Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 71 (2008) 693–703
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Rate versus time representation of high-frequency spectral notches in the peripheral auditory system: A computational modeling study

Enrique A. Lopez-Poveda[a,c,*], Ana Alves-Pinto[a], Alan R. Palmer[b], Almudena Eustaquio-Martín[a]

[a]*Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Avda. Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain*
[b]*MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK*
[c]*Dto. Informática y Automática, Facultad de Ciencias, Universidad de Salamanca, Plaza de los Caídos s/n, 37008 Salamanca, Spain*

Available online 13 November 2007

Abstract

A computational model of the peripheral auditory system is used to explain the paradoxical observation that discriminating between broadband noise sounds with and without high-frequency spectral notches is more difficult at mid-intensities than at lower or higher intensities [A. Alves-Pinto, E.A. Lopez-Poveda, Detection of high-frequency spectral notches as a function of level, J. Acoust. Soc. Am. 118 (2005) 2485–2469.]. The simulations suggest that the discrimination task in question relies on comparing the timing of auditory nerve spikes, hence that high-frequency sounds are represented in the auditory nerve by a time code. They further suggest that the improvement in spectral discrimination at high intensities is associated with inherent inner hair cell nonlinearities.

Keywords: Auditory profile analysis; Auditory spectral discrimination; Time code; Place code; Phase locking; Auditory nonlinearity; Auditory dynamic range; Inner hair cell nonlinearity; Spectral notch

1. Introduction

High-frequency (>4 kHz) spectral features such as peaks and notches of broadband auditory stimuli are useful for sound localization (reviewed by Ref. [4]) as well as for speech (e.g. Refs. [12,36]) and music perception. The peripheral auditory system processes auditory signals in a nonlinear, compressive manner and it is likely that the auditory nerve representation of such high-frequency spectral features degrades gradually with increasing sound intensity. A recent psychophysical report, however, undermines this idea by showing that discrimination of high-frequency spectral notches is more difficult at intensities around 70–80 dB SPL than at lower or higher intensities [1]. The present study was undertaken to explain this paradoxical result and, hence, to clarify how high-frequency spectral features may be represented in the auditory nerve response.

1.1. The nature of the paradox

Auditory nerve fibers (ANFs) transmit acoustic information from the cochlea to the central auditory system in the form of spike trains. Every ANF is most sensitive to a particular sound frequency (the characteristic frequency or CF) and responds to a restricted range of frequencies around its CF. Therefore, every ANF can be viewed as a spectral filter with a certain CF and the population of ANFs as a bank of filters with CFs spanning approximately the frequency range of hearing ([10]; reviewed by Ref. [20]).

There are at least two ways in which the spectrum of a signal may be obtained from the output of a filter-bank system. First, it may be obtained by plotting the amplitude of the output signals from the filters as a function of the filters' center frequencies. We refer to this representation as the *excitation pattern* (Fig. 1C). An alternative way would

*Corresponding author at: Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Avda. Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain. Tel.: +34 923294500; fax: +34 923294750.
E-mail address: ealopezpoveda@usal.es (E.A. Lopez-Poveda).

0925-2312/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2007.07.030

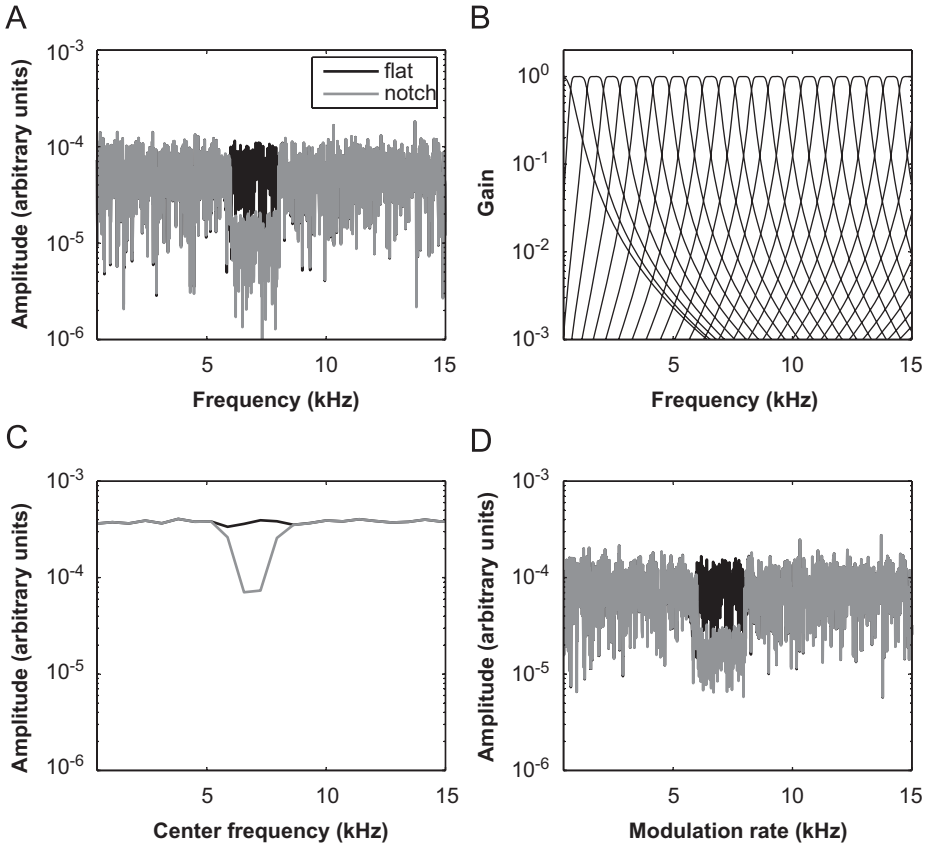

Fig. 1. Two possible representations of the stimulus spectrum at the output of a linear filter-bank. (A) Example spectra of flat-spectrum (dark trace) and notch (light trace) frozen-noise stimuli. Both noises have a broad-band spectrum (20–16 kHz). The notch is centered at 7 kHz and has 2-kHz bandwidth. Its depth is 15 dB re spectrum level in notch side bands. The two spectra are identical except for the amplitude in the notch band. (B) An example linear filter-bank. (C and D) The excitation and modulation-rate patterns of the two stimuli at the filter-bank output, respectively.

be to estimate the spectra of the output signals from all the filters and add them together to obtain the total filter-bank output spectrum. We will refer to this representation as the *modulation-rate pattern* (Fig. 1D). For a linear filter-bank, both the excitation and the modulation-rate patterns resemble the coarse stimulus spectrum, assuming a large number of perfectly overlapping filters (compare Fig. 1A with Fig. 1C and D).

The central auditory system may obtain the stimulus spectrum from the response of ANFs in two corresponding ways: from the rate profile and from the discharge times for the population of ANFs. The rate profile is a representation of the average discharge rate of ANFs as a function of their CFs [31]. This is directly related to the excitation pattern. ANFs fire in synchrony with the individual cycles of the driving stimulus waveform (but see the following text), a property called phase locking. Therefore, the timing of spikes for the ANF population carries information about the stimulus waveform and hence of its spectrum. This representation of the stimulus spectrum relates closely to the modulation-rate pattern. Unfortunately, these two representations may *not* always be available to encode high-frequency spectral features in the auditory nerve.

Phase locking rolls off rapidly with increasing frequency above 2 kHz mostly due to low-pass filtering in the inner hair cells (IHCs) [15,27]. Consequently, it is widely accepted that high-frequency spectral features (>4 kHz) are likely to be encoded in the rate profile of ANFs solely. However, the frequency response of cochlear filters, and thus of ANFs, broadens gradually with increasing sound intensity (reviewed by Ref. [32]). In addition, several important peripheral auditory elements (the basilar membrane, the IHC, and the auditory nerve) have strongly compressive input/output characteristics (reviewed by Refs. [20,32]). For instance, the majority of ANFs have low discharge-rate thresholds (<15 dB SPL) and only a narrow range of levels (20–30 dB) over which their discharge rate varies according to sound intensity [18]. While the remaining ANFs have higher thresholds and wider dynamic ranges, only a small proportion of fibers are still unsaturated at high levels [26]. As a result, the quality of the rate profile representation of high-frequency spectral features deteriorates gradually with increasing intensity ([19,31], but see Ref. [24]).

All this suggests that the auditory nerve representation of high-frequency spectral features should degrade gradually with increasing sound intensity and, consequently, that discriminating between two auditory stimuli with different high-frequency spectra should be increasingly more difficult with increasing stimulus intensity. Paradoxically, this

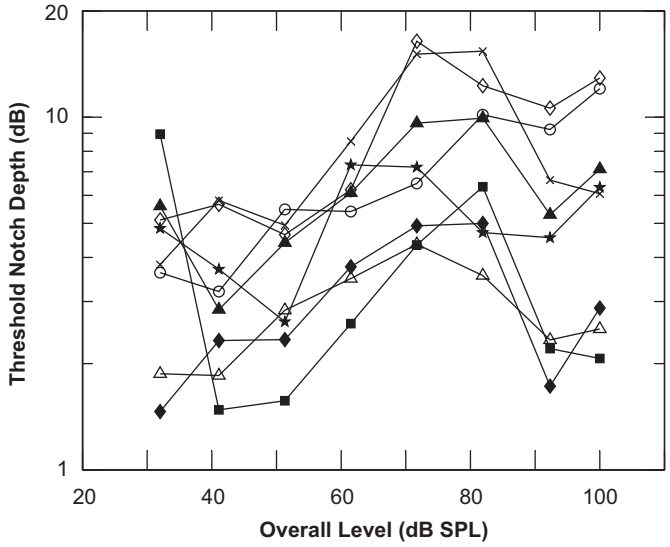

Fig. 2. Threshold notch depths for discriminating between a flat-spectrum noise and a notch noise as a function of stimulus overall intensity (adapted from Alves-Pinto and Lopez-Poveda [1]). The notch depth is in dB re the spectrum level in the notch side bands. Each symbol illustrates the results for a different listener. The notch bandwidth was 4 kHz. For all but one of the listeners (circles), the threshold notch depth is a nonmonotonic function of intensity; i.e., it is larger around 70–80 dB SPL than at lower or higher intensities.

does *not* seem to be the case. Alves-Pinto and Lopez-Poveda [1] showed that discriminating between broadband noise bursts with and without a high-frequency (8 kHz) spectral notch is more difficult at intensities around 70–80 dB SPL than at lower or higher intensities (Fig. 2).

1.2. Possible explanations

Alves-Pinto and Lopez-Poveda [1] proposed several explanations for their result. They conjectured that it was consistent with the existence of two types of ANFs with different thresholds and dynamic ranges, as explained above. Discrimination might be worst at 70–80 dB SPL because this would be the intensity at which the transition between the dynamic ranges of the two fiber types occurs (see also Refs. [38,39]). Direct evidence in support of this explanation (see [2]) is inconclusive. Alternative explanations were also discussed by Alves-Pinto and Lopez-Poveda, all of which were based on the assumption that high-frequency spectral notches must be represented in the rate profile of the auditory nerve.

On the other hand, a computational modeling study [13] has suggested that the just-noticeable difference in frequency between two tones even at high frequencies (up to 10 kHz) could be explicable in terms of encoding in the discharge times of ANFs. In support of this result, recent physiological studies have shown that, contrary to common belief, significant phase locking can occur for frequencies as high as 14 kHz [37]. Based on these ideas, an alternative explanation of the result of Alves-Pinto and Lopez-Poveda [1] could be that high-frequency spectral notches may still be encoded in the temporal pattern of auditory nerve discharge and that discrimination may rely on comparing the auditory nerve modulation-rate pattern representations of the spectra (cf. Ref. [23]).

1.3. Aims

The aim of the present study was to investigate which of the two representations of high-frequency spectral features, the excitation (rate) or the modulation-rate (time) patterns, was more consistent with the observation of Alves-Pinto and Lopez-Poveda [1]. A computational model of the peripheral auditory system was used to simulate receptor potential waveforms (the signals driving ANF firing) for a bank of IHCs with different CFs in response to the stimuli used by these authors. The simulations suggest that only the modulation-rate pattern representation is consistent with the behavioral results of Alves-Pinto and Lopez-Poveda. Additionally, they suggest that the improvement in spectral discrimination for intensities above 70–80 dB SPL is mainly associated with inherent IHC nonlinearities.

2. Methods

The IHC instead of the auditory nerve representation of the stimulus spectrum was considered for two reasons. First, the receptor potential drives auditory nerve spiking. Therefore, the quality of the spectrum representation in the IHCs imposes a limit on the quality of its associated auditory-nerve representation. In other words, the auditory-nerve representation of the spectrum cannot be better than that observed at the level of the IHCs. For example, it is likely that the excitation pattern representation of the stimulus spectrum degrades at high sound levels because of the saturation of the receptor potential (e.g. Ref. [34]). If this were the case, it would undermine the idea of Alves-Pinto and Lopez-Poveda that the peak in the threshold notch depth vs. level function (Fig. 2) relates to the transition between the dynamic ranges of low- and high-threshold ANFs.

A second reason for using an IHC model is that it outputs a continuous, deterministic signal (the IHC receptor potential), which is easier to analyze than trains of stochastic spikes.

2.1. The model

The model consisted of four signal-processing stages, each of which represents a physiological stage of the peripheral auditory system (Fig. 3). The input to the model was a digital sound waveform (in units of Pa) assumed to represent an auditory stimulus at the eardrum.[1]

[1]The effect of the outer-ear was not included in the model because Alves-Pinto and Lopez-Poveda (2005) delivered their stimuli through Etymotic ER2 earphones, which are designed to give a flat-frequency response at the eardrum.

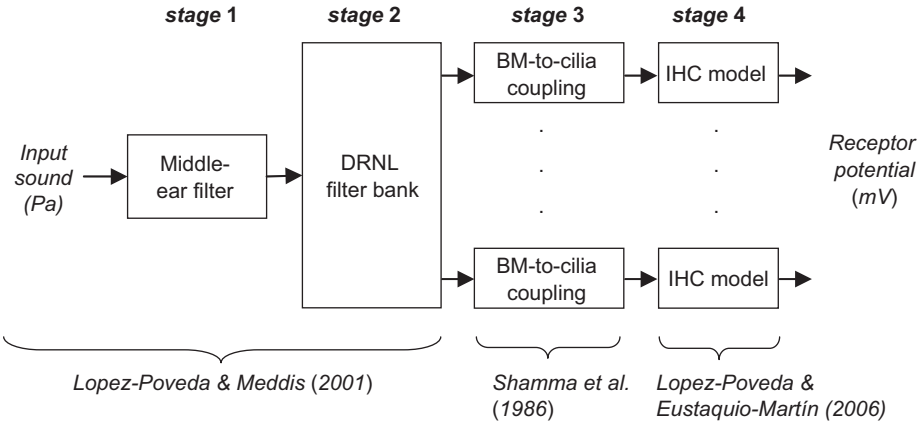

Fig. 3. Model structure. See main text for a description.

The model output was a collection of waveforms representing the receptor potential for a bank of IHCs with different CFs.

The first model stage simulated the middle-ear transfer function. Its output was the stapes velocity (in units of m/s). The second model stage was a bank of dual-resonance nonlinear (DRNL) filters [25] that simulated human cochlear frequency selectivity. The output from each filter was a waveform representing the velocity of vibration (in units of m/s) of a given point along the human basilar membrane. A first-order low-pass filter (cut-off frequency of 1 Hz) was applied to the filter-bank output signals to transform their units from velocity to displacement (in units of m), as required for the next model stage. The third model stage was a high-pass filter that coupled basilar-membrane displacement to the displacement of IHC stereocilia (in units of m) [35]. The last model stage was a biophysical model of the IHC that outputs the IHC receptor potential (in V) as a function of the cell's stereocilia displacement [21]. Details of these models can be found in the original publications and will not be reproduced here.

It is noteworthy that the DRNL filter-bank and the IHC models were specifically and carefully designed to simulate the frequency-response and the nonlinear (compressive) characteristics of their physiological counterparts.

2.2. Implementation and parameters

The model was implemented in Matlab R14 as a dynamic signal-processing chain. The code is available from the authors upon request. The middle-ear and filter-bank stages were implemented as described in Lopez-Poveda and Meddis [22]. A bank of 100 filters was considered with the lowest and the highest CF of 100 Hz and 20 kHz, respectively. We used the parameter values provided by the same authors for an average human listener (see Fig. 2B and Table 2 in Lopez-Poveda and Meddis [22]). These were obtained by linear regression fits to the values optimized at six CFs, from 250 Hz to 8 kHz in octave steps.

The basilar membrane-to-cilia coupling stage was implemented as described in Shamma et al. ([35]; Eq. (1)) (coupling gain $C = 2$; cut-off frequency of 530 Hz). We used the *in vivo* IHC model of Lopez-Poveda and Eustaquio-Martín [21] with the parameter values provided by these authors (see their Table 1). It was assumed that the parameter values for these two model stages were identical across frequency channels.

2.3. Evaluations

The model was evaluated in the time domain (sampling frequency = 44,100 Hz) in response to broadband (20 Hz to 16 kHz) noise bursts with flat-spectrum (dark trace in Fig. 1A) and with a rectangular spectral notch centered at 7 kHz (light trace in Fig. 1A). These were generated as described in Alves-Pinto and Lopez-Poveda [1]. The bandwidth of the notch was 2 kHz, and its depth 15 dB re the spectrum level on the notch side bands (Fig. 1). The spectra of the two noises were identical except for the amplitude in the notch band (Fig. 1). The noise bursts had a total duration of 110 ms including raised cosine rise/fall ramps of 10 ms. The model was evaluated for overall sound intensities from 38 to 98 dB SPL in 10-dB steps, which correspond to spectrum levels from −2 to 58 dB SPL. These were identical to those considered by Alves-Pinto and Lopez-Poveda.

2.4. Output analysis and estimates of the spectral-discrimination sensitivity

For any given stimulus, the model output was used to obtain the excitation pattern and the modulation-rate pattern representations of the stimulus spectrum. The average discharge rate of ANFs is thought to be proportional to the average receptor potential of IHCs [5]. Therefore, the excitation pattern was obtained by plotting the average receptor potential of each IHC as a function of the cell's CF (e.g. Fig. 4A). For each IHC and each stimulus condition, the average receptor potential was calculated by summing the samples in the

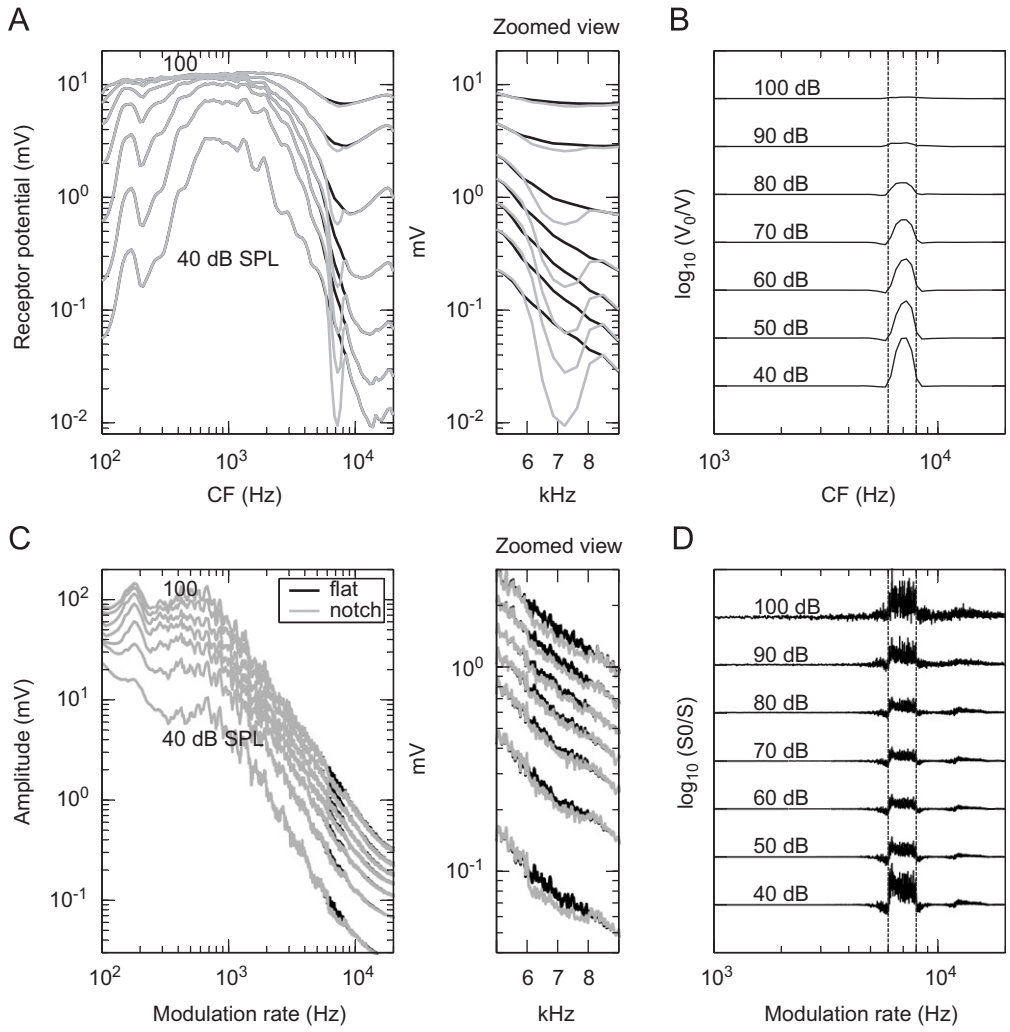

Fig. 4. (A) IHC excitation pattern representation of the flat-spectrum (dark traces) and notch noises (light traces). The numbers next to each trace indicate stimulus intensity in dB SPL. (B) The difference excitation patterns [$\log_{10}(V_0/V)$] normalized to the maximum value across CFs and intensities. Traces are vertically displaced to ease the visual comparison of the amplitudes. The numbers next to each trace indicate stimulus intensity in dB SPL. (C) IHC modulation-rate pattern representation of the flat-spectrum (dark traces) and notch noises (light traces). (D) The difference modulation rate patterns [$\log_{10}(S_0/S)$] normalized to the maximum value across modulation-rates and intensities. The vertical dashed lines in (B) and (D) indicate the notch frequency band.

receptor potential waveform with amplitudes greater than zero and dividing the result by the total number of samples in the waveform [5]. The modulation-rate pattern representation was obtained by applying a fast Fourier transform (FFT) to the receptor potential waveform of each IHC and adding the resulting modulation-rate spectra of all IHCs in the frequency domain (e.g. Fig. 4C). A number of nonlinear elements affected the simulated receptor potential (i.e. compressive DRNL filters, IHC half-wave rectification and the saturation of the receptor potential) and hence one should not expect the modulation-rate pattern to be identical to the stimulus spectrum.

For either representation, the *ad hoc assumption* was made that behavioral spectral discrimination would be based on the Euclidean distance between the logarithms of the IHC representations of the flat-spectrum and notch noise spectra; that is, on the squared difference between the dark and light traces in Fig. 4A and C. The logarithm of the magnitude was used because in the absence of compressive processes, the IHC receptor potential grows approximately exponentially with increasing sound intensity in decibels (e.g. Ref. [28]). Consequently, at any given intensity, L, the discrimination sensitivity based on a comparison of the excitation patterns of the two stimuli, $S_{EP}(L)$, was estimated as

$$S_{EP}(L) = \left\{ \sum_{i=1}^{N} \log_{10} \left[\left(\frac{V_{0i}(L)}{V_i(L)} \right) \right]^2 \right\}^{0.5}, \quad (1)$$

where V_{0i} and V_i denote the average receptor potential for the i-th IHC in response to the flat-spectrum and notch noise bursts, respectively, and N is the number of CF channels.

Likewise, the discrimination sensitivity based on the modulation-rate patterns of the two stimuli, $S_{MP}(L)$, was estimated as

$$S_{MP}(L) = \left\{ \int_0^{\infty} \left[\log_{10}\left(\frac{S_0(f,L)}{S(f,L)} \right) \right]^2 \mathrm{d}f \right\}^{0.5}, \quad (2)$$

where $S_0(f,L)$ and $S(f,L)$ denote the modulation-rate patterns in response to the flat-spectrum and notch noise bursts, respectively.

Given that the peripheral auditory system is highly compressive, S_{EP} and S_{MP} were expected to have different values for different sound intensities and thus were evaluated for every sound intensity considered (cf. Section 2.3). It was pointless to compare the numerical values of S_{EP} and S_{MP} because they depend on unrelated arbitrary parameters; S_{EP} depends on the number of CFs and the CF range (Eq. (1)), while S_{MP} depends on the sampling frequency, df (Eq. (2)). It was more important, however, to compare how they varied with stimulus intensity. To facilitate such comparisons, their values were normalized relative to their respective maximum values across intensities.

3. Results

3.1. Excitation patterns

Fig. 4A illustrates the IHC excitation pattern representation of the spectra of the flat-spectrum (dark traces) and the notch (light traces) stimuli for the seven stimulus intensities considered. For the lowest intensity, the excitation pattern peaks around 1 kHz due to the band-pass characteristics of the middle-ear filter (cf. Fig. 2 in Lopez-Poveda and Meddis [22]). The pattern becomes all-pass at high stimulus intensities. This is due to a combination of basilar-membrane compression (cf. Fig. 4 in Lopez-Poveda and Meddis [22]) and the saturation of the IHC receptor potential at high intensities (cf. Figs. 9 and 10 in Lopez-Poveda and Eustaquio-Martín [21]). Fig. 4A also shows that the quality of the excitation-pattern representation of the spectral notch degrades gradually with increasing stimulus intensity. This is more clearly seen in the zoomed view and in Fig. 4B. The latter illustrates the difference excitation patterns [i.e., $\log_{10}(V_0/V)$] at each intensity normalized to the maximum difference found across CFs and intensities. Note that differences occur for IHCs with CFs within or around the notch band only and that the largest difference occurs for the lowest intensity (40 dB SPL).

3.2. Modulation-rate patterns

Fig. 4C illustrates the simulated modulation-rate patterns for the flat-spectrum (dark traces) and the notch (light traces) noise stimuli. Each pair of traces is for a different stimulus intensity. The noisiness of the patterns reflects the noisiness of the stimulus spectrum (Fig. 1A). To facilitate the visual distinction between the patterns for the two noises, Fig. 4C illustrates five-point running averages of the original patterns.

Unlike the excitation patterns in Fig. 4A, the modulation-rate patterns appear as *low*-pass at all levels. This reflects the low-pass transfer characteristics of IHCs, which are approximately intensity independent [34]. Note that despite its high frequency, the spectral notch is still observed in the patterns, particularly at the lowest and highest intensities (zoomed view).

Fig. 4D shows the normalized difference modulation-rate patterns (i.e. $\log_{10}[S_0(f)/S(f)]$) for the seven intensities considered. Clearly, the difference is smaller at mid intensities than at lower or higher intensities. Furthermore, significant differences occur for modulation rates outside those corresponding to the notch frequency band (denoted by vertical dashed lines). Specifically, negative differences occur for modulation rates adjacent to the notch frequency band that are almost certainly due to suppression introduced by the DRNL filters [25]. Additionally, positive differences occur for modulation rates exactly an octave above the notch frequency band that relate to IHC half-wave rectification. At the highest intensities, positive differences occur over a range of modulation rates much wider than the notch frequency band. In the following text, it will be discussed that these differences relate to inherent IHC nonlinearities.

3.3. Spectral-discrimination sensitivity

The sensitivity for discriminating between the flat-spectrum and the notch noises was assumed to be directly proportional to the areas under the *squared* difference excitation and modulation-rate patterns in Fig. 4B and D [Eqs. (1) and (2)]. Fig. 5 shows how these two sensitivity measures depend on sound intensity. The sensitivity based on the difference excitation pattern (S_{EP}) decreases monotonically with increasing sound intensity (triangles in Fig. 5). By contrast, the sensitivity based on the difference modulation-rate pattern (S_{MP}) is a nonmonotonic function of intensity (circles in Fig. 5). S_{MP} is smallest at mid-levels and higher at the lowest/highest intensities tested. Only S_{MP} is consistent overall with the results of Alves-Pinto and Lopez-Poveda [1] (Fig. 2).

The model fails to account for the tendency of threshold notch depth to increase again with increasing level near 100 dB SPL (Fig. 2). This increase could relate to post-IHC (possibly synaptic) mechanisms, which are not included in the model.

3.4. Why does S_{EP} decrease monotonically with increasing sound intensity?

S_{EP} decreases gradually with increasing sound intensity due to a combination of three effects: (1) the gradual broadening of the DRNL filters with increasing level

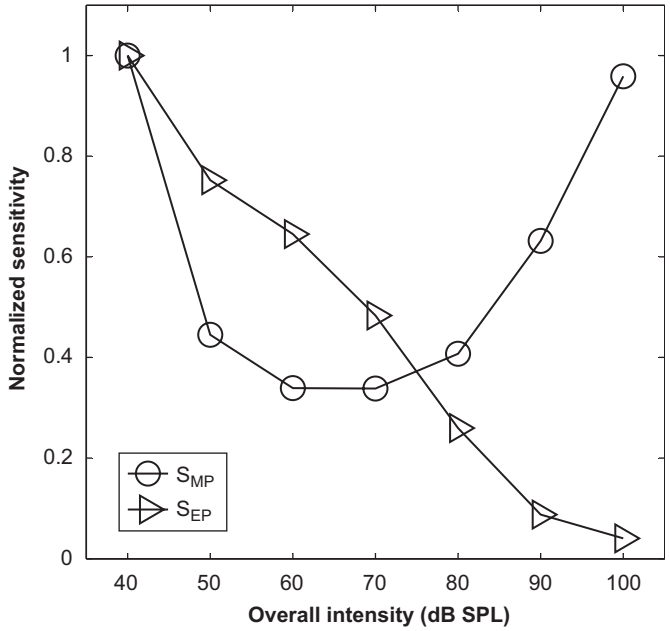

Fig. 5. Normalized sensitivity to spectral notches as a function of overall sound intensity for the two models tested, S_{EP} and S_{MP}.

(cf. Fig. 7 in Ref. [22]), (2) the compression applied by the DRNL filters at moderate to high intensities (cf. Fig. 4 in [22]), (3) the saturation of the IHC receptor potential at high intensities (cf. Fig. 12 in Ref. [21]).

3.5. Why does S_{MP} vary nonmonotonically with sound intensity?

This question is particularly important, because it may clarify why spectral discrimination improves at high levels in the results of Alves-Pinto and Lopez-Poveda [1]. To answer it, S_{MP} was evaluated for four different model configurations with different nonlinear characteristics. The one considered so far (Fig. 3) was denoted DRNL/IHC. In this configuration, both the DRNL and the IHC models were nonlinear (see Ref. [21,22]). In another configuration (denoted LIN/IHC), the DRNL filter-bank was replaced by a bank of *linear* asymmetric gammatone filters [8]. This was obtained from the DRNL filter-bank simply by setting its parameter '*a*' to zero (see Ref. [22]). The resulting filter-bank may be thought of as the cochlear filter-bank for a listener with total outer-hair-cell dysfunction, hence without the "active mechanism" [32]. Only the IHC was nonlinear in this configuration. In a third configuration (denoted DRNL/HWR), the IHC model was replaced by a half-wave rectification stage followed by a first-order Butterworth low-pass filter (cut-off frequency of 1400 Hz). The idea was to preserve the low-pass, half-wave rectification characteristics of the IHC without the nonlinearities associated with the IHC transduction; specifically, without the expansive and compressive growth of the dc receptor potential with increasing intensity that occurs at low and high intensities, respectively [21]. There were two sources of nonlinearity in this configuration: the DRNL filter-bank and the half-wave rectification. The last model configuration (denoted LIN/HWR) consisted of a linear filter-bank followed by a half-wave rectification/low-pass filter stage, thus the half-wave rectification was the only nonlinearity in the model.

Fig. 6 illustrates the resulting normalized S_{MP} vs. level functions for the four model configurations as indicated in the legend. The figure shows some striking results. First, the sensitivity is intensity independent for the linear-like model (LIN/HWR; triangles in Fig. 6). This suggests that it is the nonlinearities (other than that associated with the half-wave rectification), which cause spectral discrimination sensitivity to be intensity-dependent. Second, when the DRNL filter-bank is the only compressive nonlinear stage (DRNL/HWR), sensitivity decreases with increasing intensity up to 80 dB SPL and then increases slightly. In other words, the filter-bank compressive nonlinearity causes sensitivity to decrease for the most part of the intensity range. Third, when the IHC is the only source of nonlinearity (LIN/IHC), the sensitivity increases rapidly with increasing intensity above 60 dB SPL. Indeed, maximum sensitivity occurs at the highest intensity considered, at which the saturation of the IHC receptor potential is maximum.

Combined, these results indicate that (1) basilar membrane compression is responsible for the decrease in sensitivity with increasing intensity up to around 60–70 dB SPL, and (2) the improvement in spectral discrimination at high intensities reported by Alves-Pinto and Lopez-Poveda [1] is strongly associated with inherent IHC nonlinearities. How this might happen is discussed in the following section.

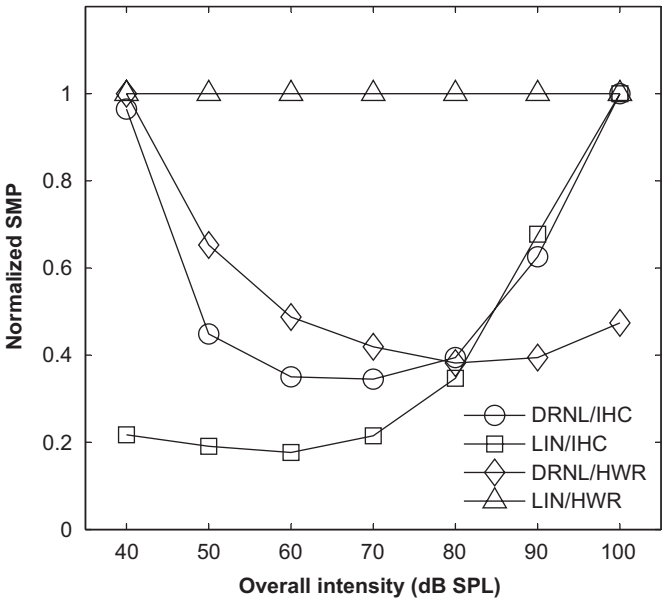

Fig. 6. Intensity dependence of the discrimination sensitivity based on the modulation rate pattern representation of the spectrum for different model configurations with different sources of compression. See main text for details.

4. Discussion

We aimed to clarify how high-frequency spectral features are represented in the auditory nerve. Our approach consisted in modeling the paradoxical result of Alves-Pinto and Lopez-Poveda. They found that discriminating between broadband noises with and without high-frequency spectral notches is more difficult at 70–80 dB SPL than at lower or higher intensities. We tested two possibilities: (a) that discrimination is based on the difference between the IHCs excitation patterns for the two stimuli, a representation related to the auditory nerve difference rate profile; and (b) that it is based on the difference between IHC modulation-rate patterns, a representation related to the difference in the temporal pattern of auditory nerve discharges. The results support the latter. The results also indicate that the improvement at high intensities in the spectral discrimination task of Alves-Pinto and Lopez-Poveda [1] is associated with inherent IHC nonlinearities different from mere half-wave rectification.

4.1. Model limitations

The highest CF in our filter-bank (20 kHz) was above the highest CF (8 kHz) for which the filter-bank of Lopez-Poveda and Meddis was designed. Therefore, the model representations of spectral features above around 8–9 kHz may be inaccurate to some (uncertain) extent. It is unlikely, however, that this affected the representation of the notch, as it extended from 6 to 8 kHz.

We have assumed that the IHC model of Lopez-Poveda and Eustaquio-Martín [21] is adequate to simulate *human* IHCs. Additionally, we have assumed that its parameters have identical values across CFs, even though this is not the case. At least the endocochlear potential and the relative size of two of the IHC ionic currents are known to vary along the length of the cochlea [6,17]. The impact of these assumptions on the IHC representations of the stimulus spectrum is uncertain.

The modulation-rate pattern model (Fig. 5, circles) suggests that spectral-discrimination sensitivity is lowest at 60–70 dB SPL, while according to the data of Alves-Pinto and Lopez-Poveda [1] (Fig. 2) it should be the lowest at 70–80 dB SPL. This 10-dB discrepancy may be due to a mismatch between the model and experimental SPL calibration. Alternatively, the gain of one or several model stages may be inadequate. We made no attempt to vary the original model parameters to optimize the fit between model and experimental results. Indeed, the 10-dB discrepancy could be solved simply by reducing the middle-ear gain by 10 dB (results not shown).

4.2. Implications of the results

The quality of the IHC excitation pattern representation of the spectral notch, and hence the associated discrimination sensitivity, decreases gradually with increasing sound intensity (Figs. 4A and 5). This implies that the quality of the auditory nerve rate profile must necessarily decrease with increasing intensity, regardless of the threshold and dynamic range of ANFs. This undermines the conjecture put forward by Alves-Pinto and Lopez-Poveda [1] and Alves-Pinto et al. [2] that the peaks in the behavioral threshold notch depth vs. level functions (Fig. 2) reflect the transition between the dynamic ranges of ANFs with low and high thresholds. This is not to negate the existence of two types of ANFs in the human auditory nerve or to challenge that two fiber types with different properties can account for other psychoacoustical phenomena (e.g. Refs. [38,39,41]).

The difference between the IHC modulation-rate pattern representations for the flat-spectrum and notch noises decreases with increasing intensity up to around 60–70 dB SPL and increases again at higher intensities (Fig. 4D). Therefore, the associated spectral-discrimination sensitivity is a nonmonotonic function of intensity (circles in Fig. 5), which is consistent with the observations of Alves-Pinto and Lopez-Poveda [1]. In other words, the information available in the IHC receptor potential *waveforms* is sufficient to explain (at least qualitatively) the paradoxical results of Alves-Pinto and Lopez-Poveda.

Auditory nerve discharges follow IHC receptor potential waveforms in a probabilistic manner, with absolute and relative refractory periods. Therefore, the presence of high-frequency information in the IHC receptor potential waveform does not guarantee that the *same* information be available in corresponding auditory nerve spike trains. However, the remaining information might be sufficient and useful. On the one hand, recent studies demonstrate that despite the rapid roll-off of phase locking with increasing frequency above 4 kHz, significant phase locking still occurs in mammals for frequencies as high as 14 kHz [30,37]. On the other hand, the high-frequency information in the IHC receptor potential waveform might be conveyed to the central auditory system even in the absence of auditory nerve phase locking to high frequency stimuli. Indeed, it has been mathematically proven that any signal which is sparse in either the frequency or the time domain may be reconstructed to a useful extent based on signal samples taken at random times [9]; a principle that has been successfully applied to the problem of identifying some (noisy) sounds [3]. It is also known that it is possible to obtain a statistical approximation of the signal spectrum from unevenly sampled data (see pp. 575–584 in Ref. [29]).

In other words, with or without phase locking to high-frequency stimuli, it is possible that the high-frequency information in the IHC waveforms might be recovered to some extent from auditory nerve discharges. This implies that spectral information up to 14 kHz may be encoded in the timing of auditory nerve spikes and that the discrimination task of Alves-Pinto and Lopez-Poveda [1] may be based on comparisons of the internal representations of the spectra obtained by precise analysis of the timing of auditory nerve spikes, at least above 70–80 dB SPL. Heinz

et al. [13] reached a similar conclusion based on a computational-modeling study of the just-noticeable difference in frequency between two pure tones.

The actual mechanism that would allow the central auditory system to extract the stimulus spectrum from the ANF discharges is unknown but it might be similar in effect to a Fourier transform. Indeed Young and Sachs [40] showed that a method of this kind allows recovering vowel formants from ANF responses even at very high intensities at which most fibers are saturated. One could argue that the results of Young and Sachs are not surprising because vowel formants are in the low-frequency region, where phase locking is strong, and significant phase locking occurs even when the fibers discharge at saturation [7,16,33]. What is most interesting, perhaps, is that the present results suggest that a similar mechanism may apply to *nonperiodic*, *high*-frequency stimuli.

4.3. What is the reason for the nonmonotonic aspect of the threshold notch depth vs. level function?

It has been shown (Fig. 6) that the improvement in spectral discrimination sensitivity at high intensities almost certainly relates to inherent IHC nonlinearities different from mere half-wave rectification. The actual mechanism is uncertain but may be related to the saturation of the receptor potential at high intensities. The saturation of the IHC transducer current would alter the receptor potential waveform at high intensities relative to the corresponding waveform at low intensities. For example, a perfectly sinusoidal receptor-potential waveform at low intensities would become akin to a square waveform at high intensities. The receptor potential waveform drives neurotransmitter release from the IHC to the synapses between the cell and its neighbor ANFs. As a result, the latter would show different patterns of phase-locked auditory-nerve spikes at low and high intensities. Indeed, assuming perfect phase-locking, at low intensities spikes would occur with a periodicity equal to that of the driving sinusoidal receptor potential. At high intensities, by contrast, spikes would occur not only with the frequency of the sinusoidal waveform, but also with those of its harmonics. Of course, perfect phase locking is unlikely to occur, but the principle would still apply.

The nonlinearity associated with the IHC transducer current occurs instantaneously (at least in the present model). Therefore, its effects would also alter the shape of complex receptor potential waveforms like the ones produced by the flat-spectrum and notch noises considered here. Consider, now an IHC with a CF equal to the notch center frequency. At low intensities, both the flat-spectrum and the notch noises would drive the IHC into a linear region. Discriminating between the two would be purely based on the amplitude of the receptor potential produced (or equivalently on the associated number of spikes) evoked by the two noise stimuli. At high intensities, however, the flat-spectrum noise would drive the IHC in question more into saturation than the less-energetic notch noise stimulus. As a result, discriminating between the two stimuli could be based not only on the amplitude of the receptor potential (or the number of spikes), but also on the timing of the spikes. The spike trains in response to the flat-spectrum noise would contain distortion frequencies not present (or present with less amplitude) in the spike trains evoked by the notch noise. In other words, harmonic distortion associated with the saturation of the receptor potential could explain why the differences between the modulation-rate patterns in Fig. 4D extend to frequencies outside the notch band at high intensities.

The effect just described would affect primarily IHCs (or ANFs) with CFs within the notch frequency band, but would not be restricted to them. IHCs (or ANFs) with CFs above the notch band would have a broad-enough frequency response at high intensities to be affected by the energy difference between the two stimuli and translate this into spike trains with different high-frequency timing.

This proposed mechanism is physiologically plausible because isolated IHCs introduce distortion [14].

5. Conclusions

A computational model has been used to simulate two different representations of the spectrum of broadband sounds at the level of the IHC: the excitation pattern and the modulation-rate pattern. Only the latter appears consistent with the observation of Alves-Pinto and Lopez-Poveda that it is more difficult to discriminate between broadband noise stimuli with and without a high-frequency spectral notch at 70–80 dB SPL than at lower or higher intensities. This suggests that discrimination between two sounds with different high-frequency spectra relies on differences in the timing of auditory-nerve spikes, rather than on spike rates, evoked by the two stimuli. Additionally, the simulations suggest that the improvement in spectral discrimination above 70 dB SPL is associated with inherent IHC nonlinearities, possibly due to the saturation of the receptor potential at high intensities.

Acknowledgments

We thank two anonymous reviewers for their insightful comments. Work supported by the Spanish Ministry of Education and Science (PROFIT ref. CIT-390000-2005-4 and MEC ref. BFU2006-07536) and by IMSERSO (ref. 131/06).

References

[1] A. Alves-Pinto, E.A. Lopez-Poveda, Detection of high-frequency spectral notches as a function of level, J. Acoust. Soc. Am. 118 (2005) 2458–2469.

[2] A. Alves-Pinto, E.A. Lopez-Poveda, A.R. Palmer, Auditory nerve encoding of high-frequency spectral information, in: Proceedings of IWINAC 2005, Lecture Notes in: Computer Science, vol. 3561, Springer, Berlin, 2005, pp. 223–232.
[3] R. Arora, R.A. Lutfi, Sound source identification by compressive sensing, J. Acoust. Soc. Am. 121 (2007) 3184.
[4] S. Carlile, R. Martin, K. McAnally, Spectral information in sound localization, Int. Rev. Neurobiol. 70 (2005) 399–434.
[5] M.A. Cheatham, P. Dallos, Inner hair cell response patterns: implications for low-frequency hearing, J. Acoust. Soc. Am. 110 (2001) 2034–2044.
[6] J.W. Conley, M.L. Bennett, Turn-specific differences in the endocochlear potential between albino and pigmented guinea pigs, Hear. Res. 65 (1993) 141–150.
[7] N.P. Cooper, D. Robertson, G.K. Yates, Cochlear nerve fiber responses to amplitude modulated stimuli: variations with spontaneous rate and other response characteristics, J. Neurophysiol. 70 (1993) 370–386.
[8] E. de Boer, Synthetic whole-nerve action potentials for the cat, J. Acoust. Soc. Am. 58 (1975) 1030–1045.
[9] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (2006) 1289–1306.
[10] E.F. Evans, The frequency response and other properties of single fibers in the guinea-pig cochlear nerve, J. Physiol. (London) 226 (1972) 263–287.
[12] K.W. Grant, B.W. Walden, Spectral distribution of prosodic information, J. Speech Hear. Res. 39 (1996) 228–238.
[13] M.G. Heinz, H.S. Colburn, L.H. Carney, Evaluating auditory performance limits. I. One-parameter discrimination using a computational model for the auditory nerve, Neural Comput. 13 (2001) 2273–2316.
[14] F. Jaramillo, V.W. Markin, A.J. Hudspeth, Auditory illusions and the single hair cell, Nature 364 (1993) 527–529.
[15] D.H. Johnson, The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones, J. Acoust. Soc. Am. 68 (1980) 1115–1122.
[16] P.X. Joris, T.C. Yin, Responses to amplitude-modulated tones in the auditory nerve of the cat, J. Acoust. Soc. Am. 91 (1992) 215–232.
[17] T. Kimitsuki, K. Kawano, K. Matsuda, A. Haruta, T. Nakajima, S. Komune, Potassium current properties in apical and basal inner hair cells from guinea-pig cochlea, Hear. Res. 180 (2003) 85–90.
[18] M.C. Liberman, Auditory-nerve response from cats raised in a low-noise chamber, J. Acoust. Soc. Am. 63 (1978) 442–455.
[19] E.A. Lopez-Poveda, The physical origin and physiological coding of pinna-based spectral cues, Ph.D. Thesis, Department of Human Sciences, Loughborough University, UK, 1996.
[20] E.A. Lopez-Poveda, Spectral processing by the peripheral auditory system: facts and models, Int. Rev. Neurobiol. 70 (2005) 7–48.
[21] E.A. Lopez-Poveda, A. Eustaquio-Martín, A biophysical model of the inner hair cell: the contribution of potassium current to peripheral compression, JARO—J. Assoc. Res. Otolaryngol. 7 (2006) 218–235.
[22] E.A. Lopez-Poveda, R. Meddis, A human nonlinear cochlear filterbank, J. Acoust. Soc. Am. 110 (2001) 3107–3118.
[23] E.A. Lopez-Poveda, A. Alves-Pinto, A. Palmer, Psychophysical and physiological assessment of the representation of high-frequency spectral notches in the auditory nerve, in: B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, J. Verhey (Eds.), Hearing: From Sensory Processing to Perception, Springer, Heidelberg, 2007.
[24] B.J. May, Physiological and psychophysical assessments of the dynamic range of vowel representations in the auditory periphery, Speech Commun. 41 (2003) 49–57.
[25] R. Meddis, L.P.O. O'Mard LPO, E.A. Lopez-Poveda, A computational algorithm for computing non-linear auditory frequency selectivity, J. Acoust. Soc. Am. 109 (2001) 2852–2861.
[26] A.R. Palmer, E.F. Evans, Cochlear fibre rate-intensity functions: no evidence for basilar membrane nonlinearities, Hear. Res. 2 (1980) 319–326.
[27] A.R. Palmer, I.J. Russell, Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells, Hear. Res. 24 (1986) 1–15.
[28] R. Patuzzi, P.M. Sellick, A comparison between basilar membrane and inner hair cell receptor potential input–output functions in the guinea pig cochlea, J. Acoust. Soc. Am. 74 (1983) 1734–1741.
[29] W.H. Press, S.A. Teukolsky, W.T. Vettering, B.P. Flanney, Numerical Recipes in C, The Art of Scientific Computing, second ed., Cambridge University Press, New York, 2002.
[30] A. Recio-Spinoso, A.N. Temchin, P. van Dijk, Y.-H. Fan, M.A. Ruggero, Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers, J. Neurophysiol. 93 (2005) 3615–3634.
[31] J.J. Rice, E.D. Young, G.A. Spirou, Auditory-nerve encoding of pinna-based spectral cues: rate representation of high-frequency stimuli, J. Acoust. Soc. Am. 97 (1995) 1764–1776.
[32] L. Robles, M.A. Ruggero, Mechanics of the mammalian cochlea, Physiol. Rev. 81 (2001) 1305–1352.
[33] J.E. Rose, J.F. Brugge, D.J. Anderson, J.E. Hind, Phase-locked response to low-frequency tones in single auditory nerve fibers of squirrel monkey, J. Neurophysiol. 30 (1967) 769–793.
[34] I.J. Russell, P.M. Sellick, Intracellular studies of hair cells in the mammalian cochlea, J. Physiol. 284 (1978) 261–290.
[35] S.A. Shamma, R.S. Chadwick, W.J. Wilbur, K.A. Morrish, J. Rinzel, A biophysical model of cochlear processing: intensity dependence of pure tone responses, J. Acoust. Soc. Am. 80 (1986) 133–145.
[36] A. Simpson, H.J. McDermott, R.C. Dowell, Benefits of audibility for listeners with severe high-frequency hearing loss, Hear. Res. 210 (2005) 42–52.
[37] A.N. Temchin, A. Recio-Spinoso, P. van Dijk, M.A. Ruggero, Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and frozen noise and comparison to basilar-membrane vibrations, J. Neurophysiol. 93 (2005) 3635–3648.
[38] N.F. Viemeister, Intensity discrimination at high frequencies in the presence of noise, Science 221 (1983) 1206–1208.
[39] N.F. Viemeister, Intensity coding and the dynamic-range problem, Hear. Res. 34 (1988) 267–274.
[40] E.D. Young, M.B. Sachs, Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers, J. Acoust. Soc. Am. 66 (1979) 1381–1403.
[41] F.-G. Zeng, C.W. Turner, E.M. Relkin, Recovery from prior stimulation. II. Effects upon intensity discrimination, Hear. Res. 55 (1991) 223–230.

Enrique A. Lopez-Poveda received his B.Sc. in Physics from the University of Salamanca (Spain) in 1993 and his Ph.D. in Human Sciences from Loughborough University (UK) in 1996. He is a "Ramón y Cajal" fellow and director of the Auditory Computation and Psychoacoustics Unit of the Institute of Neurosciences of Castilla y León, at the University of Salamanca (Spain). His research interests are computational modelling of the auditory system, psychoacoustics, speech processors and auditory implants.

Ana Alves-Pinto received her B.Sc. in Physics in 2000 from the University of Lisbon (Portugal) and her Neuroscience Ph.D. in 2007 from the University of Salamanca (Spain). She is presently working as a Research Fellow at the Medical Research Council Institute for Hearing Research (Nottingham UK).

Alan Palmer received his first degree in Biological Sciences from the University of Birmingham UK in 1972 and his Ph.D. in Communication and Neuroscience from the University of Keele UK in 1977. Following four years at the National Institute for Medical Research in London he was a Royal Society University Research Fellow before taking up his current position in 1986. He is currently Assistant Director of the Medical Research Council Institute for Hearing Research and holds a Special Professorship in Neuroscience at the University of Nottingham UK.

Almudena Eustaquio-Martín received her B.Sc. in Physics in 1994 from the University of Salamanca (Spain) and her M.Phil. in Electronic Engineering from Loughborough University in 1995. For four years she worked as Research Assistant in the Speech and Hearing Laboratory of Loughborough University (UK); as Multimedia Engineer for Sistemas de Imagen y Palabra S.A. (Spain); and as Project Engineer for GMV. S.A. (Spain). In 1998, she became Subdirector of the Medical Education Unit at the University of Castilla-La Mancha (Spain). She took her current position as Research Assistant at the Auditory Computation and Psychoacoustics Laboratory of the University of Salamanca (Spain) in 2004.