JARO
Journal of the Association for Research in Otolaryngology

Isoresponse Versus Isoinput Estimates of Cochlear Filter Tuning

ALMUDENA EUSTAQUIO-MARTÍN[1] AND ENRIQUE A. LOPEZ-POVEDA[1]

[1]Unidad de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Calle Pintor Fernando Gallego 1, 37007, Salamanca, Spain

ABSTRACT

The tuning of a linear filter may be inferred from the filter's isoresponse (e.g., tuning curves) or isoinput (e.g., isolevel curves) characteristics. This paper provides a theoretical demonstration that for nonlinear filters with compressive response characteristics like those of the basilar membrane, isoresponse measures can suggest strikingly sharper tuning than isoinput measures. The practical significance of this phenomenon is demonstrated by inferring the 3-dB-down bandwidths (BW_{3dB}) of human auditory filters at 500 and 4,000 Hz from behavioral isoresponse and isoinput measures obtained with sinusoidal and notched noise forward maskers. Inferred cochlear responses were compressive for the two types of maskers. Consistent with expectations, low-level BW_{3dB} estimates obtained from isoresponse conditions were considerably narrower than those obtained from isolevel conditions: 69 vs. 174 Hz, respectively, at 500 Hz, and 280 vs. 464 Hz, respectively, at 4,000 Hz. Furthermore, isoresponse BW_{3dB} decreased with increasing level while corresponding isolevel estimates remained approximately constant at 500 Hz or increased slightly at 4 kHz. It is suggested that comparisons between isoresponse supra-threshold human tuning and threshold animal neural tuning should be made with caution.

Keywords: auditory masking, cochlear nonlinearity, human, auditory filter, frequency selectivity

Correspondence to: Enrique A. Lopez-Poveda · Unidad de Audición Computacional y Psicoacústica, Instituto de Neurociencias de Castilla y León · Universidad de Salamanca · Calle Pintor Fernando Gallego 1, 37007, Salamanca, Spain. Telephone: +34-923-294500; fax: +34-923-294750; email: ealopezpoveda@usal.es

INTRODUCTION

The mammalian cochlea acts like an auditory prism, separating the frequency components of incoming sounds so that they stimulate different populations of auditory cells. Each cochlear region is thus described as acting as a filter tuned to a particular sound frequency, known as the characteristic frequency (CF), with a certain bandwidth (BW) and quality factor (Q=CF/BW). A related topographical organization of sound frequency is maintained throughout the entire auditory system, from cochlea to cortex. It is widely believed that our ability to perceptually separate out the multiple frequency components of a complex sound directly reflects cochlear filtering (Moore 2007). This study investigates the differences in characterizing the shape and tuning of cochlear filters using isoinput versus isoresponse conditions.

The tuning of any filter may be described based on its response amplitude to fixed-level input sinusoids of varying frequency. The resulting curve is known as the filter shape, frequency response, or *isolevel* curve of the filter. An alternative description consists of representing the input level required for sinusoids of varying frequency to maintain a fixed output level. The resulting curve is referred to as an *isoresponse* or tuning curve.

Isolevel and tuning curves are both used, sometimes interchangeably, to describe the tuning characteristics of the basilar membrane (BM), cochlear hair cells, peripheral, and central auditory neurons (Pickles 2008, p. 43) or human auditory filters (Moore 2007, p. 70–78). For linear systems, isolevel and tuning curves are merely the inverse of each other. Therefore, it is justified to treat them equivalently in assessing tuning. The response characteristics of the cochlea and its innervating auditory neurons are, however, nonlinearly related with sound level. This fact is well known and acknowledged (e.g., Robles and

Ruggero 2001). The implications of this nonlinearity for the relationship between isolevel and isoresponse curves and for our inferences about the tuning of cochlear 'filters' are, by contrast, less well recognized.

In physiology, the level-dependent tuning of mammalian BM responses is typically assessed directly from isolevel curves (Robles and Ruggero 2001) while that of auditory neurons is typically assessed from isoresponse (tuning) curves, possibly due to the difficulty in measuring isolevel curves due to the limited dynamic range of these neurons. It is rather surprising that some aspects of the relationship between tuning estimates obtained with either method have remained unnoticed until recently. For example, auditory nerve fibers with low spontaneous rates (LSR) have threshold tuning curves whose tips are more sharply tuned than those of high spontaneous rate (HSR) fibers. It has been shown only recently that this due to LSR fibers having higher thresholds combined with "a previously unnoticed nonmonotonic dependence on iso-velocity criterion of the frequency tuning of BM vibrations" (Temchin et al. 2008).

The tuning of human cochlear responses may not be measured directly, and so indirect approximate estimates have been obtained, often using behavioral masking methods (e.g., Moore 2007). Auditory masking is a nonlinear process, hence estimating tuning by fixing the masker level and varying the probe level (akin to an isoinput condition) is not equivalent to do it by fixing the probe level and varying masker level (akin to an isoresponse condition; e.g., Verschuure 1980, 1981; Vogten 1978). Indeed, there has been some controversy about which of the two approaches best determines the shape and tuning of auditory filters (e.g., Baker et al. 1998; Baker and Rosen 2006; Glasberg and Moore 2000; Rosen and Baker 1994; Rosen et al. 1998). The issue has always been investigated using simultaneous notched noise (NN) maskers. Results suggest that the power spectrum model of masking accounts best for the data when it is assumed that the filter shape is determined by the probe level (Glasberg and Moore 2000; Rosen and Baker 1994; Rosen et al. 1998). Simultaneous masking, however, appears inadequate to address this issue because it causes masker–probe interactions (e.g., distortion, suppression, or beating) that confound the interpretation of the results (Oxenham and Shera 2003; Heinz et al. 2002). Furthermore, an explicit model that takes into account the nonlinearity of the BM to predict tuning estimates obtained with either approach (i.e., fixed masker level vs. fixed probe level) is yet to be provided. Specifically, the fact that the level of the varying stimulus (whether it is the probe or the masker) may depend not only on the degree of frequency selectivity to the masker but also on BM compression has not been considered.

The present study demonstrates that for nonlinear filters with compressive BM-like response characteristics, isoresponse measures can suggest strikingly sharper tuning than isoinput measures. The first part of the manuscript provides a theoretical demonstration of this phenomenon based on a simple, idealized model of BM responses. The second part of the paper demonstrates the practical significance of this phenomenon by experimentally inferring the 3-dB-down bandwidths (BW_{3dB}) of human auditory filters at 500 and 4,000 Hz from behavioral isoresponse and isolevel measures obtained with sinusoidal and NN *forward* maskers. Tuning is first inferred using isoresponse measures for several different time intervals between the forward masker and a fixed-level probe. Separate measurements of temporal masking curves then allow the isoresponse tuning measurements to be transformed into isolevel tuning. It is shown that auditory filters are compressive (nonlinear) over the measured range of levels for the two types of maskers. Altogether, the results show that isoresponse tuning is narrower than isolevel tuning and that these differences, predicted by theory, arise due to the compressive nonlinearity of the auditory periphery.

NONLINEAR EFFECTS ON TUNING: THEORETICAL CONSIDERATIONS

Consider the idealized isolevel curves of a linear system illustrated in Figure 1A (adapted from Fig. 3.12 of Pickles 2008). The curves are asymmetric, with a steeper slope on the high-frequency side of the peak response, resembling the asymmetric frequency response of basal cochlear sites. The BW has been set arbitrarily and remains constant across levels. The abscissa represents stimulus frequency relative to the CF. Different curves are for different arbitrary input levels. As would be expected for a linear system, increasing the input level produces a comparable increase in the output level across frequencies, hence the vertical shift of the curves. Figure 1B illustrates the same data but plotted as input/output (I/O) functions to illustrate the linear behavior more clearly. Note that the slope of all I/O curves is 1 dB/dB.

Figure 1C illustrates corresponding tuning curves for different response criteria. Clearly, these tuning curves are vertically flipped versions of the isolevel curves. Therefore, the BW or Q for this linear filter could be equivalently inferred from the isolevel or the tuning curves and the estimates obtained from both curves would be identical and level independent.

This is far from true for nonlinear filters. Figure 1D–F illustrates idealized response characteristics for a nonlinear filter with frequency-asymmetric compression. Note how an increase of 10 dB in the
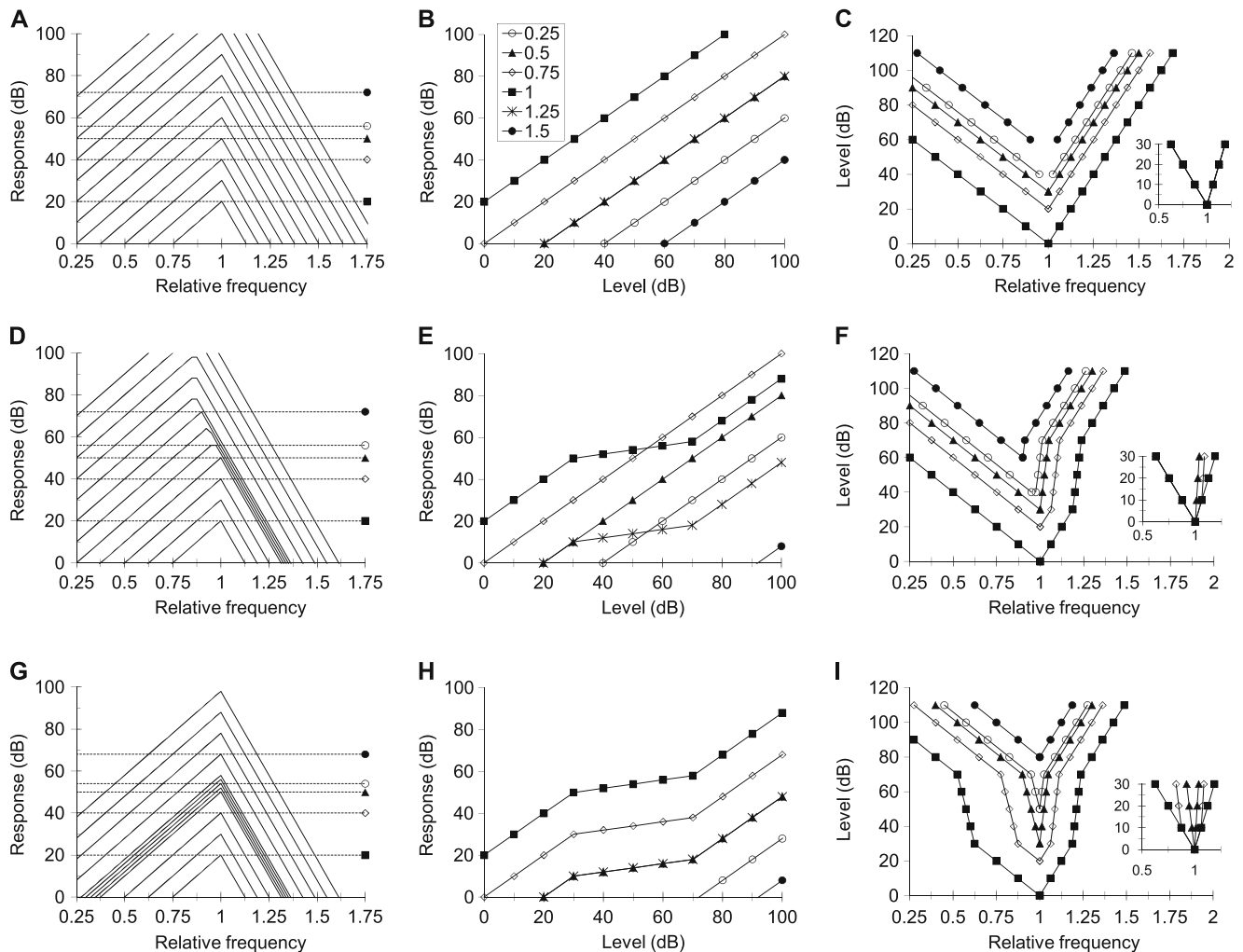
FIG. 1. Idealized response characteristics of linear and nonlinear filters. A Isolevel responses for a linear filter with level independent bandwidth. Each line is for a different input level from 0 to 110 dB in 10-dB steps. The abscissa illustrates frequency relative to CF. *Horizontal lines* and *symbols* illustrate the response criteria used to infer isoresponse curves on the right column. B I/O curves for the filter in A and for various stimulus frequencies, as indicated by the inset in units relative to CF. Note they all have slopes of 1 dB/dB and hence a linear growth. C Isoresponse (tuning) curves calculated from the isolevel curves shown in A for the response criteria depicted by the *horizontal dashed lines* and *symbols* in A. The inset illustrates the sharpness of tuning of the isoresponse curves for the three lower response criteria. They are normalized so that their tips are equal to zero. Note that the three curves overlap for a linear filter. D–E As for A–C but for a nonlinear system, with a frequency-asymmetric and compressive nonlinearity. Maximum compression is arbitrarily set to 5:1 (slope of 0.2 dB/dB). G–I As for A–C for a system with a broadband, 5:1 compressive nonlinearity (panels A and C are adapted from Fig. 3.12 of Pickles 2008).

input level produces a comparable increase in the response level for frequencies well below the CF only. For frequencies at or above the CF, the response is linear for low and high levels only. For intermediate levels, the response grows compressively with increasing input level; i.e., the corresponding I/O curve shows a segment with a slope of <1 dB/dB (Fig. 1E). This form of nonlinearity is broadly characteristic of basal cochlear sites (Ruggero et al. 1997; Robles and Ruggero 2001).

Figure 1F illustrates tuning curves for this type of nonlinear filter and for the response criteria depicted by the horizontal lines and symbols in Figure 1D. The tip of the tuning curve for the lowest response criterion (filled squares) resembles closely a vertically flipped version of an isolevel response curve for low-level input. For frequencies above CF and levels between 30 and 70 dB; however, the tuning curve has a steeper slope than the corresponding isolevel curves. This is because the increase in level required for maintaining a fixed output response is much greater for frequencies where compression occurs than for frequencies evoking linear responses.

The effect of compression on tuning is even more evident for higher response criteria. Consider, for example, the tuning curve for a response criterion of ~50 dB (filled triangles). This curve appears much more narrowly tuned than the curve for a lower response criterion. Indeed, the inset in Figure 1F shows that the sharper tuning is associated with a

steeper high-frequency slope. The reason is that compressive responses occur for frequencies immediately above the tip of the tuning curve.

Consider now Figure 1G–I. They illustrate the response characteristics for a nonlinear filter with broadband compression. That is, in this case, compressive responses extend to frequencies well below and above the CF (Fig. 1H). This is meant to qualitatively represent the response characteristics of apical cochlear sites in idealized form (Rhode and Cooper 1996; Lopez-Poveda et al. 2003).

The associated tuning curves (Fig. 1I) are striking. The tip of the curve for the lowest response criterion (filled squares) is a vertically flipped version of the tip of the low-level isolevel curves. For frequencies where compressive responses occur, however, both the high- and low-frequency slopes of this tuning curve are much steeper than the slopes of the corresponding isolevel curves. For frequencies well below and above the CF, the slope of the tuning curve becomes shallower and equal to the slope of the isolevel curves. This is because responses are linear over the range of levels required to meet the response criterion (Fig. 1G).

Consider now the tuning curve for a response criterion of ~50 dB (filled triangles in Fig. 1I). This curve is strikingly sharper than suggested by the isolevel curves (see the inset in Fig. 1I). The reason is that compressive responses occur at CF (thus at the tip of this curve) and over a range of frequencies around CF. Therefore, larger increases in input level are required to meet the response criterion.

The nonlinear phenomena described in Figure 1 are for a particular type of compressive filter where frequency-independent compression occurs *before* filtering; that is, for a system where compression occurs above a certain filter input level. For systems where frequency-independent compression takes place *after* filtering (i.e., for systems where compression occurs above a certain filter response level), isolevel and isoresponse curves would suggest identical estimates of tuning (results not shown). It is unlikely that actual BM responses conform exactly to either of these two idealized cases (i.e., compression before or after filtering). Indeed, actual BM responses have been accurately modeled by including a frequency-independent compressive gain between two filters (e.g., Meddis et al. 2001). In any case, there is strong experimental evidence that actual BM and auditory nerve (isoresponse) tuning curves sharpen with increasing response criterion, at least for basal cochlear sites and units and over a range of levels (see Figs. 8–9 of Temchin et al. 2008), even though corresponding isolevel responses broaden with increasing level. There is also evidence that psycho-physical tuning curves (PTCs) sometimes sharpen with increasing level (e.g., Lopez-Poveda et al. 2007). Furthermore, it will be shown below that the general principle put forward by Figure 1 (i.e., that for some nonlinear filters, isoresponse curves may suggest sharper tuning than isolevel curves) also applies to behavioral estimates of cochlear tuning.

In summary, insofar as the idealized plots of Figure 1 are qualitatively representative of BM responses, they show that isoresponse (tuning) and isolevel curves may *not* be used equivalently to estimate cochlear tuning, particularly over a level range where responses are compressive. They also show that isoresponse curves may suggest significantly sharper tuning than their corresponding isolevel responses. They further show that isoresponse tuning may be strongly dependent on level even when corresponding isolevel tuning is constant across level. In other words, isoresponse tuning is much more sensitive than isolevel tuning to small changes in level or, correspondingly, to response criterion. Only when responses are perfectly linear may isolevel and isoresponse curves be used equivalently to estimate the degree of tuning.

Implications for the interpretation of behavioral estimates of human cochlear tuning

Behavioral methods to estimate human cochlear tuning have evolved over the years as various confounding factors and artifacts have been addressed. The PTC (Chistovich 1957; Houtgast 1973; Kidd and Feth 1981; Vogten 1978) provides a paradigm that is, in principle, most similar to physiological cochlear tuning curves. It involves measuring the level of sinusoidal maskers of various frequencies required to just mask a fixed-level sinusoidal probe. PTCs can be measured using simultaneous or forward maskers, but the latter are now preferred because they eliminate confounding factors such as distortion, suppression, or the perception of beats that may occur in simultaneous masking. A version of this technique, known as the temporal masking curve (TMC) method, has become a favored method for estimating human cochlear nonlinearity (Lopez-Poveda et al. 2003; Nelson et al. 2001; Plack and Drga 2003) and may also be used to estimate frequency selectivity (e.g., Lopez-Poveda et al. 2007; Stainsby and Moore 2006; Yasin and Plack 2003, 2005). This method and its assumptions are summarized in Appendix A. Obtaining PTCs or TMCs involves measuring the level of sinusoidal forward maskers required to just mask a fixed, low-level probe tone. Insofar as the probe level is fixed and forward masking is assumed to occur central to cochlear filtering (Harris and Dallos 1979; Meddis and O'Mard 2006), the resulting curves may be reasonably assumed to represent cochlear *isores-*

ponse functions as opposed to isolevel conditions (Nelson et al. 2001; Lopez-Poveda et al. 2007).

The NN technique is an alternative common method for behaviorally estimating frequency selectivity (Patterson 1976; Patterson and Nimmo-Smith 1980; Moore 1987; Rosen et al. 1998; Glasberg and Moore 2000). It consists of measuring the masked threshold of a sinusoidal probe in the presence of a noise with a spectral notch as a function of the notch width relative to the probe frequency. The NN method may also be applied in simultaneous and forward masking, although the latter is now preferred because it avoids undesired masker–probe interactions (Moore and Glasberg 1981). The NN method is commonly thought to be advantageous over PTCs because it minimizes off-frequency listening (Johnson-Davis and Patterson 1979; O'Loughlin and Moore 1981).

Oxenham and Shera (2003) estimated the tuning of human auditory filters between 1 and 8 kHz with a version of the NN method designed to provide a closer comparison to animal neural tuning. The method involved measuring the level of a NN forward masker required to just mask a fixed, low-level sinusoidal probe. Therefore, like PTCs (or TMCs), the resulting NN functions may also be interpreted as isoresponse conditions. In other words, PTCs and NN functions may be interpreted using the same underlying assumptions (Appendix A) except that PTCs might be more affected by off-frequency listening. Oxenham and Shera (2003) reported two surprising results: (1) filters are much more sharply tuned than previously inferred using different versions of the NN method (Glasberg and Moore 1990); and (2) the Q value of the filters increases with increasing frequency, in good agreement with all animal estimates (Shera et al. 2002), but in contrast with previous human behavioral studies that suggested approximately constant Q across frequencies above about 1 kHz (Glasberg and Moore 1990). These results are still controversial (e.g., Ruggero and Temchin 2005; Shera et al. 2010) and, to our knowledge, have not been confirmed independently.

The nonlinear effects described in Figure 1 apply not only to sinusoidal stimuli, but also to broadband stimuli, including notched noises. Indeed, for a given increase in notch width, the level of a NN masker at threshold would increase more when the masker is compressed than when it is processed linearly by the BM. Both PTCs and the NN functions of Oxenham and Shera (2003) represent isoresponse conditions. If cochlear filters turned out to be compressive over the measured range of masker levels, then direct PTCs and NN functions would suggest sharper tuning than their corresponding isolevel curves. The present study investigated this possibility.

METHODS

Assumptions

In the TMC method (Appendix A), a fixed, low level, sinusoidal probe is used in an attempt to make the cochlear region excited by the probe as narrow as possible and fixed across conditions. It is assumed that sinusoidal maskers of identical levels and different frequencies produce different degrees of excitation at the BM region excited by the probe. It is further assumed that the time course of recovery from the BM excitation caused by the masker is independent of masker frequency. Criterion masking is achieved when the masker produces an excitation in the probe region that recovers over the masker–probe time gap to some fixed value. Since the time course of recovery is independent of masker frequency, maskers that produce equal excitation in the probe frequency region will produce equal masking at any given masker–probe time gap. Conversely, for fixed-level maskers, different masker–probe time gaps are required at the threshold of detection of a fixed-level probe. In other words, the 'gap threshold' is determined by the time it takes to recover at the BM probe place from the excitation produced by a sinusoidal masker (see Moore and Glasberg 2003).

Here, it is assumed that the same applies to NN maskers. That is, due to cochlear frequency selectivity, the excitation produced by fixed-level NN maskers at the probe place will decrease with increasing notch width. Therefore, for a fixed time gap, higher levels are required for NN maskers with wider notch widths to produce the same degree of recovery at the time when the fixed-level probe occurs. Conversely, for fixed-level NN maskers, longer gaps will be required at the detection threshold of the fixed-level probe as notch width is decreased. In other words, the 'gap threshold' is directly related to the excitation produced by fixed-level NN maskers at the cochlear region excited by the probe.

On the other hand, 'gap thresholds' for sinusoidal and NN maskers will also be inherently affected by the post-cochlear rate of recovery from forward masking. Here, it is further assumed that the latter is identical for sinusoidal and NN maskers and is well described by the linear reference TMC (Appendix A). Therefore, the post-cochlear recovery rate may be accounted for by transforming gap thresholds into BM output levels (in arbitrary decibel units) using the linear reference TMC.

In summary, for both sinusoidal and NN maskers, it is assumed that isogap curves correspond with isoresponse conditions while isolevel curves (i.e., transformed 'gap-threshold' curves) correspond with isoinput conditions.

Listeners

Three normal-hearing listeners participated in the study. Their ages were 38 (S1), 24 (S2), and 40 (S3) years. Their absolute hearing thresholds for pure tones of the two test frequencies considered (0.5 and 4 kHz) and of various durations (10 and 300 ms) are given in Table 1. They were experienced in psychoacoustical tasks. Listener S1 was author AEM.

Stimuli

Masker levels at the masked detection threshold of a fixed-level sinusoidal probe were measured in forward masking. The probe level was always fixed at 10 dB above the listener's absolute threshold for the probe (Table 1). The durations of the probe and the masker were 10 and 400 ms, respectively, including 5-ms cosine-squared onset and offset ramps (i.e., probes had no steady state). Masker-probe time gaps were defined as the silent period between the masker offset and the probe onset. Five gap values were considered: 2, 10, 30, 50, and 70 ms.

PTCs were obtained using sinusoidal maskers with various frequencies around the probe frequency (f_P): 0.5, 0.7, 0.8, 0.9, 0.95, 1, 1.05, 1.1, 1.2, 1.3, and $1.6f_P$. In the NN conditions, the masker consisted of two bands of Gaussian noise centered below and above f_P, each with a bandwidth of $0.25f_P$. Both bands were generated in the spectral domain (with quasi-infinite skirts) and their corresponding waveforms were obtained by an inverse fast Fourier transform. The independent variable was the relative notch width ($g = |f - f_P|/f_P$). This was defined as the spectral distance between f_P and the closer edge of the noise spectrum, normalized to f_P. Six relative notch widths were used for each f_P: 0, 0.05, 0.1, 0.2, 0.3, and 0.4. The stimuli and conditions were identical to those used by Oxenham and Shera (2003), except that they used only a masker–probe gap of 5 ms and that an additional notch width of 0.05 was used here.

The time course of recovery from forward masking for the three participants was inferred by

measuring a linear reference TMC for a 1.6-kHz masker and a 4-kHz probe (Lopez-Poveda and Alves-Pinto 2008). The probe level was fixed at 10 dB SL. The masker–probe gaps ranged from 10 to 100 ms in 10-ms steps with an additional initial interval of 2 ms.

Procedure and equipment

A two-interval, two-alternative, forced-choice adaptive paradigm with feedback was employed to estimate the masker level at threshold. Two sound intervals were presented to the listener in each trial. One of the intervals contained the masker only and the other contained the masker followed by the probe. The presentation order of the two intervals was random and the listener was instructed to indicate the interval that contained the probe. The time period between the two intervals was 500 ms. The initial masker level was set so that listeners could easily detect both the masker and the probe. The masker level was then changed according to a two-up, one-down adaptive procedure to estimate the masker level that produced 70.7% correct responses (Levitt 1971). The masker level was altered in 6-dB steps until two reversals occurred and in 2-dB steps thereafter. A total of 12 reversals were obtained for each run. Threshold was calculated as the mean masker level of the last 10 reversals in a run. At least three threshold estimates were obtained per condition and their mean was taken as the masker level at threshold. If their standard deviation exceeded 6 dB, a fourth estimate was obtained and included in the mean.

Stimuli were generated with a personal computer and an RME Fireface 400 soundcard at a sampling rate of 44.1 kHz and 24-bit resolution. All stimuli were played monaurally to the listener via the soundcard headphone connection through the same pair of Sennheiser HD-580 circumaural headphones. Listeners sat in a double-walled sound-attenuating booth during the measurements. Calibration was performed at 1 kHz and the obtained sensitivity was used at all other frequencies. Note that Oxenham and Shera (2003) used Etymotic Research ER2 insert ear phones designed to give an approximately flat response at the eardrum.

Experimental procedures were approved by the Human Experimentation Ethics Committee of the University of Salamanca (Spain).

Deriving auditory filter bandwidths from isogap NN functions

Auditory filter bandwidth estimates were obtained from isogap NN functions with the power spectrum

TABLE 1

Listeners' absolute hearing thresholds (in dB SPL) for pure tones of different frequencies and durations

Frequency (Hz)	500	500	4000	4000
Duration (ms)	10	300	10	300
S1	24.8	13.1	19.6	16.3
S2	18.4	4.8	12.2	1.8
S3	17.7	4.8	16.9	9.6
Mean	20.3	7.6	16.2	9.2

model. The basic filter shape assumed was a symmetric double rounded exponential, roex(p,w,t):

$$W(g) = (1 - w)(1 + pg) \exp(-pg) + w(1 + pg/t)$$
$$\times \exp(-pg/t), \tag{1}$$

where g denotes frequency relative to f_p, and p, t, and w are free parameters ($0 \le w \le 1$). Optimal model parameters were sought considering random initial values. BW_{3dB} estimates were obtained from the resulting filter shapes. Methods were identical to those employed by Oxenham and Shera (2003) and will not be reproduced here. The only difference was the combined outer/middle-ear filter, which was configured here for the headphones employed during data collection and for the human middle-ear response of Goode et al. (1994, Fig. 1, set: 104 dB SPL).

[The outer ear (headphone-to-eardrum) frequency response was measured by placing the headphones on a KEMAR equipped with a Zwislocki DB-100 coupler connected to a Brüel & Kjaer sound level meter (mod. 2188). Three amplitude frequency response functions were measured using pure tones for each one of four different pinnae. For each measurement, headphones were removed/replaced from the KEMAR. The mean of the 12 measurements was used to estimate the headphone-to-eardrum frequency response.]

Estimating auditory filter bandwidths from PTCs

Each PTC was regarded as having two sides around its tip. Each side was fitted with a roex(p, w, t) function (Eq. 1) and BW_{3dB} estimates were obtained directly from the fits (e.g., Lopez-Poveda et al. 2007; Yasin and Plack 2005). [Note that roex(p, w, t) functions were used in two different ways: as the presumed shape of auditory filters (when inferring filter shapes using the NN method) and also as convenient functions to fit raw PTCs and NN functions.]

Inference of cochlear I/O curves for sinusoidal and NN stimuli

Cochlear I/O curves for sinusoids and NN stimuli were inferred from the PTCs and NN functions by plotting the levels of the linear reference TMC against the levels of any sinusoidal or NN masker paired according to masker–probe time gap (Nelson et al. 2001).

Inference of filter isolevel responses from PTCs

Isolevel curves were obtained by choosing a fixed masker level and finding the points (gap, frequency) in the roex functions fitted to the PTCs at which the chosen level was achieved. As explained above (see "Methods" section), the resulting gap values were assumed to reflect the combined BM response to the fixed-level masker *and* the different degree of (post-cochlear) masking recovery at each time gap. To account for the latter, gap values were transformed into output levels with the linear reference TMC. Isolevel curves were obtained for masker levels from 45 to 85 in 5-dB steps.

Inference of filter isolevel responses from NN functions

Gap thresholds for fixed-level NN maskers of various notch widths were obtained by interpolation of the measured isogap NN functions to intermediate notch widths and time gaps. For convenience, roex(p,w,t) functions were used for these interpolations. Isolevel NN functions were obtained by transforming the interpolated gap thresholds into output levels with the linear reference TMC.

Auditory filter isolevel curves were then obtained from isolevel NN functions based on two assumptions: (1) auditory filter shape can be described by a symmetrical roex(p,w,t); and (2) the output excitation to the fixed-level NN maskers was proportional to the frequency integral (the area) of the product of masker spectrum times the presumed isolevel frequency response of the filter. Therefore, the filter's isolevel response was estimated by adjusting the filter's parameters to minimize the root-mean-square (RMS) difference between the predicted and the measured excitation functions. This carried the implicit assumption that the filter operates linearly for each NN stimuli; that is, for instance, that the two sidebands of the noise did not suppress each other.

RESULTS

Sinusoidal maskers: PTCs

Figure 2 illustrates the individual and mean linear reference TMCs. According to the assumptions of the TMC method of inferring BM responses (Appendix A), these curves describe the time course of post-cochlear recovery from a forward masker that is processed linearly by the BM. Also shown in Figure 2 is a least-squares straight line fit to the mean linear reference. This fit was good ($R^2=0.98$, RMS error$=0.29$ dB) and so the fitted function was used instead of the actual observations throughout. [N.B.: The linear reference TMC of subject S2 did not look like a straight line. This, however, was the result of one data point only (for a 20-ms gap). When the individual linear reference in question was fitted with a straight line, the fit error for a
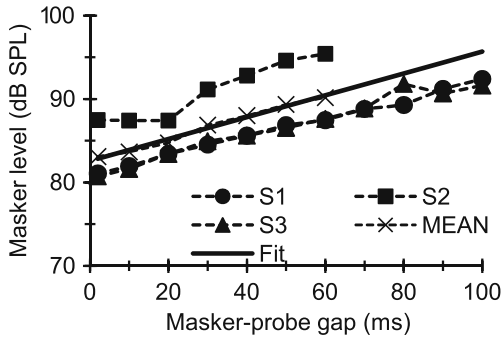
FIG. 2. Individual and mean linear reference TMCs. These were for probe and masker frequencies of 4 and 1.6 kHz, respectively, and for a probe level of 10 dB SL. Also shown is a *straight line* fit to the mean TMC with extrapolated values.

20-ms gap was small (2 dB only) and equal to the masker level step used in our adaptive procedure. Therefore, it was within the accuracy of our threshold estimates. Indeed, when the point in question was omitted, the linear reference of subject S2 looked like a straight line whose slope was close to the corresponding slope for other subjects or for the mean linear reference.]

Individual and mean PTCs are shown in Figure 3. Lines illustrate roex fits to the data. These fits were excellent (RMS errors for the MEAN curves were 2.49 and 1.21 dB for 500 Hz and 4 kHz, respectively). The pattern of results was qualitatively similar for the three subjects and so the analysis focuses on the MEAN responses. Corresponding I/O and isolevel curves for the MEAN responses are shown in Figure 4.

At 4 kHz, the MEAN PTCs (Fig. 3) appeared more narrowly tuned for a 30-ms gap than for shorter or longer gaps. This is more easily seen in the bottom panels of Figure 3, which show a zoomed-in view of the lower part of the MEAN PTCs, normalized to have their tips at 0 dB. At the same time, the corresponding I/O and isolevel curves (Fig. 4B and D, respectively) show compressive responses over the range of levels tested for frequencies at and above the probe frequency but primarily linear responses for frequencies below it. As a result, the Q_{3dB} inferred from the MEAN PTCs (shown as gray circles in Fig. 5) varied nonmonotonically with PTC-tip level, being greater at 35 dB SPL than at lower or higher levels.

Results at 500 Hz were different. First, MEAN PTCs (Fig. 3) became slightly narrower with increasing masker–probe gap. Second, I/O and isolevel curves
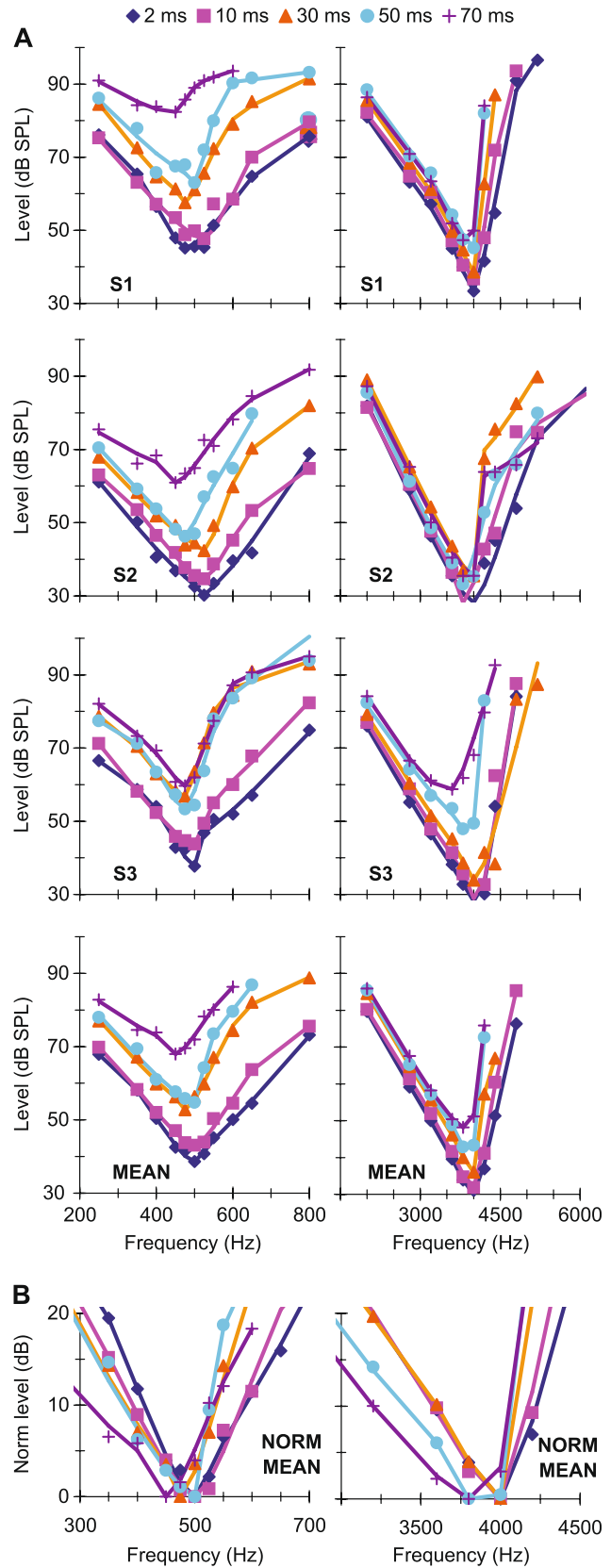
FIG. 3. A PTCs for probe frequencies of 500 (*left*) and 4,000 Hz ▶ (*right*). Each panel illustrates results for individual subjects or the mean, as indicated in the bottom-left corner of each panel. *Different symbols* illustrate results for different masker–probe gaps, as indicated by the *inset* (in ms). *Lines* illustrate roex fits to the experimental values. B Zoomed-in view of the lower part of MEAN PTCs normalized to have their tips at 0 dB.
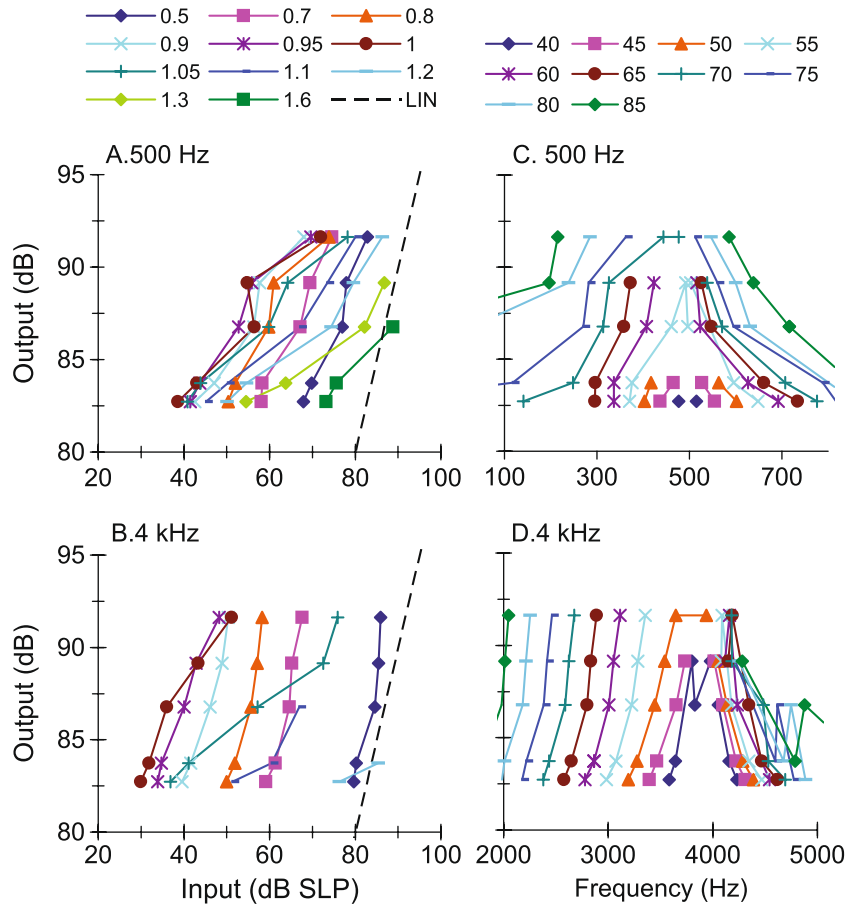
FIG. 4. A, B Inferred I/O curves for sinusoidal stimuli at 500 Hz and 4 kHz. Each *line* is for a different stimulus frequency, as indicated by the inset in units relative to the probe frequency. *Dashed lines* illustrate linear responses with zero gain. C, D Corresponding isolevel curves. Each *line* is for a different input level from 40 to 85 dB SPL in 5-dB steps, as indicated by the inset.

(Fig. 4A and C, respectively) showed compressive responses for all frequencies tested, except 250 Hz. As a result, Q_{3dB} inferred from the MEAN PTCs (gray squares in Fig. 5) increased slightly with increasing PTC tip level over the range of levels tested.
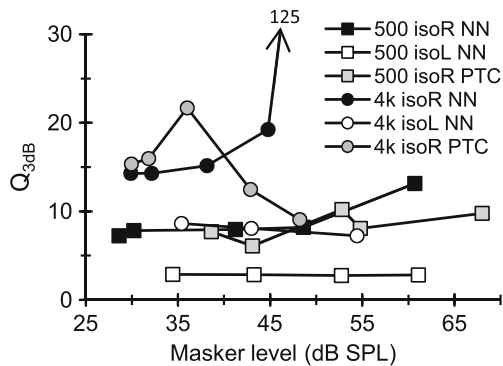

FIG. 5. Q_{3dB} as a function of level. *Gray symbols* represent estimates obtained from the mean PTCs as a function of PTC tip level. *Dark* and *open symbols* represent estimates obtained from isogap (isoresponse) and isolevel NN functions, respectively, as a function of the lowest masker level (i.e., the masker level for $g=0$). isoR and isoL refer to isoresponse and isolevel conditions, respectively. The *arrow* indicates that the last Q_{3dB} for the series was 125, hence well beyond the range denoted by the ordinate axis.

The overall pattern of results at 4 kHz was qualitatively consistent with the idealized responses of Figure 1D–F, except that experimental isolevel responses (Fig. 4) became broader at higher levels, something not seen in Figure 1D. The nonmonotonic level-dependency of Q_{3dB} likely reflects the combined broadening of the filter with compression for frequencies at and above CF. For intermediate levels and as a result of compression, a larger increase of input level was required for a given increase in frequency to achieve the required response criterion of each PTC. This explains the increase in Q_{3dB} with increasing level from ~30 to ~35 dB SPL (Fig. 5). At higher levels (~47 dB SPL), compression affects on-frequency responses only (Fig. 4B) and the filter becomes broader (Fig. 4D). This explains why PTCs were broadest for the longest gap and thus why Q_{3dB} was smallest at the highest tip level (Fig. 5).

The pattern of results at 500 Hz was qualitatively more similar to the idealized responses of Figure 1G–I insofar as they both corresponded to a filter having compressive responses below and above CF. The degree of compression in the experimental responses was, however, less than in the idealized responses of Figure 1. As a result of this broadband compression, PTCs almost certainly suggested narrower tuning

(higher Q_{3dB} values) than the corresponding isolevel functions (Fig. 4C).

That human BM compressive responses extend to a broader frequency range (relative to CF) at 500 Hz than at 4 kHz is not a new finding; it has been previously reported using the TMC and other behavioral methods (e.g., Lopez-Poveda et al. 2003; Lopez-Poveda and Alves-Pinto 2008; Place and Drga 2003). Also, it is consistent with physiological BM (Rhode and Cooper 1996) and auditory nerve (Temchin and Ruggero 2010) responses. Neither is it new that PTCs are sometimes narrower at supra-threshold than at threshold levels (e.g., Lopez-Poveda et al. 2007), a result supported by BM tuning curves (Ruggero et al. 1997) and auditory nerve tuning curves (Buser and Imbert 1992, pp. 163–166; Evans 1975). Those earlier studies, however, failed to provide a convincing explanation for the narrower tuning of some supra-threshold tuning curves. The present report provides an explanation based on the compressive characteristics of BM responses. A similar rationale has been proposed recently as an explanation for the relative sharper tuning of LSR versus HSR auditory nerve fibers (Temchin et al. 2008). These authors demonstrate that LSR fibers show sharper tuning curves because their threshold is higher and hence the tips of their tuning curves fall within the (high-side) compressive region of BM responses.

In the present data, compressive responses to some frequencies occurred over virtually the full range of levels that could be measured at 500 Hz and 4 kHz (Fig. 4A and B). This is despite our exercising special care to use a very low-level probe (10 dB SL) and long-duration forward maskers (400 ms) to obtain masker levels as low as possible, for which it is more likely to get linear BM responses (Plack and Skeels 2007). Considering that PTCs presumably reflect isoresponse conditions and in light of the analysis described in "Nonlinear effects on tuning: theoretical considerations" section, the present results suggest that the BWs of isolevel 500-Hz and 4-kHz filters are likely broader than those inferred from the present PTCs (Fig. 5). Furthermore, it is almost certain that isolevel curves would suggest a different dependency of Q_{3dB} with level from that illustrated by the gray symbols in Figure 5. This was indeed suggested from isolevel functions of Figure 4C–D. Unfortunately, these curves were incomplete and did not allow a confirmation of this conjecture.

Notched noise maskers

Figure 6 illustrates individual and mean NN functions (symbols) for the five different masker–probe time gaps (referred to as isogap NN functions). Lines illustrate model fits to each data series obtained using
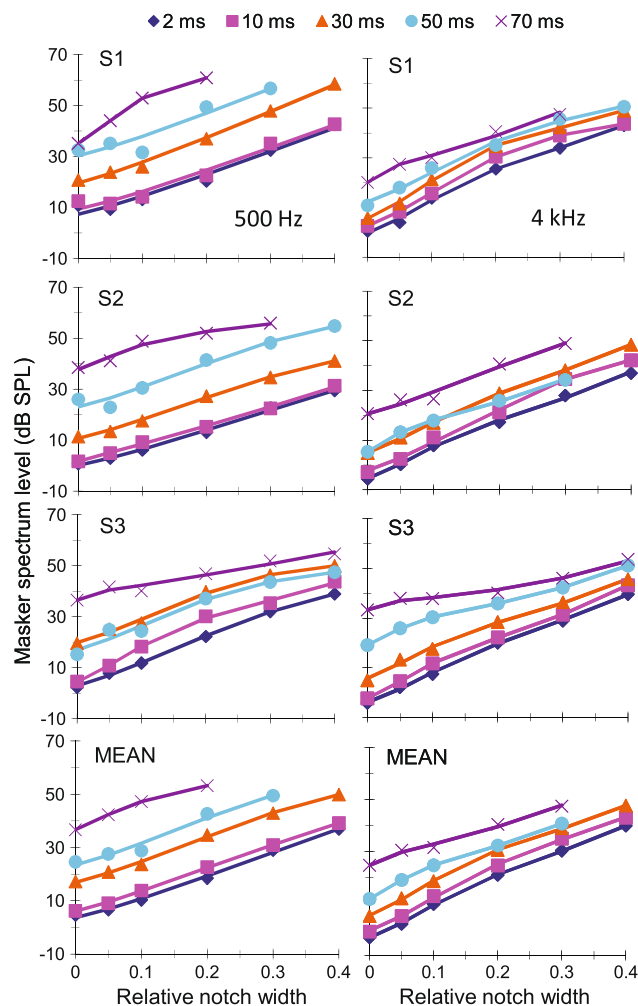

FIG. 6. Isogap NN functions for probe frequencies of 500 Hz (*left*) and 4 kHz (*right*). Each panel illustrates results for individual subjects or the mean, as indicated in the *top-left corner* of each panel. Different *symbols* illustrate results for different masker–probe gaps, as indicated by the *inset at the top* (in ms). *Lines* illustrate corresponding fits to the experimental data using the power spectrum model assuming symmetrical roex filter shapes.

the power spectrum model as applied by Oxenham and Shera (2003). The pattern of results was similar across listeners and so model parameters and RMS fit errors are shown in Table 2 for the MEAN data only. The RMS errors were small and so the fits were reasonable. Note that the value of k (the detector efficiency) decreased with increasing masker–probe gap to reflect that higher masker levels were required to mask the fixed-level probe.

Q_{3dB} values from model fits to the isogap functions of Figure 6 are illustrated as dark symbols in Figure 5. Values are plotted as a function of the lowest masker level of each isogap curve (i.e., the level for the narrowest notch width). Q_{3dB} values generally were larger for the 4-kHz than for the 500-Hz filter, which

Gap (ms)	p	t	w	k (dB)	RMS (dB)	BW_{3dB} (Hz)
			500 Hz			
2	24.52	n/a	0.000	8.84	1.55	69
10	26.58	1.58	0.030	6.76	0.56	64
30	26.83	3.74	0.001	−3.98	1.33	63
50	27.67	n/a	0.000	−13.15	3.37	61
70	14.42	0.31	0.948	−24.51	0.00	38
			4,000 Hz			
2	49.29	1.96	0.009	4.18	0.58	280
10	48.48	2.27	0.002	1.45	1.01	280
30	51.12	2.20	0.005	−4.28	0.61	264
50	69.37	2.43	0.057	−10.86	0.08	208
70	478.22	17.13	0.044	−17.57	1.78	32

t is not applicable (n/a) when $w=0$ (Eq. 1)

indicates that the 4-kHz filter is relatively more sharply tuned than the 500-Hz filter. Q_{3dB} for the lowest levels (~29 dB SPL) were 7.25 and 14.2 and at 500 Hz and 4 kHz, respectively.

Interestingly, the dark symbols of Figure 5 and the BW_{3dB} values in Table 2 both suggest that the two filters get narrower with increasing level; the slope of this dependence was steepest for the 4 kHz filter. This trend may seem surprising at first considering that physiological BM and auditory nerve isolevel responses remain approximately constant at 500 Hz or broaden with increasing level at 4 kHz (Cooper 2004; Robles and Ruggero 2001). There is, by contrast, theoretical (see "Nonlinear effects on tun-

ing: theoretical considerations" section), physiological (Evans 1975; Ruggero et al. 1997) and behavioral (Lopez-Poveda et al. 2007) evidence that isoresponse curves become narrower with increasing level when responses are compressive. Given that isogap NN functions may be regarded as isoresponse conditions, the trend in question would not be unreasonable if peripheral responses to the NN maskers were compressive over the range of masker levels considered.

Indeed, this seemed to be the case. I/O curves for the NN stimuli were inferred using the method of Nelson et al. (2001). The results are shown in Figure 7A, B. Although the degree of compression varied across conditions, responses were compres-
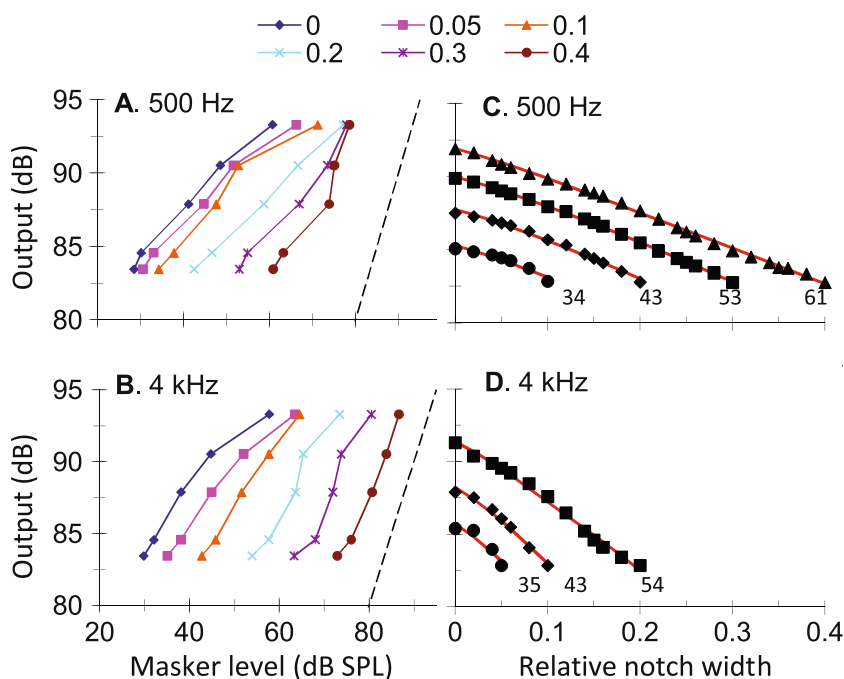

FIG. 7. A, B Inferred peripheral I/O curves for notched-noise stimuli with various relative notch widths, as indicated by the inset at the top. *Dashed lines* illustrate linear responses (slope=1 dB/dB). C, D Inferred isolevel curves as a function of relative notch width. Each curve is for a different masker overall level as indicated by the numbers next to each curve (in dB SPL). *Lines* illustrate model fits assuming symmetrical roex filter shapes.

sive (slope<1 dB/dB) for the two probe frequencies over the whole range of input levels and for all notch widths, except perhaps for a notch width of 0.4 at 4 kHz. Based on the theoretical analysis of "Nonlinear effects on tuning: theoretical considerations" section, this suggests that the Q_{3dB} estimates obtained from isogap NN functions almost certainly overestimate the tuning of corresponding isolevel curves.

To investigate this possibility, isolevel curves were inferred from MEAN isogap NN functions by data interpolation to intermediate gaps and notch widths. For convenience, roex(p,w,t) functions were used for these interpolations. This choice seemed reasonable due to the small RMS fit errors it yielded. Average fit errors were 0.78 and 1.86 dB for isogap and iso-notch width curves, respectively, at 4 kHz; average errors were 1.47 and 2.68 dB, respectively, at 500 Hz. To account for the inherent recovery from forward masking, interpolated gap values were transformed into output response levels using the linear reference TMC.

Resulting isolevel curves are illustrated by the symbols of Figure 7C, D. Recall that these curves are assumed to represent the response (in arbitrary logarithmic units) of a BM region (which is presumably fixed and narrow because the probe level is low and fixed) for fixed-level NN maskers with different notch widths. As a result of cochlear frequency selectivity, output levels unsurprisingly decrease with increasing notch width. For a given masker level, output level decreases faster with increasing notch width for the 4-kHz than for the 500-Hz filter. This suggests that the 4-kHz filter is relatively more narrowly tuned than the 500-Hz filter.

Lines in Figure 7C, D illustrate model fits assuming symmetrical roex(p,w,t) filters. Table 3 gives the model parameters. Note the small RMS errors. No tail component was necessary ($w=0$, t=n/a) to fit the data for some masker levels, which is not surprising considering that for the lower levels (34–35 dB SPL), the output varied over a narrow decibel range and

so could be modeled with the tip exponential component of the roex(p,w,t). In the power spectrum model, k represents the detector efficiency, or the threshold signal-to-masker ratio at the filter output. In the present context, by contrast, k reflected the ratio (in dB) between the BM output levels to fixed-level NN maskers, as determined by the linear reference TMC, and the output of a unit-gain roex filter. Its value increased with increasing masker level. According to the assumptions of the TMC method, this increase reflected the gain of the BM response which is not present in the unit-gain roex filter.

Table 3 shows BW_{3dB} estimates obtained from isolevel functions. These can be compared with those of Table 2, which were obtained directly from the raw isogap (isoresponse) NN functions using the power-spectrum model of masking as applied by Oxenham and Shera (2003). The differences are striking. First, estimates obtained from isolevel functions were much broader than those from isogap (isoresponse) NN functions. Second, in contrast with inferences made from isogap NN functions, which suggested increasing BW_{3dB} with increasing level at both frequencies, isolevel BW_{3dB} estimates hardly changed with level at 500 Hz but increased slightly with level at 4 kHz, at least over the range of levels considered (shown in Fig. 5). The latter trends are qualitatively more consistent with those of physiological isolevel BM responses at the two frequencies (e.g., Figs. 2.3 and 2.4 in Cooper 2004). Finally, Figure 5 also demonstrates that 4-kHz filters are comparatively more narrowly tuned than 500-Hz filters, regardless of which data (isolevel or isogap/isoresponse) are used to make the inferences.

Comparison of isolevel responses inferred from PTCs and NN functions

BW_{3dB} estimates inferred from isolevel curves (Table 3) for NN maskers could not be validated by

TABLE 3

Model parameters derived from the isolevel NN curves (Fig. 7C, D) assuming symmetrical roex(p,w,t) filters

Level (dB SPL)	p	t	w	k (dB)	RMS (dB)	BW_{3dB} (Hz)
			500 Hz			
34	9.68	n/a	0.000	−56.82	0.47	174
43	9.62	n/a	0.000	−50.27	0.49	175
53	9.26	n/a	0.000	−43.41	0.28	181
61	10.03	2.22	0.090	−36.40	0.62	177
			4,000 Hz			
35	28.97	n/a	0.000	−61.11	0.67	464
43	27.39	n/a	0.000	−56.27	0.51	496
54	24.61	n/a	0.000	−46.68	0.92	552

The first column informs of the masker overall levels. t is not applicable (n/a) when w=0 (see Eq. 1)
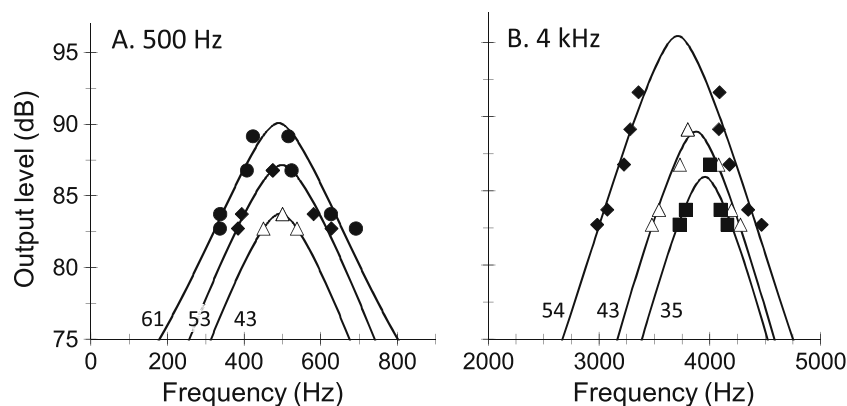
FIG. 8. Comparison of isolevel filter shapes inferred from PTCs (*symbols*) and NN functions (*lines*). A 500 Hz. B 4 kHz. Numbers next to each curve denote the input level in dB SPL. Output level is in arbitrary decibel units.

direct comparison with corresponding estimates from isolevel curves for sinusoidal maskers because the latter were 'incomplete' (Fig. 4C, D). In an attempt to corroborate these estimates, isolevel curves inferred with sinusoidal and NN maskers were compared directly, as shown in Figure 8. Symbols illustrate isolevel curves for sinusoidal stimuli (re-plotted from Fig. 4C and D) and lines illustrate symmetrical roex filters for the parameters of Table 3. To facilitate the comparison, roex filters (lines) were arbitrarily shifted vertically and horizontally to maximize the correspondence with the symbols. This shifting procedure seemed reasonable considering that roex filters were originally assumed to be centered at the probe frequency and to have unit gain at their center frequency, which need not be the case for cochlear filters or the isolevel functions of Figure 4C, D. The shifting parameters and the RMS differences between the two sets of isolevel curves are shown in Table 4. The match between the two sets of curves is remarkable, particularly for lower levels (Table 4), which supports the BW_{3dB} estimates inferred from isolevel curves for NN maskers (Table 3).

Off-frequency listening

The present data incidentally serves to estimate the significance of off-frequency listening at the two frequencies considered. Based on existing evidence, (e.g., O'Loughlin and Moore 1981) one would expect off-frequency listening to be more significant for sinusoidal than for NN maskers, resulting in BW_{3dB} estimates from PTCs that are narrower than those from NN functions. This was found to be the case at 500 Hz but not at 4 kHz, as shown in Figure 9. PTC-based BW_{3dB} were systematically narrower than NN-based estimates at 500 Hz, with average values of 49.6 and 66.7 Hz, respectively. This difference was significant ($p=0.005$, two-tailed, paired Student's t test). At 4 kHz, by contrast, differences occurred in opposite directions for individual listeners and mean values were not statistically different ($p=0.65$). This suggests that off-frequency listening was more important at 500 Hz than at 4 kHz. The reason for this difference is uncertain. Maybe energy splatter facilitated the detection of the brief (10 ms) 500-Hz probe more than that of the 4-kHz probe, and so masker levels at masked threshold were relatively higher at 500 Hz than at 4 kHz. There is evidence that off-frequency listening is negligible also at 1 kHz (Nelson et al. 2001).

On the other hand, off-frequency listening needs not be the only explanation for the difference between BW_{3dB} estimates for sinusoidal and NN maskers at 500 Hz. The difference might be due to measuring NN functions with symmetrical notches even though PTCs were not perfectly symmetrical (Fig. 3). Alternatively, it might be due to mutual suppression between the two side bands of the NN maskers, something that did not affect PTCs.

The present data also serves to minimize any concern about estimating the tuning of a 500-Hz filter from PTCs measured with a 10-ms probe. In general, it would seem inappropriate to characterize the tuning of a given filter from PTCs measured with a probe having a broader BW than the targeted filter. If

TABLE 4

Parameters to maximize the correspondence between isolevel curves inferred using sinusoidal and NN maskers

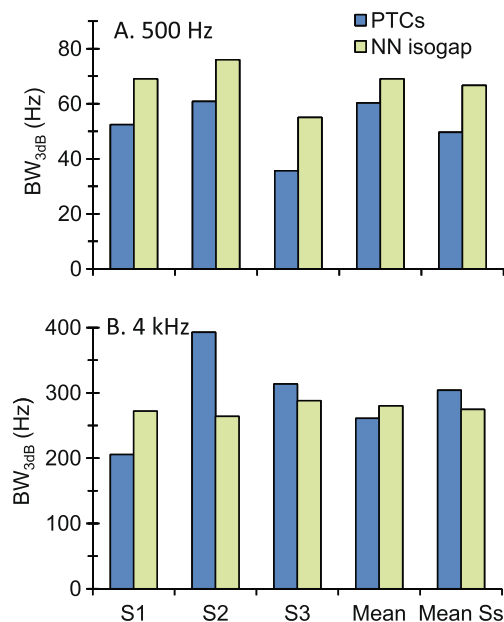| Level (dB)	CF (Hz)	Gain (dB)	RMS diff. (dB)
	500 Hz		
43	494.2	83.7	0.0
53	498.6	87.1	1.1
61	490.0	90.1	3.7
	4 kHz		
35	3,953.5	85.9	1.8
43	3,873.7	89.0	1.4
54	3,709.0	95.4	3.2

See text for details

FIG. 9. Individual and mean BW_{3dB} estimates at 500 Hz (A) and 4 kHz (B) estimated from PTCs and isogap NN functions for a masker–probe gap of 2 ms. *Mean* estimates from mean data curves. *Mean Ss* mean across individual BW_{3dB} estimates.

the probe was detected by averaging the response of a number of filters around the targeted filter, then masker threshold might stay similar for various masker frequencies. As a result, PTCs might appear broader than the underlying filters. This criticism would not apply to the same extent to tuning estimates obtained with NN maskers because the noise on both sides of the probe would maximize probe detection from the response of the filter at the center of the spectral notch (Patterson 1976).

The BW_{3dB} of the 10-ms probes used here was approximately 150 Hz. This probe duration would be adequate for estimating the tuning of 4-kHz cochlear filters from PTCs because there is strong physiological (e.g., Robles and Ruggero 2001) and behavioral (e.g., Glasberg and Moore 1990; Oxenham and Shera 2003) evidence that these filters are broader than 150 Hz. By contrast, one might think that the present 10-ms probe could have been too short for characterizing the tuning of the 500-Hz filters if they were as narrowly tuned as suggested by the trend in the data of Oxenham and Shera (2003), or the present isoresponse PTCs (Fig. 5) or NN functions (Table 2). The present results minimize this concern for two reasons. First, inconsistent with this idea, isoresponse BW_{3dB} estimates were actually narrower (not broader) for sinusoidal than for NN maskers at 500 Hz (Fig. 9). Second, isolevel NN maskers suggested low-level average BW_{3dB} estimates of ~174 Hz (Table 3), which are broader than the probe BW_{3dB} (~150 Hz). Therefore, a 10-ms probe is still appropriate for estimating the tuning of a 500-Hz filter from PTCs, at least for low input levels.

DISCUSSION

The present study had two aims: (1) to theoretically demonstrate that for compressive filters, isoresponse (tuning) curves may be strikingly sharper than isolevel curves over the range of input levels where compression occurs; (2) to experimentally demonstrate how this affects behavioral estimates of auditory filter tuning at 500 Hz and 4 kHz obtained with sinusoidal and NN forward maskers.

The main findings of the study were:

1. Isoresponse and isolevel curves do not give similar estimates of tuning for compressive filters. Over the range of input levels where compression occurs, isoresponse curves always suggest narrower tuning than isolevel curves (Fig. 1).
2. Behaviorally inferred BM I/O responses were compressive for sinusoidal and NN maskers at 500 Hz and 4 kHz (Figs. 4 and 7).
3. BW_{3dB} estimates inferred from average isoresponse (isogap) NN functions (Table 2) were much narrower than those obtained from corresponding isolevel NN functions (Table 3).
4. BW_{3dB} estimates inferred from average isoresponse NN functions decreased with increasing masker–probe time gap or, correspondingly, with increasing masker level (Table 2 and Fig. 5). Corresponding estimates inferred from isolevel NN functions remained approximately constant at 500 Hz or increased slightly at 4 kHz with increasing level.
5. Q_{3dB} values were higher at 4 kHz than at 500 Hz, regardless of whether they were inferred using sinusoidal or NN maskers, or from isoresponse or isolevel curves (Fig. 5).
6. There is a reasonably good correspondence between auditory filter isolevel curves inferred using sinusoidal and NN maskers (Fig. 8).
7. BW_{3dB} estimates inferred from isoresponse curves at low levels were narrower for sinusoidal than for NN stimuli at 500 Hz but not at 4 kHz (Fig. 9). This suggests that off-frequency listening is more important at 500 Hz than at 4 kHz.

Assumptions

The present behavioral results were obtained using the same linear reference TMC throughout (Fig. 2). This procedure carries four implicit assumptions about the post-cochlear recovery from forward masking:

1. it is a linear, level-independent process;
2. it is well described by the TMC for a sinusoidal masker that is presumably processed linearly by the BM, i.e., by the so-called linear reference TMC;

3. it is independent of probe and masker frequency;
4. it is similar for sinusoidal and NN maskers.

The first three assumptions are commonly used when inferring peripheral I/O curves from TMCs (see Appendix A). Regarding assumption #1, Wojtczak and Oxenham (2009) have reported that the rate of recovery from forward masking may be slower for levels of the linear reference above ~85 dB SPL and that this could lead to overestimate the degree of compression. It is unlikely, however, that this has affected the present results because the slope of mean linear reference employed here remained identical for levels below/above 85 dB SPL (Fig. 2).

As for assumption #3, Stainsby and Moore (2006) suggested that the rate of recovery from forward masking could be faster for low than for high-probe frequencies. This could have lead to an overestimate of compression at 500 Hz. Assumption #3, however, is supported by studies showing that identical apical compression estimates are obtained with and without assuming a linear reference condition (Lopez-Poveda and Alves-Pinto 2008; Plack et al. 2008). Furthermore, there is reasonable concern that compression may not have been totally absent in the apical regions of the subjects employed by Stainsby and Moore (Lopez-Poveda and Alves-Pinto 2008). These issues are still under debate.

Finally, assumption #4 is, to our knowledge, yet to be confirmed. This assumption, however, seems reasonable for two reasons. First, the post-cochlear recovery from forward masking ultimately depends on the response of auditory nerve fibers and these will depend on cochlear excitation rather than the type of stimulus. Second, an excellent correspondence was observed between isolevel filter shapes inferred using assumption #4 for sinusoidal and NN maskers (Fig. 8).

In any case, the present results and ideas hold true so long as responses to sinusoidal and NN maskers are compressive at 500 Hz and 4 kHz over the range of levels considered and good evidence in support of this has been obtained elsewhere using behavioral methods that do not rely on a linear reference TMCs (e.g., Lopez-Poveda and Alves-Pinto 2008; Plack and Drga 2003) as well as with physiological methods (e.g., Gorga et al. 2007; Johannesen and Lopez-Poveda 2008; Williams and Bacon 2005).

BW estimates at low levels from isoresponse and isolevel curves

The main result of the present study is a demonstration that isoresponse and isolevel curves suggest very different estimates of tuning for compressive filters. Specifically, low level BW_{3dB} estimates inferred from isoresponse and isolevel curves for NN maskers were

69 vs. 174 Hz, respectively, at 500 Hz, and 280 vs. 464 Hz, respectively, at 4 kHz (Tables 2 and 3). Overall, these results were supported by the present responses to sinusoidal maskers, even considering that the latter could be affected by off-frequency listening at 500 Hz (Fig. 9).

The present BW_{3dB} estimates inferred from isoresponse curves for NN maskers are in close agreement with those of Oxenham and Shera (2003) at 4 kHz and follow the trend of their data at 500 Hz (compare open and filled squares in Fig. 10). This is not surprising considering that the two sets of results were obtained using identical methods. [Note that Oxenham and Shera (2003) reported equivalent rectangular bandwidths (ERBs) rather than BW_{3dB}. Approximate values for the latter were obtained here assuming that the ERBs are approximately 11% larger than the corresponding BW_{3dB} (Moore 2007, p. 55)].

Interestingly, present BW_{3dB} estimates from corresponding isolevel curves were broader and in close agreement with the estimates of Glasberg and Moore (1990) at 4 kHz (Fig. 10), which were obtained using maskers with a fixed spectrum level and so reflected an isolevel condition. This makes it tempting to conjecture that the estimates of Oxenham and Shera (2003) were much sharper than those of Glasberg and Moore (1990) possibly because the former provide an isoresponse measure while the latter provide an isolevel measure of auditory filter tuning. This, however, is not directly supported by the present data at 500 Hz, where present isolevel NN estimates were much broader than those of Glasberg and Moore (Fig. 10). Furthermore, Oxenham and Shera (2003) showed that the difference between their tuning estimates and those of Glasberg and Moore (1990)
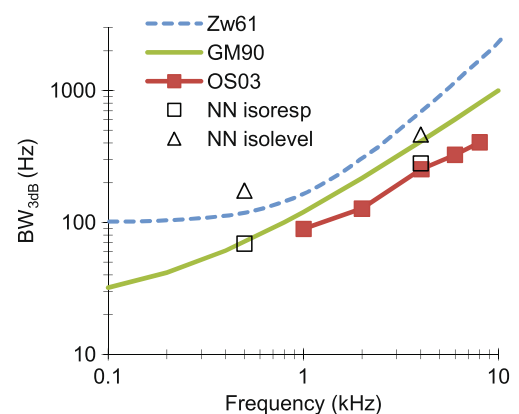
FIG. 10. Comparison of present BW_{3dB} estimates with results from earlier representative studies. Zw61: Zwicker (1961); GM90: as described by the Glasberg and Moore (1990) formula: $BW_{3dB} \sim 0.89 ERB_N = 24.7(0.0043F + 1)$, where F denotes frequency in Hz; OS03: Oxenham and Shera (2003). NN-isoresp and NN-isolevel: present estimates based on isoresponse and isolevel functions for NN maskers.

may be accounted for to a large extent by suppression effects because the methods employed in the two studies (i.e. fixed probe level vs. fixed masker level) provide closer results when they were applied using simultaneous maskers (Fig. 7 of Oxenham and Shera 2003). In any case, in light of the present results, it seems misleading to compare isoresponse with isolevel estimates of tuning, as was done by Oxenham and Shera (2003).

The present results also suggest that care must be exercised when comparing behavioral measures of human auditory tuning with physiological measures of animal frequency selectivity. Direct comparisons should be made only for corresponding measures of tuning, whether isoresponse or isolevel. Given that auditory nerve tuning is typically characterized using tuning curves (isoresponse conditions), human tuning should also be estimated from behavioral isoresponse (as opposed to isolevel) measures (e.g., Shera et al. 2002; Oxenham and Shera 2003). Even in this case, however, direct comparisons are difficult. Due to the compressive nature of cochlear responses, isoresponse tuning becomes sharper with increasing level over a range of levels (see Figs. 1F, I, 3 and 5; see also Figs. 8 and 9 of Temchin et al. 2008, and Figs. 1 and 2 of Lopez-Poveda et al. 2007). Therefore, behavioral–physiological isoresponse tuning comparisons should be made only for matching stimulus levels. It would be misleading to compare behavioral measures of tuning, which can be made only at *supra*-threshold levels, with neural *threshold* tuning curves, unless they are for matching levels. Actually, this might explain, at least in part, the across-species differences in tuning reported by Shera et al. (2002). The issue requires further investigation.

Implications for the notched-noise method for determining auditory filter shape

It is now widely accepted that behavioral auditory frequency selectivity reflects cochlear frequency selectivity (e.g., Evans 2001; Shera et al. 2002). Designing a behavioral method to characterize the *shape* and tuning of human cochlear filters has been a long-standing goal. (Note that filter shape is, by definition, an isolevel concept.) The NN method has been widely used for this purpose. It is now clear that it is best applied in forward rather than in simultaneous masking to prevent nonlinear masker–probe interactions (Oxenham and Shera 2003; Moore and Glasberg 1981). There has been some controversy, however, about whether it should be applied using fixed masker levels or fixed probe levels (e.g., Rosen and Baker 1994; Baker et al. 1998; Baker and Rosen 2006; Glasberg and Moore 2000). Traditionally, when the masker level is fixed the signal level is varied and vice versa. Given the present evidence and the

nonlinear character of mammalian cochlear responses, both approaches appear inappropriate when a level-independent filter shape (like the roex) is assumed, which is commonly the case. Fixing the probe level (as in Oxenham and Shera 2003) resembles an isoresponse condition, which in light of the present evidence, might lead to overestimating tuning. As explained above, the resulting isoresponse NN functions (Fig. 6) reflect two concurrent effects: the decrease of filter response to the NN maskers with increasing notch width and BM compression of the masker. Hence, the present isoresponse NN functions (Fig. 6) appear steeper than they would have been if the NN maskers had not been compressed (Fig. 7A, B), which may be erroneously interpreted as indicative of narrower tuning. Therefore, it is inappropriate to infer filter shapes and tuning from fixed-probe NN functions assuming a level-independent filter shape. By extension, it would be inappropriate to characterize level-dependent human cochlear filter tuning using different probe levels (e.g., as done by Oxenham and Simonson 2006).

Fixing the masker level would resemble an isoinput condition, which in light of the present theoretical evidence would allow a more direct assessment of cochlear filter tuning. Unfortunately, varying the probe level could facilitate probe detection through filters other than the targeted filter due to level-dependent changes in the vibration pattern of the BM (i.e., broader excitation patterns and peak shifts at high levels). If probe detection was based on the average response of a number of adjacent filters, these complications would affect tuning estimates even for NN maskers; particularly at high masker levels and for wide notch widths. Furthermore, and perhaps most importantly, in principle masked probe thresholds may be affected by the effects of BM compression on the probe. That is, they would reflect the combined effect of decreased filter input with increasing notch width and BM compression on the probe. As a result, the decrease in masked probe threshold with increasing notch width may be larger than would occur if the probe were not compressed, which may be erroneously interpreted as indicative of sharper tuning. Therefore, it seems inappropriate to infer filter shapes and tuning from fixed-masker NN functions considering a level-independent filter, like the roex.

An alternative NN procedure has been proposed here that overcomes these problems by using fixed-level maskers (to guarantee an isoinput condition) and a fixed, low-level probe (to confine probe excitation to a fixed, narrow cochlear region), and measuring the masker–probe time gap at the detection threshold. The assumption is that the gap threshold is proportional to the BM response to the NN masker (see 'Methods' section). The resulting isolevel functions (Fig. 7C, D) reflect the decrease in filter's response with increasing notch width for a

fixed level forward masker and a fixed level probe. Therefore, these functions allow a more accurate estimation of filter shape and tuning when assuming a level-independent filter shape (like the roex). Filter shapes and tuning estimates obtained with this procedure may be compared more directly with BM isolevel curves. It has been shown that tuning estimates obtained with this method change with stimulus level in a form that is qualitatively consistent with the level-dependent shape of physiological BM isolevel curves (Figs. 5 and 8).

This is not to say that fixed-masker or fixed-probe NN data reported in earlier studies are inadequate to infer the shape and tuning of human cochlear filters. Such inferences are indeed possible but should be made using a nonlinear filter that incorporates the concomitant effects of tuning and compression. Furthermore, inferences should always be made by comparing model and experimental responses for corresponding isolevel or isoresponse conditions and for matching levels. It is also reasonable to compare behavioral estimates of human cochlear tuning inferred from isoresponse NN functions with animal neural tuning curves because the two will be similarly affected by compression (e.g., Oxenham and Shera 2003; Shera et al. 2002). That said, such comparisons should be made only for matching stimulus levels for the reasons explained in the preceding section.

Implications for auditory physiology

The present evidence suggests that care must be exercised when explaining the tuning of auditory nerve fibers or of higher auditory neurons with the same best frequency but different thresholds. Potential differences in their frequency response areas may be due simply to the effects of cochlear compression described in "Nonlinear effects on tuning: theoretical considerations" section (Fig. 1) rather than to differences in other physiological properties or mechanisms.

CONCLUSION

Isolevel and isoresponse curves are two complementary ways of describing the tuning of cochlear filters but may not be used interchangeably. The latter may suggest much narrower tuning than the former, depending on compressive characteristics of the cochlear responses.

ACKNOWLEDGMENTS

APPENDIX A. ASSUMPTIONS OF THE TMC METHOD

The TMC method consists of measuring the level of a sinusoidal masker required to just mask a sinusoidal probe as a function of the time gap between the masker and the probe. The probe level is fixed a few decibels above the absolute threshold for the probe. The masker level at threshold increases with increasing masker–probe time gap. It is assumed that the rate of increase reflects both the time course of recovering from the internal (post-compressive) masker effect and the degree of cochlear compression that affects the masker at a cochlear site with a CF approximately equal to the probe frequency. Because the probe level is fixed at all times, the method is assumed to measure the masker level (input) required to generate a fixed recovery from masker excitation. This is the reason that the resulting functions are referred to as isoresponse TMCs. See Nelson et al. (2001) for a full justification of these assumptions.

There is evidence that the time course of recovery from the internal masker effect is independent of masker frequency over a reasonable range of masker levels ($\leq$85 dB SPL) (Wojtczak and Oxenham 2009). This makes it possible to infer BM I/O curves by plotting the level of a masker that is processed linearly by the BM, the so-called linear reference TMC, against the level of the masker of interest, paired according to time gap (Nelson et al. 2001). Because basal BM responses to tones an octave or so below the CF are approximately linear, it is common to use the TMC for a masker frequency an octave or so below that of a high-frequency probe as the linear reference (Lopez-Poveda et al. 2003; Lopez-Poveda and Alves-Pinto 2008).

The TMC method of deriving BM responses is based on the additional assumption that masker level at threshold for any given masker–probe time gap depends on the degree of excitation caused by the maskers at a cochlear site tuned to the probe frequency (Nelson et al. 2001). Therefore, a plot of masker level against masker frequency for a given time gap is assumed to be related to the degree of frequency selectivity at the cochlear site in question and is called a PTC (e.g., Kidd and Feth 1981). Masker levels at threshold increase with increasing masker–probe time interval. Consequently, frequency selectivity at different levels can be estimated from PTCs for different masker–probe time intervals (e.g., Kidd and Feth 1981; Lopez-Poveda et al. 2007).

Probe level is fixed a few decibels above threshold in an attempt for results to reflect the response characteristics of a single cochlear site with a CF approximately equal to the probe frequency (Nelson et al. 2001). Fixing the probe at a very low level almost guarantees that the region of the cochlea under study is the same for different masker input levels.

REFERENCES

BAKER RJ, ROSEN S (2006) Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. J Acoust Soc Am 119:454–462

BAKER RJ, ROSEN S, DARLING AM (1998) An efficient characterization of human auditory filter filtering across level and frequency that is also physiologically plausible. In: Palmer AR, Rees A, Summerfield AQ, Meddis R (eds) Psychophysical and physiological advances in hearing. Whurr, London, pp 81–88

BUSER P, IMBERT M (1992) Audition (English edition). MIT press, Cambridge

CHISTOVICH LA (1957) Frequency characteristics of masking effect. Biophysics (Engl Transl) 2:743–755

COOPER NP (2004) Compression in the peripheral auditory system. In: Bacon SP, Fay RR, Popper AN (eds) Compression: from cochlea to cochlear implants. Springer, New York, pp 18–61

EVANS EF (2001) Latest comparisons between physiological and behavioral frequency selectivity. In: Breebaart J, Houtsma AJM, Kohlrauschl A, Prijs R, Schoonhoven R (eds) Physiological and psychophysical bases of auditory function. Shaker, Maastricht, pp 382–387

EVANS EF (1975) Cochlear nerve and cochlear nucleus. In: Keidel WD, Neff WD (eds) Handbook of sensory physiology, vol. 5/2. Springer, Berlin, pp 1–108

GLASBERG BR, MOORE BCJ (1990) Derivation of auditory filter shapes from notched noise data. Hear Res 47:103–138

GLASBERG BR, MOORE BCJ (2000) Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. J Acoust Soc Am 108:2318–2328

GOODE RL, KILLION M, NAKAMURA K, NISHIHARA S (1994) New knowledge about the function of the human middle ear: development of an improved analog model. Am J Otol 15:145–154

GORGA MP, NEELY ST, DIERKING DM, KOPUN J, JOLKOWSKI K, GROENENBOOM K, TAN H, STIEGEMANN B (2007) Low-frequency and high-frequency cochlear nonlinearity in humans. J Acoust Soc Am 122:1671–1680

HARRIS DM, DALLOS P (1979) Forward masking of auditory nerve responses. J Neurophysiol 42:1083–1107

HEINZ MG, COLBURN MS, CARNEY LH (2002) Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates. J Acoust Soc Am 111:996–1011

HOUTGAST T (1973) Psychophysical experiments on "tuning curves" and "tow-tone inhibition". Acustica 29:168–179

JOHANNESEN PT, LOPEZ-POVEDA EA (2008) Cochlear nonlinearity in normal hearing subjects as inferred psychophysically and from distortion-product otoacoustic emissions. J Acoust Soc Am 124:2149–2163

JOHNSON-DAVIS D, PATTERSON RD (1979) Psychophysical tuning curves: restricting the listening band to the signal region. J Acoust Soc Am 65:765–770

KIDD G, FETH LL (1981) Patterns of residual masking. Hear Res 5:49–67

LEVITT H (1971) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49:467–477

LOPEZ-POVEDA EA, ALVES-PINTO A (2008) A variant temporal-masking-curve method for inferring peripheral auditory compression. J Acoust Soc Am 12:1544–1554

LOPEZ-POVEDA EA, PLACK CJ, MEDDIS R (2003) Cochlear nonlinearity between 500 and 8,000 Hz in normal-hearing listeners. J Acoust Soc Am 113:951–960

LOPEZ-POVEDA EA, BARRIOS LF, ALVES-PINTO A (2007) Psychophysical estimates of level-dependent best-frequency shifts in the apical region of the human basilar membrane. J Acoust Soc Am 121:3646–3654

MEDDIS R, O'MARD LP (2006) A computer model of the auditory nerve response to forward-masking stimuli. J Acoust Soc Am 117:3787–3798

MEDDIS R, O'MARD LP, LOPEZ-POVEDA EA (2001) A computational algorithm for computing non-linear auditory frequency selectivity. J Acoust Soc Am 109:2852–2861

MOORE BCJ (1987) Distribution of auditory filter bandwidths at 2 kHz in young normal listeners. J Acoust Soc Am 81:1633–1635

MOORE BCJ (2007) Cochlear hearing loss. Physiological, psychological and technical issues, 2nd edn. Wiley, Chichester

MOORE BCJ, GLASBERG BR (1981) Auditory filter shapes derived in simultaneous and forward masking. J Acoust Soc Am 70:1003–1014

MOORE BCJ, GLASBERG BR (2003) Behavioral measurement of level-dependent shifts in the vibration pattern on the basilar membrane at 1 and 2 kHz. Hear Res 175:66–74

NELSON DA, SCHRODER AC, WOJTCZAK M (2001) A new procedure for measuring peripheral compression in normal-hearing and hearing impaired listeners. J Acoust Soc Am 110:2045–2064

O'LOUGHLIN BJ, MOORE BCJ (1981) Off-frequency listening: effects on psychoacoustical tuning curves obtained in simultaneous and forward masking. J Acoust Soc Am 69:1119–1125

OXENHAM AJ, SHERA CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. J Assoc Res Otolaryngol 4:541–554

OXENHAM AJ, SIMONSON A (2006) Level dependency of auditory filters in nonsimultaneous masking as a function of frequency. J Acoust Soc Am 119:444–453

PATTERSON RD (1976) Auditory filter shapes derived with noise stimuli. J Acoust Soc Am 59:640–654

PATTERSON RD, NIMMO-SMITH I (1980) Off-frequency listening and auditory filter asymmetry. J Acoust Soc Am 67:229–245

PICKLES JO (2008) An introduction to the physiology of hearing, 3rd edn. Emerald, Bingley

PLACK CJ, DRGA V (2003) Psychophysical evidence for auditory compression at low characteristic frequencies. J Acoust Soc Am 113:1574–1586

PLACK CJ, SKEELS V (2007) Temporal integration and compression near absolute threshold in normal and impaired ears. J Acoust Soc Am 122:2236–2244

PLACK CJ, OXENHAM AJ, SIMONSON AM, O'HANLON CG, DRGA V, ARIFIANTO D (2008) Estimates of compression at low and high frequencies using masking additivity in normal and impaired ears. J Acoust Soc Am 123:4321–4330

RHODE WS, COOPER NP (1996) Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo. Aud Neurosci 3:101–121

ROBLES L, RUGGERO MA (2001) Mechanics of the mammalian cochlea. Physiol Rev 81:1305–1352

ROSEN S, BAKER R (1994) Characterizing auditory filter nonlinearity. Hear Res 73:231–243

ROSEN S, BAKER RJ, DARLING A (1998) Auditory filter nonlinearity at 2 kHz in normal hearing listeners. J Acoust Soc Am 103:2539–2550

RUGGERO MA, TEMCHIN AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. Proc Natl Acad Sci 102:18614–18619

RUGGERO MA, RICH NC, RECIO A, NARAYAN SS, ROBLES L (1997) Basilar-membrane responses to tones at the base of the chinchilla cochlea. J Acoust Soc Am 101:2151–2163

SHERA CA, GUINAN JJ JR, OXENHAM AJ (2010) Otoacoustic estimation of cochlear tuning: validation in the chinchilla. J Assoc Res Otolaryngol 11:343–365

SHERA CA, GUINAN JJ JR, OXENHAM AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Nat Acad Sci 99:3318–3323

STAINSBY TH, MOORE BCJ (2006) Temporal masking curves for hearing-impaired listeners. Hear Res 218:98–111

TEMCHIN AN, RUGGERO MA (2010) Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics. J Assoc Res Otolaryngol 11:297–318

TEMCHIN AN, RICH NC, RUGGERO MA (2008) Threshold tuning curves of chinchilla auditory nerve fibers. II. Dependence on spontaneous activity and relation to cochlear nonlinearity. J Neurophysiol 100:2899–2906

VERSCHUURE J (1980) Pulsation patterns and nonlinear auditory tuning. Hear Res 2:397–405

VERSCHUURE J (1981) Pulsation patterns and nonlinearity of auditory tuning. Acustica 49:288–295

VOGTEN LLM (1974) Pure-tone masking: a new result from a new method. In: Zwicker E, Terhardt E (eds) Facts and models in hearing. Springer, New York, pp 142–155

VOGTEN LLM (1978) Simultaneous pure-tone masking: the dependence of masking asymmetries on intensity. J Acoust Soc Am 63:1509–1519

WILLIAMS EJ, BACON SP (2005) Compression estimates using behavioral and otoacoustic emission measures. Hear Res 201:44–54

WOJTCZAK M, OXENHAM AJ (2009) Pitfalls in behavioral estimates of basilar membrane compression in humans. J Acoust Soc Am 125:270–281

YASIN I, PLACK CJ (2003) The effects of a high-frequency suppressor on tuning curves and derived basilar-membrane response functions. J Acoust Soc Am 114:322–332

YASIN I, PLACK CJ (2005) Psychophysical tuning curves at very high frequencies. J Acoust Soc Am 118:2498–2506

ZWICKER E (1961) Subdivision of the audible frequency range into critical bands (Frequenzgruppen). J Acoust Soc Am 33:248