Detection of high-frequency spectral notches as a function of level[a)]

Ana Alves-Pinto and Enrique A. Lopez-Poveda[b)]
Unidad de Computación Auditiva y Psicoacústica: Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Avenida Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain

(Received 17 November 2004; revised 14 July 2005; accepted 18 July 2005)

High-frequency spectral notches are important cues for sound localization. Our ability to detect them must depend on their representation as auditory nerve (AN) rate profiles. Because of the low threshold and the narrow dynamic range of most AN fibers, these rate profiles deteriorate at high levels. The system may compensate by using onset rate profiles whose dynamic range is wider, or by using low-spontaneous-rate fibers, whose threshold is higher. To test these hypotheses, the threshold notch depth necessary to discriminate between a flat spectrum broadband noise and a similar noise with a spectral notch centered at 8 kHz was measured at levels from 32 to 100 dB SPL. The importance of the onset rate-profile representation of the notch was estimated by varying the stimulus duration and its rise time. For a large proportion of listeners, threshold notch depth varied *nonmonotonically* with level, increasing for levels up to 70–80 dB SPL and decreasing thereafter. The nonmonotonic aspect of the function was independent of notch bandwidth and stimulus duration. Thresholds were independent of stimulus rise time but increased for the shorter noise bursts. Results are discussed in terms of the ability of the AN to convey spectral notch information at different levels. © *2005 Acoustical Society of America.* [DOI: 10.1121/1.2032067]

PACS number(s): 43.66.Ba, 43.66.Fe, 43.66.Qp [GDK] Pages: 2458–2469

I. INTRODUCTION

High-frequency spectral notches generated by the filtering action of the pinna (Lopez-Poveda and Meddis, 1996; Shaw, 1974; Shaw and Teranishi, 1968) are important cues for sound localization (Butler and Belendiuk, 1977; Butler and Humanski, 1992; Hebrank and Wright, 1974). Particularly important is the notch whose center frequency increases gradually from around 6500 to 10000 Hz as the vertical location of the sound source moves from −40° to +60° relative to the horizontal plane (for a review see Lopez-Poveda, 1996). The bandwidth of this notch at its 5-dB-down points ranges from 1000 Hz at −40° elevation to 4000 Hz at +10° elevation (c f. Shaw and Teranishi, 1968; Chap. 4 in Lopez-Poveda, 1996).

The auditory nerve (AN) is the only transmission path of acoustic information to the central auditory system. Therefore, the ability to detect and employ high-frequency spectral notches for sound localization requires that they be adequately represented in the AN response. In the AN, spectral features can be represented in the temporal pattern of AN fiber response across characteristic frequencies (CFs) or as rate profiles; that is, in terms of the fibers' discharge rate as a function of their CF (Sachs and Young, 1979). However, spectral features with frequencies above the cutoff of phase locking [>4000 Hz; Johnson (1980); Palmer and Russell (1986)] are most likely to be represented in terms of AN discharge rate alone (Rice *et al.*, 1995).

AN fibers have been classified in either two (Evans and Palmer, 1980; Kiang *et al.*, 1965) or three (Liberman, 1978; Winter *et al.*, 1990) types on the basis of their spontaneous rate. Fibers with high spontaneous rates (HSR) (>15 spikes/s) amount to approximately 61% of the population. These have low thresholds (<10 dB SPL) and dynamic ranges of approximately 30–40 dB (Sachs and Abbas, 1974; Evans and Palmer, 1980). The remaining 40% of the fibers are of a medium- (MSR) or low-spontaneous rate (LSR) type. These have higher thresholds (>15 dB SPL) and wider dynamic ranges (~50–60 dB; Evans and Palmer, 1980; Sachs and Abbas, 1974).

The existence of at least two fiber types with distinct thresholds and dynamic ranges have led several investigators (Delgutte and Kiang, 1984a; Rice *et al.*, 1995; Sachs and Young, 1979) to suggest that the *high-frequency* spectral characteristics of a stimulus could be conveyed to the central auditory system in the rate profile of HSR and LSR fibers at low and high levels, respectively. However, the apparent quality[1] of rate profiles degrades as the stimulus level is increased even when the rate profile of LSR fibers is considered separately from that of HSR fibers (Delgutte and Kiang, 1984a, b; Rice *et al.*, 1995; Sachs and Young, 1979). Among the factors that may contribute to the deterioration of the quality of the rate profiles at high levels are the broadening of the fibers frequency response with level (Rose *et al.*, 1971), the saturation of their discharge rate[2] (Sachs and Young, 1979), and the fiber-to-fiber variation in rate (Rice *et al.*, 1995).

Therefore, if the perception of high-frequency spectral features is based on analyzing the shape of AN rate profiles, the detection of high-frequency spectral notches should be-

[a)]Portions of this work were presented at the 27th Midwinter Meeting of the Association of Research in Otolaryngology, Daytona Beach, FL, February 2004.
[b)]Author to whom correspondence should be addressed; electronic mail: ealopezpoveda@usal.es

come increasingly more difficult as the sound level increases. The aim of this study is to test this hypothesis. In the main experiment, the threshold notch depth necessary to discriminate between a flat spectrum broadband noise and a similar noise with a spectral notch centered at 8 kHz was measured at levels from 32 to 100 dB SPL (Fig. 1). It will be shown that, contrary to the above-noted hypothesis, the threshold notch depth for discrimination is a *nonmonotonic* function of stimulation level.

In addition to the role of LSR fibers, other mechanisms have been proposed to explain how high-frequency spectral features are encoded in the AN response at high levels. For instance, Delgutte and Kiang (1984a, b) suggested that at high levels, when the adapted (steady-state) response of many of the fibers is saturated, the stimulus spectrum may still be represented reasonably well in the *onset* rate profiles. Their suggestion is based on the fact that AN fibers have a wider dynamic range over the first few (5–20) milliseconds of their response (Smith and Brachman, 1980).

Direct evidence is lacking (to the present authors' knowledge) that the dynamic range of AN fibers gets wider the shorter the stimulus rise time. However, this is likely to be the case as the maximum onset rate of AN fibers depends on the stimulus rise time [see Fig. 3 in Delgutte (1980)] and models of AN adaptation suggest that the level at which the maximum onset rate occurs gets higher the shorter the rise time [e.g., Fig. 2 in Meddis (1988)]. Therefore, following Delgutte and Kiang's suggestion (1984a, b), it is reasonable to hypothesize that the AN rate-profile representation of the stimulus spectrum should be clearer for sounds with abrupt onsets, particularly at high stimulus levels. Moreover, if at high levels an important proportion of the spectral information were indeed conveyed in the onset AN rate profile, spectral shape discrimination should hardly depend on the stimulus duration.

These hypotheses were tested by repeating the above-presented experiment but using different noise burst durations (20 vs 220 ms) and rise/fall times (2, 10, and 30 ms). It will be shown that contrary to the hypotheses, threshold notch depth is hardly affected by the stimulus rise time and depends strongly on stimulus duration, even at high levels.

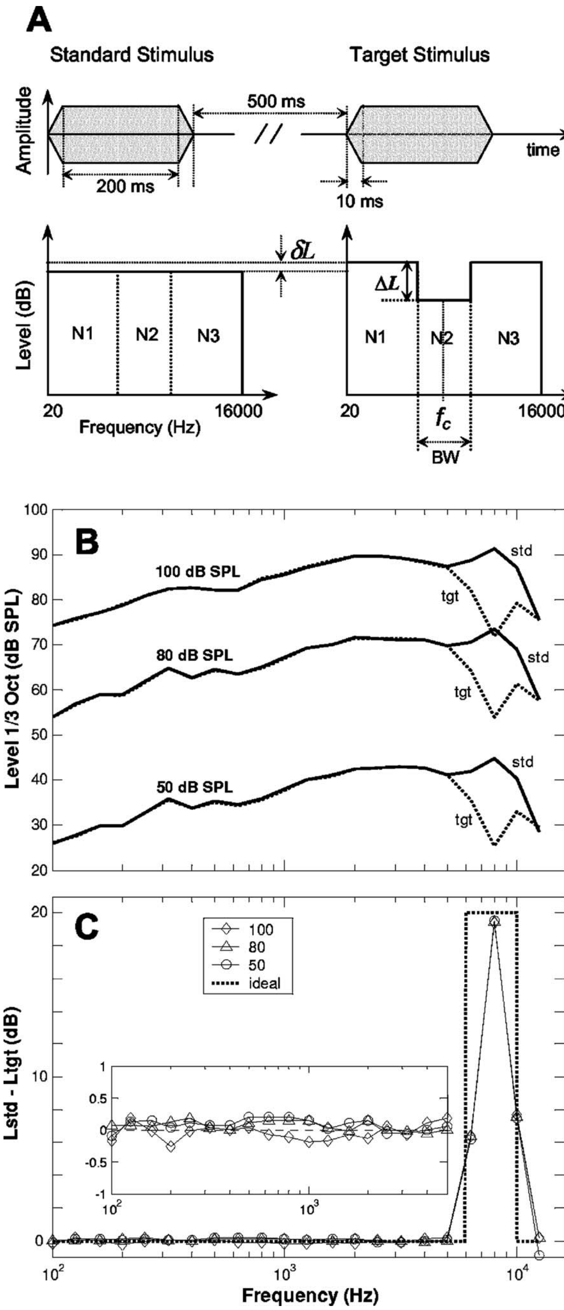

FIG. 1. (a) Wave forms and spectra of the standard and target noises. (b) Actual SPLs for the flat-spectrum (continuous lines) and the notched (dotted lines) noises delivered by the Etymotic ER2 transducer (see footnote 4).[4] The notched noise contained a 4000-Hz-wide, 20-dB-deep spectral notch. The reference spectrum level was the same for the flat-spectrum and the notched noises, and such that the overall SPL for the flat-spectrum noise was equal to 50, 80, and 100 dB SPL (indicated by the number next to each line). (c) SPL differences between the flat-spectrum and the notched noises for stimuli with overall levels of 50, 80, 100 dB SPL. The inset provides greater detail on the differences. The thick dotted line illustrates an ideal difference spectrum. Note that the actual difference profiles are almost identical at the three overall levels tested. This demonstrates that the cues available for the discrimination are identical at overall levels of 50, 80, and 100 dB SPL, at least for frequencies up to 12500 Hz.

II. METHODS

A. Procedure

Notch depths at threshold were measured using a three-interval, three-alternative forced-choice paradigm. In two of the intervals (standard intervals), the stimulus consisted of a noise burst with a flat spectrum from 20 to 16000 Hz. In the other (target) interval, the stimulus consisted of a burst of noise with a notch in its spectrum (Fig. 1). The three intervals were presented in random order to the listener who was instructed to detect the odd one out. The silence period between interval presentations was 500 ms.

The initial notch depth (ΔL in Fig. 1) was always fixed at 20 dB below the reference spectrum level of the noise. An adaptive procedure was employed to estimate the 70.7% correct point in the psychometric function (Levitt, 1971). During the adaptive procedure, notch depth decreased after two consecutive correct responses and increased after an incorrect response. Notch depth increased or decreased by 6 dB for the first four turn points and by 1 dB thereafter. Sixteen turn points were recorded for each measurement and the threshold estimate was taken as the mean of the notch depths for the last 12 turn points. When the corresponding standard deviation (SD) exceeded 6 dB, the measurement was discarded (invalid threshold) and a new threshold was obtained. The thresholds reported in this study correspond to the *geometric* mean of at least three valid measurements (see Sec. III A for more details). Sometimes detecting the notch was impossible (the adaptive procedure did not converge). Whenever this happened and to prevent endless sessions, the automatic adaptive paradigm was set to stop after 80 stimulus presentations.

Notch depths at threshold were measured as a function of the overall level of the stimulus. Overall levels ranged from 32 to 100 dB SPL (these correspond to spectrum levels of −10 to 58 dB SPL for the flat-spectrum noise). 10-dB steps were used for levels from 32 to 92 dB SPL. Sessions were organized in blocks of eight runs; each run for a single overall level. Within each block, levels were ordered quasirandomly, although the stimuli with the higher-levels (>70 dB SPL) were presented at the end of the block. This aimed at minimizing any effect that temporary threshold shifts may have on spectral discrimination of low-level sounds.

Listeners were tested individually in a double-wall sound-attenuating chamber. Interaction was provided by means of a response box. Lights were used to mark the presentation of the stimuli and to give listeners trial-by-trial feedback on their responses. Stimuli were delivered monaurally to the listener via Etymotic Research ER2 insert earphones. This earphone is designed to have a flat frequency response at the eardrum up to approximately 16000 Hz. The sound pressure levels (SPLs) reported in the following are calibrated values [for a flat-spectrum broadband noise (20 to 16000 Hz), the calibrated SPL output at a microphone coupled to a Zwislocki occluded ear simulator was 100 dB SPL for 2.3 V_{rms}].

B. Stimuli

The center frequency (f_c) of the spectral notch was fixed at 8000 Hz. The reference condition was with a notch bandwidth (BW) of 2000 Hz, and a stimulus duration of 220 ms, including 10-ms cosine-squared rise/fall ramps. The effect of notch bandwidth was investigated by testing notch BWs of 1000, 2000, and 4000 Hz. The effect of the stimulus duration was investigated by testing stimulus durations of 20 and 220 ms, both including 10-ms rise/fall times. The effect of the stimulus rise time was investigated by setting the stimulus duration at 220 ms and using rise/fall times of 2, 10, and 30 ms.

1. Noise generation

The standard and the target noise bursts were made by adding three separately generated noises with different narrowband flat spectra (N1, N2, and N3 in Fig. 1): N1 with frequency components from 20 to $f_c-0.5$ BW Hz; N2 with a spectrum centered at $f_c=8000$ Hz and a bandwidth equal to the notch bandwidth; and N3 with components ranging from $f_c+0.5$ BW to 16000 Hz.

Each of the three noises, N1, N2, and N3, were generated digitally in the time domain (sampling frequency =48 828 Hz) by adding sinusoids of equal amplitude but random phases uniformly distributed between 0 and 2π rad. The frequencies of the sinusoids spanned the spectral bandwidth of each of the noises in steps of 1 Hz. The wave forms of N1, N2, and N3 were then added together on a sample-by-sample basis in a TDT™ System 3 psychoacoustics workstation. The noise burst corresponding to a standard interval was generated by adding N1, N2, and N3, all with identical spectrum levels. The notched-noise burst (target interval), however, was generated by attenuating N2 by ΔL dB (the notch depth) as required (Fig. 1), and adding the resulting wave form to those of N1 and N3. Notch depths (ΔL in Fig. 1) are, therefore, expressed in dB relative to the spectrum level of N1 or N3.

Before delivering the stimuli, the overall level of the combined noise (N1+N2+N3) was set digitally by attenuating the signal as required. Signal clipping at high levels was avoided by ensuring that no sample of the combined digital noise reached the maximum output voltage of the TDT system (±10 V). This is an important point because the distortion associated with clipping could smear the spectrum of the notched noise in the target interval, hence making notch detection and the interpretation of the results more difficult. Furthermore, transducer distortion did not affect the results at high levels, as shown in Figs. 1(b) and 1(c), because the voltage required (2.3 V_{rms}) to produce the maximum level considered in this study (100 dB SPL) was within the nominal operational limits (2.5 V_{rms}) of the Etymotic ER2 transducer.

2. Strategies for making level differences unreliable discrimination cues

The presence of the notch in the target stimulus makes its overall level lower than that of the standard stimuli. For example, a level difference of 1.23 dB occurs for the broadest (4000-Hz-wide) notch considered and with the maximum notch depth measured (18.4 dB as shown in Fig. 2). Such a level difference could be used as a cue in a task aimed at measuring the listeners' ability to discriminate between spectral shapes. Hence, it could complicate the interpretation of the results.

Two different strategies were employed to reduce this unwanted effect. The first strategy consisted of presenting the standard and the target stimuli with *equal* overall levels. This was achieved by reducing the spectrum level of the standard noise as necessary (δL in Fig. 1) to make its overall level identical to that of the notched-noise (target) stimulus. It can be shown that the necessary reduction (in dB) is equal to

$$\delta L(\mathrm{dB}) = 10 \cdot \log_{10}\left[1 + \frac{\mathrm{BW}_{\mathrm{N2}}}{\mathrm{BW}_T}\left(10^{-\Delta L/10} - 1\right)\right], \qquad (1)$$

where $\mathrm{BW}_{\mathrm{N2}}$ is the bandwidth of noise N2 (Hz); BW_T is the total bandwidth (Hz) of the noise (15980 Hz); and ΔL is the notch depth (dB). To get an idea of the necessary reduction in spectrum level, δL equals −0.58 dB for a 2000-Hz-wide, 27-dB-deep notch. Notice that δL [Eq. (1)]

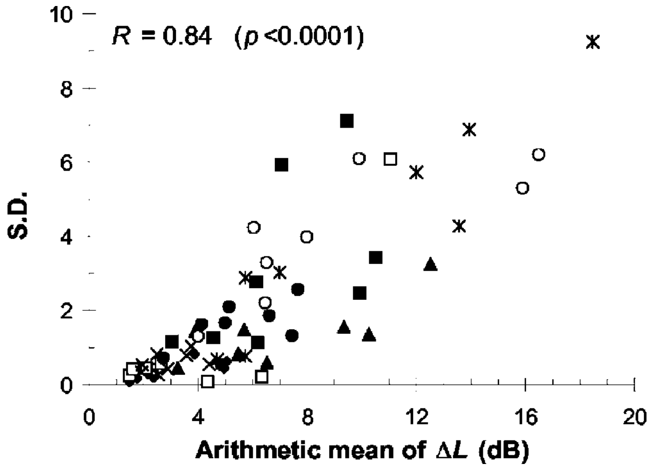

FIG. 2. Standard deviation of threshold notch depths (SDs) as a function of their associated arithmetic means. Data are pooled from Fig. 4 and thus are for eight listeners and the whole range of levels. Different symbols illustrate data for different listeners. The strong significant correlation (Pearson's $R = 0.84$, $p < 0.0001$) between the arithmetic mean and SD suggests that threshold notch depth does *not* conform to an "equal-variance" distribution.

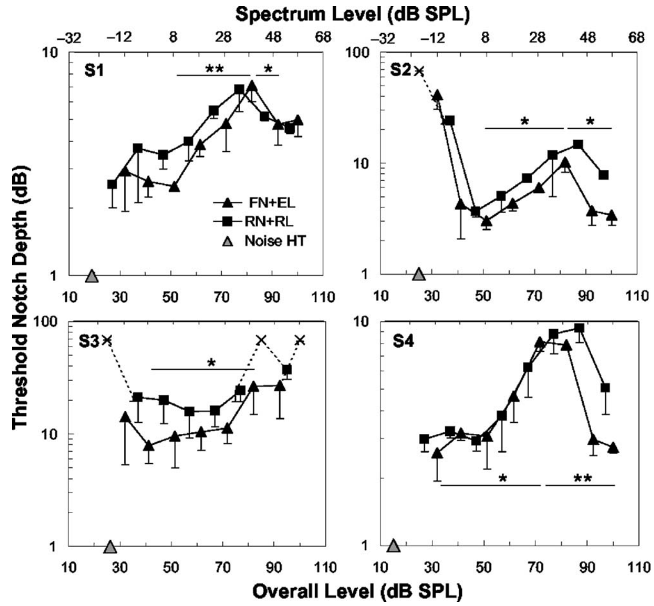

FIG. 3. Threshold notch depths for discriminating between a flat-spectrum noise and a notched noise plotted as a function of the stimulus overall level (the corresponding spectrum level scale is also shown at the top of the upper panels). The notches were centered at 8000 Hz and had a bandwidth of 2000 Hz. Each panel illustrates data for a single listener as indicated in the upper-left corner of each panel (note the different scales on the y axis of each panel). Each data point corresponds to the geometric mean of at least three measurements. Error bars illustrate the 68% confidence interval (see the text for details). Crosses indicate conditions for which no discrimination between the target and the standard stimuli was possible. Shaded triangles on the abscissa indicate the listener's absolute hearing threshold (HT) for the standard noise. Thresholds were measured in two experimental conditions: in the frozen noise/equalized level (FN-EL) condition and in the random noise/roving level (RN-RL) condition. Horizontal bars and asterisks apply to the FN-EL condition only, and denote statistically significant differences between the maximum threshold notch depth that occurs around 70–80 dB SPL and the lowest values that occur on each side of the maximum (single-tailed paired Student's t-test; one and two asterisk denote $p < 0.1$ and $p < 0.01$, respectively). The main result is the *nonmonotonic* variation of threshold notch depth with stimulus level observed for S1, S2, and S4 in the two conditions tested.

is also equal to the overall level difference between the standard, flat-spectrum noise of bandwidth BW_T and the target noise with a spectral notch of bandwidth BW_{N2} and depth ΔL.

Whenever this equalized-level strategy was employed, the same noise tokens N1, N2, and N3 were used for the standard and the target intervals and for every block of every condition. This condition will be hereafter referred to as *frozen-noised/equalized-level* (FN-EL).

The second strategy to prevent the use of level differences as discrimination cues consisted of presenting each of the intervals in a given trial at a different overall level (level roving). The idea is to make differences in overall level *and also* level differences within any given frequency band unreliable cues for detecting the spectral notch, forcing stimulus discrimination to be based solely on differences in overall spectral shape. This method of level randomization has been used in other notch detection experiments (Moore *et al.*, 1989) and is a common practice in spectral-shape discrimination tasks [for a review see Green (1988)]. Here, for any given trial, the overall level of the stimuli was randomly attenuated by an amount between 0 and 10 dB with respect to the corresponding reference level (uniform distribution, 0.25-dB steps). As a result of this, for any given reference spectrum level, the actual spectrum level for this condition was on average 5 dB lower. This will be taken into account when illustrating the results.

Even with a roving level, a notch that is deep enough could in principle be detected on the basis of comparing the overall level across intervals (cf., Green, 1988). According to Eq. (1), the overall level difference between the target and the standard stimuli increases asymptotically with increasing notch depth (ΔL in dB). Such a difference approaches 1.25 dB for the broadest (4000-Hz-wide) notch considered in the present study. Given that level was roved between 0 and 10 dB with a uniform distribution, this implies that spectral discrimination based on overall level comparisons can still occur despite the level roving, but only in 12.5% of the occasions and in the least favorable case; that is, for the broadest and deepest notch. In summary, the present level roving strategy guarantees that the listeners' responses be based on spectral shape discrimination in more than 87.5% of the occasions.

Whenever this level-roving strategy was employed, different (*random*) N1, N2, and N3 noise tokens were employed for different intervals and for different conditions. This was intended to investigate to what extent the use of stimuli with different temporal structures makes notch detection more difficult relative to the FN-EL condition. This condition will be hereafter referred to as *random-noise/roving-level* (RN-RL).

No significant differences were observed in the results obtained with either strategy (see Fig. 3 and Sec. III B). For this reason and for convenience, the frozen-noise/equalized-level strategy was more generally used in the experiments reported in the following.

C. Listeners

Data were collected for eight listeners with ages ranging from 20 to 40 years, although some of the listeners were not tested in all conditions. Their absolute thresholds were within 20 dB re. ANSI 3.6-1996 (Specifications for Audiometers) at the audiometric frequencies (250–8000 Hz). All listeners were given at least one training session in the task. Only S1 (author A.A.P.) had previous experience in psychoacoustic tasks.

III. RESULTS

A. The statistical distribution of notch depth

Figure 2 illustrates that there is a strong correlation ($R = 0.84, p < 0.0001$) between the *arithmetic* mean and the corresponding SD of threshold notch depth. Furthermore, threshold notch depths are limited to values $\geqslant 0$ dB. These properties demonstrate that threshold notch depth does *not* conform to an "equal-variance" distribution (cf., Bland and Altman, 1996a). Indeed, the data in Fig. 2 conform to a *lognormal* distribution with a probability of $p = 0.94$ (two-tailed Kolmogorov-Smirnoff test). For this reason, the results are hereafter illustrated as the *geometric* mean plotted on a *logarithmic* scale (Bland and Altman, 1996a, b, c). Furthermore, the variability of the results is illustrated as the 68% confidence interval of the geometric mean (equivalent to the more conventional arithmetic mean plus and minus one SD) (Bland and Altman 1996b, c, d).

This procedure is similar to and inspired by a previous level-discrimination study (Buus and Florentine, 1991) where it is suggested that level-discrimination thresholds are appropriately represented by plotting the difference limen for a change in sound level on a logarithmic scale. In that case, the argument in favor of this procedure was that this type of representation "provides a straightforward relation between measurements of discrimination thresholds and the sensitivity of the auditory system" as estimated in terms of *d*-prime (Buus and Florentine, 1991, p. 1379 and their Fig. 7).

B. Discrimination is based on detecting differences in overall spectral shape

Figure 3 illustrates threshold notch depth versus level functions for four representative listeners (S1–S4) for a notch bandwidth of 2000 Hz. Each panel illustrates data for a different listener (as indicated in the upper left corner of the panels). Shaded triangles on the abscissa illustrate the listener's absolute threshold for the flat-spectrum (standard) noise. Crosses indicate conditions for which the notch was undetected consistently in three or more trials. Note that the *y*-axis scale differs across panels.

Threshold notch depth values for the random-noise/roving-level condition (squares; RN-RL) are comparable or slightly larger than those for the frozen noise/equalized-level condition (triangles, FN-EL). Except for listener S3, the poorest performer, the main differences between the results for the two conditions occur at the highest overall levels. For the most part, these differences may be attributed to roving the level rather than to using random noise. If the use of frozen noise facilitated notch detection as a result of the listeners memorizing its temporal pattern, depths would be considerably lower at *all* levels for the FN-EL condition (triangles in Fig. 3) and this is not the case. Furthermore, all participants reported that level randomization was an important distracter, particularly at high levels. This agrees with the observations of Moore *et al.* (1989), who showed a deterioration in the detection of notches centered at 8000 Hz with level randomization. Moore *et al.* attributed this deterioration to a reduction of the listener's attention towards the fainter spectral features when more prominent perceptual changes are introduced.

Given that the observed differences are small and that randomizing the level for each stimulus presentation makes energy within or outside the notch band an unreliable cue for stimulus discrimination, it can be reasonably concluded that even in the frozen-noise/equalized-level condition listeners discriminate between the target and the standard stimuli by detecting a notch in a flat reference spectrum rather than by monitoring the level over certain frequency regions (i.e., in the notch band or elsewhere).

C. Notch depth increases *nonmonotonically* with level

Figure 3 illustrates that for three of the four listeners (S1, S2, and S4), threshold notch depth increases *nonmonotonically* with level. Consider, for example, the frozen-noise/equalized-level condition (triangles). Depth values increase from approximately 3 dB at 40 dB SPL to approximately 9 dB at 70–80 dB SPL and then decrease again to 3 dB (S2 and S4) or 5 dB (S1). For these three listeners, the differences between the largest and the smallest threshold notch depth values are statistically significant, as indicated by the horizontal bars and their associated asterisks in Fig. 3 [single-tailed paired Student's *t*-test on the ln-transformed data (Bland and Altman, 1996d); one asterisk: $p < 0.1$; two asterisks: $p < 0.01$]. The other listener (S3) was the "poorest" performer and reported that the task was very difficult. Her threshold depth values are larger across levels than those for the other three listeners. They increase from 8 dB at 42 dB SPL to 10 dB at 72 dB SPL, and then increase more rapidly with level until the notch becomes undetectable at 100 dB SPL (indicated by a cross). Also, the variability of her results is overall larger than those for the other three listeners. This may occur as a result of her having worse frequency resolution and/or her finding it more difficult to follow an appropriate spectral discrimination cue.

It is noteworthy that for listeners S2 and S3, threshold notch depth values are larger at the lowest overall level (32 dB SPL) and decrease rapidly for the next level tested (42 dB SPL). It is likely that this result reflects simply that 32 dB SPL is only 4 and 5 dB above the noise absolute threshold of listeners S2 and S3, while it is 11 and 15 dB above the absolute threshold of listeners S1 and S4, respectively.

To confirm that nonmonotonic threshold notch depth versus level functions are the norm, the experiment was carried out for eight listeners (S1–S8) for the frozen-noise/equalized-level condition. This time, however, the notch bandwidth was set to 4000 Hz because it is an easier condition. The results are illustrated in Fig. 4. Although the actual values vary considerably across listeners, notch depth increases nonmonotonically with level for *all* of them except S3, at least qualitatively. Quantitatively, however, the differ-

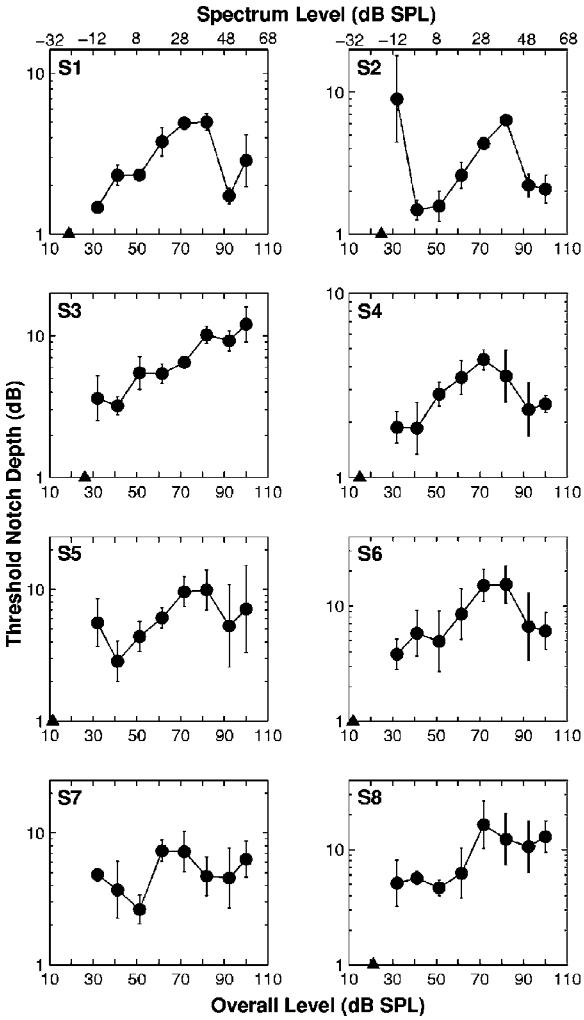

FIG. 4. Threshold notch depth vs. level functions for a spectral notch centered at 8000 Hz and a bandwidth of 4000 Hz. Except for the notch bandwidth, the experimental conditions were the same as for the FN-EL condition of Fig. 3. Each panel illustrates data for a single listener as indicated in the upper-left corner of each panel (note the different scales on the *y* axis of each panel). Error bars illustrate the 68% confidence interval of the geometric mean. Horizontal bars and asterisks denote statistically significant differences between the maximum threshold notch depth that occurs around 70–80 dB SPL and the lowest values that occur on each side of the maximum (single-tailed paired Student's *t*-test; one and two asterisk denote $p<0.1$ and $p<0.01$, respectively). Shaded triangles illustrate the hearing thresholds for the standard noise. For four of the eight listeners (S1, S2, S4, and S6), threshold notch depth is a clear nonmonotonic function of level. For three other listeners (S5, S7, and S8), the function is qualitatively nonmonotonic but the result is statistically not significant. S3 is the only listener for whom threshold notch depth increases monotonically with increasing level.

ence between the largest threshold notch depth value and the lowest one that occurs at a higher level is statistically significant for four (S1, S2, S4, and S6) out of the eight listeners only (single-tailed paired Student's *t*-test on the ln-transformed data; one asterisk: $p<0.1$; two asterisks: $p<0.01$). A maximum in the nonmonotonic functions occurs at approximately the same level (70–80 dB SPL) for *all* listeners except S3.

D. Notch depths are comparable at overall levels of 42 and 100 dB SPL

The quality of the rate-profile representation of the spectral notch is expected to deteriorate at high levels as a result of the broadening of the fibers' frequency response with level (Rose *et al.*, 1971), the saturation of their discharge rate (Sachs and Young, 1979), and/or the fiber-to-fiber variation in rate (Rice *et al.*, 1995). However, the nonmonotonic character of the threshold notch depth versus level functions suggests that the spectral notch is equally well represented in the AN response at 42 and 100 dB SPL. To investigate whether this is actually the case, threshold notch depths at 100 dB SPL were statistically compared with those at 42 dB SPL for each listener (single-tailed, paired Student's *t*-test on the ln-transformed data). Results are illustrated in Fig. 5 for the two notch bandwidths tested (2000 and 4000 Hz) and for the frozen-noise/equalized-level condition. Threshold notch depth values are larger at 100 dB SPL for approximately 50% of the listeners. However, statistically significant ($p<0.05$) differences are rare (denoted by asterisks in Fig. 5). The significance of this result is discussed later (Sec. IV A)

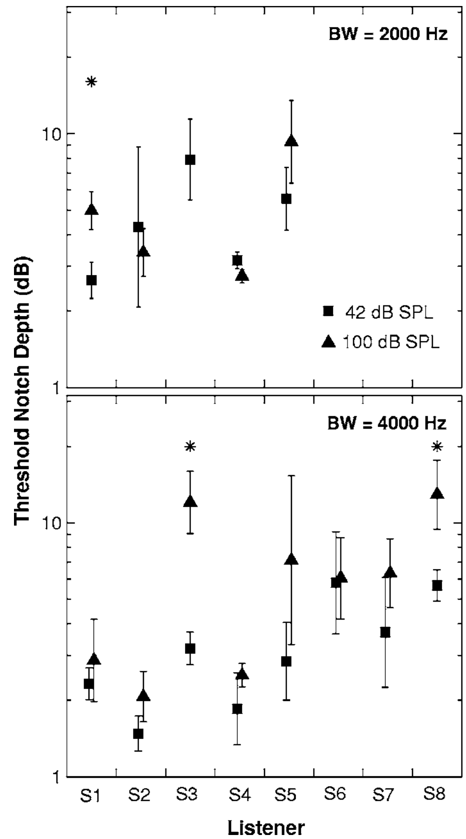

FIG. 5. Comparison of individual threshold notch depths at overall levels of 42 and 100 dB SPL, for notch bandwidths of 2000 (top panel) and 4000 Hz (bottom panel). To facilitate the comparison between the results at the two sound levels, data points for any given listener appear slightly displaced horizontally as necessary. Error bars indicate the 68% confidence interval of the geometric mean. Asterisks indicate significant differences between the thresholds at 42 and 100 dB SPL (single-tailed, paired Student's *t*-test; $p<0.05$). Although notch depths are generally larger at 100 dB SPL, statistically significant differences are rare.

E. The effect of notch bandwidth

Figure 6 illustrates threshold notch depth versus level functions for the frozen-noise/equalized-level condition for three notch bandwidths: 1000, 2000, and 4000 Hz. Crosses illustrate conditions for which notch detection became erratic and three valid measurements could not be obtained.

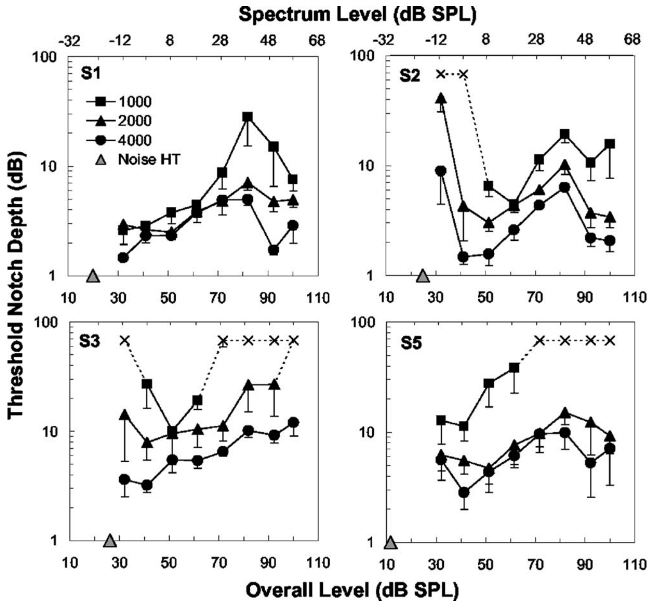

FIG. 6. Threshold notch depth vs. level functions for notch bandwidths of 1000, 2000, and 4000 Hz (as indicated by the inset in the top left panel). Error bars indicate the 68% confidence interval of the geometric mean. Data were obtained using the FN-EL strategy. Stimuli had a total duration of 220 ms, including rise/fall ramps of 10 ms. Each panel illustrates the results for a single listener. Shaded triangles on the abscissa illustrate the HTs for the standard noise.

Results vary widely across listeners. Generally, however, threshold notch depth increases as the notch bandwidth decreases. Interestingly, a significant increase in threshold notch depth occurs for the narrowest, 1000-Hz-wide notch. This is true particularly for listeners S3 and S5 and for over-

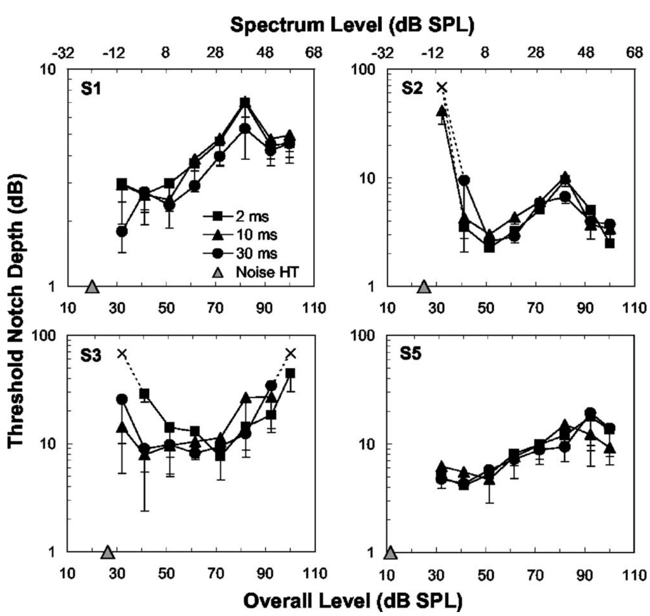

FIG. 7. Threshold notch depth vs. level functions for stimuli with different rise times (as indicated by the inset in the top left panel). Error bars indicate the 68% confidence interval of the geometric mean. In all conditions, the notch had a bandwidth of 2000 Hz and stimuli had a total duration of 220 ms including the rise/fall times. Data were obtained using the FN-EL strategy. Each panel illustrates the results for a single listener. Shaded triangles on the abscissa illustrate the hearing thresholds (HTs) for the standard noise.

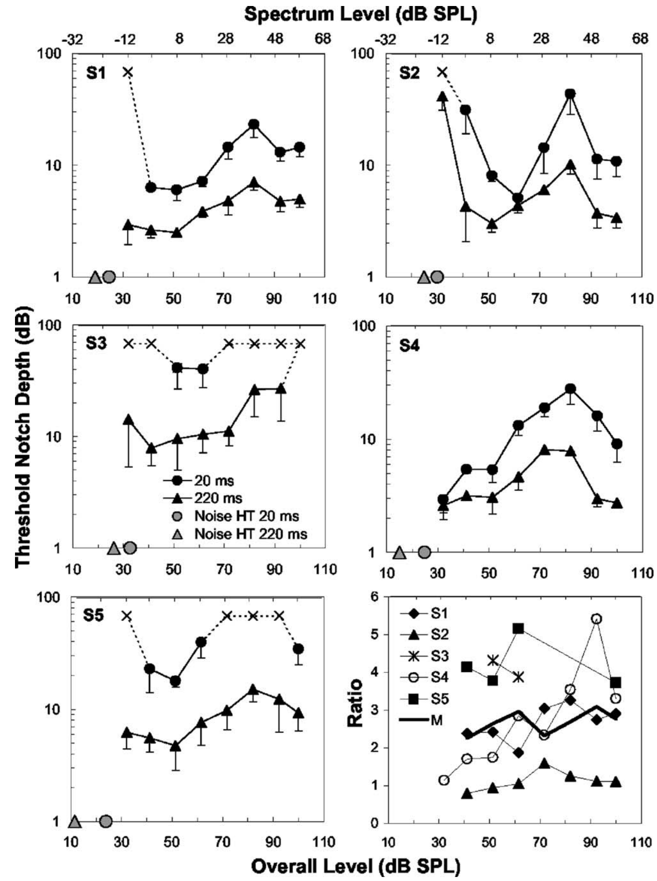

FIG. 8. Threshold notch-depth vs. level functions for stimuli with different durations (as indicated by the inset in the mid left panel). Error bars indicate the 68% confidence interval of the geometric mean. Stimuli rise/fall times were equal to 10 ms in both conditions. Notches had a bandwidth of 2000 Hz. Data were obtained using the FN-EL strategy (see the text for details). Shaded triangles and circles on the abscissa indicate the listener's absolute HTs for the standard noises of 220 and 20 ms of duration, respectively. Each panel illustrates the results for a single listener. The bottom right panel illustrates the ratio of threshold notch depths for the short and long (short/long) stimuli. Each symbol corresponds to a different listener. Note that on average (thick continuous line) the ratio is independent of stimulus level.

all levels higher than 70 dB SPL. For these listeners and conditions, obtaining a threshold notch depth was impossible.

Similar results have been described elsewhere. Moore *et al.* (1989) reported that none of their listeners could detect 1000-Hz-wide notches centered at 8000 Hz while they had no problem detecting wider notches. Heinz and Formby (1999) also reported that detecting energy decrements in the spectrogram of a random-level noise was possible only if the spectral bandwidth of the decrement was $\geqslant$500 Hz.

It should also be noted in Fig. 6 that the nonmonotonic effect, when present, is comparable for different notch bandwidths.

F. The effect of stimulus rise time and stimulus duration

Figures 7 and 8 illustrate the effects of stimulus rise time (in four listeners) and of stimulus duration (in five listeners), respectively. The data are for a notch bandwidth of 2000 Hz and for the frozen-noise/equalized-level condition. Clearly, the stimulus rise time has no significant or systematic effect on notch depth at any of the levels tested (Fig. 7). On the other hand, stimulus duration has a clear effect (Fig. 8). Threshold notch depth values are significantly larger for the short (20-ms) than for the long (220-ms) noise bursts at *all* levels with very few exceptions. The bottom right panel of Fig. 8 illustrates the *ratios*[3] of threshold notch depths for the long and the short stimuli for every listener. Although the ratio varies widely across listeners, its average value equals 2.5 and is approximately independent of level. This shows that the effect of level is comparable for the long and the short stimuli.

IV. DISCUSSION

The aim of the study was to investigate how the detection of spectral notches with center frequencies well above the cutoff of phase locking depends on stimulus level. The effects of other factors such as the notch bandwidth, the stimulus duration and the stimulus rise time have also been investigated. Two hypotheses were tested. First, since the AN rate-profile representation of the stimulus spectrum deteriorates at high levels, threshold notch depths for discrimination should increase monotonically with increasing stimulus level. Second, if at high levels the stimulus spectrum were encoded mostly in the onset rate-profile, threshold notch depths should be smaller for stimuli with abrupt onsets (short rise times) and comparable for long and short stimuli.

Results show that none of these hypotheses hold. For notches 2000- and 4000-Hz wide, although the actual values vary widely across listeners, threshold notch depth is a *nonmonotonic* function of level (Figs. 3 and 4). A clear maximum in the function occurs for overall levels around 70–80 dB SPL for most listeners. Furthermore, threshold notch depths are hardly affected by the stimulus rise time (Fig. 7) and are clearly larger for noise bursts of 20 ms of duration than for bursts 220-ms in duration (Fig. 8).

A. Physiological interpretation

1. The nonmonotonic character of the threshold notch depth versus level functions

Nonmonotonic threshold notch depth versus level functions have been observed for both the equalized and roving level conditions (Fig. 3). For the reasons explained earlier (Sec. II B 2), this suggests that discriminating between the flat-spectrum and the notched noise bursts is based on detecting differences in overall spectral shape, rather than on detecting energy differences in the frequency region of the notch or elsewhere. The notches tested span a region of frequencies higher than the cutoff of phase locking (Johnson, 1980). Hence, they are unlikely to be encoded in the temporal aspects of the AN response. Instead, they must be encoded in the rate of the response, possibly as a rate profile. Therefore, the nonmonotonic character of the threshold notch depth versus level functions suggest that the quality of the AN rate profile representation of the stimulus spectrum must be a nonmonotonic function of level. In other words, the stimulus spectrum must be reasonably well encoded in the AN rate profile at low and high levels, but more poorly represented at 70–80 dB SPL.

Considering the classification of AN fibers suggested by Winter *et al.* (1990), this result may be interpreted as an indication that humans have HSR and LSR fibers but *no* MSR fibers, at least in the 8000-Hz region. In that case, it is possible that the notch be encoded in the rate profile of HSR fibers at low levels and in that of LSR fibers at high levels. The maximum that occurs for overall levels around 70–80 dB SPL (spectrum levels of 28–38 dB SPL) would reflect the transition between the dynamic ranges of the two. That is, it would occur at levels for which HSR fibers are almost saturated and LSR fibers are just below threshold. MSR fibers have thresholds that are intermediate between those of HSR and LSR fibers (Winter *et al.*, 1990, p. 199). Furthermore, over the range of levels (100 dB) used by Winter *et al.*, their dynamic range appears to be wider than that of HSR and LSR fibers as a result of their showing sloping saturation (cf. Fig. 2 of Winter *et al.*, 1990). Therefore, if MSR fibers were present in the human AN, they could probably encode for the spectral notch at moderate-to-high levels. As a result, the threshold notch depth versus level functions would be unlikely to show a maximum and instead threshold notch depth would probably appear independent of level.

The exact way that the two fiber populations translate into nonmonotonic threshold notch depth versus level functions is unclear. However, the above-given interpretation imposes the restriction that the saturation threshold of HSR fibers *and* the rate threshold of LSR fibers should both occur around the level for which the maximum threshold notch depth occurs. That is, both thresholds would be around an overall level of 70–80 dB SPL or its equivalent spectrum level of 28–38 dB SPL (Figs. 3 and 4). For this range of spectrum levels, the effective SPL seen by an AN fiber with a CF=8000 Hz and an effective bandwidth of 1000 Hz would be approximately equal to 38–48 dB SPL. These values agree well with those estimated from pure-tone AN rate-level functions for other mammalian species (e.g., Fig. 1 of Winter and Palmer, 1991).

The above-noted interpretation also agrees with the conclusions of Zeng *et al.* (1991). They studied the effects of prior stimulation upon intensity discrimination and found unusually large just-noticeable differences for midlevel sounds. Zeng *et al.* (1991) argued that their result is consistent with humans having only two types of auditory nerve fibers, HSR and LSR, each with a different time of recovery from adaptation and a different threshold [see Zeng *et al.* (1991) for a full discussion].

It has been shown (Fig. 5) that threshold notch depths are, with some exceptions, comparable for stimulation levels of 42 and 100 dB SPL so long as the notches are wide enough (⩾2000 Hz). This is somewhat surprising considering that the frequency response of AN fibers broadens with increasing sound level (Rose *et al.*, 1971), hence the spectral notch should be less clearly represented in the AN rate profile at the highest level tested. One possible explanation for the result is that peripheral suppression may enhance the spectral notch (as represented in the AN rate profile) at high levels but less so (or nothing at all) at low levels, somehow compensating for the broadening of the fibers' frequency response at high levels. This is a likely explanation because, as discussed by Poon and Brugge (1993), suppression would be expected to cause a decrease in the discharge rate of AN fibers with CFs in the notch region, hence "…enhancing the contrast between the energy level at the spectral notch frequency and that at the surrounding frequencies," and the amount of suppression is more prominent at high levels (e.g., Schalk and Sachs, 1980).

Other explanations exist for the nonmonotonic character of threshold notch depth versus level functions. For example, discrimination based on AN rate responses is likely to be directly proportional to the rate difference between the two stimuli but inversely proportional to their corresponding rate variances (e.g., Young and Barta, 1986; Winter and Palmer, 1991). Consequently, small rate variances may lead to discriminable percepts even when the associated-rate difference is small. The variance-to-mean ratio of AN rate responses decreases with increasing mean rate, that is, with increasing stimulus level (e.g., Winter and Palmer, 1991; Young and Barta, 1986). Therefore, it is possible that the improvement in spectral discrimination observed at high levels results from a reduction in the fibers' variance-to-mean ratio in combination with (or perhaps instead of) the above-mentioned spectral-enhancing effects of suppression.

Basilar membrane (BM) compression may also contribute to the nonmonotonic character of the threshold notch depth versus level functions. Some studies (though not all; see Robles and Ruggero, 2001, pp. 1308–1309) suggest that the basal region of the BM responds linearly at low (<40 dB SPL) and high (<90 dB SPL) levels, but compressively at intermediate levels. If this were the case, the spectral notch would be more clearly represented in the BM excitation pattern at low and high levels, where linear responses occur, than at moderate levels, for which the notch would appear shallower in the BM excitation pattern as a result of compression. Additionally, cochlear distortion may negatively affect the BM representation of the spectral notch as a result of distortion-product energy traveling from remote BM sites, where they are generated, to those with CFs within the notch band (cf. Robles *et al.*, 1997). This effect is likely to be maximal at moderate levels for which compression is greatest. On the other hand, the effects of BM compression are unlikely to explain by themselves the nonmonotonic character of the function. BM compression is approximately constant over a wide range of input levels (~40-90 dB SPL, e.g., Ruggero *et al.*, 1997); hence, should the previous explanation be correct, it would lead to a plateau-shaped nonmonotonic function rather than to the peak-shaped curve that is more common in the present results.

It must be acknowledged that all the above-noted interpretations implicitly assume that listeners use the same cues at all levels and that the internal representation of the spectral notch degrades at levels around 70–80 dB SPL. However, an *alternative* explanation could be that the cues used for spectral discrimination are different at low and high levels. It could well happen that the cues that are salient at low levels degrade as the level increases and that new cues improve with increasing level to become most salient at high levels. The nature of these cues is uncertain at this time. It is unlikely that spectral discrimination at high levels be facilitated by distortion in the stimuli because, as shown in Fig. 1(c), the actual difference spectrum between the target and the standard noises are comparable for overall levels of 50, 80, and 100 dB SPL.

It should be noted that the above-discussed explanations are not mutually exclusive. They can all contribute simultaneously to explain the data. On the other hand, it is not the purpose of this report to provide a convincing explanation of the nonmonotonic effect. Physiological and modeling work is currently under way to clarify this issue (Alves-Pinto *et al.*, 2005a, b).

2. The effect of notch bandwidth and its interaction with level

Detection of spectral notches with bandwidths equal or greater than 2000 Hz was generally possible at all levels. On the other hand, detection of notches 1000-Hz wide was clearly much more difficult, particularly at high levels.

Notch detection must be based on spectral discrimination. Given the nature of the stimuli, the most salient information for spectral discrimination is likely to be provided by AN fibers with CFs within and near the notch band (cf. Poon and Brugge, 1993). Those fibers will be driven with less energy by the target, notched stimulus than by the reference, flat-spectrum one. The effective bandwidth of those fibers is around 1000 Hz at low SPLs (e.g. Fig. 10 of Evans, 1975) and increases as the SPL increases (Rose *et al.*, 1971). Therefore, at high levels, the fibers in question will also respond to energy in frequency regions adjacent to the 1000-Hz-wide notch (see Chap. 8 in Lopez-Poveda, 1996). Hence, their response to the notched stimulus will be only slightly smaller than that to the flat-spectrum stimulus, making discrimination between the two more difficult even when the notch is very deep. In psychophysical terms, this relates to the fact that the critical band at 8000 Hz is comparable to the bandwidth of the 1000-Hz-wide notch (Glasberg and Moore, 1990) and increases with level (e.g. Baker *et al.*, 1998). Furthermore, the interlistener variability observed in the present results, particularly for the narrowest notch (Fig. 6), is likely to relate to the wide interlistener variability in auditory filter bandwidth (Patterson and Moore, 1986). As noted earlier (Sec. III B), this may account for the poorer performance of listener S3.

Spectral discrimination is also likely to be affected, though to a lesser extent, by the number of fibers with CFs within the notch band. This number increases as the notch bandwidth increases and this might explain, at least in part, that threshold notch depths be lowest for stimuli with the widest notch. In any case, the shape of the threshold notch depth versus level function is comparable for different notch bandwidths (see Fig. 6), which suggests that the nonmonotonic effect is independent of the number of fibers signaling for the notch.

3. The effect of stimulus duration

AN fibers show a wider dynamic range at the onset of the stimulus (Smith and Brachman, 1980). For this reason, Delgutte and Kiang (1984a, b) suggested that at high levels, when the rate response of the majority of AN fibers is saturated, the stimulus spectrum may still be conveyed in the onset rate profile. Earlier, it has been shown that maximizing the effects of the stimulus onset by reducing the stimulus rise time hardly affected threshold notch depths (Fig. 6). It has also been shown that threshold notch depths were considerably lower for a stimulus duration of 220 ms than for a duration of 20 ms, even at the highest levels tested (Fig. 7). If at high levels notch detection were based *solely* on the onset representation of the stimulus spectrum, no differences would be observed between the long and the short stimuli. Therefore, although the present results do *not* rule out a possible contribution of the onset rate profile to the encoding of the stimulus spectrum, they suggest that the contribution in question is less important than previously suggested (Delgutte and Kiang, 1984a, b; Lopez-Poveda, 1996).

Since both the stimulus and the nature of the AN response are stochastic, the rate-profile representation of the stimulus spectrum surely varies over time. The fact that threshold notch depths are smaller for the longer stimuli across *all* levels tested (Fig. 8) suggests that the longer stimulus provides a clearer representation of the stimulus spectrum. It is possible that the brain integrates (or averages) the rate profile over time with a relatively long time window (Green, 1960). Alternatively, it is possible that the decision be based on taking multiple "looks" at rate profiles computed over a much shorter time window, as suggested by Viemeister and Wakefield (1991). Either of the two mechanisms serves to explain (at least qualitatively) the smaller values of notch depth for the longer stimulus.

Furthermore, it is noteworthy that the effect of level is comparable for the short and the long stimuli when threshold notch depths are plotted on a logarithmic scale. This indicates that the nonmonotonic character of the threshold notch depth versus level function is independent of stimulus duration and, consequently, that it is not determined by temporal integration or multiple looks.

B. Relation with level discrimination results

The dynamic-range problem has been widely investigated by previous psychophysical studies of level discrimination. Indeed, the present results resemble those of level discrimination experiments in a number of respects. First, here it has been shown (Figs. 3 and 4) that threshold notch depth varies nonmonotonically as a function of level. Similarly, the threshold for detecting a level change in a high-frequency stimulus is a nonmonotonic function of level, both for tones (e.g. Carlyon and Moore, 1984; Florentine, 1986; Florentine *et al.*, 1987) and noise (e.g., Buus, 1990; Nizami *et al.*, 2001). Second, the nonmonotonic aspect of the threshold notch depth versus level function appears more obvious for the short (20 ms) than for the long (220 ms) stimuli when the ordinate is on a linear scale, rather than on the logarithmic scale used in Fig. 8. The same is true for the level-change versus level function (e.g., Carlyon and Moore, 1984; Nizami *et al.*, 2001). Finally, here it has been shown that dynamic responses to the signal onset play little role in spectral discrimination (Fig. 7) and this seems to be the case also for level discrimination (Carlyon and Moore, 1984).

The results of level-discrimination experiments led Carlyon and Moore (1984) to suggest a bimodal distribution of thresholds in human primary auditory neurons. Such a suggestion is in line with the "two types of fibers" interpretation proposed earlier for the nonmonotonic aspect of the threshold notch depth versus level function. Of course, detecting a level change is not the same as detecting a spectral notch. Furthermore, as discussed earlier, the detection of the spectral notch is unlikely to be based on comparing the levels of the standard and the target noises (neither the overall level or the level within or outside the notch band). Instead, it is likely to depend on the quality with which the spectral notch is represented in the overall auditory nerve rate profile. Nevertheless, the operation range of auditory nerve fibers will determine, in a similar way, the level-dependence of the quality of the internal representation of the notch (i.e., its depth in the rate profile) and the fibers' difference response to a change in the stimulus level.

C. Relation with sound localization results

Sound localization relies partly on accurately detecting the spectral content of the head-related transfer function (HRTF). Specifically, it has been reported that spectral notches with sound source elevation-dependent center frequencies constitute prominent cues for judging the vertical location of sound sources (e.g., Butler and Belendiuk, 1977; Butler and Humanski, 1992). On the other hand, the bandwidth and the depth of HRTF notches vary widely across listeners [see, for instance, Fig. 3 in Shaw (1982) or Chap. 3 in Lopez-Poveda (1996)], possibly reflecting differences in the shapes and sizes of the listeners' ears (Lopez-Poveda and Meddis, 1996). The present study shows that the ability to detect high-frequency spectral notches varies widely across listeners (Fig. 4), and depends on the notch bandwidth (Fig. 6) as well as on the stimulus level and duration (Figs. 3 and 8). As a result, the ability of listeners to actually use the notches in their individual HRTFs as cues for sound localization must depend on a complex combination of their level of performance in notch detection tasks, the shape of their ears, and the characteristics of the stimulus (duration and level). In any case, since vertical localization relies on detecting HRTF notches, the present results suggest that, in general, localization judgments should be more precise for long stimuli than for short ones and for levels below 60–70 dB SPL than for levels around 70–80 dB SPL and this is indeed the case (Hartmann and Rakerd, 1993; Macpherson and Middlebrooks, 2000; Vliegen and Van Opstal, 2004). Furthermore, in light of the present results an improvement in localization accuracy should occur for levels higher than 80 dB SPL, although this remains to be tested.

Despite the emphasis given here to spectral notches, it must be acknowledged that other authors (e.g., Blauert, 1969/70; Humanski and Butler, 1988) have suggested that spectral peaks may be as important for sound localization as spectral notches, if not more. The auditory nerve rate-profile representation of spectral peaks at high levels would be also negatively influenced by the limited dynamic range of most nerve fibers. On the other hand, the broadening of the fiber's frequency response with increasing level would not deteriorate the rate-profile representation of peaks. If anything it would enhance it by spreading the energy of the peak to fibers with a wider range of CFs. Consequently, comparing the present results with those of experiments aimed at discriminating between flat-spectrum and peaked noise bursts may help elucidating the relative contribution of both mechanisms (i.e., limited dynamic range and filter broadening) to the auditory nerve representation of spectral features.

V. CONCLUSIONS

The extent that the detection of high-frequency spectral notches is affected by stimulus level was investigated by measuring the threshold notch depth necessary to discriminate between a flat-spectrum wideband noise and a similar noise with a rectangular spectral notch centered at 8000 Hz. The main conclusions are:

(1) High-frequency spectral notches are detected by detecting differences in the overall spectral shape of the stimulus rather than by detecting level differences over certain frequency regions.
(2) For a large proportion of listeners, threshold notch depth is clearly a *nonmonotonic* function of stimulation level. It increases for levels up to 70–80 dB SPL and decreases for higher levels; thus a maximum in the function occurs at levels around 70–80 dB SPL. Interpretations of this result have been discussed based on the physiological response properties of the mammalian auditory nerve.
(3) The nonmonotonic character of the threshold notch depth versus level function is independent of notch bandwidth and stimulus duration.
(4) Threshold notch depths are hardly affected by the stimulus rise time and depend strongly on the stimulus duration. Notch depth values are larger for shorter stimuli. These results suggest that the onset rate-profile contributes less than previously thought to encoding for the stimulus spectrum at high levels. It is also consistent with the idea that the detection of the stimulus spectrum involves either integrating the AN rate profile over a relatively long time window or multiple looks at the rate profiles computed over relatively narrow time windows.

ACKNOWLEDGMENTS

The authors thank David A. Eddins for useful suggestions on a poster version of this work, and Ray Meddis and Alan R. Palmer for their comments on a draft of this paper. The authors are indebted to two anonymous reviewers for their very helpful comments, particularly for suggesting that we use the geometric mean instead of the arithmetic mean and to plot the results on a logarithmic scale. Work supported by the Spanish Fondo de Investigaciones Sanitarias (Grant Nos. PI02/0343 and G03/203) and by European Regional Development Funds.

[1]In this context, the word "quality" must be understood to refer to the degree of similarity between the stimulus spectrum and the AN rate-profile representation.
[2]Some studies report *nonsaturating* LSR fibers (e.g., Winter *et al.*, 1990). The existence of this type of fiber seems to be species and frequency specific but the issue is still controversial.
[3]Since threshold notch depth conforms to a log-normal distribution (Sec. III A), the *ratio* is more appropriate than the *difference* to quantify the effect of duration or of any other stimulus parameter.

Alves-Pinto, A., Lopez-Poveda, E. A., and Palmer, A. R. (**2005a**). "Auditory nerve encoding of high-frequency spectral information," in *Mechanisms, Symbols, and Models Underlying Cognition*, edited by J. Mira and J. R. Álvarez, Lecture Notes in Computer Science **3561**, 223–232.

Alves-Pinto, A., Lopez-Poveda, E. A., and Palmer, A. R. (**2005b**). "Auditory nerve response to broadband noise with high-frequency spectral notches," Abstract 4pPP16, *149th Meeting of the Acoustical Society of America*, Vancouver, Canada, 16–20 May; J. Acoust. Soc. Am. **117**, 2654.

Baker, R. J., Rosen, S., and Darling, A. M. (**1998**). "An efficient characterization of human auditory filtering across level and frequency that is also physiologically reasonable," in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, Q. Summerfield, and R. Meddis (Whurr, London), pp. 81–88.

Bland, J. M., and Altman, D. G. (**1996a**). "Measurement error and correlation coefficients," Br. Med. J. **313**, 106.

Bland, J. M., and Altman, D. G. (**1996b**). "Transformations, means and confidence intervals," Br. Med. J. **312**, 1079.

Bland, J. M., and Altman, D. G. (**1996c**). "Transforming data," Br. Med. J. **312**, 770.

Bland, J. M., and Altman, D. G. (**1996d**). "The use of transformation when comparing two means," Br. Med. J. **312**, 1153.

Blauert, J. (**1969/70**). "Sound localization in the median plane," Acustica **22**, 205–213.

Butler, R. A., and Belendiuk, K. (**1977**). "Spectral cues utilized in the localization of sound in the median sagittal plane," J. Acoust. Soc. Am. **61**, 1264–1269.

Butler, R. A., and Humanski, R. A. (**1992**). "Localization of sound in the vertical plane with and without high-frequency spectral cues," Percept. Psychophys. **51**, 182–186.

Buus, S. (**1990**). "Level discrimination of frozen and random noise," J. Acoust. Soc. Am. **87**, 2643–2654.

Buus, S., and Florentine, M. (**1991**). "Psychometric functions for level discrimination," J. Acoust. Soc. Am. **90**, 1371–1380.

Carlyon, R. P., and Moore, B. C. J. (**1984**). "Intensity discrimination: A severe departure from Weber's law," J. Acoust. Soc. Am. **76**, 1369–1376.

Delgutte, B. (**1980**). "Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers," J. Acoust. Soc. Am. **68**, 843–857.

Delgutte, B., and Kiang, N. Y. S. (**1984a**). "Speech coding in the auditory nerve. III. Voiceless fricative consonants," J. Acoust. Soc. Am. **75**, 887–896.

Delgutte, B., and Kiang, N. Y. S. (**1984b**). "Speech coding in the auditory nerve. IV. Sounds with consonant-like dynamic characteristics," J. Acoust. Soc. Am. **75**, 897–907.

Evans, E. F. (**1975**). "Cochlear nerve and cochlear nucleus," in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer, Berlin), Vol. **5/2**, pp. 1–108.

Evans, E. F., and Palmer, A. R. (**1980**). "Relationship between the dynamic range of cochlear nerve fibers and their spontaneous activity," Exp. Brain Res. **40**, 115–118.

Florentine, M. (**1986**). "Level discrimination of tones as a function of duration," J. Acoust. Soc. Am. **79**, 792–798.

Florentine, M., Buus, S., and Mason, C. R. (**1987**). "Level discrimination as a function of level for tones from 0.25 to 16 kHz," J. Acoust. Soc. Am. **81**, 1528–1541.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched noise data," Hear. Res. **47**, 103–138.

Green, D. M. (**1960**). "Auditory detection of a noise signal," J. Acoust. Soc. Am. **32**, 121–131.

Green, D. M. (**1988**). *Profile Analysis* (Oxford University Press, New York).

Hartmann, W. M., and Rakerd, B. (**1993**). "Auditory spectral discrimination and the localization of clicks in the sagittal plane," J. Acoust. Soc. Am. **94**, 2083–2092.

Hebrank, J., and Wright, D. (**1974**). "Spectral cues used in the localization of sound sources on the median plane," J. Acoust. Soc. Am. **56**, 1829–1834.

Heinz, M. G., and Formby, C. (**1999**). "Detection of time- and band-limited increments and decrements in a random-level noise," J. Acoust. Soc. Am. **106**, 313–326.

Humanski, R. A., and Butler, R. A. (**1988**). "The contribution of the near and far ear toward localization of sound in the sagittal plane," J. Acoust. Soc. Am. **83**, 2300–2310.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," J. Acoust. Soc. Am. **68**, 1115–1122.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (**1965**). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT, Cambridge).

Levitt, H. (**1971**). "Transformed up down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," J. Acoust. Soc. Am. **63**, 442–455.

Lopez-Poveda, E. A. (1996). "The physical origin and physiological coding of pinna-based spectral cues," Ph.D. thesis. Loughborough University, Loughborough, UK.

Lopez-Poveda, E. A., and Meddis, R. (**1996**). "A physical model of sound diffraction and reflections in the human concha," J. Acoust. Soc. Am. **100**, 3248–3259.

Macpherson E. A., and Middlebrooks, J. C. (**2000**). "Localization of brief sounds: Effects of level and background noise," J. Acoust. Soc. Am. **108**, 1834–1848.

Meddis, R. (**1988**). "Simulation of auditory-neural transduction: Further studies," J. Acoust. Soc. Am. **83**, 1056–1063.

Moore, B. C. J., Oldfield, S. R., and Dooley, G. J. (**1989**). "Detection and discrimination of spectral peaks and notches at 1 and 8 kHz," J. Acoust. Soc. Am. **85**, 820–836.

Nizami, L., Reimer, J. F., and Jesteadt, W. (**2001**). "The intensity-difference limen for Gaussian-enveloped stimuli as a function of level: Tones and broadband noise," J. Acoust. Soc. Am. **110**, 2505–2515.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," Hear. Res. **24**, 1–15.

Patterson, R. D., and Moore, B. C. J. (**1986**). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 123–177.

Poon, P. W. F., and Brugge, J. F. (**1993**). "Sensitivity of auditory nerve fibers to spectral notches," J. Neurophysiol. **70**, 655–666.

Rice, J. J., Young, E. D., and Spirou, G. A. (**1995**). "Auditory-nerve encoding of pinna-based spectral cues: Rate representation of high-frequency stimuli," J. Acoust. Soc. Am. **97**, 1764–1776.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Robles, L., Ruggero, M. A., and Rich, N. C. (**1997**). "Two-tone distortion on the basilar membrane of the chinchilla cochlea," J. Neurophysiol. **77**, 2385–2399.

Rose, J. E., Hind J. E., Anderson, D. J., and Brugge, J. F. (**1971**). "Some effects of stimulus intensity on response of auditory-nerve fibers in the squirrel monkey," J. Neurophysiol. **34**, 685–699.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," J. Acoust. Soc. Am. **101**, 2151–2163.

Sachs, M. B., and Abbas, P. J. (**1974**). "Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli," J. Acoust. Soc. Am. **56**, 1835–1847.

Sachs, M. B., and Young, E. D. (**1979**). "Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate," J. Acoust. Soc. Am. **66**, 470–479.

Schalk T. B., and Sachs, M. B. (**1980**). "Nonlinearities in auditory-nerve fiber responses to bandlimited noise," J. Acoust. Soc. Am. **67**, 903–913.

Shaw, E. A. G. (**1974**). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," J. Acoust. Soc. Am. **56**, 1848–1861.

Shaw, E. A. G., and Teranishi, R. (**1968**). "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," J. Acoust. Soc. Am. **44**, 240–249.

Shaw, E. A. G. (**1982**). "External ear response and sound localization," in *Localization of Sound: Theory and Applications*, edited by R. W. Gatehouse (Amphora Press, Groton), pp. 30–41.

Smith, R. L., and Brachman, M. L. (**1980**). "Operating range and maximum response of single auditory nerve fibers," Brain Res. **184**, 499–505.

Viemeister, N. F., and Wakefield, G. H. (**1991**). "Temporal integration and multiple looks," J. Acoust. Soc. Am. **90**, 858–865.

Vliegen, J., and Van Opstal, A. J. (**2004**). "The influence of duration and level on human sound localization," J. Acoust. Soc. Am. **115**, 1705–1713.

Winter, I. M., and Palmer, A. R. (**1991**). "Intensity coding in low-frequency auditory-nerve fibers of the guinea pig," J. Acoust. Soc. Am. **90**, 1958–1967.

Winter, I. M., Robertson, D., and Yates, G. K. (**1990**). "Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres," Hear. Res. **45**, 191–202.

Young, E. D., and Barta, P. E. (**1986**). "Rate responses of auditory nerve fibers to tones in noise near masked threshold," J. Acoust. Soc. Am. **79**, 426–442.

Zeng, F.-G., Turner, C. W., and Relkin, E. M. (**1991**). "Recovery from prior stimulation. II. Effects upon intensity discrimination," Hear. Res. **55**, 223–230.

[4]The SPLs were measured with a Brüel & Kjaer sound level meter (model 2238A) coupled to a Zwislocki occluded ear simulator over a frequency range from 100 to 12500 Hz in 1/3 octave bands.