



## THE REPRESENTATION OF HIGH-FREQUENCY SPECTRAL NOTCHES IN THE PERIPHERAL AUDITORY SYSTEM: A COMPUTATIONAL MODEL STUDY

PACS: 43.64.Bt

Lopez-Poveda, Enrique A.<sup>1</sup>; Alves-Pinto, Ana<sup>1</sup>; Palmer, Alan R.<sup>2</sup>; Eustaquio-Martín, A.<sup>1</sup>

<sup>1</sup>Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Avda. Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain; [ealopezpoveda@usal.es](mailto:ealopezpoveda@usal.es)

<sup>2</sup>MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, United Kingdom.

### ABSTRACT

Discriminating between broadband noises with and without high-frequency spectral notches appears to be more difficult at 70-80 dB SPL than at lower or higher intensities [Alves-Pinto and Lopez-Poveda (2005) J. Acoust. Soc. Am. 118: 2458-2469]. The present study attempts to shed light on this paradoxical observation using a computational inner hair cell (IHC) model. Two possibilities were tested: (a) that discrimination relies on the difference between the IHC excitation patterns of the two stimuli, a representation related to the auditory nerve difference rate profile; and (b) that it relies on the difference IHC modulation-rate pattern, a representation related to the difference in the temporal pattern of auditory nerve discharges. The simulations support the latter. This suggests that spectral features as high in frequency as 8 kHz may be encoded in the temporal pattern of auditory nerve responses despite the rapid decay of phase locking above 2 kHz. Work supported by the Spanish Ministry of Education and Science (CIT-390000-2005-4 and BFU2006-07536) and by IMSERSO (131/06).

### INTRODUCTION

There are at least two ways in which the spectrum of a sound may be recovered from the output signals 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 will refer to this representation as the *excitation pattern*. An alternative way would be to estimate the spectra of the output signals from all filters and add them together to obtain the total filter-bank output spectrum. We will refer to this representation as the *modulation-rate pattern*. For a linear filter-bank, both the excitation and the modulation-rate patterns resemble the coarse stimulus spectrum, assuming a large number of overlapping filters.

The auditory nerve may be regarded as a bank of filters with characteristic frequencies (CFs) spanning approximately the frequency range of hearing (reviewed by [6]). Therefore, in principle the central auditory system could obtain the stimulus spectrum from the response of auditory nerve fibers (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 [14]; hence it 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 notches in the auditory nerve.

Phase locking rolls off rapidly with increasing frequency above 2 kHz due to low-pass filtering in the inner hair cells (IHCs) [12]. Consequently, it is widely accepted that high-frequency spectral features (above ~4 kHz) are likely to be encoded in the rate profile of ANFs solely ([14];[10];[5]). On the other hand, the frequency response of cochlear filters, and thus of ANFs, broadens gradually with increasing sound intensity and several important peripheral auditory elements (the basilar membrane, the IHC, and the auditory nerve) have strongly-compressive input/output characteristics (reviewed by [6]). 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. As a result, the quality of the rate profile representation of high-frequency spectral notches deteriorates gradually with increasing intensity ([5];[14]; but see [10]).

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 does not seem to be the case. Alves-Pinto and Lopez-Poveda [1] have shown 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.

The present study aimed at investigating 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. 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 discrimination results. A more comprehensive report on this work has been submitted for publication elsewhere [7].

## 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 auditory-nerve representation of the spectrum cannot be better than observed at the level of the IHCs. Second, the IHC model outputs a continuous, deterministic signal (the IHC receptor potential), which is easier to analyze than trains of stochastic spikes.

### The model

The model consisted of four signal-processing stages, each of which represents a physiological stage of the peripheral auditory system (Fig. 1). The input to the model was a digital sound waveform (in units of Pascal) assumed to represent an auditory stimulus at the eardrum. The model output was a collection of waveforms representing the receptor potential for a bank of IHCs with different CFs.

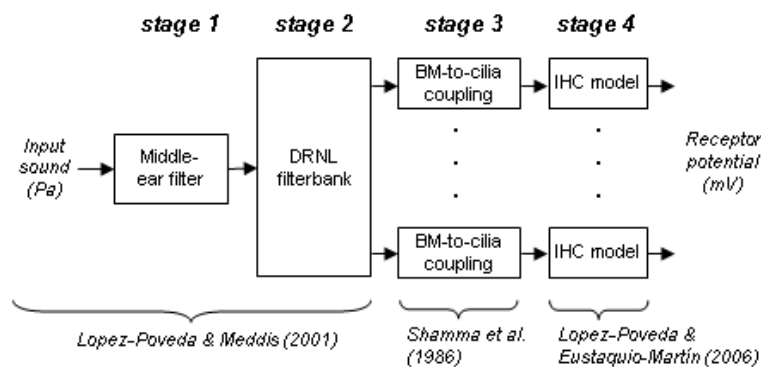


Figure 1. Model structure.

Stage 1 simulated the middle-ear transfer function. Its output was the stapes velocity (in units of m/s). Stage 2 was a bank of 100 dual-resonance (DRNL) nonlinear filters (CFs from 100 Hz to 20 kHz) that simulated human cochlear frequency selectivity [9]. 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. Stage 3 was a high-pass filter (scalar gain = 2; cut-off frequency = 530 Hz) that coupled basilar-membrane displacement to the displacement of IHC stereocilia (in units of m) (Eq. 1 in [16]). Stage 4 was a biophysical model of the IHC that

outputs the IHC receptor potential (in volts) as a function of the cell's stereocilia displacement [8]. Details on these models can be found in the original publications.

### Evaluations

The model was implemented in Matlab and evaluated in the time domain (sampling frequency = 44100 Hz) in response to broadband (20 Hz to 16 kHz) noise bursts with flat-spectrum and with a rectangular spectral notch centered at 7 kHz. The notch had a bandwidth of 2 kHz and a depth of 15 dB re the spectrum level on the notch side bands. The spectra of the two noises were identical except for the amplitude in the notch band. 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 stimuli were almost identical to those used by Alves-Pinto and Lopez-Poveda [1] and were generated as described by these authors.

### 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 excitation pattern was obtained by plotting the average receptor potential of each IHC as a function of the cell's CF (Fig. 2A). This was calculated by summing the samples in the receptor potential waveform with amplitudes greater than zero and dividing the result by the total number of samples in the waveform [2]. The modulation-rate pattern was obtained by applying a fast Fourier transform (FFT) to the receptor potential waveform of each IHC and adding the resulting spectra of all IHCs in the frequency domain (Fig. 2C). 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 Figs. 2A and 2C.

## RESULTS

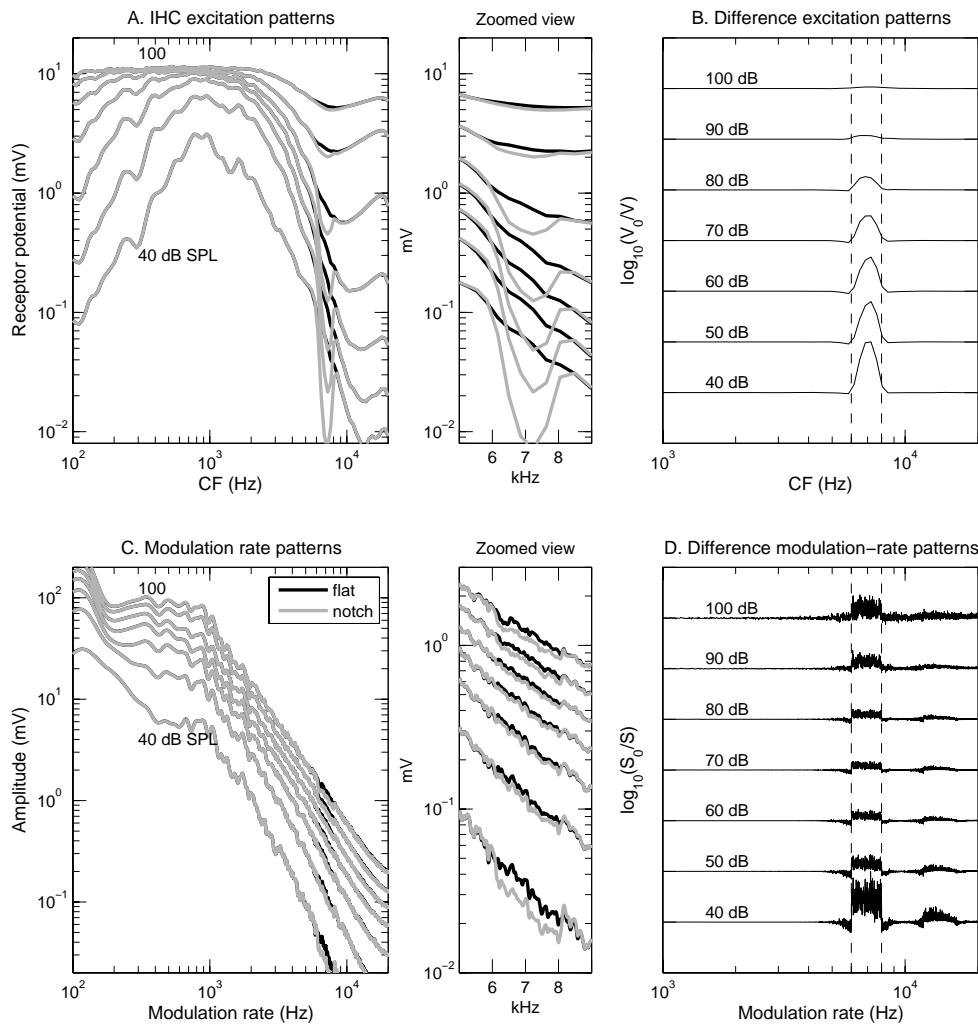
### Excitation patterns

Figure 2A 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 [9]). The pattern becomes almost all-pass at high stimulus intensities. This is due to a combination of basilar-membrane compression (cf. Fig. 4 in [9]) and the saturation of the IHC receptor potential at high intensities (cf. Figs. 9 and 10 in [8]). Figure 2A 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. 2B. The latter illustrates the difference excitation patterns (i.e.,  $\log_{10}[V_{\text{flat}}(\text{CF})/V_{\text{notch}}(\text{CF})]$ ) 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).

### Modulation-rate patterns

Figure 2C 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. To facilitate the visual distinction between the patterns for the two noises, Fig. 2C illustrates 5-point running averages of the original patterns.

Unlike the excitation patterns of Fig. 2A, 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 [15]. Note that despite its high frequency, the spectral notch is still observed in the patterns, particularly at the lowest and highest intensities (zoomed view).



**Figure 2.** **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 [ $\log_{10}(V_{\text{flat}}/V_{\text{notch}})$ ] 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 [ $\log_{10}(S_{\text{flat}}/S_{\text{notch}})$ ] normalized to the maximum value across modulation-rates and intensities. The vertical dashed lines in B and D indicate the notch frequency band.

Figure 2D shows the normalized difference modulation-rate patterns (i.e.,  $\log_{10}[S_{\text{flat}}(f)/S_{\text{notch}}(f)]$ ) 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 the DRNL filters [11]. 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 a more comprehensive report [7], we provide evidence that these differences relate to inherent IHC nonlinearities.

## DISCUSSION

The quality of the IHC excitation pattern representation of the spectral notch, and hence the hypothetical associated discrimination sensitivity, decreases gradually with increasing sound

intensity (Fig. 2A). Average auditory nerve discharge rates are proportional to the average receptor potential [2]. Therefore, 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] that the peak in the behavioral threshold notch depth vs. level function reflects the transition between the dynamic ranges of ANFs with low and high thresholds. Furthermore, it challenges the notion that high-frequency spectral notches must be conveyed to central auditory system in the average-rate profile of ANFs [10; 14].

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. 2D). Therefore, the associated spectral-discrimination sensitivity is a non-monotonic function of intensity, which would be consistent with the observations of Alves-Pinto and Lopez-Poveda. Since auditory nerve discharge follows the IHC receptor potential, this supports two important conjectures. First, spectral discrimination may be based on comparisons of internal representations of the spectra obtained by precise analysis of the timing of auditory nerve spikes. The actual mechanism that allows the central auditory system to extract such a representation is unknown but may be similar in effect to a Fourier transform. Other authors have suggested that this could be the case for low-frequency periodic stimuli [17]. What is most interesting, perhaps, is that the present results suggest that a similar mechanism may apply to nonperiodic, high-frequency stimuli. This leads us to the second conjecture.

The present results also indicate that the high-frequency spectrum (at least up to 8 kHz) is possibly encoded in the timing of auditory-nerve spike occurrences. Heinz et al. [3] reached a similar conclusion based on a computational model of the just-noticeable difference in frequency between two pure tones. The present study extends the conclusion of Heinz *et al.* to broad-band, nonperiodic stimuli. Whatever the stimuli, however, this would be possible only if ANFs fired synchronously with modulation rates of their driving receptor potential waveforms higher than 4 kHz. While this seemed improbable in view of early reports [12], recent studies demonstrate that significant phase locking occurs for frequencies as high as 14 kHz [13]. Therefore, the present conclusions as well as those of Heinz et al. are physiologically plausible.

#### **What is the reason for the non-monotonic aspect of the threshold notch depth vs. level function?**

In an extended version of this report [7], evidence is given 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.

The nonlinearity associated to 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 to the saturation of the receptor potential could explain why the differences between the modulation-rate patterns of Fig. 2D 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 [4].

## CONCLUSIONS

The present simulations suggest that discriminating 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.

## REFERENCES

1. Alves-Pinto A, Lopez-Poveda EA. Detection of high-frequency spectral notches as a function of level. *J. Acoust. Soc. Am.* 2005; 118 (4):2458-2469
2. Cheatham MA, Dallos P. Inner hair cell response patterns: implications for low-frequency hearing. *J. Acoust. Soc. Am.* 2001; 110 (4):2034-2044
3. Heinz MG, Colburn HS, Carney LH. Evaluating Auditory Performance Limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neural Computation* 2001; 13 (10):2273-2316
4. Jaramillo F, Markin VS, Hudspeth AJ. Auditory Illusions and the Single Hair Cell. *Nature* 1993; 364 (6437):527-529
5. Lopez-Poveda, E. A. The physical origin and physiological coding of pinna-based spectral cues. 1996. Ph.D. Thesis. Loughborough University. U.K.
6. Lopez-Poveda EA. Spectral processing by the peripheral auditory system: Facts and models. *Int. Rev. Neurobiol.* 2005; 70:7-48
7. Lopez-Poveda EA, Alves-Pinto A, Palmer AR, Eustaquio-Martín A. Rate versus time representation of high-frequency spectral notches in the peripheral auditory system: A computational modeling study. *Neurocomputing*. submitted 2007.
8. Lopez-Poveda EA, Eustaquio-Martín A. A biophysical model of the inner hair cell: The contribution of potassium currents to peripheral auditory compression. *JARO-J. Assoc. Res. Otolaryngol.* 2006; 7 (3):218-235
9. Lopez-Poveda EA, Meddis R. A human nonlinear cochlear filterbank. *J. Acoust. Soc. Am.* 2001; 110 (6):3107-3118
10. May BJ. Physiological and psychophysical assessments of the dynamic range of vowel representations in the auditory periphery. *Speech Communication* 2003; 41:49-57
11. Meddis R, O'Mard L, Lopez-Poveda EA. A computational algorithm for computing nonlinear auditory frequency selectivity. *J. Acoust. Soc. Am.* 2001; 109 (6):2852-2861
12. Palmer AR, Russell IJ. Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear. Res.* 1986; 24:1-15
13. Recio-Spinoso A, Temchin AN, van Dijk P, Fan YH, Ruggero MA. Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. *J. Neurophysiol* 2005; 93 (6):3615-3634
14. Rice JJ, Young ED, Spirou GA. Auditory-nerve encoding of pinna-based spectral cues: Rate representation of high-frequency stimuli. *J. Acoust. Soc. Am.* 1995; 97 (3):1764-1776
15. Russell IJ, Sellick PM. Intracellular studies of hair cells in the mammalian cochlea. *J. Physiol.* 1978; 284:261-290
16. Shamma SA, Chadwick RS, Wilbur WJ, Morrish KA, Rinzell J. A biophysical model of cochlear processing: Intensity dependence of pure tone responses. *J. Acoust. Soc. Am.* 1986; 80 (1):133-145
17. Young ED, Sachs MB. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *J. Acoust. Soc. Am.* 1979; 66 (5):1381-1403