JAIST Repository
https://dspace.jaist.ac.jp/

Title	Estimate of auditory filter shape using notched-noise masking for various signal frequencies
Author(s)	Unoki, Masashi; Ito, Kazuhito; Ishimoto, Yuichi; Tan, Chin-Tuan
Citation	Acoustical science and technology, 27(1): 1-11
Issue Date	2006
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/4017
Rights	日本音響学会, Masashi Unoki, Kazuhito Ito, Yuichi Ishimoto, and Chin-Tuan Tan, Acoustical science and technology, 27(1), 2006, 1-11.
Description	

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Acoust. Sci. & Tech. **27**, 1 (2006)

PAPER

Estimate of auditory filter shape using notched-noise masking for various signal frequencies

Masashi Unoki[1,*], Kazuhito Ito[1,2,†], Yuichi Ishimoto[1,‡] and Chin-Tuan Tan[1,§]

[1]*School of Information Science, Japan Advanced Institute of Science and Technology, 1–1 Asahidai, Tatsunokuchi, Nomi, Ishikawa, 923–1292 Japan*
[2]*Living Informatics Group, Institute of Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology, 1–8–31 Midorigaoka, Ikeda, 563–8577 Japan*

(*Received 27 December 2004, Accepted for publication 2 August 2005*)

Abstract: In this paper, the masked threshold of a sinusoidal signal in the presence of a notched-noise masker was measured experimentally for five normal-hearing subjects. The frequencies of sinusoidal signals used in the measurement were 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz. The conditions and procedure in our measurement were the same as those used by Glasberg and Moore (2000), with additional measurements at 125 and 6,000 Hz. Uniformly excited noise (UEN) was not used in our measurements. The measured data was used to estimate the parameters of a double roex auditory filter as presented in Glasberg and Moore (2000). Basically, this filter is the sum of a tip filter and a tail filter, with its gain controlled by a schematic family of input-output functions. The PolyFit procedure was used to fit the filter to the measured data. An individual auditory filter was fitted at each of the signal frequencies in our measurements. The results showed that auditory filter shape varied with level. The gain of the filters centered at frequencies between 125 Hz and 1,000 Hz, increased as the center frequency increased. Above 1,000 Hz, the gain of the filters remained at a constant value. These results are consistent with the results in Baker *et al.* (1998) and Glasberg and Moore (2000).

Keywords: Frequency selectivity, Notched-noise masking, Auditory filter shape, Asymmetry, Compression

PACS number: 43.66.Ba, 43.66.Dc [DOI: 10.1250/ast.27.1]

1. INTRODUCTION

Fundamentally, the human auditory system analyzes sound in the time-frequency domain. This analysis of sound can be conceptualized as a series of overlapping bandpass filters, which is often referred to as the auditory filterbank. It is widely believed that the nature of frequency selectivity in the human auditory system can be characterized through the analytical nature of the auditory filterbank [1,2].

Over the past 30 years, many studies have been done in efforts to achieve better estimation of the auditory filterbank with behaviorally measured data obtained from both simultaneous and non-simultaneous masking experiments using a notched-noise masker (e.g., [3–17]). Recently, Baker *et al.* (1998) and Glasberg and Moore (2000) measured masked thresholds for detecting sinusoidal signals simultaneously presented with notched-noise maskers over a wide range of signal frequencies and levels encountered in everyday hearing [18,19]. The notch in the masker was positioned both symmetrically and asymmetrically around the signal frequency. They developed the PolyFit procedure to better fit the auditory filters to the measured data using roex filter functions [2]. Several critical observations regarding the resultant auditory filters were made through by these two studies: (i) the shape of the auditory filter changes with the signal level at all center frequencies; (ii) the gain of the filter at the peak frequency increases nonlinearly as the signal level decreases (i.e., compression occurs); and (iii) for the filters centered between 250 Hz and 1,000 Hz, the degree of compression of the filter increases with increasing frequency [18,19]. However, similar studies on auditory filters fitted to measured data from forward masking [3,6,7,9] and temporal masking [5,12] experiments have led to a different observation. The tuning of auditory filters fitted to the data measured in a non-simultaneous masking experiment was sharper than

*e-mail: unoki@jaist.ac.jp
†e-mail: keith.ito@aist.go.jp
‡e-mail: y-ishi@jaist.ac.jp
§e-mail: chintuan@jaist.ac.jp

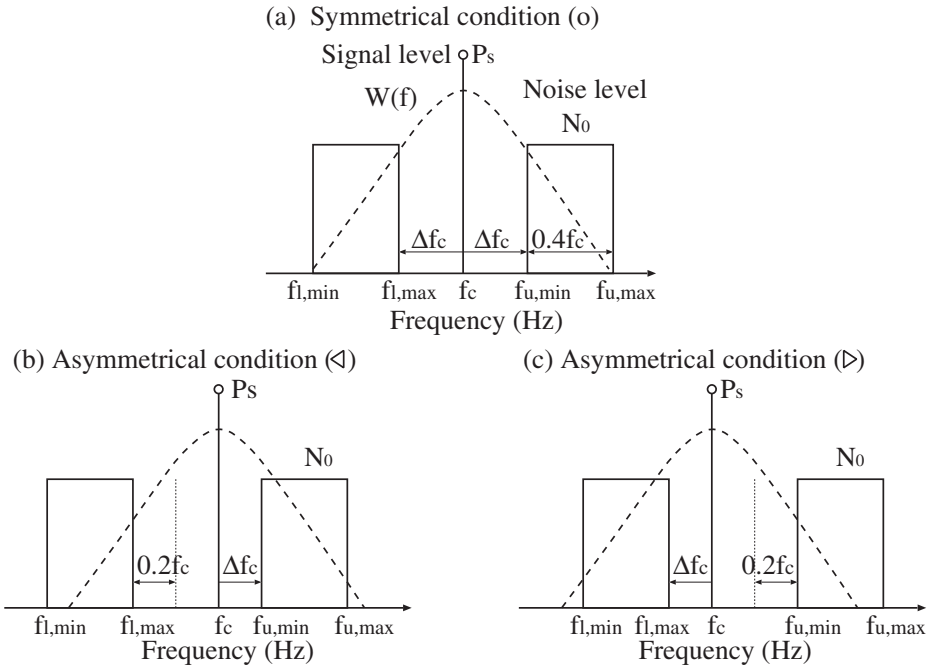

Fig. 1 Stimulus shape used in notched-noise masking measurement. f_c, Δf_c, P_s, and N_0 are signal frequency (Hz), notch width (Hz), signal level (dB SPL), and noise masker level (dB SPL/Hz), respectively. $W(f)$ is the weighting function in the power spectrum model, corresponding to the auditory filter shape: (a) symmetrical notch condition, and asymmetrical notch condition in the (b) lower and (c) upper sides.

that when filters were fitted to data measured in a simultaneous masking experiment. Studies have shown that the difference in the tuning of the auditory filters derived from simultaneous and forward masking data can be caused by the suppression effect.

The past studies suggest that simultaneous and non-simultaneous masking experiments should influence the results of the present auditory filter modeling in significantly different ways. Since different experimental setups were used in the previous studies for both the simultaneous and forward masking experiments, a sound comparison study to investigate the difference between the findings these two groups of experiments seems unlikely. A complete and systematic measurement of masked thresholds through both simultaneous and forward masking experiments using one common experimental setup is necessary for more comprehensive auditory filter modeling. With this objective, we attempt to perform a series of simultaneous and non-simultaneous experiments, which systematically measure the masked threshold with various possible combinations of maskers and signals covering the range of frequencies and levels encountered in everyday hearing, and within the ranges of variation observed in the shape of the auditory filters reported from previous studies.

In this paper, we present the first part of our study, which is the simultaneous masking experiment. We measured masked thresholds in simultaneous notched-noise masking and estimated the auditory filter shape from the measured thresholds. We compared our results to those reported by Glasberg and Moore [19] and Baker *et al.* [18]. The organization of this paper is as follows. Section 2 describes the simultaneous masking experiment where we used a notched-noise masker. Section 3 describes the estimation of the auditory filter shape using the measured thresholds. Section 4 gives a summary.

2. NOTCHED-NOISE MASKING MEASUREMENT

2.1. Stimuli

Figure 1 shows the shape of the stimulus used in this notched-noise masking experiment. A listener was required to detect a brief sinusoidal signal (referred to as the "signal"), in the presence of a noise with a spectral notch designed to be placed within the frequency region of the signal (referred to as the "notched-noise masker"). The level of the brief sinusoidal signal at which it becomes just audible to the listener in the presence of the notched-noise masker was referred to as the "masked threshold". In the following explanations of the stimulus, we use the following symbols: f_c denotes signal frequency (Hz), P_s denotes signal level (dB SPL), N_0 denotes masker noise level (dB SPL/Hz), and Δf_c denotes the notch width from the signal frequency (Hz).

The notched-noise used in the experiment was digitally created using Matlab (ver. 6.5.1. Mathworks) at a sampling rate of 24.41 kHz on a PC Linux computer. The notched-

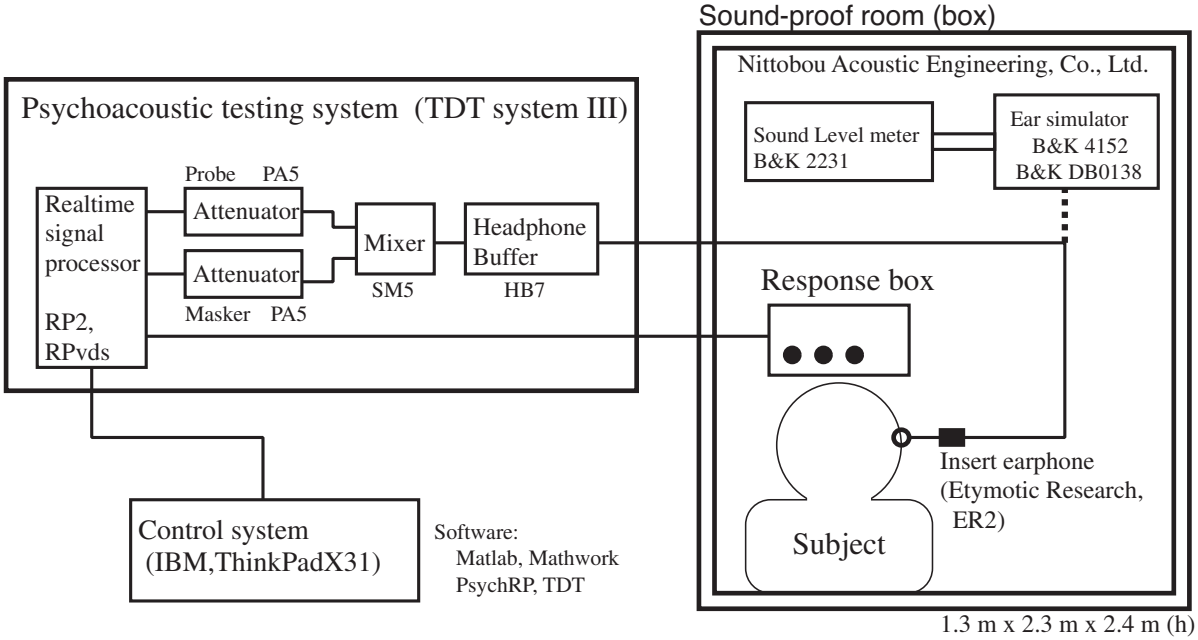

Fig. 2 Environment for the masking experiments: Psychoacoustical testing system, control system, and sound-proof room.

noise comprised two bands of noise and each band had a bandwidth of $0.4f_c$ (Fig. 1). A spectral notch was created between these two bands of noise. Bandwidths of the lower and upper bands are denoted as $f_{l,max} - f_{l,min}$ and $f_{u,max} - f_{u,min}$, respectively. Subscripts l and u refer to the lower and upper bands, and max and min refer to the high and low frequency cutoff positions of each band. These two bands of noise were shifted in such a way that a notch was symmetrically or asymmetrically placed around the signal frequency f_c. This notched-noise was created using fast Fourier transforms (FFT). The amplitudes of the spectral components (spaced at intervals of 0.187 Hz) within the two bands of the noise were set to equal and non-zero values. The phases of these spectral components were randomized from the range of 0 to 2π in radian. The amplitudes of the remaining spectral components beyond the boundaries of the two bands of noise were set to zero. An inverse FFT was applied to this constructed spectrum to recreate the waveform of the notched-noise masker in a time domain with a length of 5.4 s.

The relative notch width for each notched-noise masker centered at f_c was defined as the normalized deviation of each edge of the normalized notch from f_c, denoted as $\Delta f_c/f_c$. There were seven conditions in which the notch was symmetrically placed about the signal: the values of $\Delta f_c/f_c$ were 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6. There were twelve conditions in which the notch was asymmetrically placed about the signal. The combination of the upper and lower normalized deviations $\Delta f_c/f_c$ were $(0.1, 0.3)$, $(0.3, 0.1)$, $(0.2, 0.4)$, $(0.4, 0.2)$, $(0.3, 0.5)$, $(0.5, 0.3)$, $(0.4, 0.6)$, $(0.6, 0.4)$, $(0.5, 0.7)$, $(0.7, 0.5)$, $(0.6, 0.8)$, and $(0.8, 0.6)$.

At a fixed noise level N_0, we measured masked thresholds at signal frequencies (f_c) of 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz with the notch width of the masker and the signal level P_s varying. The noise level N_0 was chosen as follows: (1) at $f_c = 125$ Hz, $N_0 = 37.3$, 47.3, and 57.3 dB SPL/Hz; (2) at $f_c = 250$, 500, 1,000, and 2,000 Hz, $N_0 = 27.3$, 37.3, and 47.3 dB SPL/Hz; (3) at $f_c = 4{,}000$ and 6,000 Hz, $N_0 = 17.3$, 27.3, and 37.3 dB SPL/Hz. At $f_c = 125$ Hz, the minimum value of the low-frequency edge of the lower noise band $f_{l,min}$ was kept at 10 Hz. Whereas at other values of f_c, the minimum value of the low-frequency edge of the lower noise band $f_{l,min}$ was kept at 40 Hz.

Figure 2 shows the setup of our testing system using the Tucker-Davis Technologies (TDT) system III. Both the sinusoidal signal and the notched-noise masker were reproduced via a real-time signal processor (TDT RP2) at a sampling rate of 24.41 kHz using RPvds software. In each trial, three bursts of the notched-noise masker were presented, of which one of them (randomly selected) was accompanied with a sinusoidal signal. The notched-noise masker was 200 ms of notched-noise extracted from the 5.4 seconds of notched-noise previously created. The masker was reshaped with a 180-ms steady-state portion and 10-ms raised-cosine ramps. The ramps on the masker were achieved through the RPvds gating function via the TDT RP2. The inter-stimulus interval was kept at 500 ms (Fig. 3). The sinusoidal signal was reproduced using the TDT RP2 with RPvds and was presented simultaneously with one of the three maskers randomly selected. The signal duration was equal to the masker duration.

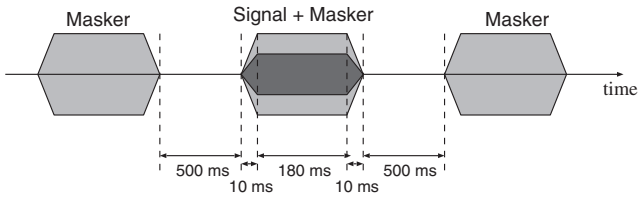

Fig. 3 Stimuli position in notched-noise masking.

Both the sinusoidal signal and the notched-noise masker were individually attenuated by two separate TDT PA5s to achieve their desired levels. They were then mixed at the TDT SM5 and passed into a headphone buffer (TDT HB7) before being presented to the subject via the insert earphone (Etymotic Research ER2). The ER2 earphone has a flat frequency response at the eardrum up to about 14 kHz. The levels of the stimuli were verified using a B&K 4152 Artificial Ear Simulator with a 2 cm^3 coupler (B&K DB 0138) and a B&K 2231 Modular Precision Sound Level Meter. The entire experiment was conducted in a double-walled sound-attenuating booth box (Nittobou Acoustical Engineering Co., Ltd.). The conditions and procedure used in the notched-noise masking experiment were the same as those used by Glasberg and Moore [19] with the exception of two additional measurements at frequencies of 125 Hz and 6,000 Hz. Uniformly exciting noise (UEN) was not used in this experiment.

2.2. Subjects

Five subjects (AH, MT, YY, MU, and YI) were tested. Two of the subjects (MU and YI) are authors of this paper. The absolute thresholds of all subjects, measured through a standard audiometric tone test using a RION AA-72B audiometer, were 10 dB HL or less for both ears at octave frequencies between 125 Hz and 8,000 Hz. The output signal levels (in dB SPL) of the AD-02 headphone with a RION AA-72B audiometer at various dB HL and frequencies were verified using a B&K 4153 Artificial Ear Simulator and a B&K 2231 Modular Precision Sound Level Meter. These verified signal levels were used to convert the absolute thresholds from dB HL to dB SPL. The mean of the absolute thresholds for the five subjects were 32.0, 22.7, 11.8, 8.4, 10.0, 12.9, and 11.3 dB SPL at 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz, respectively. Only the better ear of the subject was tested. Three subjects (AH, YY, and MU) were tested on their right ears while the other subjects (MT and YI) were tested on their left ears. The subject ages ranged from 23 to 35 years old. All subjects were given at least two hours of practice.

2.3. Procedure

Masked thresholds were measured using a three-alternative forced-choice (3AFC) three-down one-up procedure that tracks the 79.4% point on the psychometric function [20]. The procedure was controlled by TDT PsychRP software. Three intervals of stimuli were presented sequentially in each trial. Subjects were asked to identify the interval which carried the signal using the numbered push-buttons on the response box. Feedback was provided by lighting up LEDs on the response box, when the correct interval was identified. A run was terminated after twelve reversals. The step size was 5 dB for the first four reversals and 2 dB thereafter. The threshold was defined as the mean signal level at the last eight reversals. Each subject was tested with at least two runs for all conditions. When the difference between the thresholds measured in the two runs for a condition was more than 8 dB, additional runs were done until any two of the runs had a threshold difference within 8 dB. The masked threshold of the condition is the average value of the thresholds of the two runs with a minimum threshold difference.

2.4. Results

The mean masked thresholds of the five subjects are shown in Fig. 7. In each panel of Fig. 7, the abscissa shows the smaller of the two values of $\Delta f_c/f_c$ and the ordinate shows the masked threshold. Circles ("○") denote the mean masked thresholds under the symmetric notched-noise conditions. Left-pointing triangles ("◁") denote the mean masked thresholds under the asymmetric notched-noise conditions where $\Delta f_c/f_c$ for the lower noise band was 0.2 greater than $\Delta f_c/f_c$ for the upper noise band. Right-pointing triangles ("▷") denote the mean masked threshold under the asymmetric notched-noise conditions where $\Delta f_c/f_c$ for the upper noise band was 0.2 greater than $\Delta f_c/f_c$ for the lower noise band. In each plot of Fig. 7, there are three groups of lines (each group having three lines connecting the symbols ("○", "◁", and "▷") arranged from top to bottom. The decreasing height of these three groups of lines indicates that the mean masked thresholds were measured at a decreasing masker level N_0. In plot 7(a), $N_0 = 57.3$, 47.3, and 37.3 dB SPL/Hz; in plots 7(b)–(e), $N_0 = 47.3$, 37.3, and 27.3 dB SPL/Hz; and in plots 7(f) and (g), $N_0 = 37.3$, 27.3, and 17.3 dB SPL/Hz.

In this paper, the range of levels of the masked thresholds for each masker level is referred to as the dynamic range of the masked thresholds. We found that the distribution of the dynamic range measured at different signal frequencies was similar to that found in a previous study [19]. The dynamic ranges of the masked thresholds measured at signal frequencies of 125 and 250 Hz were narrower than those of the masked thresholds measured at the other frequencies (500, 1,000, 2,000, 4,000, and 6,000 Hz). That is the masked thresholds measured at signal frequencies other than 125 and 250 Hz had a wider range of values. The dynamic ranges of the masked

thresholds measured at signal frequencies of 500, 1,000, 2,000, 4,000, and 6,000 Hz were approximately the same while the dynamic range at 125 Hz was smaller than the rest. For all signal frequencies, the masked thresholds measured under asymmetric notch conditions ("◁" and "▷") differed from the corresponding masked thresholds measured under the symmetric notch condition ("○") shown in Fig. 1. This indicates that the auditory filter shapes were asymmetrical.

In Fig. 7, an asterisk "*" on the vertical axis of each plot denotes the mean absolute threshold of the subjects. In general, the mean absolute threshold of the subjects should have been the lowest level in each plot, and all measured masked thresholds would be higher in value than the mean absolute threshold of the subjects. The masked threshold decreased as the width of the notch increased, and approached the level of the mean absolute threshold as the notch widened further.

3. FILTER SHAPE ESTIMATION

3.1. Power Spectrum Model of Masking

If the roll-off of the noise band is as steep as in Fig. 1, it is possible to write a function that relates the signal level at the masked threshold to the integral of the auditory filter. In this work, we used this relationship to estimate the auditory filter shape [1,10]. If the auditory filter shape is represented as the weighting function, $W(f)$, then the masked threshold predicted by the power spectrum model of masking is given by

$$P_s = K + N_0 + 10\log_{10}\left\{\int_{f_{l,min}}^{f_{l,max}} W(f)df + \int_{f_{u,min}}^{f_{u,max}} W(f)df\right\}, \quad (1)$$

where P_s is the power of the signal at the masked threshold, N_0 is the spectrum level of the noise, and K is a constant related to the efficiency of the detection mechanism following the auditory filter. The limits on the filter integrals are from $f_{l,min}$ to $f_{l,max}$ for the lower noise band and from $f_{u,min}$ to $f_{u,max}$ for the upper noise band. This model is often referred to as a "power spectrum model" as it simply assumes that the fluctuations within the noise bands can be ignored.

3.2. Roex Auditory Filter

We used a standard double roex (rounded-exponential) filter, roex(p, w, t), proposed by Glasberg and Moore (2000) to estimate the shape of the auditory filter. $W(f)$ is represented as

$$W(g) = \begin{cases} (1-w)(1+p_l g)e^{-p_l g} + w(1+tg)e^{-tg}, & f \le f_c \\ (1+p_u g)e^{-p_u g}, & f > f_c \end{cases} \quad (2)$$

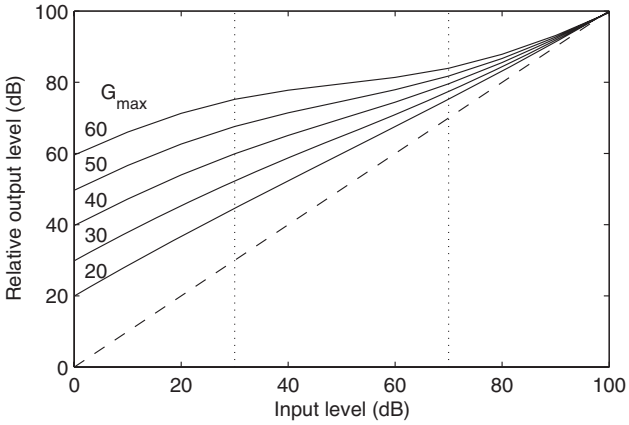

Fig. 4 Schematic I/O functions modeled by G_{dB} for values of G_{max} from 20 to 60 dB in 10-dB steps.

where the normalized frequency g is defined as $g = |f - f_c|/f_c$ and the relative gain w is defined as $w = 1/(G_{lin} + 1)$ [19]. The value of G_{lin} specifies the gain of the tip filter relative to that of the tail filter. The upper frequency side of the filter is modeled as one single roex function (tip filter), whereas the lower frequency side of the filter is modeled as the sum of two roex functions (tip and tail filters). p_l and p_u are the parameters determining the sharpness of the tip filter. t is the parameter determining the shallowness of the tail filter on the lower frequency side.

If we assume the gain of the tip filter varies with input level as defined in Eq. (3), the relative gain of the roex filter is specified by the gain of the schematic I/O function, G_{dB} [19], which models the I/O function of the basilar membrane [19,21]. G_{dB} is defined as

$$G_{dB} = 0.9L + A + B\left(1 - \frac{1}{1 + e^{-0.05(L-50)}}\right) - L, \quad (3)$$

where L is the input level, $A = -0.0894G_{max} + 10.894$, $B = 1.1789G_{max} - 11.789$, and G_{max} is a parameter that can determine the filter characteristics of the I/O function, as shown in Fig. 4. The subscripts "lin" and "dB" with regard to G denote linear- and log-scales, and $G_{lin} = 10^{G_{dB}/10}$.

The schematic I/O function was nearly linear for very low input levels (i.e., the slope of the I/O function was close to 1 on a dB/dB scale), but was compressive (i.e., the slope of the I/O function was less than 1 dB/dB) for mid-range input levels. When the input level was more than 100 dB, the slope of the I/O function remained at 1 dB/dB. Further details regarding the schematic family of I/O functions are available elsewhere [21].

3.3. Fitting Procedure

The PolyFit procedure [4,18,19,22] was used to fit a double roex filter to the notched-noise masking data

obtained as described in Sect. 2. Four filter parameters (p_l, t, G_{max}, and p_u) and two non-filter parameters (K and Abs) were used. The parameter Abs was used as the low-level limit on estimated thresholds [19]. Abs and the signal level at the threshold in Eq. (1) were used in the form of $10\log_{10}(10^{P_s/10} + 10^{Abs/10})$ in estimating the masked thresholds. In this study, the six parameters of the roex filter were optimized by nonlinearly minimizing the root mean square (rms) error between the masked thresholds and the estimated thresholds of the auditory filter, as in [22].

As for the signal level dependent (SLD) model [19], the effective input level L was assumed to equal the signal level at the masked threshold. A value of -8.2 dB, which is equivalent to a factor of 0.15 in linear power units, was used to offset the gain of the tip filter from that of the tail filter at very high signal levels, where G_{dB} became zero. In addition, two refinements were made to incorporate the outer and middle ear effect [19] and the off-frequency listening [4,22] effect into the fitting procedure. MidEar correction [23] was used in a way similar to [18,19] so as to include the effect of transmission in precochlear processing. The effect of off-frequency listening was included by locating the auditory filter that produced the best signal-to-noise ratio when the thresholds were estimated. Since the ER2 insert earphone has a flat frequency response, correction for the effect of the insert earphone was not necessary.

Alternatively, another double roex filter roex(p, w, t), whose relative gain $w = 10^{w_{dB}/10}$ was a polynomial function of the input level, $w_{dB} = w_{dB}^{(0)} + w_{dB}^{(1)} \cdot L$ (in dB), was fitted to the same set of data using the same procedure described in the two previous paragraphs. There are seven parameters in this double roex filter: five filter parameters (p_l, t, $w_{dB}^{(0)}$, $w_{dB}^{(1)}$, and p_u) and two non-filter parameters (Abs and K). The purpose of fitting the second roex filter was to highlight the role of G_{dB}, as explained further in Sect. 3.4. To facilitate the discussion below, we refer to the double roex filter whose relative gain is a polynomial function as "SLD($w$lin)" model, while the double roex filter whose relative gain is a G_{dB} function is referred to as the "SLD" model [19].

3.4. Results

The shapes of the fitted auditory filters (SLD model) centered at the following signal frequencies of 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz are plotted in Fig. 5. Figures 5(a)–(e) show the auditory filter shapes for the five subjects (AH, MT, YY, MU, and YI) with the abscissa showing the ERB rate [4]. The five curves (from top to bottom) centered at each ERB rate illustrate the changes in the auditory filter shape when the signal level increased from 30 to 70 dB in 10-dB steps. The signal level L for estimating the filter shape in Fig. 5 was directly set to the values of 30 to 70 dB without using the iterative level determination in [19]. The filter shapes for signal levels that were not among the measured signal levels in Fig. 7 were estimated through interpolation or extrapolation.

The differences in the filter shapes obtained for each individual subject were small. The pattern of variation in these auditory filter shapes seems to reveal a similar trend in all five subjects. Figure 5(f) shows the auditory filters obtained from mean masked thresholds for all five subjects. The optimized values for the six parameters of the double roex filter and the rms error at each frequency are tabulated in Table 1. The double roex filter in this study fit excellently with the simultaneous notched-noise masking data collected from the five subjects.

The optimized values of t, p_l, and p_u seemed to increase linearly when the signal frequency increased. Similar increasing trends in the values of these parameters were reported by Baker *et al.* [18]. In line with their observations, the linear variation of these parameters seemed to account for the consistently similar shape of the asymmetry for these auditory filters centered over the range of signal frequencies in their measurements. The patterns of these auditory filter shapes at 250, 1,000, and 4,000 Hz were similar to those in Glasberg and Moore (Figs. 7–9) [19].

The optimized values of the seven parameters of the roex filter and the rms error in the SLD($w$lin) model and the SLD model were separately tabulated in Tables 1 and 2. The differences in the rms errors between the SLD and SLD($w$lin) models were small at all signal frequencies. This shows that both models could account for the mean masked thresholds. However, the SLD model has the advantage of needing one less free parameter than the SLD($w$lin) model. Furthermore, the I/O function will require a higher order polynomial description, which will mean more parameters for the SLD($w$lin) model. Therefore, we will restrict our further discussion of the double roex filter to the SLD model only.

3.5. Discussion

3.5.1. Filter shape

The thresholds estimated by the auditory filters shown in Fig. 5(f) are represented in Fig. 7 by three types of lines. In Fig. 7, solid lines show the estimated thresholds corresponding to the symmetric notch condition ("○"). Dotted lines and dashed lines show the estimated thresholds corresponding to the asymmetric notch conditions. The left-pointing triangles "◁" on the dotted lines denote conditions of which the notch was skewed towards a lower frequency, and the right-pointing triangles "▷" on the dashed lines denote conditions of which the notch was skewed towards a higher frequency.

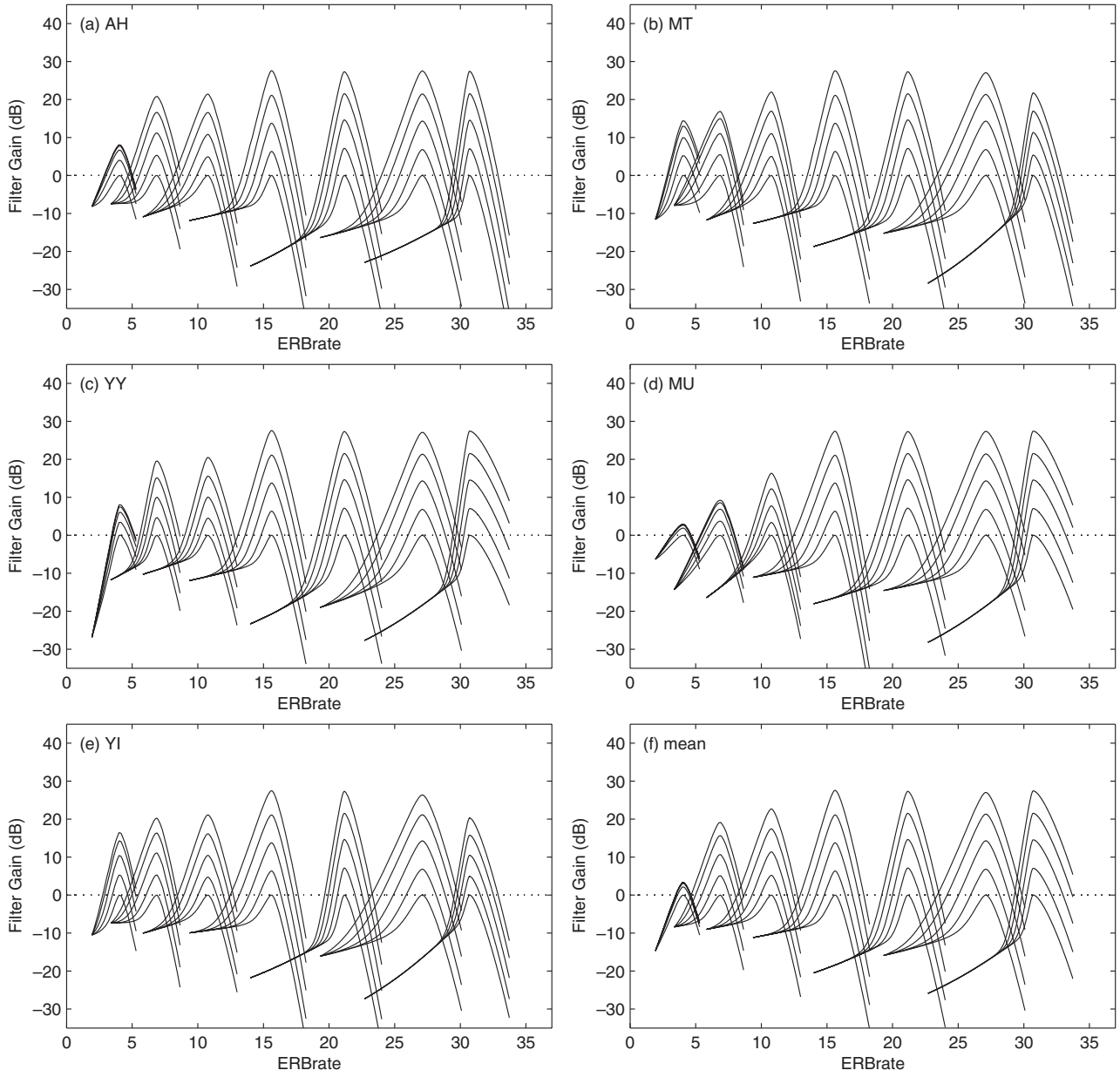

Fig. 5 Families of the double roex filter, roex(p, w, t), at all signal frequencies f_c = 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz: (a) subject AH, (b) subject MT, (c) subject YY, (d) subject MU, (e) subject YI, (f) mean. All seven curves from top to bottom show how the functions changed as the signal level increased from 30 to 70 dB in 10-dB steps.

Table 1 Filter/non-filter coefficients of the parameters and rms error value in the individual fit for each signal frequency.

f_c (Hz)	p_l	t	G_{max} (dB)	p_u	Abs (dB)	K (dB)	rms (dB)
125	10.6	5.6	44.4	10.0	27.8	2.7	0.94
250	16.2	1.6	43.9	16.2	17.2	0.3	1.35
500	21.3	2.6	47.1	21.3	12.0	−2.9	1.36
1,000	28.6	3.2	59.5	26.4	7.8	−1.4	1.57
2,000	40.3	6.2	58.8	26.1	4.9	−0.7	1.36
4,000	24.5	4.8	58.3	23.3	10.5	0.7	1.76
6,000	58.7	9.0	58.3	17.9	5.8	0.7	2.08

Table 2 Filter/non-filter coefficients of the parameters in the SLD($w$lin) model and rms error value in the individual fit for each signal frequency when using $w_{dB} = w_{dB}^{(0)} + w_{dB}^{(1)} \cdot L$ instead of G_{dB}.

f_c (Hz)	p_l	t	$w_{dB}^{(0)}$	$w_{dB}^{(1)}$	p_u	Abs (dB)	K (dB)	rms (dB)
125	10.6	5.2	−48.0	0.65	10.0	27.1	2.7	1.25
250	16.4	1.8	−36.8	0.30	16.5	17.2	0.9	1.32
500	24.9	9.0	−25.1	0.31	21.3	12.0	−2.5	1.36
1,000	32.4	8.7	−36.2	0.37	26.4	7.6	0.2	1.54
2,000	45.9	11.8	−36.7	0.47	26.1	4.7	−0.7	1.23
4,000	31.7	13.8	−24.3	0.40	23.3	10.8	0.6	1.51
6,000	63.5	13.0	−32.3	0.44	21.9	6.0	0.5	1.42

When masker level N_0 was high, the levels of the left-pointing triangle ("◁") were significantly lower than the levels of the right-pointing triangles ("▷"), indicating that the auditory filters were asymmetrical with a steeper high frequency slope. When masker level N_0 was low, the levels of the left-pointing triangles ("◁") are almost the same as the levels of the right-pointing triangles ("▷") indicating that the auditory filters are more symmetrical. However, at $f_c = 6{,}000$ Hz, the levels of the right-pointing triangles ("▷") were higher than the levels of the left-pointing triangles ("◁"), indicating that the auditory filter were asymmetrical with a steeper low frequency slope.

The auditory filters were well fitted to the results with small rms errors. As expected from the above observations, the shape of the fitted auditory filters in Fig. 5 was found to vary with the signal levels at all center frequencies; asymmetrical at higher levels and symmetrical at lower levels. The values of parameter t (Table 1) at lower frequencies were lower than those values presented in Table 2 of [19]. Hence, the low frequency slopes of the auditory filters centered at 250, 500, and 1,000 Hz were shallower than those corresponding results presented in [19]. Apart from these few differences, the auditory filter shapes obtained in this study were similar to those reported by Glasberg and Moore [19].

3.5.2. Filter bandwidth

Equivalent rectangular bandwidth (ERB) is a commonly used measure for evaluating the tuning of an auditory filter. There are two methods of calculating ERB: the direct-calculation method, which measures an equivalent rectangular bandwidth from the auditory filter shape in Eq. (2), and the approximation method which assumes the level of the signal is small and G_{lin} is large. In Fig. 6, the ERB values of the auditory filters were directly calculated from filter shapes of the auditory filter roex(p, w, t) as shown in Fig. 5(f), and plotted as a function of signal frequency at three signal levels of 50, 60, and 70 dB SPL. This figure shows that the ERB increased when the signal level increased, as was similarly reported from previous studies [18,19,22]. In the same figure, the ERBs of the auditory filters derived at moderate signal levels [4] were plotted as dotted lines with asterisks for comparison. The ERBs of the auditory filters at signal level of 70 dB SPL were wider than the ERBs of the auditory filters at stimulus levels of 50 and 60 dB SPL, indicating that the auditory filter had a finer tuning at lower signal levels. However, the ERBs at 50 and 60 dB SPL were still wider than the ERBs calculated in [4]. The ERBs at 30 and 40 dB SPL were almost the same as at 50 dB SPL, thus they are omitted for clarity.

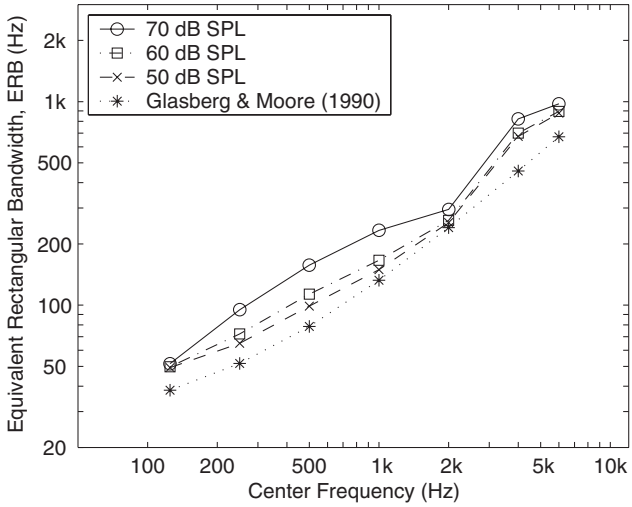

Fig. 6 Equivalent rectangular bandwidth, ERB, of the double roex filter as a function of signal frequency on a log-frequency scale. ERBs of the filters are drawn with dashed, dot-dashed, and solid lines as a parameter is the signal level (50, 60, and 70 dB SPL, respectively). ERBs at 30 and 40 dB SPL were the same as that at 50 dB SPL so they are omitted for clarity. The dotted line shows the ERB of Glasberg and Moore (1990).

Alternatively, the ERBs for the auditory filters centered at frequencies of 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz, were approximated by $(2/p_{\mathrm{l}} + 2/p_{\mathrm{u}})f_c$ [11] and are tabulated in Table 3 together with the equivalent values calculated in the previous studies. The first row of Table 3 shows the ERBs of the roex filters in Glasberg and Moore (1990) [4]. The second and third rows show the ERBs of the roex filters of Glasberg and Moore [19] and Baker *et al.* [18]. The last row shows the ERBs of the double roex filter obtained in this study. The values obtained in this study were greater than the values obtained in [4]; the ERBs at the lower signal level were approximately 1.2 times the ERBs in [4]. As shown, the approximations agreed with the results of Fig. 6. The ERBs calculated by both methods agreed, suggesting that the widening of the auditory filters derived from the simultaneous masking experiment might

Table 3 Bandwidths of the roex filter at the lower level. Equivalent rectangular bandwidths were approximately calculated using $(2/p_{\mathrm{l}} + 2/p_{\mathrm{u}})f_c$.

Signal frequency, f_c (Hz)	125	250	500	1,000	2,000	3,000	4,000	6,000
Roex (Glasberg and Moore, 1990)	38	52	79	132	240	349	456	672
Roex (Glasberg and Moore, 2000)	—	67	94	157	242	—	548	—
Roex (Baker *et al.*, 1998)	—	55	84	118	263	413	452	717
Roex (present study)	49	62	94	146	252	—	670	875

Table 4 Slope values for the input/output functions of the roex filter for signal frequencies from 125 Hz to 6,000 Hz.

Signal frequency, f_c (Hz)	125	250	500	1,000	2,000	3,000	4,000	6,000
Roex (Glasberg and Moore, 2000)	—	0.73	0.70	0.39	0.56	—	0.57	—
Roex (Baker *et al.*, 1998)	—	0.51	0.50	0.45	0.44	0.37	0.39	0.36
Compressive GC (Patterson *et al.*, 2003)	—	0.61	0.51	0.43	0.39	0.38	0.37	0.37
Roex (present study)	0.91	0.52	0.43	0.32	0.32	—	0.33	0.32

be complicated by the issue of suppression.

3.5.3. Slope of the I/O function

Physiologically, compression can be described in terms of the slope of the input-output (I/O) function of the basilar membrane [24]. The analogous measure in psychophysical term is the I/O function of the auditory filter. The results of previous studies [18,19,22] are included in Table 4 for the purpose of comparison. The general trend of the results of the previous studies shows that the slope of the I/O functions of the filter decreases when the center frequencies increase from 250 Hz to 1,000 Hz; the range of decrease is approximately from 0.6 dB/dB to 0.4 dB/dB. For filters centered above the frequency of 1,000 Hz, the slopes of the I/O functions remain at 0.4 dB/dB. This suggests that the filters centered at higher frequency are more compressive than the filters centered at lower frequency.

The last row of Table 4 shows the slope of the I/O function of the auditory filter as shown in Fig. 5(f). The slope of the I/O function of the auditory filter was calculated by dividing the gain difference of the filter at two input levels of 30 dB and 70 dB, with an input range of 40 dB. The maximum gain of the filter increased from about 4 dB at 125 Hz to about 27 dB at 1,000 Hz, and thereafter remained at about 27 dB, as was also observed in the previous studies [18,19,22].

In general, the patterns of the slope of the I/O function across signal frequency were similar across all four studies. Similarly, the present results show that the slope of the I/O function of the filter decreases when the center frequency increases from 125 Hz to 1,000 Hz; the range of decrease was approximately from 0.9 dB/dB to about 0.3 dB/dB. For filters centered above the frequency of 1,000 Hz, the slopes of the I/O functions remained at about 0.3 dB/dB. Extrapolating from the general trend of the other studies, the expected slope of the I/O function of the filter centered at 125 Hz is greater than 0.6 dB/dB. The present results show that this filter centered at 125 Hz is less compressive than the other filters with a slope of 0.9 dB/dB.

3.6. Further Work

On the whole, the results of this study and previous studies [18,19] agreed and showed that auditory filters centered at low frequencies between 125 and 1,000 Hz are less compressive when derived through a simultaneous notched-noise masking experiment. The corresponding auditory filters derived in a non-simultaneous notched-noise masking experiment were more compressive [5]. Furthermore, the compression ratios measured in a physiological study of the basilar membrane [24] were higher than the corresponding values measured in these simultaneous masking studies. According to Glasberg and Moore (2000), the discrepancy is partially due to the nonlinearity in the cochlea.

In general, the roex auditory filter fit excellently with the masked thresholds measured in the notched-noise masking experiments. However, the roex auditory filter described its filter characteristics in the lower and the upper sides of the frequency region separately, and with its shape defined by either one or both of the tip and tail filters. Therefore, the roex auditory filter would not have an impulse response realized as a time-domain auditory filter, which would be a disadvantage in modeling the nonlinear process of the cochlea. To capture the influence of cochlear nonlinearity in auditory filter modeling, Irino and Patterson [25] proposed a compressive gammachirp auditory filter. The compressive gammachirp filter has a well defined impulse response and can be easily realized in the time domain. In their subsequent work [22], they demonstrated that the compressive gammachirp auditory filter could account for the influence of nonlinearity on the auditory filter. However, the compressive gammachirp auditory filter would require data from both simultaneous and non-simultaneous experiments, which would be compatible with our overall objective.

In the immediate future, we will measure the masked threshold in a forward notched-noise masking experiment and compare the aspects of filter shape, filter bandwidth, and compression ratio with the present data. Ultimately, we will fit the compressive gammachirp auditory filter with the data in both simultaneous and non-simultaneous masking experiments for a more comprehensive study on nonlinear auditory filter modeling.

4. SUMMARY

In this paper, we performed a simultaneous masking experiment as a step towards establishing a common platform for both simultaneous and non-simultaneous masking experiments. Masked thresholds were measured for five normal-hearing listeners in a simultaneous masking experiment using a noise masker with a varying spectral

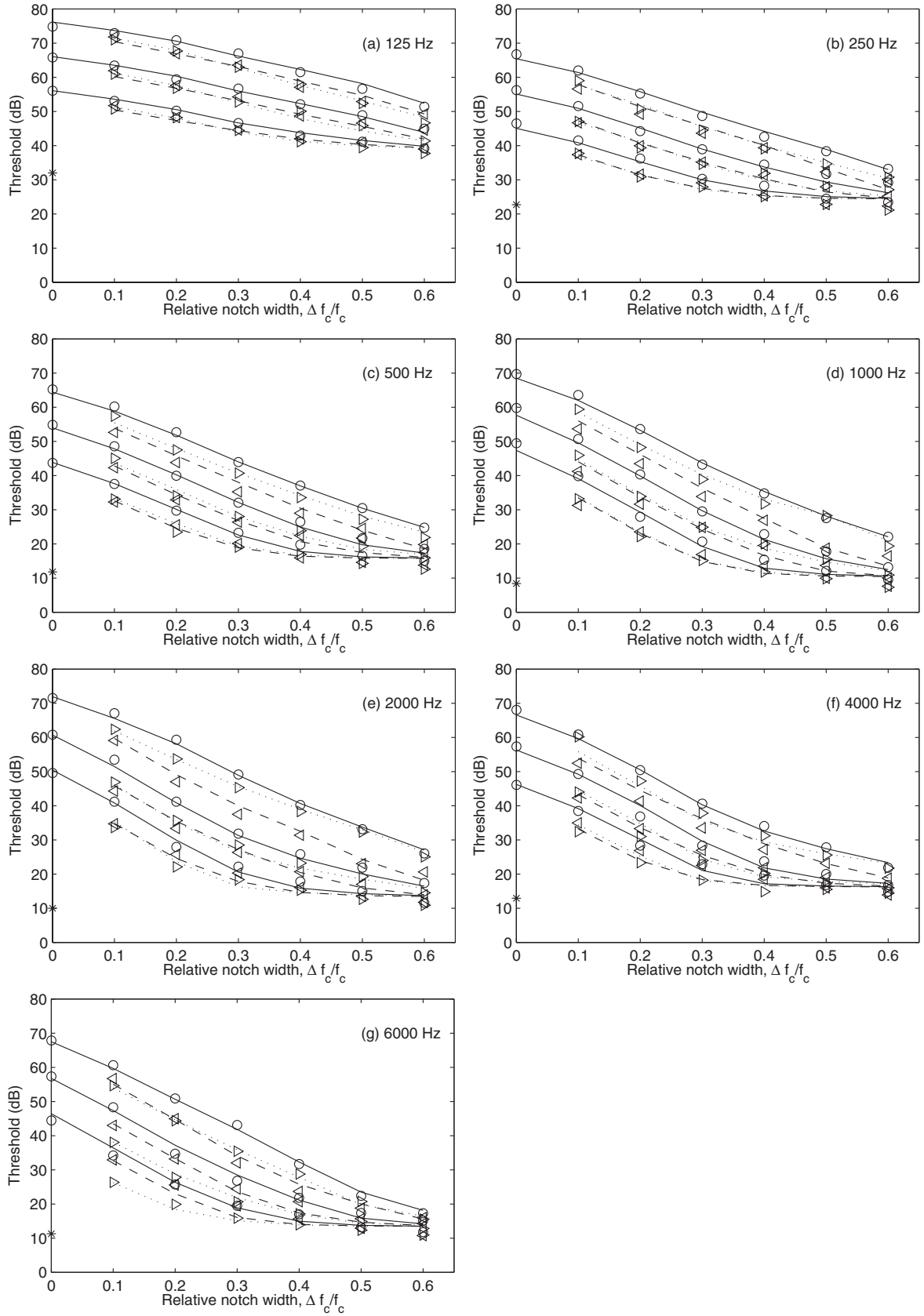

Fig. 7 Signal level at mean masked thresholds (dB SPL) in the notched-noise masking method measured at (a) f_c = 125 Hz, (b) 250 Hz, (c) 500 Hz, (d) 1,000 Hz, (e) 2,000 Hz, (f) 4,000 Hz, and (g) 6,000 Hz. Symbols ○, ◁, and ▷ show the masked threshold under the symmetrical condition, the asymmetrical condition on the lower side, and the asymmetrical condition on the upper side, respectively. Solid, dashed, and dotted lines show the thresholds estimated by using a double roex filter under the symmetrical condition (○), and asymmetrical conditions (◁ and ▷). "*" shows the averaged absolute threshold of the subjects.

notch, centered at signal frequencies of 125, 250, 500, 1,000, 2,000, 4,000, and 6,000 Hz. Basically, the conditions and procedure of the measurement were the same as in Glasberg and Moore [19]. Two additional measurements at signal frequencies of 125 and 6,000 Hz were made. UEN was not used. The double roex filter, roex(p, w, t), was fitted to the measured data to estimate the auditory filter shapes. This filter was modeled as the sum of the tip filter and tail filter, and the gain of the tip filter was assumed to be a function of the signal level. The fitting procedure was also the same as in [19].

Most of the results in this study were consistent with the previous results [18,19]. The following points summarized our results:

(1) The patterns of the masked threshold measured at the seven signal frequencies were similar across the subjects. The pattern of the mean masked threshold was almost the same as that in a previous study [19].
(2) The double roex auditory filter, roex(p, w, t), fit well with the measured data. The shapes of the fitted auditory filters for each subject were almost the same with little individual difference.
(3) In general, the shapes of the auditory filter centered at different frequencies do not vary much. However, the shape of each individual auditory filter did vary with the signal level. The bandwidth of the auditory filter widened as the signal level increased.
(4) The slope of the I/O function of the auditory filter decreased from 0.9 dB/dB to about 0.3 dB/dB when the center frequency increased from 125 Hz to 1,000 Hz. For filters centered above the frequency of 1,000 Hz, the slopes of the I/O function remained at about 0.3 dB/dB.

ACKNOWLEDGEMENTS

We thank Brian Moore and Brian Glasberg for their help in setting up this notched-noise masking experiment. We also thank Roy D. Patterson and Masato Akagi for their helpful comments. This work was supported by special coordination funds for promoting science and technology (supporting young researchers with fixed-term appointments).

REFERENCES

[1] H. Fletcher, "Auditory patterns," *Rev. Mod. Phys.*, **12**, 47–61 (1940).
[2] R. D. Patterson and B. C. J. Moore, "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, B. C. J. Moore, Ed. (Academic, London, 1986), pp. 123–177.
[3] B. R. Glasberg and B. C. J. Moore, "Auditory filter shapes in forward masking as a function of level," *J. Acoust. Soc. Am.*, **71**, 946–949 (1999).
[4] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, **47**, 103–138 (1990).
[5] E. A. Lopez-Poveda, C. J. Plack and R. Meddis, "Cochlear nonlinearity between 500 and 8,000 Hz in listeners with normal hearing," *J. Acoust. Soc. Am.*, **113**, 951–960 (2003).
[6] B. C. J. Moore and B. R. Glasberg, "Psychophysical tuning curves measured in simultaneous and forward masking," *J. Acoust. Soc. Am.*, **63**, 524–523 (1978).
[7] B. C. J. Moore and B. R. Glasberg, "Auditory filter shapes derived in simultaneous and forward masking," *J. Acoust. Soc. Am.*, **70**, 1003–1014 (1981).
[8] B. C. J. Moore, R. W. Peters and B. R. Glasberg, "Auditory filter shapes at low center frequencies," *J. Acoust. Soc. Am.*, **88**, 132–140 (1990).
[9] A. J. Oxenham and C. A. Shera, "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," *J. Assoc. Res. Otolaryngol.*, **4**, 541–554 (2003).
[10] R. D. Patterson, "Auditory filter shapes derived with noise stimuli," *J. Acoust. Soc. Am.*, **59**, 640–654 (1976).
[11] R. D. Patterson and I. Nimmo-Smith, "Off-frequency listening and auditory-filter asymmetry," *J. Acoust. Soc. Am.*, **67**, 229–245 (1980).
[12] C. J. Plack, A. J. Oxenham and V. Drga, "Linear and nonlinear processes in temporal masking," *Acustica*, **88**, 348–358 (2002).
[13] S. Rosen and D. Stock, "Auditory filter bandwidths as a function of level at low frequencies," *J. Acoust. Soc. Am.*, **92**, 773–781 (1992).
[14] S. Rosen and R. J. Baker, "Characterising auditory filter nonlinearity," *Hear. Res.*, **73**, 231–243 (1994).
[15] S. Rosen, R. J. Baker and A. M. Darling, "Auditory filter nonlinearity at 2 kHz in normal hearing listeners," *J. Acoust. Soc. Am.*, **103**, 2539–2550 (1998).
[16] M. J. Shailer, B. C. J. Moore, B. R. Glasberg, N. Watson and S. Harris, "Auditory filter shapes at 8 and 10 kHz," *J. Acoust. Soc. Am.*, **88**, 141–148 (1990).
[17] B. A. Wright, "Auditory filter asymmetry at 2000 Hz in 80 normal-hearing ears," *J. Acoust. Soc. Am.*, **100**, 1717–1721 (1996).
[18] R. J. Baker, S. Rosen and A. M. Darling, "An efficient characterisation of human auditory filtering across level and frequency that is also physiologically reasonable," in *Psychophysical and Physiological Advances in Hearing: Proc. ISH98*, A. Palmer, A. Rees, Q. Summerfield and R. Meddis, Eds. (Whurr, London, 1998), pp. 81–88.
[19] B. R. Glasberg and B. C. J. Moore, "Frequency selectivity as a function of level and frequency measured with uniformly exciting noise," *J. Acoust. Soc. Am.*, **108**, 2318–2328 (2000).
[20] H. Levitt, "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.*, **49**, 467–477 (1970).
[21] B. R. Glasberg, B. C. J. Moore and M. A. Stone, "Modelling changes in frequency selectivity with level," in *Psychophysics, Physiology and Models of Hearing*, T. Dau, V. Hohmann and B. Kollmeier, Eds. (World Scientific, Singapore, 1999).
[22] R. D. Patterson, M. Unoki and T. Irino, "Extending the domain of center frequencies for the compressive gammachirp auditory filter," *J. Acoust. Soc. Am.*, **114**, 1529–1542 (2003).
[23] B. C. J. Moore, B. R. Glasberg and T. Bear, "A model for the prediction of thresholds, loudness and partial loudness," *J. Audio Eng. Soc.*, **45**, 224–240 (1997).
[24] S. P. Bacon, R. R. Fay and A. N. Popper, *Compression, From Cochlea to Cochlear Implants* (Springer, New York, 2004).
[25] T. Irino and R. D. Patterson, "A compressive gammachirp auditory filter for both physiological and psychophysical data," *J. Acoust. Soc. Am.*, **109**, 2008–2022 (2001).