

Estimates of Human Cochlear Tuning at Low Levels Using Forward and Simultaneous Masking

ANDREW J. OXENHAM,¹ AND CHRISTOPHER A. SHERA^{2,3}

¹Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

²Eaton–Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

³Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02115, USA

Received: 14 November 2002; Accepted: 9 May 2003; Online publication: 10 July 2003

ABSTRACT

Auditory filter shapes were derived from psychophysical measurements in eight normal-hearing listeners using a variant of the notched-noise method for brief signals in forward and simultaneous masking. Signal frequencies of 1, 2, 4, 6, and 8 kHz were tested. The signal level was fixed at 10 dB above absolute threshold in the forward-masking conditions and fixed at either 10 or 35 dB above absolute threshold in the simultaneous-masking conditions. The results show that filter equivalent rectangular bandwidths (ERBs) are substantially narrower in forward masking than has been found in previous studies using simultaneous masking. Furthermore, in contrast to earlier studies, the sharpness of tuning doubles over the range of frequencies tested, giving Q_{ERB} values of about 10 and 20 at signal frequencies of 1 and 8 kHz, respectively. It is argued that the new estimates of auditory filter bandwidth provide a more accurate estimate of human cochlear tuning at low levels than earlier estimates using simultaneous masking at higher levels, and that they are therefore more suitable for comparison to cochlear tuning data from other species. The data may also prove helpful in defining the parameters for nonlinear models of human cochlear processing.

Keywords: cochlear tuning, psychoacoustics, auditory filters, masking

INTRODUCTION

The ability to perceptually separate simultaneous sounds of different frequencies is a fundamental property of the auditory system. The filtering of sounds according to frequency and the resulting tonotopic organization found throughout the auditory pathways to the auditory cortex (Read et al. 2002) have their basis in the electromechanical properties of the cochlea. Indeed, it is widely believed that frequency selectivity, measured behaviorally, can be regarded as a direct reflection of the filtering that takes place in the cochlea.

Estimates of frequency selectivity in humans have evolved over many decades as numerous confounding factors and potential artifacts were addressed (e.g., Wegel and Lane 1924; Fletcher 1940; Zwicker et al. 1957; Bos and de Boer 1966; Houtgast 1973; Patterson and Nimmo–Smith 1980; Moore et al. 1984; Glasberg and Moore 2000). The psychophysical tuning curve (PTC) (Houtgast 1973; Moore 1978) provides a paradigm that is, in principle, most similar to physiological measures of cochlear tuning, such as the neural tuning curve. However, issues such as off-frequency listening and the detection of beats and/or distortion products can make the results from PTC experiments difficult to interpret (Johnson–Davies and Patterson 1979; O’Loughlin and Moore 1981; Patterson and Moore 1986). Over the last two decades, the notched-noise technique has become the favored method of behaviorally estimating frequency selectivity (Patterson 1976; Patterson and Nimmo–Smith 1980; Moore 1987; Rosen et al. 1998; Glasberg and Moore 2000). This technique involves measuring

Correspondence to: Andrew J. Oxenham, Ph.D. MIT • bldg. 36-763 • Cambridge, MA 02139 • Telephone: (617) 253-5995; fax: (617) 258-7003; email: oxenham@mit.edu

the masked threshold of a sinusoidal signal in the presence of a noise with a spectral notch as a function of the width and position of the notch relative to the signal. The filter functions derived from such techniques are often referred to as auditory filters and are believed to reflect the filtering properties of the cochlea (Moore 1995). One of the most comprehensive studies of frequency selectivity using notched-noise maskers provides a function that can be used to calculate the estimated equivalent rectangular bandwidth (ERB) of the auditory filter at any given frequency (Glasberg and Moore 1990). The equation for the function is

$$\text{ERB} = 24.7(4.37f/1000 + 1) \quad (1)$$

where ERB is the equivalent rectangular bandwidth and f is the filter center frequency, both in Hz. This function has been used as an estimate of human cochlear tuning in a wide range of studies and applications (e.g., Beauvois and Meddis 1996; Dau et al. 1996, 1997; Moore et al. 1997; Breebaart et al. 2001). In all these studies, the auditory filters are assumed to be linear in any given condition, although it is explicitly acknowledged that they change their shape with level.

While ignoring cochlear nonlinearities for the sake of simplicity has certain advantages, some important effects are dependent on these nonlinearities. One consequence of cochlear nonlinearity is known as suppression, whereby the neural response to one tone can be reduced by the introduction of a second, suppressor, tone (Sachs and Kiang 1968). It has been known since this effect was first investigated behaviorally that estimates of cochlear tuning can be altered by effects ascribed to suppression. In general, estimates using simultaneous masking, where the masker is thought to suppress the signal to some degree, produce wider estimates of filter bandwidth than do estimates using nonsimultaneous masking, where the masker does not peripherally interact with, or suppress, the signal (Houtgast 1973, 1974; Moore 1978; Vogten 1978; Moore et al. 1987). Consistent with these findings, Heinz et al. (2002) showed in a recent modeling study that when their nonlinear model was used to predict thresholds in a simultaneous masking notched-noise experiment, the resulting estimated auditory filter bandwidths were wider than the bandwidths of the filters actually used in the model. In other words, the nonlinearities in the model in combination with the simultaneous notched-noise method resulted in biased estimates of filter bandwidth. As their model incorporated stronger nonlinearity at high frequencies than at low frequencies (in line with relevant physiological data), the bias effects increased with increasing center frequency.

The prediction that suppression may affect tuning estimates more at high characteristic frequencies (CFs) than at low CFs is important: If suppression simply changes the effective tuning by a constant factor, regardless of frequency, then this would be easy to incorporate into current functions relating auditory filter ERB to frequency (Glasberg and Moore 1990). On the other hand, if suppression effects change with frequency then this could affect the entire shape of the ERB function. In either case, as has been pointed out previously (e.g., Glasberg and Moore 2000), it is clear that caution must be exercised when relating data from simultaneous-masking experiments to cochlear tuning, as measured in physiological experiments.

Ideally, a psychoacoustic measure of cochlear tuning would share many of the properties of neural measures of tuning. Most importantly, perhaps, the stimuli should be at a low level (where tuning is generally sharpest) and the masker and signal should not be presented simultaneously (to avoid suppression effects). There are surprisingly few modern studies of frequency selectivity that meet these criteria. Using the notched-noise method, Glasberg and Moore (1982) measured auditory filter shapes in forward masking for a fixed low-level 1-kHz signal. They found a 3-dB bandwidth of about 85 Hz. Using a rounded-exponential (roex) model, as they did, this corresponds to an ERB of about 100 Hz—about a factor of 1.3 less than the value given by Eq. (1). Using a relatively high masker level, Moore et al. (1987) compared auditory filter shapes in forward and simultaneous masking directly and also found that the resulting ERBs differed by a factor of about 1.3. To our knowledge, there are no notched-noise studies using forward masking that have measured frequency selectivity at any signal frequency other than 1 kHz. There are both physiological (Cooper and Yates 1994; Rhode and Cooper 1996) and psychophysical (Hicks and Bacon 1999; Plack and Oxenham 2000) indications that nonlinearity increases at high CFs. Furthermore, Heinz et al. (2002) showed in their model that increasing nonlinearity leads to increasingly biased estimates of cochlear tuning when using simultaneous masking. Thus, it is possible that the differences between estimates using forward masking and those using simultaneous masking will increase with increasing CF. If, as is often assumed, forward masking produces a measure that is more comparable to physiological measures of cochlear tuning, then knowledge of how frequency selectivity in forward masking changes with CF should provide valuable information on the nature of human cochlear tuning.

This study measured frequency selectivity in forward masking using the notched-noise method for

signal frequencies between 1 and 8 kHz, with the aim of providing improved estimates of human cochlear tuning as a function of frequency. For comparison, data from simultaneous-masking conditions were collected in the same subjects. In an effort to match the conditions of neural tuning curves as closely as possible, the signal was presented at a fixed low level and the masker level was varied to measure threshold. This fixed-signal method has been shown to be preferable to a fixed-noise method for notched-noise measures on both empirical and theoretical grounds (Rosen and Baker 1994; Rosen et al. 1998; Glasberg and Moore 2000).

EXPERIMENT: AUDITORY FILTER SHAPES AT LOW LEVELS IN FORWARD AND SIMULTANEOUS MASKING

Stimuli

Thresholds were measured for a 10-ms signal (5-ms raised-cosine ramps; no steady state) in the presence of 400-ms bands of noise, also gated with 5-ms raised-cosine ramps. In the forward-masking conditions, the silent interval or gap between the masker and the signal was 5 ms (defined as the duration over which the envelope voltage was 0 V). In the simultaneous-masking conditions, the onset of the signal occurred 380 ms after the masker onset. Thresholds were measured for signal frequencies (f_s) of 1, 2, 4, 6, and 8 kHz. The masker consisted of two bands of Gaussian noise centered below and above the signal frequency, each with a bandwidth of $0.25f_s$. The noises were generated in the spectral domain and were band-limited by setting all spectral components outside the desired passband to zero. In that way, the slope of the filtering was limited only by the 5-ms onset and offset ramps of the masker. The spectral notch width between the two noises provided the independent variable at each signal frequency. The notch was defined as the normalized deviation of the closer edge of each noise, Δf , from the signal frequency, i.e., $\Delta f/f_s$. There were five conditions in which the notch was placed symmetrically about the signal: values of $\Delta f/f_s$ were 0 (no spectral notch), 0.1, 0.2, 0.3, and 0.4. Two asymmetric conditions were also tested, where the upper and lower normalized deviations were 0.2 and 0.4 and vice versa. This provided a total of seven conditions at each signal frequency.

All stimuli were generated digitally at a sampling rate of 32 kHz and were played out via a LynxStudio LynxOne soundcard at 16-bit resolution. In the forward-masking conditions, where the level difference between the signal and masker could be very large, the masker and signal were passed through different programmable attenuators (TDT PA4) before being

mixed (TDT SM3) and passed through a headphone buffer (TDT HB6). In the simultaneous-masking conditions, where the level differences between the masker and signal were not as great, the masker and signal were added digitally before being passed through a single programmable attenuator and headphone buffer. The stimuli were presented monaurally in a double-walled sound-attenuating booth via Etymotic Research ER2 insert earphones, which are designed to provide a flat frequency response at the eardrum up to about 14 kHz.

Procedure

Initially, thresholds in quiet were measured for the signals at all the test frequencies. This was done using a three-interval three-alternative forced-choice method with a two-down one-up adaptive procedure that tracks the 70.7%-correct point on the psychometric function. Intervals were marked on a virtual response box on a flat-panel monitor located in the booth, responses were made via the computer keyboard or mouse, and feedback was provided after each trial. A run was terminated after 10 reversals. The step size was 8 dB for the first two reversals, 4 dB for the following two, and 2 dB thereafter. Threshold was defined as the mean signal level at the last 6 reversals. For each subject and frequency, a mean threshold was calculated from at least three repetitions of each condition.

The absolute threshold values for each listener and frequency were then used to determine the individual fixed signal level in the masking experiments. Eight subjects were tested in the forward-masking conditions. The signal was presented 10 dB above its threshold in quiet (referred to here as 10 dB SL) and the masker level was adaptively varied. One group of four subjects (who were tested earlier) were also tested in simultaneous masking, with the signal presented at 10 dB SL. The other four subjects (who participated later) were tested in simultaneous masking with the signal presented at 35 dB SL.

The presentation and tracking procedure in the masking experiments was the same as for the absolute thresholds, except that the masker level increased after two consecutive correct responses and decreased after each incorrect response. Also, a run was terminated after 12 reversals and the mean masker spectrum level at the last 8 reversals was defined as the threshold. Any run where the standard deviation (SD) of the reversal was greater than 4 dB was discarded.

After three successful runs ($SD < 4$ dB) of a given condition with a given subject were completed, the mean and SD across runs were calculated. If the SD across the three runs exceeded 4 dB, another run was

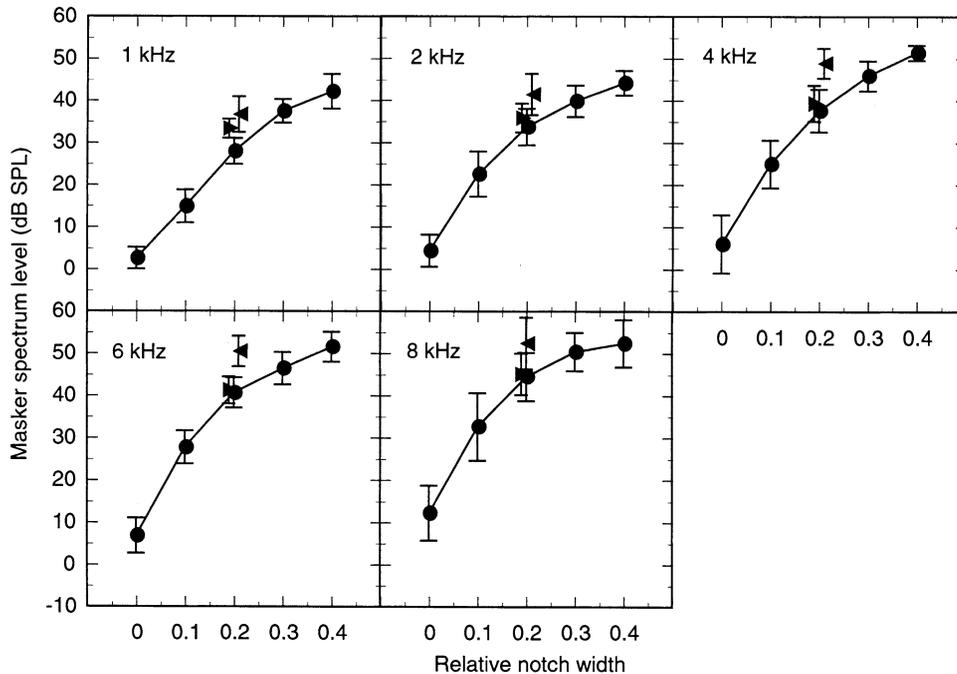


FIG. 1. Mean data from the forward-masking condition. Masker level at threshold is plotted as a function of notch width. Circles denote conditions with symmetrically placed spectral notches; triangles denote asymmetric conditions where the lower and upper edges of the notch are $0.2f_s$ and $0.4f_s$ Hz (right-pointing triangles) or $0.4f_s$ and $0.2f_s$ Hz (left-pointing triangles) from the signal frequency, respectively. Error bars denote ± 1 SD of the mean across the eight listeners. The signal was fixed at 10 dB above the threshold in quiet for each listener individually.

undertaken. The mean and SD of the last three runs was calculated. If the SD still exceeded 4 dB, this procedure was repeated until a total of six runs had been undertaken. In the very rare event of the SD of the last three runs still exceeding 4 dB, data from all six runs were combined and the mean and SD across all six runs are reported.

Subjects

Eight normal-hearing listeners (four females, four males), aged between 18 and 29, participated in these experiments. All had absolute thresholds at octave frequencies between 250 and 8000 Hz of 15 dB HL or lower. All were students or recent graduates in the Boston area and were paid for their participation. Listeners were given about 2 hours of practice, after which their performance appeared stable. The experimental protocol was approved by MIT's committee on the use of human experimental subjects and written informed consent was obtained from all listeners.

Results

The pattern of results was similar across listeners and so only the mean results are shown. Mean thresholds in quiet (and between-subject SDs in parentheses) for the 10-ms signal presented over ER2 insert earphones were 25.0 (2.9), 26.5 (4.0), 26.3 (5.3), 29.0 (2.6), and 31.7 (3.1) dB SPL at 1, 2, 4, 6, and 8 kHz, respectively. Within-subject SDs were typically less than 2 dB. Thresholds in forward masking were measured in all

listeners with the signal presented 10 dB above the individual listener's threshold in quiet. The mean results across the eight listeners are shown in Figure 1. Error bars denote ± 1 SD of the mean. As expected, the circles show that with increasing symmetric masker notch width, the masker level necessary to mask the signal increases. The left- and right-pointing triangles denote the two asymmetric notch conditions. If the auditory filters were symmetric, both conditions would produce the same threshold; larger differences suggest greater asymmetries. In Figure 1, the fact that the left-pointing triangles are consistently higher than the right-pointing triangles suggests that the auditory filters are asymmetric, with a steeper high-frequency slope.

Figure 2 shows the mean results from the subset of four listeners who were tested in simultaneous masking with a 10-dB SL signal (mean absolute thresholds in this subset were 28.8, 28.9, 26.9, 28, and 30 dB SPL at 1, 2, 4, 6, and 8 kHz, respectively). A cursory comparison of Figures 1 and 2 suggests a difference in frequency selectivity, depending on whether the signal was presented simultaneously (Fig. 2) or nonsimultaneously (Fig. 1) with the masker. First, the increase in masker level with increasing notch width appears shallower in the simultaneous-masking case, suggesting poorer effective frequency selectivity. Second, the asymmetry of masking (indicated by the relative heights of the triangles) in the simultaneous-masking case is reversed compared with the nonsimultaneous case for all frequencies. It should be noted that all references to filter asymmetry are based on only two data points in each condi-

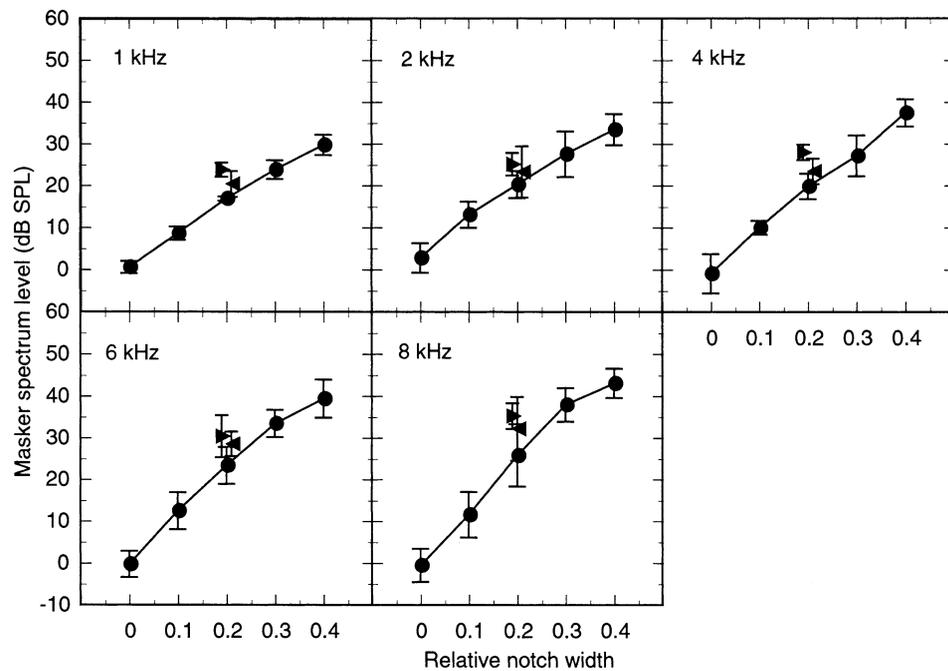


FIG. 2. Mean data from the simultaneous-masking condition, with the signal level fixed at 10 dB above listeners' thresholds in quiet. The symbols represent the mean of the four listeners who participated in this condition. Other aspects of the figure are as in Figure 1.

tion. Nevertheless, the trends discussed here were consistent across listeners.

Figure 3 shows the mean results from the subset of four listeners who were tested in simultaneous masking with a 35-dB SL signal (mean absolute thresholds in this subset of listeners were 23.3, 24.1, 25.8, 30, and 33.3 dB SPL at 1, 2, 4, 6, and 8 kHz, respectively). The slope of the masking function seems similar to that found for simultaneous masking at the lower signal level (Fig. 2) and hence shallower than in the forward-masking case (Fig. 1). However, the asymmetry of masking in Figure 3 is consistent with that found in the forward-masking data, not with that found in the simultaneous-masking data at the lower level (Fig. 2). These issues are discussed in more detail in the following section, within the context of the auditory-filter model.

DERIVING AUDITORY FILTERS FROM THE DATA

Methods

Both the individual and mean data were used to derive auditory filter shapes. Briefly, two versions of the roex(p, w, t) function (Glasberg et al. 1984) were used in conjunction with the middle ear function, described by Moore et al. (1997). These choices and their rationale are described in more detail below.

Outer and middle ear transfer functions. The insert earphones used in this study (Etymotic Research ER2) are designed to produce a flat transfer function at the eardrum for frequencies up to about 14 kHz.

For this reason no filtering was used within the model to simulate outer ear filtering. Two approaches were used in simulating the middle ear transfer function. The first was to ignore middle ear filtering completely. This approach is probably the most appropriate when comparing the results to physiological tuning curves, as most physiological studies of cochlear tuning have not taken middle ear filtering effects into account. The second approach was to use the function described by Moore et al. (1997), based in part on the middle ear functions measured by Puria et al. (1997), which is similar to that used in earlier modeling studies (Glasberg and Moore 1990) and which has also been used in subsequent studies (Glasberg and Moore 2000). Because of its previous use, this approach is preferable when comparing our data with those of earlier notched-noise studies of behavioral frequency selectivity. In practice, the inclusion of the middle ear transfer function had no systematic effect on the estimated ERBs of our data (although it often affected the estimated asymmetry of the filters), so that the overall conclusions were not affected by our choice. For simplicity, and to facilitate comparisons with earlier behavioral studies, the analyses and parameters discussed in this article refer only to those conditions in which the middle ear transfer function was included.

Filter shapes. Filter shapes were derived using methods similar to those described in many previous studies (e.g., Patterson and Nimmo-Smith 1980; Glasberg and Moore 1990, 2000). The basic assumed filter shape was the rounded exponential (roex) filter. Both the roex(p, r) (Patterson et al. 1982) and the

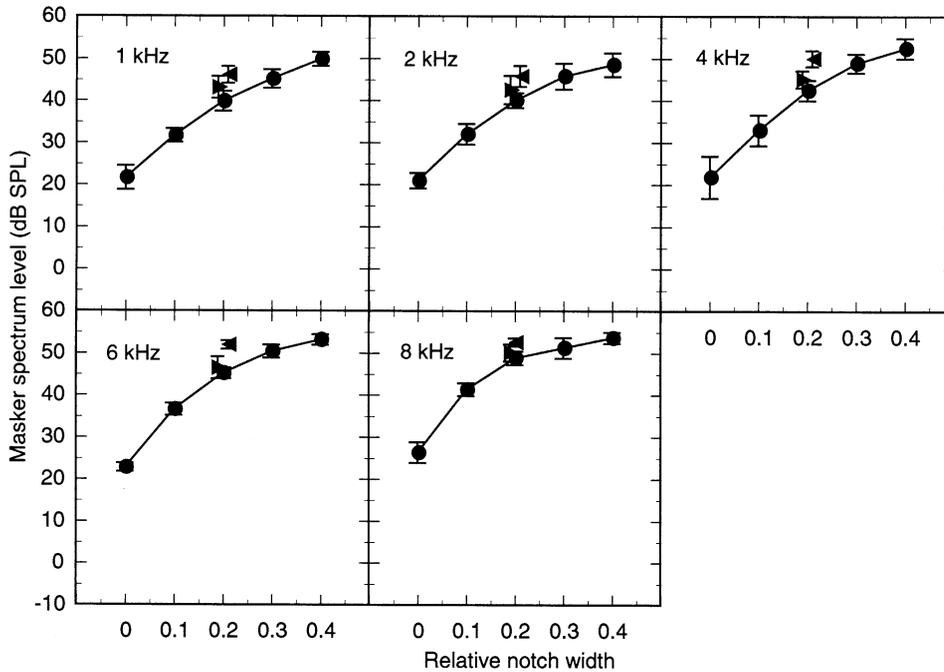


FIG. 3. Mean data from the simultaneous-masking condition, with the signal level fixed at 35 dB above listeners' thresholds in quiet. The symbols represent the mean of the four listeners who participated in this condition. Other aspects of the figure are as in Figure 1.

roex(p, w, t) (Glasberg et al. 1984) were tested. In addition, a variant of the roex(p, w, t), as recently used by Rosen et al. (1998) and Glasberg and Moore (2000), was also tested. These three filter shapes are described in turn. The equation for one side of the roex(p, r) filter is

$$W(g) = (1 - r)(1 + pg) \exp(-pg) + r \quad (2)$$

where W is the filter weighting function, g is the deviation from the center frequency as a proportion of the center frequency, p is the parameter determining the slope of the filter, and r is the parameter defining the dynamic range of the filter. As in previous studies, p was allowed to differ on either side of the filter (p_u for the upper side and p_l for the lower side), whereas the value of r was assumed to be the same on both sides of the filter. The equation for the roex(p, w, t) filter is

$$W(g) = (1 - w)(1 + pg) \exp(-pg) + w(1 + pg/t) \exp(-pg/t) \quad (3)$$

The difference between Eq. (2) and (3) is that Eq. (3) has two slopes instead of a dynamic range limiter. The parameter t determines the factor by which the second slope is shallower than the first; the parameter w determines the relative weights of the first and second slope, or the point on the filter function at which the second slope begins to dominate. Both t and w are assumed to be the same on both sides of the filter. Thus, the roex(p, w, t) function has one more free parameter than the roex(p, r) function. The variant of the roex(p, w, t) filter has a lower side as

described by Eq. (3). Its upper side, however, is described by simply one slope:

$$W(g) = (1 + pg) \exp(-pg) \quad (4)$$

This version has the same number of parameters as the original roex(p, w, t) function, with the only difference being that the w and t parameters are not used for the upper side. This version is referred to as the roex(p, w, t, p) filter.

All three filter forms were tested with both individual and mean data. In almost all cases, the roex(p, r) fits were worse than the roex(p, w, t) fits, and the overall root mean squared (rms) deviations of the model predictions from the individual data were higher by about a factor of 2, which is more than expected given simply the difference in the number of free parameters.

Qualitatively, this can be appreciated in the data by noting that masker thresholds at larger notch widths continue to increase, but at a slower rate than at the smaller notch widths. In contrast, the roex(p, r) model requires the thresholds to increase initially at a roughly constant rate and then to remain constant. Because of the poorer fits of the roex(p, r) model, both qualitatively and quantitatively, only the roex(p, w, t) fits are discussed further in this article. As shown below, the roex(p, w, t) and roex(p, w, t, p) models provided comparably good fits to the data and resulted in similar estimates of filter bandwidth.

Fitting procedure. A multidimensional nonlinear minimization routine [Nelder–Mead, as implemented in Matlab (Mathworks, Natick, MA)] was used to find the best-fitting parameters of the filters in the

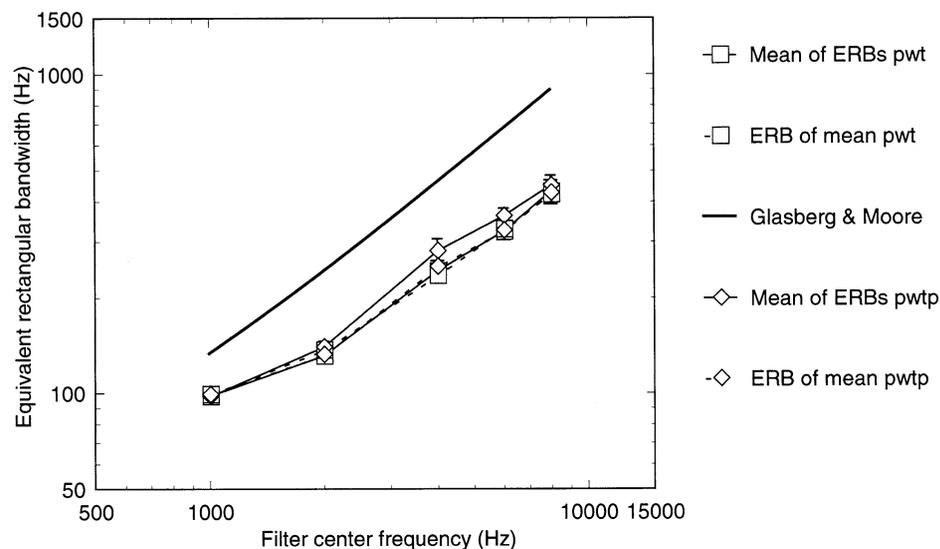


FIG. 4. Equivalent rectangular bandwidths (ERBs) of the filters derived from forward masking as a function of filter center frequency. Squares represent estimates using the $\text{roex}(p,w,t)$ model and diamonds represent estimates using the $\text{roex}(p,w,t,p)$ model; see text for details. Symbols connected with solid lines represent mean ERB estimates from the individual data, whereas symbols connected with dashed lines represent ERB estimates from the mean data. Error bars are shown only for symbols connected with solid lines and denote 1 SE of the mean estimate. The heavy solid curve shows the predicted ERBs of the function proposed by Glasberg and Moore (1990).

least-squares sense. It was assumed that the signal was detected by the filter with the best signal-to-noise ratio (SNR). In most situations this was also the filter centered at the signal frequency. However, in some cases the filter was centered somewhat away from the signal that had the best SNR, although its CF was always within 10% of the signal frequency. The “efficiency” of the detector, K , is the threshold SNR at the output of the detection filter across all conditions. The squared deviation of the predicted masker levels at threshold (in dB) with a constant K from the actual thresholds was used to drive the minimization routine.

Model predictions and filter shapes

Auditory filter shapes in forward masking. Data from the forward-masking conditions were collected from all eight listeners, and both the mean and individual data were used to produce auditory filter shapes with both the $\text{roex}(p,w,t)$ and $\text{roex}(p,w,t,p)$ filters. Both functions provided good descriptions of the data. As they each resulted in very similar overall rms errors for the individual data of around 0.9 dB, it was not possible to distinguish between them based on goodness of fit. The predicted thresholds are not shown in Figure 1, as the deviations between the predicted thresholds and the actual thresholds were always so small as to be essentially indistinguishable on the scale of the figure.

The ERB provides a convenient measure of filter tuning, although other measures, such as the 3-dB or 10-dB bandwidth, may also be used. All three measures were tried and all resulted in the same basic pattern of results and conclusions. For ease of comparison, only the ERBs are discussed here. Figure 4 shows four estimates of the ERB as a function of sig-

nal frequency. The symbols connected with solid lines denote the mean ERB values derived from the individual data, and the symbols connected with dashed lines denote the ERB values derived from fitting the function to the mean data. Squares and diamonds denote estimates using the $\text{roex}(p,w,t)$ and $\text{roex}(p,w,t,p)$ filters, respectively. Error bars for the symbols connected with solid lines, representing ± 1 standard error (SE) of the arithmetic mean, are shown where they exceed the size of the symbol. The heavy curve with no symbols shows the ERB function as predicted by the equation given in Glasberg and Moore (1990), referred to here as the GM90 function.

The close correspondence between the mean ERBs (solid lines) and the ERBs of the mean data (dashed lines) provides some evidence for the validity and reliability of the fitting procedure. In general, there are no consistent differences in ERBs between the two filter functions [squares for $\text{roex}(p,w,t)$; diamonds for $\text{roex}(p,w,t,p)$]; there is a tendency for the $\text{roex}(p,w,t,p)$ model to produce somewhat larger ERB values at 4 and 6 kHz, but the differences are never greater than 20%. An obvious aspect of Figure 4 is the large difference between the ERBs from the present data and the GM90 function. At 1 kHz, the GM90 ERB is about a factor of 1.3 larger; at 8 kHz the difference is a factor of about 2.1. These differences in tuning are illustrated further in Figure 5. Here, the ERB (in Hz) has been replaced by a dimensionless value, denoted Q_{ERB} . This is simply the filter center frequency divided by the ERB; a large Q_{ERB} value implies sharp frequency tuning. In this graph, the geometric mean Q_{ERBs} from the individual data fits are shown, together with ± 1 SE of the geometric mean. As in Figure 4, squares and diamonds represent fits using the $\text{roex}(p,w,t)$ and $\text{roex}(p,w,t,p)$ fil-

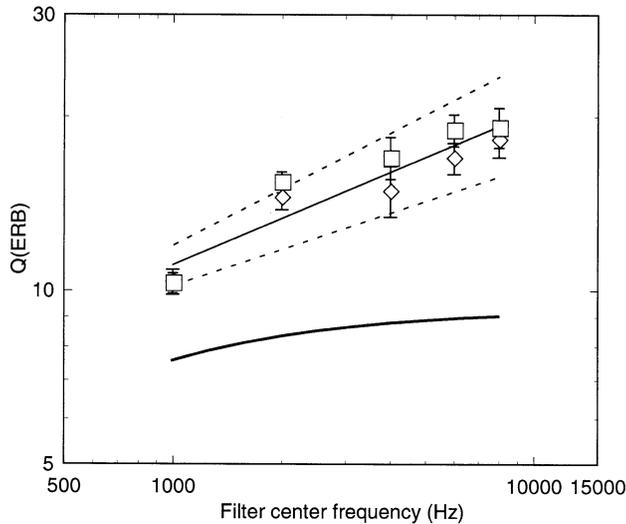


FIG. 5. Geometric mean values of Q_{ERB} from the $\text{roex}(p,w,t)$ (squares) and $\text{roex}(p,w,t,p)$ (diamonds) models. Error bars denote 1 SE of the mean. The Q_{ERB} is a dimensionless measure of tuning, defined as the center frequency divided by the ERB. The heavy solid curve shows the predicted Q_{ERB} , as defined by Glasberg and Moore (1990). The light solid line and the dashed lines represent the best-fitting power function to the pooled estimates and the 95 confidence intervals, respectively.

ters, respectively. In contrast to the tuning of the GM90 function, which stays reasonably constant over the frequency range of interest, the new data suggest tuning that increases substantially from a Q_{ERB} value of about 10 at 1 kHz to a value of nearly 20 at 8 kHz.

Shera and Guinan (2003) found that a power law well described the changes in Q_{ERB} in cat and guinea-pig auditory-nerve tuning data as a function of CF. A power law could also describe human tuning data derived indirectly from otoacoustic emissions data (Shera et al. 2002). The same type of fit was used here. The equation used is

$$Q_{\text{ERB}} = \beta F^\alpha \quad (5)$$

where F is the CF in kHz and α and β are free parameters. Given that there was no compelling reason to select one filter type over the other, individual estimates from both the $\text{roex}(p,w,t)$ and the $\text{roex}(p,w,t,p)$ models were pooled. The solid and dashed lines in Figure 5 represent the best-fitting power-law function and the 95% confidence intervals, respectively. The parameters used are $\alpha = 0.27 \pm 0.06$ and $\beta = 11.1 \pm 0.9$. On the log coordinates used in Figure 5, the function is a straight line, with α determining the slope and β determining the intercept or vertical position of the line. If sharpness of tuning did not change as a function of frequency, α would be close to zero.

The filter shapes derived from the mean data for both the $\text{roex}(p,w,t)$ model (solid lines) and the

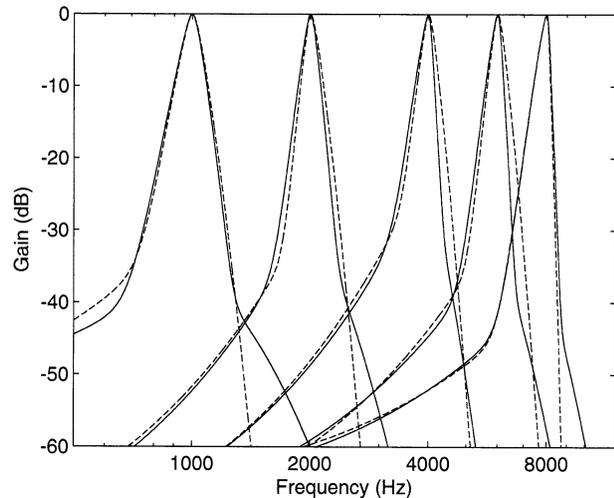


FIG. 6. Auditory filter shapes with center frequencies between 1 and 8 kHz, derived from the mean data in the forward-masking condition. Solid and dotted curves denote $\text{roex}(p,w,t)$ and $\text{roex}(p,w,t,p)$ filters, respectively.

$\text{roex}(p,w,t,p)$ model (dotted lines) are shown in Figure 6. Again, the increase in the sharpness of tuning with increasing CF is evident. The filter parameters and ERBs from the mean data are given in Table 1. Both Figure 6 and Table 1 indicate that the lack of a second slope on the upper side generally made the $\text{roex}(p,w,t,p)$ more symmetric than the $\text{roex}(p,w,t)$ around the peak (the values of p_u and p_l are more similar). No other systematic differences are apparent.

Auditory filter shapes in simultaneous masking. As with the forward-masking data, both the $\text{roex}(p,w,t)$ and the $\text{roex}(p,w,t,p)$ functions provided good fits to the simultaneous-masking data, with rms errors, pooled across all subjects, of about 0.9 and 0.6 dB for the 10-dB SL and 35-dB SL conditions, respectively. The mean Q_{ERB} values from the $\text{roex}(p,w,t,p)$ model for both the 10- and 35-dB SL conditions are shown in Figure 7 as downward- and upward-pointing triangles, respectively. Q_{ERB} values from the $\text{roex}(p,w,t)$ model (not shown) were very similar. For comparison, the Q_{ERB} s from the forward-masking condition (diamonds), also with the $\text{roex}(p,w,t,p)$ model, together with the predictions of the GM90 function (solid curve), are replotted from Figure 5. As expected, the Q_{ERB} values derived from the simultaneous masking are smaller than those derived from the forward masking, indicating broader tuning. The tuning still appears to be marginally sharper than that predicted by the GM90 function, a difference which seems to increase with increasing frequency. This may well be due to the relatively low masker levels used in the present study. It is known that filter shapes broaden at higher levels, especially at higher frequencies, and the GM90 function is specifically designed to predict

TABLE 1

Model parameters derived from the mean forward-masking data^a

Frequency (Hz)	roex(p, w, t)						roex(p, w, t, p)					
	p_l	p_u	t	$10 \log(w)$	K	ERB	p_l	p_u	t	$10 \log(w)$	K	ERB
1000	37.5	43.8	5.86	-36.8	12.7	99.2		39.8	39.9	4.60	-31.0	12.6
2000	56.1	61.6	4.01	-31.0	10.8	137		68.9	47.9	4.98	-30.7	10.9
4000	50.3	108	3.37	-25.6	6.88	235	59.9	59.0	3.84	-23.7	6.18	271
6000	56.1	106	4.79	-34.7	6.96	328	65.3	60.3	5.06	-32.3	6.49	383
8000	51.4	146*	5.74	-40.1	2.97	421	51.0	181*	6.12	-41.2	3.14	402

^aParameters are given for the roex(p, w, t) and roex(p, w, t, p) models. The filter parameters are described in Eqs. (3) and (4); K is the detector efficiency, or signal-to-noise ratio at the output of the filter; and ERB is the equivalent rectangular bandwidth of the filter.

*Such large values should not be taken literally. As discussed by Glasberg and Moore (1990), if one slope is steeper than the other by more than a factor of 2, the steeper slope is generally poorly defined.

tuning in the presence of a masker with a constant spectrum level of around 30 dB SPL. Even in our 35-dB SL signal condition, the masker spectrum level in the no-notch condition was generally around 10 dB lower than that.

At first glance it appears that the filters from the 35-dB SL condition may be somewhat sharper than those from the 10-dB SL condition, despite considerable overlap of the error bars. Closer inspection of the individual data, however, reveals that the apparent difference is probably due to individual differences in frequency selectivity: Recall that the two simultaneous-masking conditions were completed by two different groups of listeners. In fact, mean forward-masking ERBs in listeners who completed the 10-dB simultaneous-masking condition were somewhat broader than those for listeners in the other group, although not significantly so (t -test, $p > 0.05$ at all CFs). This point is illustrated in Figure 8, which shows the individual ratios of the ERBs for the simultaneous masker to the ERBs for the forward masker, using the roex(p, w, t, p) model in all cases. Filled and open symbols represent individual listeners in the 10-dB and 35-dB SL simultaneous-masking conditions, respectively. The geometric mean ratios for the 10-dB and 35-dB conditions are shown by the solid and dashed lines, respectively. It can be seen that there is no systematic difference between the ratios from the two conditions. This in turn suggests that the ERB, as measured here with simultaneous masking, does not change much with level at very low levels. This is consistent with the idea that the input-output function of the basilar membrane is approximately linear for the first 20–40 dB above threshold, as suggested by both physiological (Ruggero et al. 1997) and psychophysical studies (Oxenham and Plack 1997; Plack and Oxenham 1998). Of course, our measure of tuning, the ERB or Q_{ERB} , depends more on the shape of the filter tip than that of the

tail. This makes the measure relatively insensitive to certain changes in filter shape, such as a change in the tip-to-tail ratio (Rosen et al. 1998), which is closely related to the filter parameter w . Inspection of this parameter across the two simultaneous-masking conditions did not reveal a consistent trend, although this too may be due to the across- rather than within-subject nature of the comparison. The lack of an effect of level on the parameter w can be observed in the fits to the mean data, where the best-fitting parameters for the 10-dB and 35-dB simultaneous-masking conditions are shown in Tables 2 and 3, respectively.

One change that does seem to occur with level in the simultaneous-masking conditions relates to the asymmetry of the auditory filter. The 10-dB simultaneous-masking condition (Table 2) consistently yielded asymmetries in the opposite direction to those found for the forward-masking conditions. The asymmetries for the 35-dB simultaneous-masking conditions were less marked and, at least for the roex(p, w, t) model, were in the same direction as for the forward-masking conditions.

Values of K in forward and simultaneous masking. Within the power spectrum model, the value of K represents the detector efficiency, or the threshold signal-to-noise ratio at the filter output. While its meaning is clear for long-duration signals embedded in long-duration noise, the definition becomes problematic for short-duration signals and becomes even more so in the case of forward masking. The values of K quoted in the tables are simply the raw signal-to-noise ratios at threshold, without regard to differences in duration or temporal position. In general, there is a trend for values of K to decrease with increasing CF, which is more pronounced in the forward-masking condition than in the simultaneous-masking conditions. However, to be able to compare values of K across condition and studies, the effects of

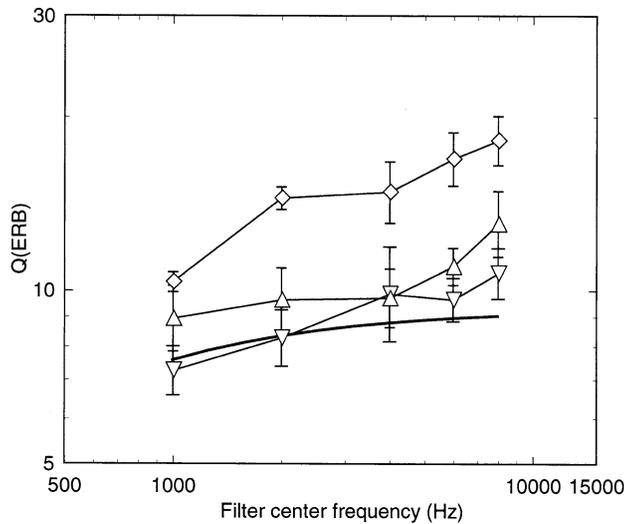


FIG. 7. Geometric mean Q_{ERB} values in simultaneous masking with the signal level fixed at 10 dB (down-pointing triangles) or 35 dB (up-pointing triangles). Error bars denote 1 SE of the mean across the four listeners in each condition. Values of $\text{roex}(p,w,t,p)$ Q_{ERB} s from the forward-masking condition (diamonds) and the Glasberg and Moore predictions (solid curve) are replotted from Figure 5 for comparison.

temporal integration would have to be accounted for, as would the decay of forward masking. Thus, without making further assumptions about the nature of processing, it is possible to make only qualitative observations about trends in K within a given condition.

DISCUSSION

The purpose of this study was to reexamine notched-noise measures of frequency selectivity by using a behavioral technique that matched physiological measures of tuning as closely as possible. This was achieved by using forward masking to avoid suppressive interactions between the masker and signal and by presenting the signal at a fixed low level (10 dB above threshold in quiet). The resulting auditory filters show substantially narrower tuning than those derived from earlier studies employing simultaneous masking, often at higher levels. Certain aspects and consequences of the findings are discussed below.

Frequency selectivity as a function of CF

Although it has been known for some time that forward masking produces sharper tuning than does simultaneous masking, the differences have not before been studied systematically as a function of frequency. Perhaps the most important finding of this study is that frequency tuning increases substantially with CF between 1 and 8 kHz. Because of this, the difference

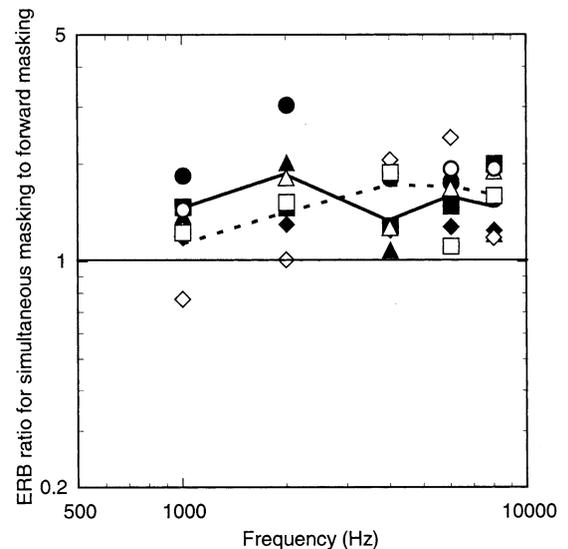


FIG. 8. Individual ratios of simultaneous-masking ERBs to forward-masking ERBs. Filled symbols denote listeners who participated in the 10-dB SL simultaneous-masking condition; open symbols denote listeners who participated in the 35-dB SL simultaneous-masking conditions. Geometric mean ratios for the 10-dB and 35-dB SL conditions are shown by the solid and dashed lines, respectively. All estimates are from the $\text{roex}(p,w,t,p)$ model.

between the frequency selectivity implied by the GM90 equation and that found here increases with increasing frequency, as shown in Figures 5 and 6.

Our estimated ERB of about 100 Hz at a signal frequency of 1 kHz is in excellent agreement with the only other study of notched-noise forward masking at low masker levels (Glasberg and Moore 1982); as mentioned in the Introduction, they also found the ERB at 1 kHz to be about 100 Hz. To our knowledge, no other notched-noise studies have examined frequency selectivity in forward masking at frequencies other than 1 kHz. One study involving psychophysical tuning curves, with a low-level notched noise added to avoid possible confounding factors such as off-frequency listening and “confusion” effects (e.g., Neff 1986), measured tuning at 0.5, 1, 2, and 4 kHz in two listeners (Moore et al. 1984). Unfortunately, although the bandwidth estimates at 1 kHz were similar to ours, at 4 kHz the estimates of tuning in the two listeners differed from each other by a factor of 2 (3-dB bandwidths of 210 and 440 Hz), making it difficult to draw conclusions regarding changes in bandwidth with frequency.

Although our forward-masking data show generally sharper tuning than our simultaneous-masking data, all three conditions (forward masking, simultaneous masking at 10 dB SL, and simultaneous masking at 35 dB SL) show an increase in tuning with increasing CF. This can be seen directly in Figure 7 and indirectly in Figure 8, in that there seems to be no consistent trend for the ratio of ERBs between forward and si-

TABLE 2

Filter parameters derived from the mean simultaneous-masking data, with the signal presented at 10 dB SL^a

Frequency (Hz)	roex(p, w, t)						roex(p, w, t, p)					
	p_l	p_u	t	$10 \log(w)$	K	ERB	p_l	p_u	t	$10 \log(w)$	K	ERB
1000	32.9	22.6	2.69	-20.2	14.4	152	39.0	21.2	4.27	-20.0	14.6	147
2000	63.9	39.8	2.31	-7.99	13.4	197	182	24.2	12.1	-20.9	14.3	189
4000	44.0	33.8	2.00	-12.8	11.3	441	81.2	25.0	5.17	-17.8	11.2	425
6000	45.8	30.9	2.56	-23.3	10.0	656	50.1	28.9	4.41	-28.4	10.2	656
8000	48.0	36.8	10.7	-49.4	11.8	768	48.7	36.9	7.89	-45.8	11.9	762

^aA subset of four from the eight listeners was tested in this condition. Parameters are given for the roex(p, w, t) and roex(p, w, t, p) models. The filter parameters are described in Eqs. (3) and (4); K is the detector efficiency, or signal-to-noise ratio at the output of the filter; and ERB is the equivalent rectangular bandwidth of the filter.

multaneous masking to change as a function of CF. Thus, the difference between our data and those of previous studies cannot be ascribed wholly to the difference between simultaneous and nonsimultaneous masking and the resultant effects of suppression. The difference may also lie in our levels of signal presentation, which were generally lower than in previous studies, and in the fact that the signal level was fixed and the masker level was varied. If changes in filter bandwidth with level are more pronounced at high than at low CFs (e.g., Hicks and Bacon 1999), then changes in tuning with CF should be more pronounced at low levels.

Filter asymmetry in forward and simultaneous masking

While the estimated ERBs and the goodness-of-fit of the roex(p, w, t) and roex(p, w, t, p) models were very similar, one difference involved the estimated amount of asymmetry at the tip of the filter. In general, roex(p, w, t) fits produced lower values of the ratio of asymmetry, p_u/p_l , indicating a steeper upper filter slope relative to the lower filter slope. This can be understood in terms of the additional, shallower slope on the upper side of the roex(p, w, t) function.

Glasberg and Moore (2000) concluded that a function equivalent to the roex(p, w, t, p) function used here could produce good fits to the data if the asymmetry ratio was fixed at unity, i.e., if the tip of the filter was assumed to be symmetric. An examination of the functions derived from the forward-masking (Table 1) and the 35-dB simultaneous-masking (Table 3) conditions suggests that the same may be true of our data: The values of p_u and p_l do not seem to deviate systematically from one another, except at 8 kHz in the forward-masking condition. Thus, the filters are somewhat asymmetric, with steeper upper than lower slopes; this asymmetry can be expressed in terms of either a filter with an asymmetric tip region,

as for the roex(p, w, t), or a filter with a symmetric tip region but an asymmetric tail region, as for the roex(p, w, t, p).

The situation is different in the 10-dB simultaneous-masking condition. Here, both filter functions suggest an asymmetry in the opposite direction: The upper slope appears to be shallower than the lower slope. Similar results have been found before in simultaneous masking (e.g., Shailer et al. 1990), and the reversal of asymmetry at low levels is included in Glasberg and Moore's (1990) analysis of auditory filter shapes [their Eq. (5)]. The fact that this asymmetry reversal is found only in simultaneous masking suggests that it too may be related to suppression effects. Both physiological and psychophysical studies of suppression have shown that high-side suppression (where the suppressor is above the signal in frequency) occurs at lower sound pressure levels than low-side suppression (e.g., Duifhuis 1980; Delgutte 1990). If suppression broadens the apparent tuning, then it should affect the upper slope of the filter at lower levels than it affects the lower slope of the filter, leading to the apparent asymmetry reversal. In conclusion, the asymmetry reversal at low levels is probably a consequence of suppression and does not reflect cochlear tuning *per se*.

Comparisons with physiological data from humans and animals

The present work was motivated by a recent study of stimulus-frequency otoacoustic emissions (SFOAEs) in humans, cats, and guinea pigs (Shera and Guinan 2003). This study found that the frequency dependence of SFOAE group delays correlates well with changes across CF in the bandwidths of auditory-nerve tuning curves in cat and guinea pig. Shera et al. (2002) showed that although SFOAE-based estimates of human cochlear tuning do not match well with the GM90 function, the otoacoustic estimates are in very

TABLE 3

Filter parameters derived from the mean simultaneous-masking data, with the signal presented at 35 dB SL^a

Frequency (Hz)	roex(p, w, t)						roex(p, w, t, p)					
	p_l	p_u	t	$10 \log(w)$	K	ERB	p_l	p_u	t	$10 \log(w)$	K	ERB
1000	30.5	41.9	3.80	-21.0	16.1	116	38.5	31.7	4.08	-17.4	16.1	118
2000	38.4	34.6	3.65	-23.7	14.6	222	40.4	31.6	4.24	-23.8	14.6	227
4000	32.5	66.0	3.30	-18.8	13.5	378	38.2	39.7	3.46	-15.8	12.9	425
6000	40.2	62.8	5.76	-26.9	15.2	494	49.2	39.8	5.91	-24.6	15.0	550
8000	44.6	68.4	6.76	-30.0	14.2	596	65.7	43.4	8.64	-28.0	14.1	615

^aThe four listeners not tested in simultaneous masking at 10 dB SL were tested in this condition. Parameters are given for the roex(p, w, t) and roex(p, w, t, p) models. The filter parameters are described in Eqs. (3) and (4); K is the detector efficiency, or signal-to-noise ratio at the output of the filter; and ERB is the equivalent rectangular bandwidth of the filter.

good agreement with the filters derived from the present data. This provides independent physiological support for the idea that the present behavioral data in forward masking reflect tuning at the level of the cochlea.

The study of Shera et al. (2002) is one of very few studies to directly compare physiological with behavioral measures of frequency selectivity. Similar work has been done in the guinea pig by Evans and colleagues (Evans et al. 1992; Evans 2001). They found good correspondence between auditory-nerve tuning curves and behavioral measures using simultaneous masking. This good correspondence may seem at odds with our assertion that the suppression effects in simultaneous masking can severely underestimate cochlear tuning. The apparent discrepancy may be resolved by considering that guinea-pig cochlear tuning is generally poorer than that found in humans by a factor of 2 or 3 (Shera et al. 2002). Furthermore, the amount of suppression found in guinea-pig auditory-nerve fibers is often rather small (e.g., Prijs 1989). Thus, differences between forward and simultaneous masking may not be as apparent in that species as they are in humans.

Implications of the revised tuning estimates

The data and analysis presented here suggest that the GM90 function does not provide a good estimate of human cochlear tuning at low levels. A comparison of the GM90 function and, for instance, auditory-nerve tuning curves is therefore not appropriate. Instead, it seems that when the effects of suppression are eliminated by using forward masking, and when low signal levels are used, human cochlear tuning is considerably sharper than previously thought. While it has been known for many years that forward masking produces sharper estimates of tuning than does simultaneous masking, the present results provide the first evidence that tuning, as measured by the Q_{ERB} ,

becomes markedly sharper with increasing frequency between 1 and 8 kHz. Although this is a new finding, it brings the pattern of human cochlear tuning in line with that found in other mammals, such as cat and guinea pig, where tuning also increases substantially with increasing CF (see Shera et al. 2002).

There are a number of situations in which tuning estimates derived from simultaneous masking, such as the GM90 function, remain valid and more appropriate than those from the present study. For instance, in situations where predictions of audibility are required for one high-level stimulus embedded in another, it is of interest whether the stimulus is masked, not whether the masking is due to suppression or excitation, or whether the cochlea is tuned more sharply at low levels. Thus, in applications such as low bit-rate coding of digital audio, where the coding noise must fall below masked or absolute threshold, the current procedures (and the GM90 function) remain valid under most everyday situations.

There are, however, a number of situations in which the use of the estimates from the present study would be more appropriate. The most obvious example involves models of human cochlear processing. The filter shapes derived from the present data provide a basis for defining the basic frequency-tuning properties of such models, just as neural tuning curve data do for cochlear models in animals. In order to account for the differences between simultaneous and forward masking, such models will probably also need to incorporate realistic implementations of suppression mechanisms. Further studies with forward masking, using either higher signal levels or longer gaps between the masker and signal, will be required to define how tuning changes with stimulus level. It is expected that filters will broaden considerably, given previous data using forward-masking psychophysical tuning curves at high levels (e.g., Nelson 1991). The technique used here is difficult to use at frequencies below 1 kHz, as the short duration

of the signal results in a bandwidth exceeding the bandwidth of the auditory filter. However, it is possible that the problem can be overcome using the pulsation-threshold method (Houtgast 1972), which has been used recently by Plack and Oxenham (2000) to provide a behavioral estimate of basilar-membrane compression at low frequencies.

Many models of auditory perception include auditory filtering as the first stage. In most cases, it is assumed that the masking patterns found in simultaneous masking reflect the excitation pattern along the basilar membrane (Zwicker 1960; Florentine and Buus 1981; Moore et al. 1997). Thus, the frequency selectivity assumed in such models is similar to the functions derived from simultaneous-masking data, such as the GM90 function. The present data support earlier assertions (Moore and Vickers 1997; Oxenham and Plack 1998) that the frequency selectivity assumed by such models is incorrect and will require revision.

Finally, the method developed here may be of use in testing the frequency selectivity of listeners with mild-to-moderate hearing loss. In previous studies of hearing-impaired listeners, it has been found that frequency selectivity is generally not greatly affected for hearing losses of 35 dB or less (e.g., Moore et al. 1999). However, it is possible that the apparent insensitivity of frequency selectivity to mild hearing loss may be due to the measurement technique used in previous studies. If suppression in the normal cochlea results in an underestimate of frequency tuning, then a reduction in suppression, concomitant with cochlear hearing loss, may compensate for the decrease in tuning that would otherwise be observed. Recently, Baker and Rosen (2002) showed, using simultaneous masking, that the dependence of frequency selectivity on level could be affected by hearing losses as small as 20 dB. It may be that forward masking will prove to be a more sensitive measure of the effects of hearing loss on frequency selectivity, such that changes in tuning are more readily apparent even for mild hearing losses. If frequency selectivity is determined primarily by outer hair cell function, such a measure may help determine the different contributions of various physiological mechanisms to an individual hearing loss.

SUMMARY

The masker levels necessary to mask a fixed low-level signal between 1 and 8 kHz were measured for eight listeners in forward and simultaneous masking with a varying notched-noise masker. The data were used to derive auditory filter shapes. The following conclusions were drawn:

- The derived tuning from the forward-masking conditions was considerably sharper than has been

found in previous studies using simultaneous masking.

- In forward masking, the dependence of tuning on CF was different from that found in earlier studies using higher-level simultaneous maskers (e.g., Glasberg and Moore 1990). Instead of near-constant relative bandwidth, the sharpness of tuning doubled over the range of frequencies tested, with Q_{ERB} values increasing from around 10 at 1 kHz to around 20 at 8 kHz. The revised estimates of human cochlear tuning as a function of CF between 1 and 8 kHz can be summarized as $Q_{\text{ERB}} = 11F^{0.27}$, where F is the filter CF in kHz.
- In simultaneous masking, tuning was broader than that found in forward masking, although the dependence of tuning on CF was similar. The difference between these results and previous results using simultaneous masking may be due to our use of a low fixed-level signal.
- The results may be of use in constructing models of the human cochlea and should facilitate comparisons of cochlear tuning between humans and other animals.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (National Institute on Deafness and other Communication Disorders grants R01 DC 03909 and R01 DC 03687). We thank Christophe Micheyl, Armin Kohlrausch, and two reviewers for helpful comments on a previous version of the manuscript. An initial analysis of the forward-masking data in this study was presented in Shera et al. (2002).

REFERENCES

- BAKER RJ, ROSEN S. Auditory filter nonlinearity in mild/moderate hearing impairment. *J. Acoust. Soc. Am.* 111:1330–1339, 2002.
- BEAUVOIS MW, MEDDIS R. Computer simulation of auditory stream segregation in alternating-tone sequences. *J. Acoust. Soc. Am.* 99:2270–2280, 1996.
- BOS CE, DE BOER E. Masking and discrimination. *J. Acoust. Soc. Am.* 39:708–715, 1966.
- BREEBAART J, VAN DE PAR S, KOHLRAUSCH A. Binaural processing model based on contralateral inhibition. I. Model structure. *J. Acoust. Soc. Am.* 110:1074–1088, 2001.
- COOPER NP, YATES GK. Nonlinear input–output functions derived from the responses of guinea-pig cochlear nerve fibres: Variations with characteristic frequency. *Hear. Res.* 78:221–234, 1994.
- DAU T, PÜSCHEL D, KOHLRAUSCH A. A quantitative model of the ‘effective’ signal processing in the auditory system. II. Simulations and measurements. *J. Acoust. Soc. Am.* 99:3623–3631, 1996.
- DAU T, KOLLMEIER B, KOHLRAUSCH A. Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers. *J. Acoust. Soc. Am.* 102:2892–2905, 1997.

- DELGUTTE B. Two-tone suppression in auditory-nerve fibers: Dependence on suppressor frequency and level. *Hear. Res.* 49:225–246, 1990.
- DUIFFHUIS H. Level effects in psychophysical two-tone suppression. *J. Acoust. Soc. Am.* 67:914–927, 1980.
- EVANS EF. Latest comparisons between physiological and behavioural frequency selectivity. In: Evans J, Breebaart AJM, Houtsuma A, Kohlrauschl VF, Puijs R (Eds.) *Physiological and Psychophysical Bases of Auditory Function*. Shaker, Maastricht, pp 382–387, 2001.
- EVANS EF, PRATT SR, SPENNER H, COOPER NP. Comparisons of physiological and behavioural properties: Auditory frequency selectivity. In: Cooper Y, Cazals K, Horner L (Eds.) *Auditory Physiology and Perception*. Pergamon, Oxford, pp 159–169, 1992.
- FLETCHER H. Auditory patterns. *Rev. Mod. Phys.* 12:47–65, 1940.
- FLORENTINE M, BUUS S. An excitation-pattern model for intensity discrimination. *J. Acoust. Soc. Am.* 70:1646–1654, 1981.
- GLASBERG BR, MOORE BCJ. Auditory filter shapes in forward masking as a function of level. *J. Acoust. Soc. Am.* 71:946–949, 1982.
- GLASBERG BR, MOORE BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear. Res.* 47:103–138, 1990.
- GLASBERG BR, MOORE BCJ. Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. *J. Acoust. Soc. Am.* 108:2318–2328, 2000.
- GLASBERG BR, MOORE BCJ, PATTERSON RD, NIMMO-SMITH I. Dynamic range and asymmetry of the auditory filter. *J. Acoust. Soc. Am.* 76:419–427, 1984.
- HEINZ MG, COLBURN HS, CARNEY LH. Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates. *J. Acoust. Soc. Am.* 111:996–1101, 2002.
- HICKS ML, BACON SP. Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing. *J. Acoust. Soc. Am.* 105:326–338, 1999.
- HOUTGAST T. Psychophysical evidence for lateral inhibition in hearing. *J. Acoust. Soc. Am.* 51:1885–1894, 1972.
- HOUTGAST T. Psychophysical experiments on “tuning curves” and “two-tone inhibition. *Acustica* 29:168–179, 1973.
- Houtgast T. Lateral suppression in hearing. Ph.D. thesis, Free University of Amsterdam. 1974.
- JOHNSON-DAVIES D, PATTERSON RD. Psychophysical tuning curves: Restricting the listening band to the signal region. *J. Acoust. Soc. Am.* 65:765–770, 1979.
- MOORE BCJ. Psychophysical tuning curves measured in simultaneous and forward masking. *J. Acoust. Soc. Am.* 63:524–532, 1978.
- MOORE BCJ. Distribution of auditory-filter bandwidths at 2 kHz in young normal listeners. *J. Acoust. Soc. Am.* 81:1633–1635, 1987.
- MOORE BCJ. Frequency analysis and masking. In: Moore BCJ (Eds.) *Handbook of Perception and Cognition*, Volume 6. Hearing. Academic Press, Orlando, FL, pp 161–205, 1995.
- MOORE BCJ, VICKERS DA. The role of spread of excitation and suppression in simultaneous masking. *J. Acoust. Soc. Am.* 76:2284–2290, 1997.
- MOORE BCJ, GLASBERG BR, ROBERTS B. Refining the measurement of psychophysical tuning curves. *J. Acoust. Soc. Am.* 76:1057–1066, 1984.
- MOORE BCJ, POON PWF, BACON SP, GLASBERG BR. The temporal course of masking and the auditory filter shape. *J. Acoust. Soc. Am.* 81:1873–1880, 1987.
- MOORE BCJ, GLASBERG BR, BAER T. A model for the prediction of thresholds, loudness, and partial loudness. *J. Aud. Eng. Soc.* 45:224–240, 1997.
- MOORE BCJ, VICKERS DA, PLACK CJ, OXENHAM AJ. Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *J. Acoust. Soc. Am.* 106:2761–2778, 1999.
- NEFF DL. Confusion effects with sinusoidal and narrowband-noise forward maskers. *J. Acoust. Soc. Am.* 79:1519–1529, 1986.
- NELSON DA. High-level psychophysical tuning curves: Forward masking in normal hearing and hearing-impaired listeners. *J. Speech Hear. Res.* 34:1233–1249, 1991.
- O’LOUGHLIN BJ, MOORE BCJ. Off-frequency listening: Effects on psychoacoustical tuning curves obtained in simultaneous and forward masking. *J. Acoust. Soc. Am.* 69:1119–1125, 1981.
- OXENHAM AJ, PLACK CJ. A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. *J. Acoust. Soc. Am.* 101:3666–3675, 1997.
- OXENHAM AJ, PLACK CJ. Suppression and the upward spread of masking. *J. Acoust. Soc. Am.* 104:3500–3510, 1998.
- PATTERSON RD. Auditory filter shapes derived with noise stimuli. *J. Acoust. Soc. Am.* 59:640–654, 1976.
- PATTERSON RD, MOORE BCJ. Auditory filters and excitation patterns as representations of frequency resolution. In: Moore BCJ (Eds.) *Frequency Selectivity in Hearing*. Academic Press, London, pp 123–177, 1986.
- PATTERSON RD, NIMMO-SMITH I. Off-frequency listening and auditory filter asymmetry. *J. Acoust. Soc. Am.* 67:229–245, 1980.
- PATTERSON RD, NIMMO-SMITH I, WEBER DL, MILROY R. The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *J. Acoust. Soc. Am.* 72:1788–1803, 1982.
- PLACK CJ, OXENHAM AJ. Basilar-membrane nonlinearity and the growth of forward masking. *J. Acoust. Soc. Am.* 103:1598–1608, 1998.
- PLACK CJ, OXENHAM AJ. Basilar-membrane nonlinearity estimated by pulsation threshold. *J. Acoust. Soc. Am.* 107:501–507, 2000.
- PRIJS VF. Lower boundaries of two-tone suppression regions in the guinea pig. *Hear. Res.* 42:73–81, 1989.
- PURIA S, PEAKE WT, ROSOWSKI JJ. Sound-pressure measurements in the cochlear vestibule of human-cadaver ears. *J. Acoust. Soc. Am.* 101:2754–2770, 1997.
- READ HL, WINER JA, SCHREINER CE. Functional architecture of auditory cortex. *Curr. Opin. Neurobiol.* 12:433–440, 2002.
- RHODE WS, COOPER NP. Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo. *Aud. Neurosci.* 3:101–121, 1996.
- ROSEN S, BAKER RJ. Characterising auditory filter nonlinearity. *Hear. Res.* 73:231–243, 1994.
- ROSEN S, BAKER RJ, DARLING A. Auditory filter nonlinearity at 2 kHz in normal hearing listeners. *J. Acoust. Soc. Am.* 103:2539–2550, 1998.
- RUGGERO MA, RICH NC, RECIO A, NARAYAN SS, ROBLES L. Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J. Acoust. Soc. Am.* 101:2151–2163, 1997.
- SACHS MB, KIANG NYS. Two-tone inhibition in auditory nerve fibers. *J. Acoust. Soc. Am.* 43:1120–1128, 1968.
- SHAILER MJ, MOORE BCJ, GLASBERG BR, WATSON N, HARRIS S. Auditory filter shapes at 8 and 10 kHz. *J. Acoust. Soc. Am.* 88:141–148, 1990.
- SHERA CA, GUINAN JJ. Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. *J. Acoust. Soc. Am.* 113:2762–2772, 2003.
- SHERA CA, GUINAN JJ, OXENHAM AJ. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc. Nat. Acad. Sci. USA.* 99:3318–3323, 2002.
- VOGTEN LLM. Low-level pure-tone masking: A comparison of “tuning curves” obtained with simultaneous and forward masking. *J. Acoust. Soc. Am.* 63:1520–1527, 1978.
- WEGEL RL, LANE CE. The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear. *Phys. Rev.* 23:266–285, 1924.
- ZWICKER E. Ein Verfahren zur Berechnung der Lautstärke. *Acustica* 10:304–308, 1960.
- ZWICKER E, FLOTTORP G, STEVENS SS. Critical bandwidth in loudness summation. *J. Acoust. Soc. Am.* 29:548–557, 1957.