencies. This was also true for the data reported by Dunn and White (1940), Cox, et al. (1988), and Byrne, et al. (1994). This means that the part of the dynamic range of speech that lies above the RMS level is slightly greater in extent for high than for low center frequencies.

For exceedances >20%, the EECs were often not horizontal. The rise in the EECs for frequencies above 10 kHz for both male and female talkers, and for the lowest center frequency for the female talkers, was probably caused by noise in the recordings; when the RMS level in a band was very low, this noise prevented the level dropping below a certain “floor.” However, the nonflat EECs for center frequencies between 0.2 and 10 kHz cannot be accounted for by noise in the recordings. For exceedances between 90 and 50%, the EECs often showed a broad dip around 0.75 to 1.25 kHz and a peak in the range 2 to 6 kHz. Again, similar patterns were reported by Dunn and White (1940) and by Cox, et al. (1988). These results indicate that the part of the dynamic range of speech below the RMS level is widest around 0.75 to 1.25 kHz, and narrowest around 2 to 6 kHz, with some variation across talkers.

Our results differ from some earlier results in the position of the EECs for high exceedance values. For example, the 99% EECs in our data mostly lay below –40 dB for center frequencies from 0.2 to 10 kHz. In contrast, in the mean data of Byrne, et al. (1994) the 99% EEC lay at about –30, –30, and –23 dB for center frequencies of 0.4, 1, and 4 kHz, respectively. The difference again almost certainly reflects the influence of microphone and recording noise on the data of Byrne, et al. Neither Dunn and White (1940) nor Cox, et al. (1988) showed a 99% EEC because of acknowledged limitations in the analysis or recording equipment; Dunn and White used an analog “live” analysis method with a limited dynamic range, and the digital recording system used by Cox, et al. had only 12-bit resolution.

Figures 6 and 7 are similar to Figures 4 and 5, but they show the outcomes when the window duration was decreased from 125 msec to 10 msec. This latter value is comparable with the width of the auditory temporal window, as determined in experiments using combined forward and backward masking (Moore, et al., 1988; Oxenham & Moore, 1994; Plack & Moore, 1990). The shorter window is also comparable with the frame sizes used for signal

processing in digital hearing aids (Hamacher, et al., 2005). Decreasing the window duration had only a small effect on the distribution of levels above the RMS value; the 1% EEC increased by 1 to 2 dB. The effect was greater for levels below the RMS value. For example, the 90% EEC was about 15 dB lower for the 10-msec windows than for the 125-msec windows.

The limits of the dynamic range of speech are sometimes taken as defined by the 99% EEC at the lower end and the 1% EEC at the upper end (Rhebergen, 2006). However, the 99% EEC may well reflect the brief pauses in the speech. Inspection of our edited files suggested that about 5% of the total duration consisted of pauses. It seems more realistic to take the 95% EEC as the lower end of the dynamic range. With this definition, our data suggest that the dynamic range extends from about -40 to $+11$ dB relative to the RMS value in a given band if 125-msec windows are used. If 10-msec windows are used, the dynamic range increases, extending from about -52 to $+12$ dB. The dynamic range for a center frequency of 10 kHz was roughly the same as for a center frequency of 1 kHz.

It is of interest to compare the physical dynamic range measured here with the “perceptual dynamic range” (Pavlovic, 1993) that is used for calculating the articulation index (ANSI, 1969) or speech intelligibility index (ANSI, 1997). The dynamic range in the standards was chosen to give good predictions of data on speech intelligibility based on the assumption that each part of the dynamic range is equally important (Pavlovic & Studebaker, 1984). This assumption is only an approximation (Rhebergen, 2006). The upper limit of the dynamic range found here is similar to that assumed in the standards. However, the part of the dynamic range below the RMS level is wider than the value of -15 to -18 dB assumed in the standards (ANSI, 1969; 1997). Some data on speech intelligibility suggest that the perceptual dynamic range of speech in each frequency band is closer to 50 dB than to 30 dB (Dubno & Ahlstrom, 1995; Dubno & Schaefer, 1995; Studebaker & Sherbecoe, 2002; Studebaker, et al., 1999). If this is the case, then there may be a reasonably close correspondence between the physical and perceptual dynamic ranges.

Figure 8 compares the distribution of levels for 2-ERB_N-wide bands centered at 1.26 kHz (solid lines) and 10 kHz (dashed lines) for the female (top) and male (bottom) talkers, respectively. Here, the data are presented as histograms of levels. The “bin width” used in the analysis was 0.4 dB, but the histograms were smoothed by taking a running average across five adjacent bins. In the figure, the ordinate is scaled so that the histograms represent the percentage of frames per dB. For example, if the ordinate value is 3 at a level of -18 dB relative to the RMS value, this means that three percent of frames fall within the range -18.5 to -17.5 dB. The left panels show histograms based on the use of 125-msec windows, as was also the case for the analyses presented in Figures 2 to 5. The right panels show histograms based on 10-msec windows.

For both window durations, the main peak of the histograms is higher and sharper at 10 kHz than at 1.26 kHz. In other words, the distribution of levels is more “focused” at 10 kHz. At 10 kHz, the peaks of the histograms lie about 12 to 17 dB below the RMS level, for both genders. At 1.26 kHz, the peaks of the histograms fall at lower values, in the range 16 to 25 dB below the RMS value. For the 10-msec windows and the 10-kHz center frequency, the histograms show a secondary peak at about 50 dB below the RMS level. This peak represents noise in the recordings with a small contribution from low-level wheezing or breathing noise in the brief pauses in the speech that remained after hand editing. The part of the histograms that was probably unaffected by such noise extends from about -43 dB to $+18$ dB relative to the RMS level. Overall, these histograms confirm that the part of the dynamic range of speech below the RMS level is wider than the value of -15 dB assumed in the ANSI (1997) standard for calculating the speech intelligibility index.

Characteristics of Audiometric Thresholds at High Frequencies

Acquired hearing loss in adults typically is greater at high frequencies than at low frequencies (Davis, 1995). Hearing-impaired people who might benefit from amplification of frequencies above 6 kHz would typically have only mild or moderate losses at lower frequencies. Hence, we focused our study on the audiometric characteristics of such people. Audiometric thresholds were measured for frequencies of 0.125, 0.25, 0.5, 1, 2, 3, 4, 6, 8, 10, 12.5, and 14 kHz for both ears of 31 people selected to have mild or moderate cochlear hearing losses. These were recruited at the House Ear Institute (Los Angeles, CA), EarLens Corporation (Redwood City, CA), and the Department of Experimental Psychology (Cambridge, UK). In all cases, thresholds were measured using a Grason-Stadler GSI61 Audiometer with the high-frequency option. Either Telephonics TDH50P headphones or Sennheiser HDA200 headphones were used for audiometric frequencies up to 8 kHz. Sennheiser HDA200 headphones were always used for higher frequencies. The method was that recommended by the British Society of Audiology (2004).

The hearing loss was never greater than 70 dB for frequencies up to 4 kHz. The pure-tone-average hearing loss for the frequencies 0.5, 1, and 2 kHz had an average value of 25 dB for the left ear and 24 dB for the right ear. The means and SDs of the audiometric thresholds are shown in Table 3. Results for the frequency of 14 kHz are omitted, because thresholds at this frequency were often above the limit of what could be measured. It is noteworthy that the mean thresholds at high frequencies (10 and 12.5 kHz) were relatively high (68–78 dB HL), even though the mean thresholds for frequencies below 4 kHz were 41 dB HL or better.

Figure 9 shows examples of the audiograms for 18 people. The pattern of the audiograms at high frequencies varied markedly across listeners. In most cases, the audiometric threshold increased (worsened) as the frequency increased and the highest audiometric threshold occurred for the highest frequency. However, sometimes (e.g., S14) the audiometric thresholds for frequencies above 8 kHz were the same as or better than at 8 kHz. On average, the audiometric threshold at 10 kHz was poorer than that at 8 kHz by 11.4 dB for the left ear and 10.4 dB for the right ear. The audiometric threshold at 10 kHz was significantly correlated with the absolute threshold at 8 kHz ($r = 0.73$ for the left ear and $r = 0.71$ for the right ear, both $p < 0.01$). However, the difference in threshold between 8 and 10 kHz (i.e., the local slope of the audiogram) was only weakly and nonsignificantly correlated with the threshold at 8 kHz ($r = -0.26$ for the left ear and $r = -0.14$ for the right ear). Thus, if a high-frequency audiometer is not available, the best predictor of the audiometric threshold at 10 kHz is the audiometric threshold at 8 kHz plus 11 dB.

Of the 62 ears, nine had audiometric thresholds of 40 dB HL or better at 8 kHz, and 24 had audiometric thresholds of 50 dB HL or better at 8 kHz. Of the same ears, 15 had audiometric thresholds of 50 dB HL or better at 10 kHz, and 21 had audiometric thresholds of 60 dB HL or better at 10 kHz.

Implications for Gain and Output-Level Requirements

In what follows we focus on the gain required at 10 kHz to restore audibility of at least a part of the dynamic range of the speech to people with hearing losses of the extent measured here. As earlier, we assume that the auditory-filter bandwidth for ears with mild-to-moderate hearing loss is about 2 ERB_N and that stimulus energy is integrated, for the purpose of detection, over a frequency range of about 2 ERB_N .

Ideally, one might want to apply sufficient gain to make speech audible over most of its dynamic range in the band around 10 kHz. According to the ANSI (1997) standard for

calculating the speech intelligibility index, this would require gain such that the RMS level of the speech in the 2-ERB_N-wide band around 10 kHz falls 15 dB above the absolute threshold at 10 kHz (as the dynamic range is assumed to extend from 15 dB below to 15 dB above the RMS level). If, in reality, the dynamic range extends below -15 dB relative to the RMS level, as suggested by our measurements and by some previous work (Dubno & Ahlstrom, 1995; Dubno & Schaefer, 1995; Studebaker & Sherbecoe, 2002; Studebaker, et al., 1999), then even more gain would be required. For example, to amplify speech so that the main peak in the histogram of levels fell slightly above the absolute threshold (see Fig. 8) would require the RMS level to be about 20 dB above the absolute threshold.

In practice, it may be impractical to aim to restore full audibility of speech in the frequency band around 10 kHz, for two reasons:

1. The required gains might lead to an unpleasant, harsh, or tinny sound quality. Calculations based on the model of loudness perception described by Moore and Glasberg (2004) suggest that gain sufficient to place the RMS level of speech 15 dB above absolute threshold for frequencies around 10 kHz for a person with mild hearing loss would lead to relatively high specific loudness around 10 kHz for speech at normal conversational levels.
2. The required gains would be difficult to achieve in current devices.

Here, we take it as a realistic goal to amplify the band centered at 10 kHz so that its RMS level reaches the absolute threshold, for speech presented at an overall level of 65 dB SPL. This would restore partial audibility of speech without leading to an unpleasant sound quality, and with more modest gain requirements.

The mean speech level in the 2-ERB_N-wide band centered at 10 kHz was 37.2 dB SPL, and the corresponding level at the eardrum was 42.2 dB SPL, which is about 29 dB above the normal monaural MAP (see Table 2). For a hearing loss of X dB ($X > 29$), the insertion gain required to amplify the band so that its RMS level is at the absolute threshold is therefore $X - 29$ dB. For example, to restore audibility in this way for a hearing loss of 65 dB at 10 kHz would require an effective insertion gain of about 36 dB at 10 kHz. With this gain, audibility could be (partly) restored for 25 of the 62 ears assessed. The effective RMS output level at the eardrum required in the band around 10 kHz is about 78 (42 + 36) dB SPL. The peak level would have to be about 11 to 13 dB higher than this to avoid excessive limiting or clipping. Clipping 1% of the time has only a very small deleterious effect on sound quality for hearing-impaired people (Tan & Moore, 2008).

Discussion

This study has focused on the insertion gains that would be required to make speech audible at high frequencies for people with mild-to-moderate losses for frequencies of 4 kHz and below. A separate and still unresolved issue is the extent to which providing such audibility will be of benefit. Calculations of speech intelligibility using the articulation index (ANSI, 1969) or the speech intelligibility index (ANSI, 1997) are based on the assumption that frequency components between 500 and 5000 Hz contribute most to intelligibility, but components with frequencies up to about 8000 Hz make some contribution. There is evidence that the audibility of frequencies up to at least 8000 Hz is important for normal development of phonetic and language skills, for example, perception of sounds like /s/ and /z/ (Stelmachowicz, et al., 2001; 2007).

Some studies using adult hearing-impaired listeners have shown little or no benefit for speech intelligibility of providing amplification at high frequencies (Amos & Humes, 2007; Ching, et al., 1998; Hogan & Turner, 1998; Murray & Byrne, 1986), whereas others have

shown clear benefit (Hornsby & Ricketts, 2006; Skinner, 1980; Skinner, et al., 1982; Skinner & Miller, 1983). Vickers, et al. (2001) and Baer, et al. (2002) suggested that hearing-impaired people without dead regions (Moore, 2001; 2004) at high frequencies did benefit from amplification of high frequencies, whereas those with high-frequency dead regions obtained little or no benefit. The extent of hearing loss at high frequencies may be of importance (Moore, 2002; Rankovic, 2002), partly because high-frequency dead regions tend to be associated with hearing losses greater than 70 dB HL (Aazh & Moore, 2007; Vinay & Moore, 2007). Also, as the high-frequency hearing loss increases, it becomes increasingly difficult to provide sufficient amplification to restore audibility.

The studies referred to above all determined speech intelligibility in quiet or in noise under conditions where the speech and noise came from the same position in space, or were delivered by the same headphone. There is some evidence that, under conditions where the target speech and interfering sounds are spatially separated, there may be a clear advantage of extending the audible bandwidth provided by bilaterally fitted hearing aids (Hamacher, et al., 2005). This may happen partly because of head-shadow effects, which lead to a better signal-to-background ratio at one ear than the other whenever the target speech is on the opposite side of the head to the most prominent interfering sound. The magnitude of head-shadow effects increases progressively with increasing frequency (Bronkhorst & Plomp, 1988; 1989), so the advantage of listening with the “better” ear would be expected to increase as the audible high-frequency bandwidth increases.

Related to this is the fact that sound localization using pinna cues, which are important for judgments of elevation and for resolving front-back confusions, depends on information from high frequencies. Best, et al. (2005) showed that localization of speech was impaired when the speech was lowpass filtered at 8 kHz. Hearing-impaired listeners may have a reduced ability to use pinna cues, because their reduced frequency selectivity makes them less able to resolve peaks and dips in the spectrum. However, simulations of reduced frequency selectivity suggest that most hearing-impaired subjects should retain some ability to use pinna cues (Jin, et al., 2002).

Apart from affecting speech intelligibility, increasing audibility at high frequencies may have effects on sound quality. Snow (1931) reported that “transmission of the highest audible frequencies was needed for perfect reproduction of musical instruments,” but that lowpass filtering at 10 kHz had only a slight effect on the tone quality of most instruments. Olson (1957) stated that “To reproduce speech and music without loss of naturalism requires a frequency range of 40 to 15,000 cycles,” but he also pointed out that the preferred frequency range may differ for live and reproduced sounds, because of distortion in the reproduction equipment. Moore and Tan (2003) investigated the effect of lowpass and highpass filtering on the subjective quality (“naturalness”) of music and speech signals, using normally hearing listeners. They found that the quality of music was reduced when the upper cutoff frequency was <16.5 kHz and the quality of speech was reduced when the upper cutoff frequency was <10.8 kHz. Franks (1982) obtained paired-comparison judgments of preference for hearing-aid processed music, varying both the lower and upper cutoff frequencies. Upper cutoff frequencies were either 4 or 10 kHz. Listeners with normal hearing preferred the higher cutoff frequency, but preferences were not consistent across listeners with impaired hearing. Ricketts, et al. (2008) pointed out that the absolute thresholds of the hearing-impaired listeners used by Franks varied over a wide range, and suggested that the inconsistency across listeners might have occurred because some listeners consistently preferred the higher cutoff frequency and some consistently preferred the lower cutoff frequency.

Füllgrabe, et al. (2007) studied the subjective quality of speech (male and female talkers) and music (a piece of jazz) recorded through a wide-bandwidth hearing aid using a receiver sealed deeply in the ear canal of a dummy head (KEMAR). The stimuli were either broadband, or were lowpass filtered at 5 or 9 kHz. There were 15 participants with normal hearing, 18 with mild hearing loss, and 12 with moderate hearing loss for frequencies below 6 kHz. The frequency-gain characteristics of the hearing aid were programmed for each group using the CAMEQ fitting procedure (Moore, et al., 1999). However, because that procedure did not give recommended gains for frequencies above 6 kHz, gains for these frequencies were set to be the same as the gain at 6 kHz. The nominal input level was 65 dB SPL, which is typical of the level of normal conversational speech, but lower than the typical level of live music. A paired-comparison method was used. For the normal-hearing group, the two wider-band-width conditions (broadband and lowpass filtered at 9 kHz) consistently yielded higher preferences scores than the condition with lowpass filtering at 5 kHz. For the two hearing-impaired groups, preferences did not differ significantly across bandwidth conditions. However, subsequent analyses, partly based on the data presented in the present paper, indicated that, at least for the speech stimuli, the gains applied for frequencies above 6 kHz would not have been sufficient to restore audibility for most of the participants. This may have been partly because of the moderate level of the stimuli.

Ricketts, et al. (2008) obtained paired-comparison judgments of preference for (simulated) hearing-aid processed sounds using upper cutoff frequencies of 5.5 and 9 kHz. The sounds were a piece of music and a movie soundtrack, both chosen because they contained a relatively large amount of high-frequency energy. Both normal-hearing and hearing-impaired listeners were tested. The latter were selected to have mild hearing losses; average absolute thresholds were 38, 43, and 53 dB HL at 8, 10, and 12 kHz, respectively. The gains were adjusted for each listener using the NAL-NL1 fitting method (Byrne, et al., 2001). Because this method does not give recommended gains for frequencies above 6 kHz, gains at high frequencies were based on a form of extrapolation. Both normal-hearing and hearing-impaired listeners showed, on average, a preference for the higher cutoff frequency, but the strength of preference was greater for the former than for the latter. Preference within the hearing-impaired group did not seem to be related to average absolute thresholds, but a steep slope of the audiogram was associated with a preference for the lower cutoff frequency.

It is not obvious why some hearing-impaired listeners seemed to prefer the lower cutoff frequency. One possibility is that these listeners were unused to hearing frequencies above about 6 kHz, because of their long-standing hearing loss and the limited frequency range of their hearing aids (if any). When these high frequencies were amplified and presented in a laboratory setting, the sound quality may have seemed somewhat harsh or tinny. If this were the case, such listeners might come to prefer a higher cutoff frequency after an acclimatization period (Gatehouse, 1992). Another possibility is that the fitting rule used for the high frequencies, which was acknowledged by Ricketts, et al. (2008) to be somewhat arbitrary, may have led to excessive loudness of the high frequencies for some listeners.

Overall, evidence is accruing that increasing the highest frequency at which gain is applied in hearing aids may have beneficial effects for some listeners with mild-to-moderate hearing loss, in terms of speech intelligibility, subjective quality of speech and music, and sound localization. However, it is clear that rules for prescribing appropriate amounts of gain at high frequencies are needed. We hope that the results presented in this paper may contribute to the development of appropriate recommendations for gain at high frequencies, and will help to indicate the proportion of people with mild-to-moderate hearing loss who might benefit from an extended high-frequency response.

Conclusions

1. The spectrum level of speech falls progressively with increasing frequency. The average speech spectrum determined here is similar to the average spectrum determined by Byrne, et al. (1994) for center frequencies up to 8 kHz, but the spectrum level found here for higher center frequencies is lower than reported by Byrne, et al. We suggest that the difference reflects the influence of microphone and recording noise on the data of Byrne, et al. For speech with an overall level of 65 dB SPL, the mean 1/3-octave level at a center frequency of 10 kHz was 37 dB SPL.
2. The dynamic range of the speech was similar for center frequencies of 10 and 1 kHz. The part of the dynamic range below the RMS level was larger than reported in some previous studies.
3. For hearing-impaired people with mild losses for frequencies up to 4 kHz, the mean audio-metric thresholds at high frequencies (10 and 12.5 kHz) were relatively high (69 and 77 dB HL, respectively). The hearing loss at 10 kHz was, on average, 11 dB greater than at 8 kHz.
4. To partially restore the audibility of speech in the frequency region around 10 kHz, for a hearing loss of 65 dB at 10 kHz, would require an effective insertion gain of about 36 dB at 10 kHz. With this gain, audibility could be (partly) restored for 25 of the 62 ears (40%) assessed.

Acknowledgments

The authors thank Mead Killion and two anonymous reviewers for helpful comments on an earlier version of this paper.

This work was supported by the MRC (UK) and by EarLens Corporation (USA) and by Marie Curie Fellowship from the EU (to C.F.).

References

- Aazh H, Moore BCJ. Dead regions in the cochlea at 4 kHz in elderly adults: relation to absolute threshold, steepness of audiogram, and pure tone average. *J Am Acad Audiol*. 2007; 18:96–107.
- Amos NE, Humes LE. Contribution of high frequencies to speech recognition in quiet and noise in listeners with varying degrees of high-frequency sensorineural hearing loss. *J Speech Lang Hear Res*. 2007; 50:819–834. [PubMed: 17675588]
- ANSI. Methods for the calculation of the articulation index. American National Standards Institute; New York: 1969. ANSI S3.5.
- ANSI. Methods for the calculation of the speech intelligibility index. American National Standards Institute; New York: 1997. ANSI S3.5–1997.
- ANSI. Procedure for the computation of loudness of steady sounds. American National Standards Institute; New York: 2007. ANSI S3.4-2007.
- Baer T, Moore BCJ, Kluk K. Effects of lowpass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies. *J Acoust Soc Am*. 2002; 112:1133–1144. [PubMed: 12243160]
- Best V, Carlile S, Jin C, et al. The role of high frequencies in speech localization. *J Acoust Soc Am*. 2005; 118:353–363. [PubMed: 16119356]
- British Society of Audiology. Pure tone air and bone conduction threshold audiometry with and without masking and determination of uncomfortable loudness levels. British Society of Audiology; Reading, UK: 2004.
- Bronkhorst AW, Plomp R. The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J Acoust Soc Am*. 1988; 83:1508–1516. [PubMed: 3372866]

- Bronkhorst AW, Plomp R. Binaural speech intelligibility in noise for hearing-impaired listeners. *J Acoust Soc Am.* 1989; 86:1374–1383. [PubMed: 2808911]
- Byrne D. The speech spectrum—some aspects of its significance for hearing aid selection and evaluation. *Br J Audiol.* 1977; 11:40–46. [PubMed: 922224]
- Byrne D, Dillon H, Ching T, et al. NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *J Am Acad Audiol.* 2001; 12:37–51. [PubMed: 11214977]
- Byrne D, Dillon H, Tran K, et al. An international comparison of long-term average speech spectra. *J Acoust Soc Am.* 1994; 96:2108–2120.
- Ching T, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am.* 1998; 103:1128–1140. [PubMed: 9479766]
- Cornelisse LE, Seewald RC, Jamieson DG. The input/output formula: a theoretical approach to the fitting of personal amplification devices. *J Acoust Soc Am.* 1995; 97:1854–1864. [PubMed: 7699167]
- Cox RM, Matesich JS, Moore JN. Distribution of short-term rms levels in conversational speech. *J Acoust Soc Am.* 1988; 84:1100–1104. [PubMed: 3183202]
- Cox RM, Moore JN. Composite speech spectrum for hearing and gain prescriptions. *J Speech Hear Res.* 1988; 31:102–107. [PubMed: 3352247]
- Davis, A. *Hearing in adults.* Whurr; London: 1995.
- Dillon, H. *Hearing aids.* Thieme; New York: 2001.
- Dubno JR, Ahlstrom JB. Masked threshold and consonant recognition in low-pass maskers for hearing-impaired and normal-hearing listeners. *J Acoust Soc Am.* 1995; 97:2430–2441. [PubMed: 7714260]
- Dubno JR, Dirks DD, Schaefer AB. Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II. Articulation index predictions. *J Acoust Soc Am.* 1989; 85:355–364. [PubMed: 2921418]
- Dubno JR, Schaefer AB. Frequency selectivity and consonant recognition for hearing-impaired and normal-hearing listeners with equivalent masked thresholds. *J Acoust Soc Am.* 1995; 97:1165–1174. [PubMed: 7876438]
- Dunn HK, White SD. Statistical measurements on conversational speech. *J Acoust Soc Am.* 1940; 11:278–283.
- Fletcher, H. *Speech and hearing in communication.* Van Nostrand; New York: 1953.
- Franks JR. Judgments of hearing aid processed music. *Ear Hear.* 1982; 3:18–23. [PubMed: 7060840]
- French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. *J Acoust Soc Am.* 1947; 19:90–119.
- Füllgrabe C, Moore BCJ, van Tasell DJ, et al. Effects of bandwidth on sound-quality preferences for hearing aids. *Bull Am Aud Soc.* 2007; 32:45.
- Gatehouse S. The time course and magnitude of perceptual acclimatization to frequency responses: evidence from monaural fitting of hearing aids. *J Acoust Soc Am.* 1992; 92:1258–1268. [PubMed: 1401514]
- Glasberg BR, Moore BCJ. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *J Acoust Soc Am.* 1986; 79:1020–1033. [PubMed: 3700857]
- Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear Res.* 1990; 47:103–138. [PubMed: 2228789]
- Glasberg BR, Moore BCJ. Prediction of absolute thresholds and equal-loudness contours using a modified loudness model. *J Acoust Soc Am.* 2006; 120:585–588. [PubMed: 16938942]
- Green DM, Kidd G, Stevens KN. High-frequency audiometric assessment of a young adult population. *J Acoust Soc Am.* 1987; 81:485–494. [PubMed: 3558966]
- Hamacher V, Chalupper J, Eggers J, et al. Signal processing in high-end hearing aids: state of the art, challenges, and future trends. *EURASIP J Appl Sig Proc.* 2005; 18:2915–2929.
- Hogan CA, Turner CW. High-frequency audibility: benefits for hearing-impaired listeners. *J Acoust Soc Am.* 1998; 104:432–441. [PubMed: 9670535]

- Hornsby BW, Ricketts TA. The effects of hearing loss on the contribution of high- and low-frequency speech information to speech understanding. II. Sloping hearing loss. *J Acoust Soc Am.* 2006; 119:1752–1763. [PubMed: 16583917]
- Jin, C.; Best, V.; Carlile, S., et al. AES 112th Convention. Munich, Germany: 2002. Speech localization; p. 1-13.
- Killion MC. Hearing aids: past, present and future: moving toward normal conversations in noise. *Br J Audiol.* 1997; 31:141–148. [PubMed: 9276096]
- Killion MC, Tillman TW. Evaluation of high-fidelity hearing aids. *J Speech Hear Res.* 1982; 25:15–25. [PubMed: 7087418]
- Kryter KD. Methods for the calculation and use of the Articulation Index. *J Acoust Soc Am.* 1962; 34:467–477.
- Kurakata K, Mizunami T, Matsushita K, et al. Statistical distribution of normal hearing thresholds under free-field listening conditions. *Acoust Sci Tech.* 2005; 26:440–446.
- Levitt, H. Speech discrimination ability in the hearing impaired: spectrum considerations. In: Studebaker, GA.; Bess, FH., editors. *The Vanderbilt hearing-aid report. Monographs in Contemporary Audiology*; Upper Darby, Pennsylvania: 1982. p. 32-43.
- Moore BCJ. Use of a loudness model for hearing aid fitting. IV. Fitting hearing aids with multi-channel compression so as to restore “normal” loudness for speech at different levels. *Br J Audiol.* 2000; 34:165–177. [PubMed: 10905450]
- Moore BCJ. Dead regions in the cochlea: diagnosis, perceptual consequences, and implications for the fitting of hearing aids. *Trends Amplif.* 2001; 5:1–34.
- Moore BCJ. Response to “Articulation index predictions for hearing-impaired listeners with and without cochlear dead regions”. *J Acoust Soc Am.* 2002; 111:2549–2550. [*J. Acoust. Soc. Am.* 111, 2545–2548 (2002)] (L).
- Moore BCJ. Dead regions in the cochlea: conceptual foundations, diagnosis and clinical applications. *Ear Hear.* 2004; 25:98–116. [PubMed: 15064655]
- Moore, BCJ. The use of a loudness model to derive initial fittings for hearing aids: validation studies. In: Anderson, T.; Larsen, CB.; Poulsen, T.; Rasmussen, AN.; Simonsen, JB., editors. *Hearing Aid Fitting.* Holmens Trykkeri; Denmark: 2005. p. 11-33.
- Moore, BCJ. *Cochlear hearing loss: physiological, psychological and technical issues.* 2nd ed. Wiley; Chichester: 2007.
- Moore BCJ, Glasberg BR. A revised model of loudness perception applied to cochlear hearing loss. *Hear Res.* 2004; 188:70–88. [PubMed: 14759572]
- Moore BCJ, Glasberg BR. Modeling binaural loudness. *J Acoust Soc Am.* 2007; 121:1604–1612. [PubMed: 17407897]
- Moore BCJ, Glasberg BR, Baer T. A model for the prediction of thresholds, loudness and partial loudness. *J Audio Eng Soc.* 1997; 45:224–240.
- Moore BCJ, Glasberg BR, Plack CJ, et al. The shape of the ear’s temporal window. *J Acoust Soc Am.* 1988; 83:1102–1116. [PubMed: 3356815]
- Moore BCJ, Glasberg BR, Stone MA. Use of a loudness model for hearing aid fitting. III. A general method for deriving initial fittings for hearing aids with multi-channel compression. *Br J Audiol.* 1999; 33:241–258. [PubMed: 10509859]
- Moore BCJ, Stone MA, Alcántara JI. Comparison of the electroacoustic characteristics of five hearing aids. *Br J Audiol.* 2001; 35:307–325. [PubMed: 11824533]
- Moore BCJ, Tan CT. Perceived naturalness of spectrally distorted speech and music. *J Acoust Soc Am.* 2003; 114:408–419. [PubMed: 12880052]
- Murray N, Byrne D. Performance of hearing-impaired and normal hearing listeners with various high-frequency cut-offs in hearing aids. *Aust J Audiol.* 1986; 8:21–28.
- Northern JL, Downs MP, Rudmose W, et al. Recommended high-frequency audiometric threshold levels (8000-16000 Hz). *J Acoust Soc Am.* 1972; 52:585–595.
- Olsen WO, Hawkins DB, van Tasell DJ. Representations of the long-term spectra of speech. *Ear Hear.* 1987; 8:100S–108S. [PubMed: 3678648]
- Olson, HF. *Elements of acoustical engineering.* Van Nostrand; New York: 1957.

- Oxenham AJ, Moore BCJ. Modeling the additivity of nonsimultaneous masking. *Hear Res.* 1994; 80:105–118. [PubMed: 7852196]
- Pavlovic C. Derivation of primary parameters and procedures for use in speech intelligibility predictions. *J Acoust Soc Am.* 1987; 82:413–422. [PubMed: 3624646]
- Pavlovic, CV. Problems in the prediction of speech recognition performance of normal-hearing and hearing-impaired individuals. In: Studebaker, GA.; Hochberg, I., editors. *Acoustical factors affecting hearing aid performance.* Allyn and Bacon; Boston: 1993. p. 221-234.
- Pavlovic CV, Studebaker GA. An evaluation of some assumptions underlying the articulation index. *J Acoust Soc Am.* 1984; 75:1606–1612. [PubMed: 6736424]
- Pearsons, KS.; Bennett, RL.; Fidell, S. *Speech levels in various environments.* report no. 3281. Bolt, Beranek and Newman; Cambridge, Massachusetts: 1976.
- Perkins R. Earlens tympanic contact transducer: a new method of sound transduction to the human ear. *Otolaryngol Head Neck Surg.* 1996; 114:720–728. [PubMed: 8643293]
- Pick, G.; Evans, EF.; Wilson, JP. Frequency resolution in patients with hearing loss of cochlear origin. In: Evans, EF.; Wilson, JP., editors. *Psychophys Physiol Hear.* Academic Press; London: 1977. p. 273-281.
- Plack CJ, Moore BCJ. Temporal window shape as a function of frequency and level. *J Acoust Soc Am.* 1990; 87:2178–2187. [PubMed: 2348022]
- Rankovic CM. Articulation index predictions for hearing-impaired listeners with and without cochlear dead regions (L). *J Acoust Soc Am.* 2002; 111:2545–2548. [PubMed: 12083183]
- Rhebergen, KS. *Modeling the speech intelligibility in fluctuating noise.* University of Amsterdam; 2006. Ph.D. Thesis
- Ricketts TA, Dittberner AB, Johnson EE. High frequency amplification and sound quality in listeners with normal through moderate hearing loss. *J Speech Lang Hear Res.* 2008; 51:160–172. [PubMed: 18230863]
- 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]
- Scollie S, Seewald R, Cornelisse L, et al. The desired sensation level multistage input/output algorithm. *Trends Amplif.* 2005; 9:159–197. [PubMed: 16424945]
- Skinner MW. Speech intelligibility in noise-induced hearing loss: effects of high-frequency compensation. *J Acoust Soc Am.* 1980; 67:306–317. [PubMed: 7354198]
- Skinner MW, Karstaedt MM, Miller JD. Amplification bandwidth and speech intelligibility for two listeners with sensorineural hearing loss. *Audiology.* 1982; 21:251–268. [PubMed: 7092728]
- Skinner MW, Miller JD. Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. *Audiology.* 1983; 22:253–279. [PubMed: 6870667]
- Snow WB. Audible frequency range of music, speech, and noise. *J Acoust Soc Am.* 1931; 3:155–166.
- Stelmachowicz PG, Beauchaine KA, Kalberer A, et al. Normative thresholds in the 8–10 kHz range as a function of age. *J Acoust Soc Am.* 1989; 86:1384–1391. [PubMed: 2808912]
- Stelmachowicz PG, Lewis DE, Choi S, et al. Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear Hear.* 2007; 28:483–494. [PubMed: 17609611]
- Stelmachowicz PG, Pittman AL, Hoover BM, et al. Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *J Acoust Soc Am.* 2001; 110:2183–2190. [PubMed: 11681394]
- Studebaker GA, Sherbecoe RL. Intensity-importance functions for bandlimited monosyllabic words. *J Acoust Soc Am.* 2002; 111:1422–1436. [PubMed: 11931319]
- Studebaker GA, Sherbecoe RL, McDaniel DM, et al. Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am.* 1999; 105:2431–2444. [PubMed: 10212424]
- Tan CT, Moore BCJ. Perception of nonlinear distortion by hearing-impaired people. *Int J Audiol.* 2008; 47:246–256. [PubMed: 18465409]
- Vickers DA, Moore BCJ, Baer T. Effects of lowpass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies. *J Acoust Soc Am.* 2001; 110:1164–1175. [PubMed: 11519583]

Vinay, Moore BCJ. Prevalence of dead regions in subjects with sensorineural hearing loss. *Ear Hear.* 2007; 28:231–241. [PubMed: 17496673]

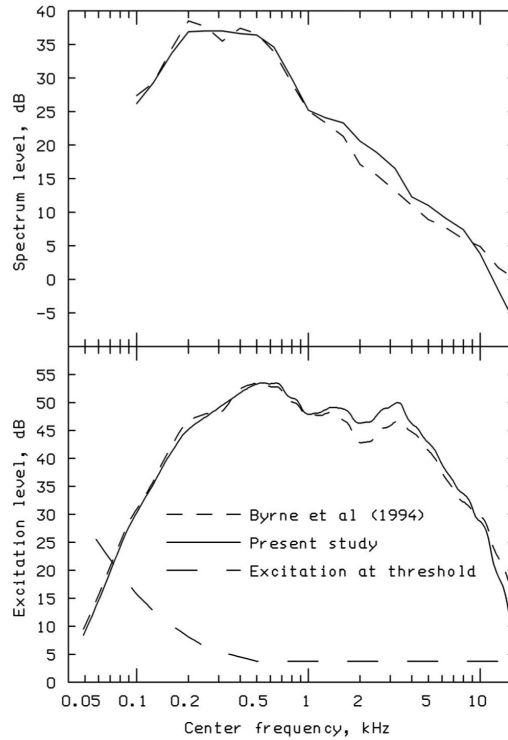


Fig. 1.

The upper panel shows the composite speech spectrum (average for the male and female talkers) determined at Cambridge (solid line), and the composite speech spectrum reported by Byrne, et al. (1994) (dashed line). In both cases, the spectrum level (re $20 \mu\text{Pa}$) is plotted against center frequency, and the overall level was adjusted to be 65 dB SPL. The lower panel shows excitation patterns calculated from these spectra, together with the excitation level required for threshold.

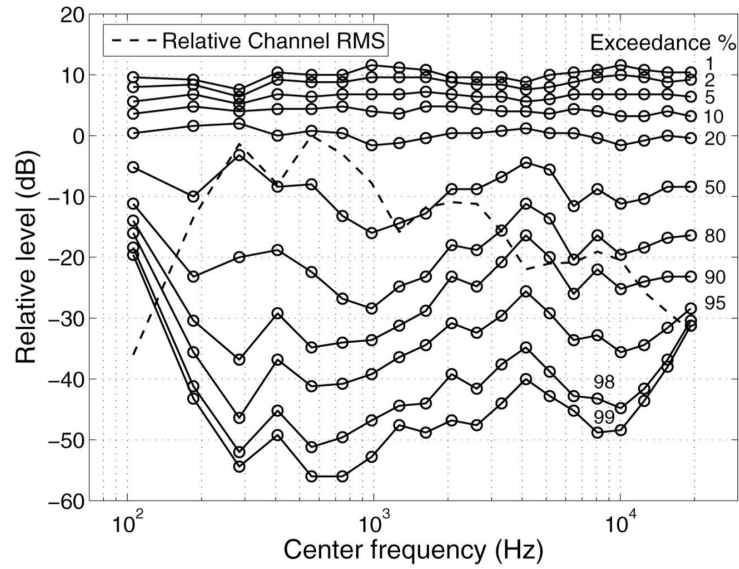


Fig. 2.

Distribution of levels for a typical female talker recorded at Cambridge. The abscissa shows the center frequency of the 2- ERB_N -wide analysis band. The ordinate shows relative level in dB. The dashed line shows the relative root-mean-square (RMS) level in each band, scaled so that the level of the band with the highest level is set to 0 dB. This line indicates the spectral shape of the speech for that specific talker. The solid lines with circles show “equal-exceedance contours” (EECs); see text for details.

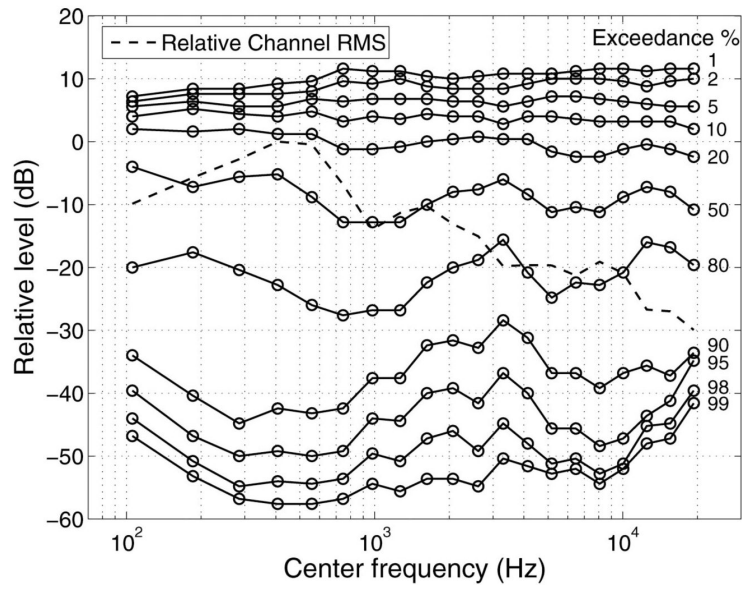


Fig. 3.
As Figure 2, but for a typical male talker recorded at Cambridge.

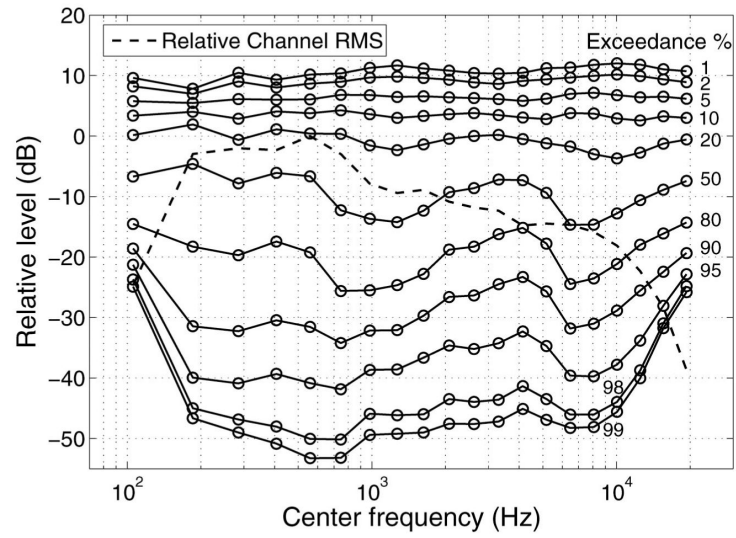


Fig. 4.
As Figure 2, but showing data averaged across the eight female talkers recorded at Cambridge.

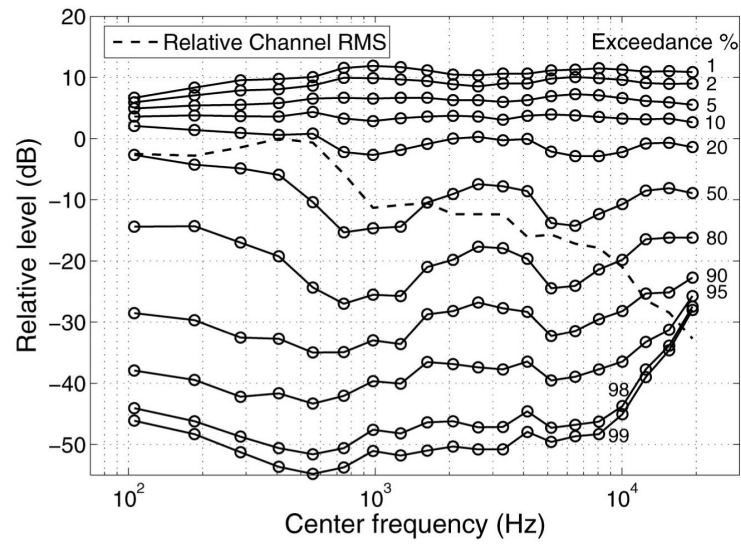


Fig. 5. As Figure 2, but showing data averaged across the six male talkers recorded at Cambridge using the 24-bit system.

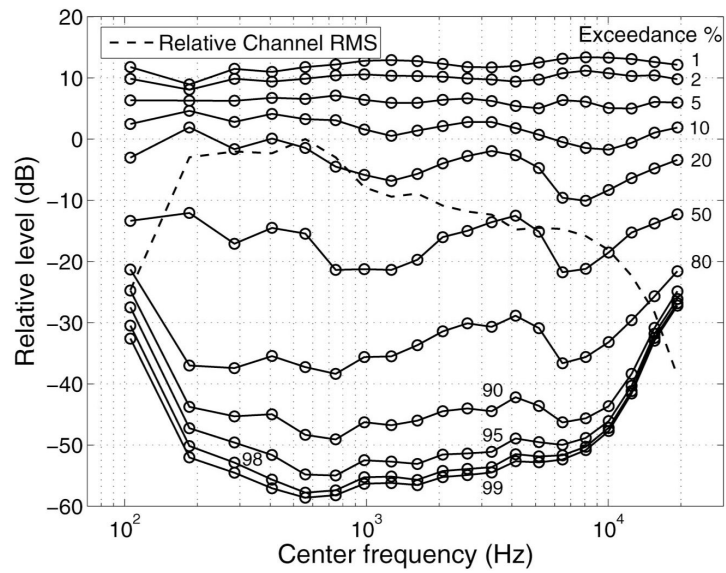


Fig. 6. Distribution of levels averaged across the eight female talkers recorded at Cambridge, obtained using 10-msec windows. Otherwise, as Figure 2.

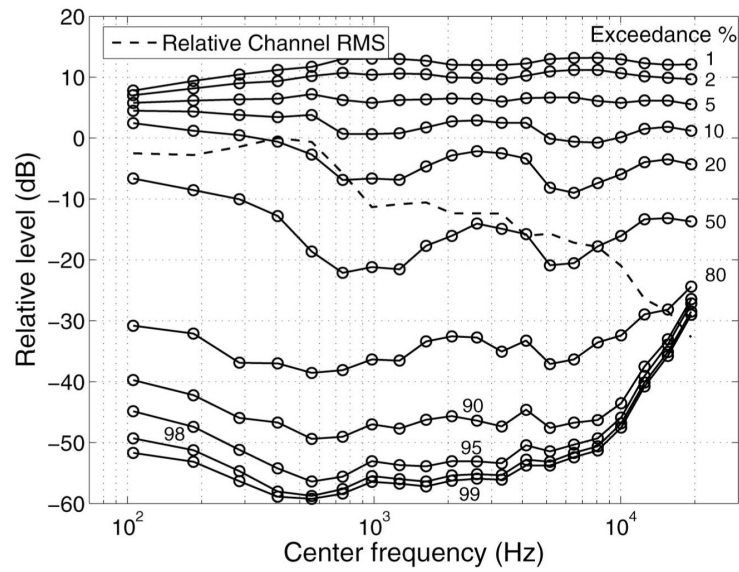


Fig. 7. Distribution of levels averaged across the six male talkers recorded at Cambridge using the 24-bit system, obtained using 10-msec windows. Otherwise, as Figure 2.

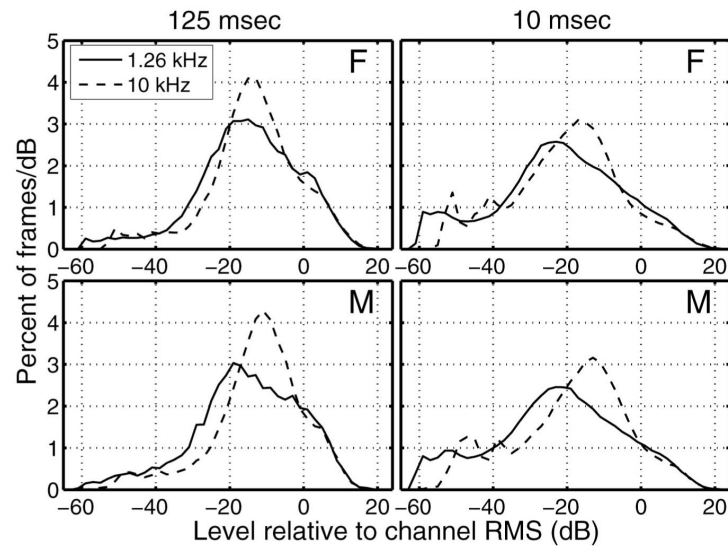


Fig. 8. Histograms showing the distribution of levels in 2- ERB_N -wide analysis bands centered at 1.26 kHz (solid lines) and 10 kHz (dashed lines). The analysis was based on 125-msec windows (left) or 10-msec windows (right). The data for the female talker (same as in Fig. 2) are shown at the top and the data for the male talker (same as in Fig. 4) are shown at the bottom.

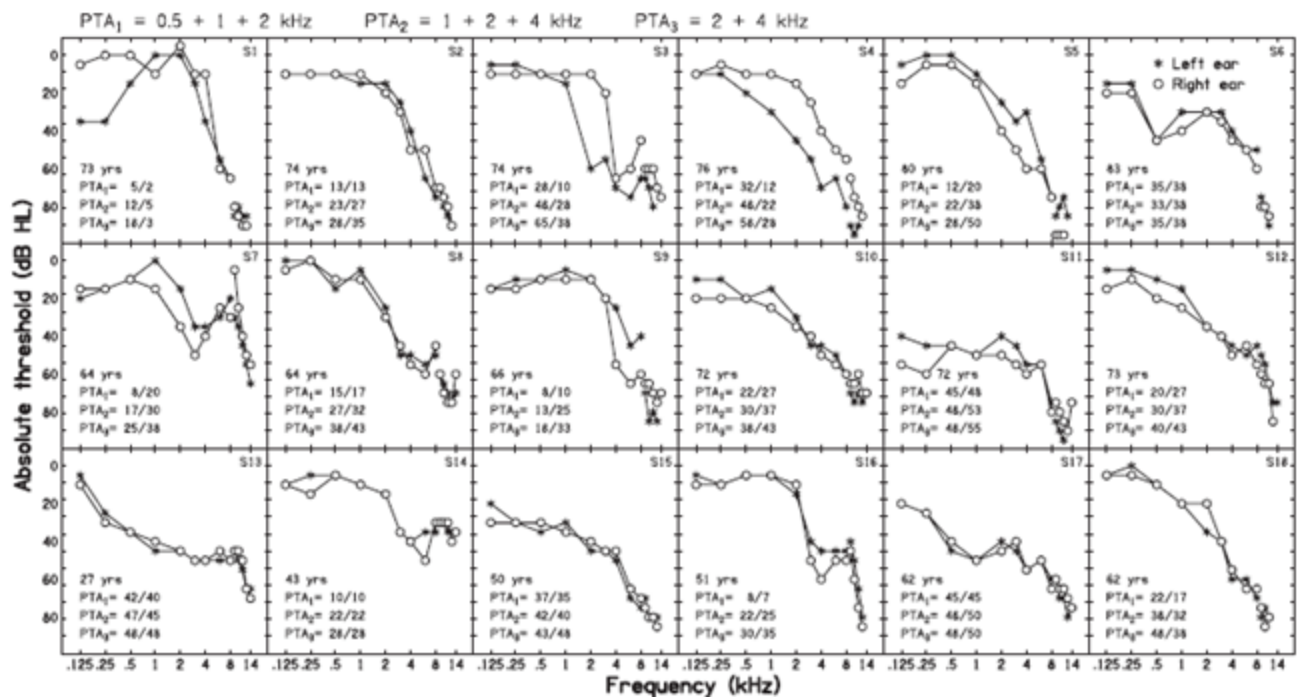


Fig. 9.

Examples of audiograms obtained from both ears of 18 people with mild-to-moderate hearing loss for frequencies up to 4 kHz. Asterisks and circles indicate thresholds for the left and right ears, respectively. The curves are broken between 8 and 10 kHz (not always visible) to indicate that thresholds for frequencies up to 8 kHz were obtained with TDH50P headphones, whereas thresholds for higher frequencies were obtained with Sennheiser HDA200 headphones. The age of each listener is indicated in the panels. Various measures of the pure-tone-average (PTA) threshold are also shown: PTA₁—average over 0.5, 1, and 2 kHz; PTA₂—average over 1, 2, and 4 kHz; PTA₃—average over 2 and 4 kHz. Numbers to the left and right of / indicate PTA values for the left and right ears, respectively.

TABLE 1
Spectrum levels (re 20 μ Pa) of speech produced at normal conversational levels when the overall speech level is scaled to be 65 dB SPL, given as a function of center frequency (CF)

CF, Hz	Male					Female				
	D&W	Pears	Arch	Byrne	This	D&W	Pears	Arch	Byrne	This
100	31.2		36.6	35.8	37.7			28.0	18.9	14.6
125	33.0	37.4	36.7	38.1	37.4	29.0	18.4	30.5	20.5	21.1
160	34.2	31.3	36.9	36.1	37.1	34.5	28.3	34.0	32.7	30.2
200	34.8	35.3	36.7	36.5	36.9	38.0	38.3	36.9	40.5	36.9
250	35.3	38.4	35.9	37.1	37.1	38.3	36.4	38.2	38.3	36.9
315	36.2	36.4	36.0	36.4	37.3	37.8	30.4	38.6	34.5	36.6
400	37.4	37.3	37.7	37.7	37.4	36.6	36.3	37.1	37.0	35.7
500	38.3	37.4	36.3	37.0	36.3	34.8	36.4	36.6	36.1	36.5
630	32.7	34.4	33.0	34.0	33.4	32.5	33.4	33.6	33.8	35.7
800	26.9	27.3	27.6	28.0	28.0	29.0	30.3	27.0	30.3	31.8
1000	26.3	25.4	24.0	24.5	23.2	26.4	25.4	22.0	25.7	27.2
1250	25.1	26.4	25.2	24.1	23.1	24.0	25.4	20.6	22.7	25.0
1600	22.2	22.3	24.5	21.6	22.1	22.0	23.3	18.4	21.0	24.4
2000	18.7	17.3	21.0	17.0	19.4	19.6	18.3	13.9	17.1	21.7
2500	15.5	17.4	19.2	16.3	17.9	13.5	15.4	11.8	14.7	19.4
3150	11.9	16.4	12.5	13.3	15.9	9.7	16.4	10.0	13.1	17.1
4000	6.1	12.3	10.2	11.3	11.2	9.1	17.3	10.7	10.6	13.4
5000	2.3	7.4	11.0	8.8	9.8	8.2	12.4	10.9	9.0	12.1
6300	0.3	7.4	8.6	6.7	7.3	6.2	11.4	13.1	8.6	10.8
8000	-3.8	3.3	9.5	4.7	5.7	2.7	8.3	15.1	7.2	9.0
10000	-8.8		7.2	3.3	1.5	0.0		10.6	6.4	6.1
12500	-12.9		-1.0	0.1	-4.4	-2.6		2.5	3.2	1.1
16000			-6.0	-0.3	-7.2			-3.9	0.4	-5.7

Data are taken from Dunn and White (1940) (D&W), Pearsons, et al. (1976) (Pears), the Archimedes CD (Arch), Byrne, et al. (1994) (Byrne), and the present study (This).

TABLE 2
Overall estimate of the mean spectrum of conversational speech with an overall level of 65 dB SPL

Frequency, Hz	Mean N0	1/3-Oct level, dB SPL	Diffuse field to eardrum, dB	1/3-Oct level at eardrum, dB SPL	MAP
100	26.2	39.8	0.0	39.8	27.5
125	29.3	43.9	0.1	44.0	23.2
160	33.6	49.3	0.3	48.8	19.2
200	36.9	53.5	0.4	53.9	15.9
250	37.0	54.6	0.5	55.1	13.3
315	36.9	55.5	1.0	56.5	11.0
400	36.5	56.2	1.6	57.8	8.8
500	36.4	57.0	1.7	58.7	7.0
630	34.5	56.1	2.2	58.3	6.4
800	29.9	52.5	2.9	55.4	5.7
1000	25.2	48.8	3.8	52.6	5.7
1250	24.0	48.6	5.3	53.9	7.6
1600	23.2	48.9	7.2	56.1	9.2
2000	20.6	47.2	10.2	57.4	11.6
2500	18.7	46.3	14.9	61.2	13.5
3150	16.5	45.1	14.4	59.5	10.1
4000	12.3	41.9	12.7	54.6	9.7
5000	10.9	41.5	10.8	52.3	10.1
6300	9.1	40.8	8.7	49.5	13.3
8000	7.3	40.0	8.5	48.5	15.3
10000	3.8	37.4	5.0	42.4	13.2
12500	-1.7	33.0	4.0	37.0	18.1
16000	-6.4	29.2	2.0	31.2	44.1

The first column shows the center frequency. The second column shows the spectrum level (N0) at that frequency (re 20 μ Pa). The third column shows the level per 1/3-octave band. The fourth column shows the diffuse-field to eardrum transfer function. The fifth column shows the estimated level per 1/3-octave band at the eardrum, obtained by adding the numbers in columns three and four. The sixth column shows the estimated normal minimum audible pressure (MAP) for monaural listening.

TABLE 3
Means (in bold) and standard deviations (SD) of the audiometric thresholds for 31 left ears and 31 right ears, for frequencies from 0.125 to 12.5 kHz

Ear	Frequency, kHz												
	0.125	0.25	0.5	1	2	3	4	6	8	10	12.5		
Left (mean)	15	16	21	22	30	41	47	53	57	68	76		
Left (SD)	12	13	15	16	15	12	13	12	16	17	14		
Right (mean)	18	17	20	23	28	41	51	55	59	69	78		
Right (SD)	12	14	15	15	18	13	12	12	14	17	14		

All subjects had mild to moderate hearing loss.