

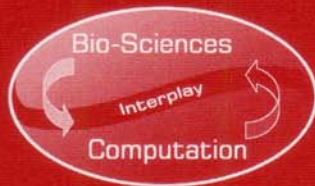
José Mira  
José R. Álvarez (Eds.)

LNC3 3561

# Mechanisms, Symbols, and Models Underlying Cognition

First International Work-Conference on the Interplay  
Between Natural and Artificial Computation, IWINAC 2005  
Las Palmas, Canary Islands, Spain, June 2005  
Proceedings, Part I

**1** Part I



 Springer

## Volume Editors

José Mira  
José R. Álvarez  
Universidad Nacional de Educación a Distancia  
E.T.S. de Ingeniería Informática  
Departamento de Inteligencia Artificial  
Juan del Rosal, 16, 28040 Madrid, Spain  
E-mail: {jmira, jras}@dia.uned.es

Library of Congress Control Number: 2005927486

CR Subject Classification (1998): F.1, F.2, I.2, G.2, I.4, I.5, J.3, J.4, J.1

ISSN	0302-9743
ISBN-10	3-540-26298-9 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-26298-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media  
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper SPIN: 11499220 06/3142 5 4 3 2 1 0

# Auditory Nerve Encoding of High-Frequency Spectral Information

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

<sup>1</sup> Instituto de Neurociencias de Castilla y León, Universidad de Salamanca,  
Avenida Alfonso X El Sabio s/n, 37007 Salamanca, Spain  
{aalvespinto, ealopezpoveda}@usal.es  
<http://web.usal.es/~ealopezpoveda/>

<sup>2</sup> MRC Institute of Hearing Research, University Park,  
Nottingham NG7 2RD, United Kingdom  
alan@ihr.mrc.ac.uk

**Abstract.** We have recently shown [1] that our ability to discriminate between a flat-spectrum noise and a similar noise with a high-frequency spectral notch deteriorates for levels around 70-80 dB SPL. The present paper explores an underlying physiological mechanism for this result. The hypothesis is that discriminability relies on the sensitivity of the auditory nerve to changes in the stimulus spectrum. A good correlation was found between behavioural results and sensitivity functions for two auditory nerve fibers with different dynamic ranges and with characteristic frequencies within the notch band. Although preliminary, these results suggest that the sensitivity of the auditory nerve to spectral notches is a non-monotonic function of stimulus level.

## 1 Introduction

Research on the functional properties of the peripheral auditory system is motivated, at least in part, by the need to develop better speech processors or robots with hearing capabilities. The underlying assumption is that any new knowledge on the functioning of the peripheral auditory system may contribute to developing better sound processors for artificial systems.

A physiological stage that is crucial for auditory perception is the auditory nerve (AN), since it is the only transmission path of auditory information from the peripheral to the central auditory system. Indeed, our ability to detect and employ the auditory information contained in the spectrum of an acoustic stimulus requires that the spectrum be adequately represented in the response of the AN.

In the AN, spectral features may be represented either in the temporal pattern of the fibers' activity or in terms of rate profiles, that is, as the distribution of activity over a population of AN fibers with different characteristic frequencies (CFs) [15]. Spectral features with frequencies above the cutoff of phase locking (>4000 Hz; [10]) must be represented in terms of AN discharge rate alone [13].

The majority of AN fibers are of the high spontaneous rate (HSR) type. These have low thresholds ( $<10$  dB SPL) and dynamic ranges of approximately 30-40 dB ([6], [14]). Fibers with low spontaneous rates (LSR) have higher thresholds ( $>15$  dB SPL) and wider dynamic ranges ( $\sim 50$ -60 dB; [6], [14]). The existence of these two fiber types with distinct thresholds and dynamic ranges have led several investigators ([5],[11],[13],[15]) to suggest that high-frequency spectral features could be encoded in the rate profile of HSR fibers at low levels and in that of LSR fibers at high levels. This would provide the auditory system with the ability to encode spectral features over a wide range of levels.

The present work investigates whether this applies to the encoding of high-frequency spectral notches. The study focuses on spectral notches because they are important cues for sound localization ([3], [4], [9]). Summarised is a psychophysical experiment [1] that shows that the threshold notch depth for discriminating between a flat-spectrum noise and a similar noise with a spectral notch at 8 kHz is a *non-monotonic* function of the stimulus level. Specifically, spectral discrimination deteriorates for levels around 70-80 dB SPL.

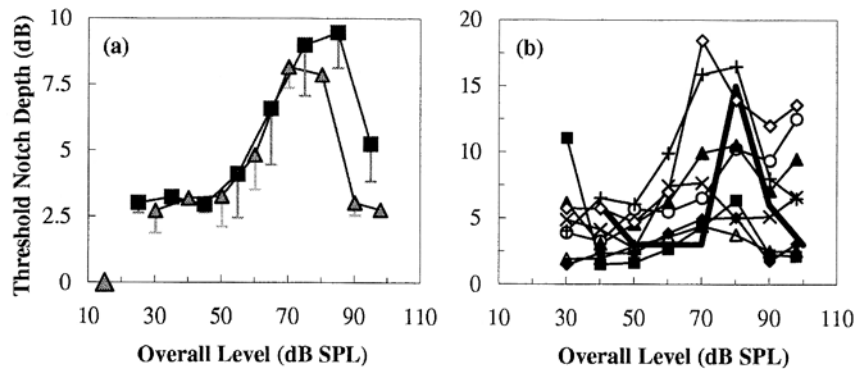
To elucidate the physiological mechanisms underlying this result, estimates of threshold notch depths based on guinea pig auditory nerve sensitivity to changes in the noise spectra are compared with the behavioral data. A strong correlation will be shown to exist between behavioral and physiological notch-depth thresholds. It is suggested that LSR fibers are responsible for the improvement in discrimination at very high levels, possibly as a result of suppression. The implications of these results for the design of bio-inspired speech processors are discussed.

## 2 Psychophysical Experiment

The ability to detect high-frequency spectral notches was studied by measuring the threshold notch depth necessary for listeners to discriminate between a broadband (20-16000 Hz) noise with a notch centered at 8000 Hz (target stimulus) and a similar noise with a flat spectrum (standard stimulus). The notch had a rectangular shape and a bandwidth of 4000 Hz. Threshold notch depths were measured for overall levels of the noise ranging from 30 to 100 dB SPL.

The presence of the notch in the target stimulus makes its overall level lower than that of the standard stimulus. To minimize for the possibility that listeners use this level difference as a cue for discrimination, the standard and the target noise bursts were presented with equal overall level. This was done by reducing the spectrum level of the flat spectrum noise as appropriate (e.g., a reduction of 0.58 dB would be required for a 2000-Hz wide, 27-dB deep notch). The details of the experimental procedure are provided elsewhere [1].

The main result is shown in Fig. 1(b). Despite the large variability of the data across listeners, the most significant feature is the non-monotonic variation of threshold notch-depth with stimulus level. As the stimulus level increases up to 70-80 dB SPL threshold notch depth also increases, meaning that spectral



**Fig. 1.** Threshold notch depths for discriminating between a flat-spectrum noise and a notch noise as a function of stimulus overall level. (a) Results for a single listener for two conditions (notch bandwidth = 2000 Hz): with equal levels for the standard and target stimuli (triangles); with levels being roved (squares). The triangle on the abscissa indicates the absolute threshold for the flat spectrum noise. (b) Threshold notch depths for eight listeners (notch bandwidth = 4000 Hz). Each symbol illustrates the results for a different listener. The continuous bold curve represents threshold notch depths computed from  $d'$  values based on the response of AN fibers to the flat-spectrum noise and the notch noise stimuli (Sec. 3.3)

discrimination becomes harder. However, for levels between 80-100 dB SPL, notch depth thresholds decrease with increasing level; that is, discrimination becomes easier.

A similar pattern was observed when the overall level of the noise was roved across trials (Fig. 1(a)). This suggests that discrimination must be based on comparing the AN rate profiles rather than by monitoring the discharge rate of specific AN fibers with CFs within the notch band. Furthermore, the non-monotonic variation of the threshold notch depth *vs.* level function suggests that the AN conveys clearer information for spectral discrimination at levels lower and higher than 70-80 dB SPL. The next section investigates the validity of this interpretation.

### 3 Auditory Nerve Response to High-Frequency Spectral Notches

The physiological mechanisms underlying the non-monotonic character of the threshold notch depth *vs.* level function (Fig. 1) were investigated by measuring the response of guinea pig AN fibers to the same stimuli used in the psychophysical experiment. Though the obvious separation between the two species (i.e., human and guinea pig) imposes a special care when establishing any relations between psychophysical and physiological results, the following data provide information that is fundamental to establish a consistent and reliable theory of how high-frequency spectral notches are encoded in the human AN.

### 3.1 Experimental Procedure

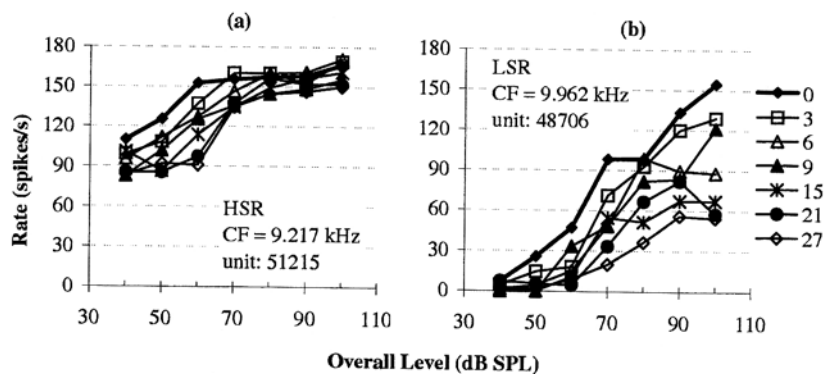
Stimuli consisted of wideband noise bursts (20-16000 Hz). Their spectra were either flat or contained a rectangular notch centered at a high frequency (7000 or 9000 Hz), with a bandwidth of 4000 Hz and with varying depth. The duration of the noise bursts was 220 ms. The notch center frequency was 7000 and 9000 Hz during the first and second halves of the stimulus duration respectively. The change between the two notch center frequencies was abrupt. The notch depths tested ranged from 3 to 9 dB in 3-dB steps and from 9 to 27 dB in 6-dB steps. Noise levels ranged from 40 to 100 dB SPL in steps of 10 dB. Stimuli were presented every 880 ms. Notch depth and noise level varied randomly between presentations. Unlike in the behavioral experiments, the level difference between the flat-spectrum and the notch noise bursts was not made equal. The surgical procedure and the recording methods are similar to those employed by [6].

AN responses were collected for at least four stimulus presentations. The results presented below are average discharge rates over the last 110 ms of the response, hence correspond to the portion of the stimulus with the spectral notch centered at 9000 Hz. The onset response is, therefore, excluded from the analysis.

### 3.2 Rate-Level Functions for HSR and LSR Fibers

Figure 2 illustrates rate *vs.* levels functions (RLFs) for two AN fibers, one HSR and one LSR. The CFs of both fibers are well within the notch band (7000-11000 Hz).

Let us first consider the RLFs for the flat-spectrum noise (filled diamonds in Fig. 2). Their shapes are clearly different for the two fibers. For the HSR fiber (Fig. 2(a)), it is representative of AN fibers of the saturating type [16]. The maximum discharge rate is reached at a level of approximately 20 dB above threshold. The RLF for the LSR fiber (Fig. 2(b)), however, clearly shows a wider



**Fig. 2.** Rate-level functions for the notch noise stimuli for two AN fibers with CFs well within the notch band. (a) HSR fiber; (b) LSR fiber. Symbols denote the notch depths in dB *re.* reference spectrum level

dynamic range and does not saturate over the range of levels tested. It is typical of the non-saturating type of AN fibers [16].

Since both fibers have CFs near the notch center frequency, they both effectively receive less energy from the notch noise stimulus than from the flat-spectrum noise stimulus. For any given notch depth, the energy difference between the two stimuli is the same regardless of the noise level. Therefore, one would expect this energy difference to shift the level at which their rate threshold occurs, but *not* the slope of the RLFs. In other words, one would expect that the effect of the notch be to shift the RLF in parallel to higher levels. This is clearly *not* the case. For the LSR fiber, the deeper the notch the shallower the associated RLF. As a result, the vertical distance between the RLF for any given notch depth and that for the flat-spectrum noise increases gradually as the overall level of the noise increases. As for the HSR fiber, its discharge rate also increases more slowly with increasing the noise level, mainly for deeper notches. Furthermore, saturation seems to occur at a lower rate and for a level higher than when stimulated with the flat-spectrum noise.

As the level of the wideband noise increases, fibers not only receive a stronger stimulation as a result of their frequency response areas becoming wider with increasing level, but they may also become gradually more affected by suppression effects [2]. Hence, it is possible that the differences between the RLFs for the flat-spectrum and the notch noise stimuli reflect the combined effect of suppression and energy reduction associated to the notch.

The rate reduction produced by the notch introduces a cue for discriminating between the flat-spectrum noise and the notch noise stimuli. The largest differences between the rates for the two stimuli occur between 40 and 60 dB SPL for the HSR fiber, and between 90-100 dB SPL for the LSR fiber (Fig. 2). This may be why spectral discrimination is easier over these ranges of level (Fig. 1). This idea is further explored in the following section.

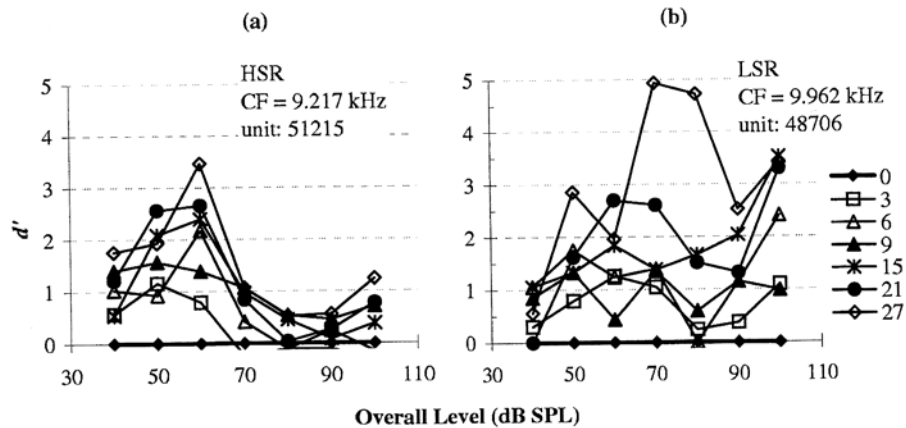
### 3.3 Neural Sensitivity to the Spectral Notch

Although the spectral notch produces a rate reduction in fibers with CFs within the notch band, it is difficult to evaluate whether the decrease is enough for discriminating between the flat-spectrum and the notch noise stimuli. A common numerical measure of the discriminability between two stimuli is the  $d'$  parameter (see, for example, [7] and [13]).  $d'$  values are computed as follows:

$$d' = \frac{R_0 - R}{\sqrt{0.5 \times (\sigma_0^2 + \sigma^2)}}$$

where  $(R_0, \sigma_0)$  and  $(R, \sigma)$  are the mean discharge rates and the associated standard deviations in response to the flat-spectrum and the notch noise stimuli, respectively, over a number of stimulus presentations.  $d'$  takes into account the variability in the response inherent to each fiber and, therefore, provides a more accurate measure of the fibers sensitivity to stimulus variations. Discrimination, whether behavioral or physiological, is considered to occur when  $|d'| \geq 1$ . There-





**Fig. 3.**  $d'$  values as a function of noise overall level, calculated from the discharge rates presented in Fig. 2, for the same (a) HSR and (b) LSR fibers. Symbols denote the notch depths in dB *re.* reference spectrum level

fore, the analysis of  $d'$  values makes it possible to compare psychophysical and physiological discrimination thresholds.

Figure 3 illustrates  $d'$  values as a function of noise overall level for the two fibers whose RLFs are illustrated in Fig. 2. For both HSR and LSR fibers,  $d'$  values change non-monotonically with noise overall level. The non-monotonic character of the function gets more obvious the deeper the notch.

The non-monotonic shape of  $d'$  *vs.* level functions suggests that the ability of the fiber to discriminate spectral differences should correspondingly vary non-monotonically with stimulus level. Discrimination based on the information provided by HSR fibers would be optimal at levels between 50-70 dB SPL, where  $|d'| \geq 1$ , and would be almost impossible from 70-90 dB SPL. LSR fibers show a less consistent trend. Their associated  $d'$  values are generally higher than corresponding values for HSR fibers and exceed unity for most notch depths at almost all levels except 40 dB SPL. Discrimination based on the information provided by LSR fibers seems to be optimal at around 100 dB SPL.

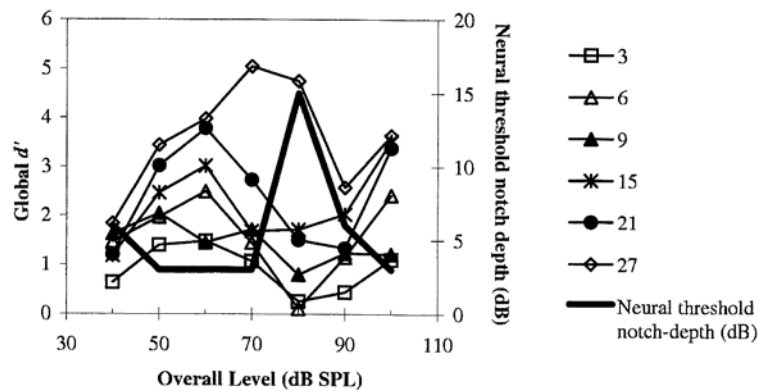
Imagine, now, a system comprised of these two AN fibers only. Its ability to discriminate between the flat-spectrum and the notch noise stimuli would be the result of combining the discrimination capabilities of each of the two fibers. This can be quantified by calculating a global  $d'$  as follows [8]:

$$d' = \sqrt{d'_{HSR}^2 + d'_{LSR}^2}$$

Figure 4 illustrates global  $d'$  values as a function of level for different notch depths. Also plotted are the minimum notch depths for which this global  $|d'|$  exceeds unity at each of the overall levels considered.

The global  $d'$  also varies non-monotonically with stimulus level, as would be expected, since it is calculated from two functions that already are non-monotonic (Fig. 3). A comparison of the global  $d'$  values (Fig. 4) with the





**Fig. 4.** Global  $d'$  values as a function of noise overall level for different notch depths (inset). The bold continuous line (right ordinate axis) illustrates the minimum notch depth, at each noise level, for which the global  $|d'|$  exceeds 1. Symbols denote the notch depths in dB *re.* reference spectrum level

individual  $d'$  values of each fiber type (Fig. 3) reveals that the global  $d'$  is essentially determined by  $d'_{HSR}$  and  $d'_{LSR}$  at low and high levels, respectively. This suggests, therefore, that discrimination based on the combined sensitivity of these two fibers, essentially relies on HSR fibers at low stimulus levels and on LSR fibers at high levels.

Approximate threshold notch depths for each stimulation level can be computed from the global  $d'$  values of Fig. 4 by calculating the minimum notch depth that produces a  $d'$  value greater than one. The results are shown in Fig. 1(b), where they can be directly compared with the behavioral notch depth thresholds. Notice the excellent agreement between the physiological and behavioral results. Not only the peak in the two functions occur around 80 dB SPL, but also the threshold notch depth estimates and the behavioural data are similar in magnitude over the range of noise levels tested.

#### 4 Discussion

Psychophysical discrimination between a flat-spectrum noise and a similar noise with a notch centered at 8000 Hz was shown to vary non-monotonically with stimulus level. The activity of guinea pig AN fibers in response to the stimuli employed in the psychophysical experiment was measured with the aim to provide a physiological explanation for the behavioral data. Physiological data have been shown for two representative fibers only rather than for a large population of fibers. Hence, the present study must be considered as preliminary and the conclusions should be taken cautiously. However, the data illustrate real properties of HSR and LSR AN fibers; properties that are likely to apply to human AN fibers.

A close agreement was found between the physiological and the psychophysical results (Fig. 1(b)). This supports the hypothesis that the main limitations

for behavioral spectral discrimination come from the sensitivity of the AN to changes in the stimulus spectrum.

For both HSR and LSR fibers, differences were observed in the RLFs in response to the flat-spectrum noise and the notch noise. These differences are likely to reflect a combined effect of suppression and a reduction of energy due to the spectral notch. However, understanding the way these two factors, especially suppression, affects the activity of these fibers in response to the notch noise stimulus needs further investigation. The decrease in the slope of the RLFs observed for both HSR and LSR fibers increases the dynamic range of the two fiber types in response to the notch noise stimuli. This is likely to be responsible for the non-monotonic variation of the global  $d'$  values with stimulus level, and by extension, for the non-monotonic variation of the physiologically-estimated threshold notch depths with stimulus level.

Suppression may affect not only the rate of fibers with CFs within the notch-band, but also the activity of fibers with CFs at the notch boundaries. Indeed, activity of the latter may be less suppressed when the notch is present than when it is absent, hence their rate may be higher than that of other fibers with CFs well away (in frequency) from the notch. As has been previously suggested (e.g. [12]), this serves to enhance the notch boundaries. The influence of these fibers on threshold notch depths has not been considered in the present study.

Suppression plays an important role for encoding not only high-frequency spectral features but also low-frequency features. Specifically, suppression linked to LSR fibers is essential for encoding the spectrum of vowels (of both their formants and troughs). As noted by [11], "...under conditions where HSR rate representation is degraded by saturation effects, LSR fibers are able to sustain a differential response to formants and troughs because their trough-driven discharge rates are suppressed by off-BF energy in the surrounding formants."

The fact that the interpretation given to psychophysical results agrees with the results derived from physiological responses of only two AN fibers does not contradict the conclusion of Alves-Pinto and Lopez-Poveda ([1]) that high-frequency spectral notches are probably represented in the AN rate profile, that is in the activity of a larger group of fibers. A global  $d'$  based on a population of fibers would probably yield physiological threshold notch depths that are closer to the behavioral data. More data are currently being collected to test this hypothesis.

Finally, it should be noted that the  $d'$  values for the LSR fiber are generally larger than those for the HSR fiber. This suggests that LSR fibers are more sensitive than HSR fibers to changes in the spectrum, which may compensate for the scarcity of LSR fibers. Having said that, it remains unclear (to us at least) whether discriminability benefits more from having few fibers producing large  $d'$  values than from having many fibers producing smaller  $d'$  values or vice versa.

#### 4.1 Implications for the Design of Bio-inspired Speech Processors

It has been shown that HSR fibers by themselves can not explain the improvement in spectral discrimination performance observed at levels higher than 70-80

dB SPL. Furthermore, it has been discussed that such improvement probably occurs as the result of suppression acting upon LSR fibers, which have higher thresholds and wider dynamic ranges. This imposes minimum requirements for the speech processors of artificial intelligence systems that aim to mimic human performance at spectral discrimination tasks, namely they must reproduce the effects of both HSR and LSR fibers as well as AN suppression effects.

There already exist computational models that simulate these properties of the peripheral auditory system (e.g., [17-19]). They, therefore, constitute a good starting point for building bio-inspired speech processors for artificial intelligence systems with human-like hearing capabilities.

## 5 Conclusions

The main conclusion is that the combined sensitivity of HSR and LSR AN fibers to the presence of a high-frequency spectral notch is a non-monotonic function of stimulus level. The good correlation between the psychophysical data and the present analysis of auditory nerve responses to flat-spectrum and notch noise stimuli suggests that spectral discrimination depends directly on information conveyed by the *difference* rate profile rather than on the quality of the individual rate profiles.

## Acknowledgments

Work supported by the Spanish *Fondo de Investigaciones Sanitarias* (grant refs. PI02/0343 and G03/203) and by European Regional Development Funds.

## References

1. Alves-Pinto A. and Lopez-Poveda, E. A.: Detection of high-frequency spectral notches as a function of level. *J. Acoust. Soc. Am.* (submitted)
2. Arthur R.M., Pfeiffer R.R., and Suga N.: Properties of 'two-tone inhibition' in primary auditory neurones. *J. Physiol.* **212** (1971) 593-609.
3. Butler, R. A., and Belendiuk, K.: Spectral cues utilized in the localization of sound in the median sagittal plane. *J. Acoust. Soc. Am.* **61** (1977) 1264-1269.
4. Butler, R. A., and Humanski, R. A.: Localization of sound in the vertical plane with and without high-frequency spectral cues. *Percept. Psychophys.* **51** (1992) 182-186.
5. Delgutte, B., and Kiang, N. Y. S.: Speech coding in the auditory nerve: III. Voiceless fricative consonants. *J. Acoust. Soc. Am.* **75** (1984) 887-896.
6. Evans, E. F., and Palmer, A. R.: Relationship between the dynamic range of cochlear nerve fibers and their spontaneous activity. *Exp. Brain Res.* **40** (1980) 115-118.
7. Green, D. M. and Swets, J.A. *Signal Detection Theory and Psychophysics*. New York; Wiley, 1966.

8. Hancock, K. E. and Delgutte, B.: A physiologically based model of Interaural Time difference discrimination. *The Journal of Neuroscience* **24** (2004) 7110–7117.
9. Hebrank, J., and Wright, D.: Spectral cues used in the localization of sound sources on the median plane. *J. Acoust. Soc. Am.* **56** (1974) 1829–1834.
10. Johnson, D. H.: The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J. Acoust. Soc. Am.* **68** (1980) 1115–1122.
11. May, B. J.: Physiological and psychophysical assessments of the dynamic range of vowel representations in the auditory periphery. *Speech Communication* **41** (2003) 49–57.
12. Poon, P. W. F., and Brugge, J. F.: Sensitivity of auditory nerve fibers to spectral notches. *J. Neurophysiol.* **70** (1993) 655–666.
13. Rice, J. J., Young, E. D., and Spirou, G. A.: Auditory-nerve encoding of pinna-based spectral cues: Rate representation of high-frequency stimuli. *J. Acoust. Soc. Am.* **97** (1995) 1764–1776.
14. Sachs, M. B., and Abbas, P. J.: Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. *J. Acoust. Soc. Am.* **56** (1974) 1835–1847.
15. Sachs, M. B., and Young, E. D.: Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate. *J. Acoust. Soc. Am.* **66** (1979) 470–479.
16. Winter, I. M., Robertson, D., and Yates, G. K.: Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres. *Hear. Res.* **45** (1990) 191–202.
17. Sumner, C., Lopez-Poveda, E. A., O'Mard, L. O. P., and Meddis, R.: A revised model of the inner hair cell and auditory nerve complex. *J. Acoust. Soc. Am.* **111** (2002) 2178–2188.
18. Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H.: A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J. Acoust. Soc. Am.* **109** (2001) 648–670.
19. Robert, A., and Eriksson, J. L.: A composite model of the auditory periphery for simulating responses to complex sounds. *J. Acoust. Soc. Am.* **106** (1999) 1852–1864.