ICSV13 - Vienna
The Thirteenth International Congress on Sound and Vibration
Vienna, Austria, July 2-6, 2006

NEW MODEL OF AUDITORY FREQUENCY SELECTIVITY IN ENVELOPE CHANGES DOMAIN

Dariusz Jan Kutzner*

Institute of Acoustics
Department of Room Acoustics and Psychoacoustics
Adam Mickiewicz University
85 Umultowska Street, 61-612 Poznań, Poland
konsbol@wp.pl

Abstract

This paper is concerned with a new model of auditory system frequency selectivity in the modulation frequency domain, i.e. non-negative impulse response modulation filters concept. The model argues that if any form of the sound envelope analysis took place at the higher stages of the auditory pathway, this process could not be described as a band-pass filtration. This is due to unipolarity of the sound envelope and its neural representation (temporal changes of neural spikes generation rate), which are functions of non-negative values. Consequently, impulse response of the hypothetical modulation filter could not be bipolar function and, therefore, its frequency characteristics could not be band-pass, as it has been assumed in the traditional modulation filters concept. Results of model simulations are compared to outcomes of psychoacoustical measurements related to some aspects of amplitude modulation perception.

INTRODUCTION

Overwhelming majority of environmental sounds are characterized by continuous changes of their parameters. Such a temporal modification of some parameter of a signal is known as modulation, for example amplitude modulation, AM. Although AM is not an immanent feature of acoustic stimuli, results of many psychophysical experiments revealed that the sound envelope temporal fluctuations are crucial for many aspects of auditory perception, for example: speech intelligibility [1]. Therefore, it is noteworthy to examine how the auditory system analyses the temporal changes of acoustic signal amplitude, i.e. the sound envelope.

So far, two main concepts of the auditory envelope processing have been proposed. According to Viemeister [2], the sound envelope is low-pass filtered at the higher stages of the auditory pathway. The main psychophysical evidence supporting this approach is the so-called temporal modulation transfer function, TMTF, which is, generally, low-pass shaped. In other words, the higher frequency of amplitude modulation is, these envelope changes become less audible. Hence, internal representation of the sound envelope changes might be regarded as being attenuated by some post-cochlear low-pass filter. The cut-off frequency of the filter is estimated to be about 64 Hz.

The second model of the auditory envelope processing assumes that, besides of the auditory filters [3], there is another stage of the acoustic stimuli filtering in the auditory system. This stage is assumed to contain the so-called modulation filterbank, MFB, i.e. an array of bandpass, overlapping, linear filters tuned to different envelope rates [4],[5]. The main task of the MFB is assumed to analyse the sound envelope by means of analysis of its spectral content. Hence, the hypothetical modulation filters are believed to act analogously to the cochlear filters, but in the modulation rate domain. There are many psychophysical and physiological evidences supporting this concept. For instance, phenomena reflecting an activity of such modulation filters were observed in the modulation rate domain: tuning and masking [4],[5], discrimination [6], perception of changes in modulator phase spectrum [7] or detection of asynchronous presentation of spectral components of a modulation wave [8]. Since all the observed relationships are qualitatively comparable to those observed in the audible frequency domain, these outcomes might be treated as reflecting an activity of post-cochlear filters tuned to the different modulation frequencies. As far as the physiological measurements are concerned, neural responses depending on the modulation rate were found in the *cochlear nucleus* [9] and *inferior colliculus* [10]. Nevertheless, it should be stressed that the MFB concept is still controversial and often called in question [11].

Since, both the mentioned concepts suggest qualitatively different processes applied to the sound envelope, i.e. low-pass filtering (Viemeister's model) or band-pass filtering (MFB concept), they might be regarded, in a sense, as opposite approaches to the auditory modulation analysis. It should be emphasised that the MFB concept itself does not question an existence of a single low-pass modulation filter suggested by Viemeister's model. On the other hand, it seems to be impossible to interpret and simulate many aspects of modulation perception, for example modulation masking, assuming an activity of a single low-pass modulation filter.

This paper deals with a new model of the frequency selectivity in a modulation rate domain. The starting point of considerations was an analysis of the sound envelope's neural representation and simulation of temporal characteristics (impulse response) of a low-pass filter and a band-pass filter. As it will be shown, there are some strong analytical arguments suggesting that band-pass modulation filters are neither theoretically nor physiologically realisable. This is due to the fact that band-pass filter produces bipolar impulse response, which is not possible in the auditory system structures, as temporal fluctuations of the sound envelope are reflected in temporal changes of action potentials generation rate. Since a rate of neural spikes generation cannot be negative, this should not be modelled by a bipolar function and, consequently, the hypothetical modulation filters cannot produce bipolar impulse responses, so they cannot be characterized by band-pass frequency characteristics. In accordance, a new model of auditory frequency selectivity in the envelope changes domain has been developed. On the basis of this model, namely non-negative impulse response, NNIR, modulation filters concept, results of some investigations related to the modulation perception are predicted and compared to the results of real psychophysical measurements.

IMPULSE RESPONSE OF LOW-PASS FILTER AND STANDARD MODULATION FILTER

The starting point of developing non-negative impulse response modulation filters concept was an analysis of impulse responses of filtering elements associated with the both mentioned above models of the auditory envelope processing. Fig. 1 presents frequency characteristics and corresponding impulse responses of a single band-pass modulation filter (the upper panels) and a low-pass filter (the bottom panels).

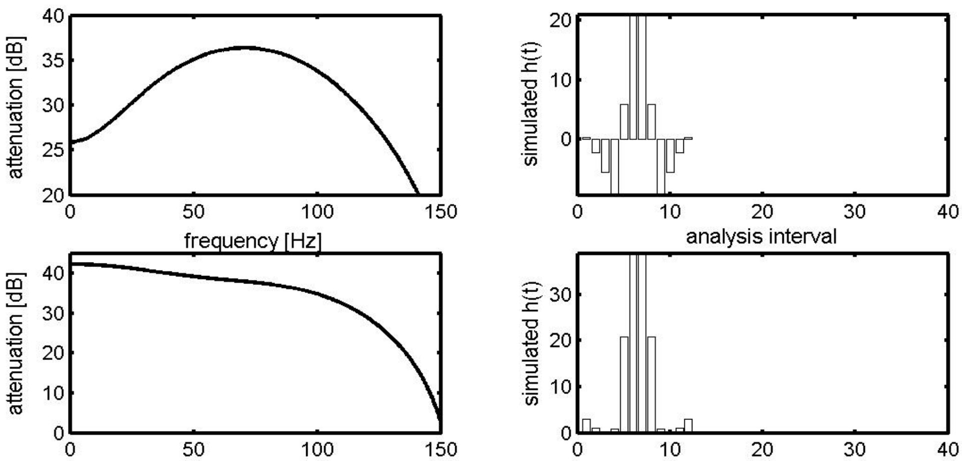

Fig. 1. Frequency and temporal characteristic of a single hypothetical modulation band-pass filter (the upper panels) and a low-pass filter (the bottom panels). The left panels present frequency characteristics; the right panels show corresponding impulse responses. Since in the auditory system AM is reflected in temporal changes of neural spike generation rate, the impulse responses simulate changes of spike generation rate (period histograms).

The main conclusion that can be drawn from inspection of Fig. 1 is that the impulse response of a band-pass modulation filter is a bipolar function. Therefore, both positive and negative rates of neural spike generation, or probabilities of spike occurrence, are predicted. Since these parameters cannot be negative, the process of bandpass envelope filtration cannot be realized physiologically. Nevertheless, it should be stressed that the classical MFB concept is very useful when predicting and interpreting results of many psychoacoustical experiments (see Introduction).
On the other hand, the temporal integrator suggested by Viemeister is realisable physiologically. In contrast to the MFB model, this approach does not reflect the auditory frequency selectivity in the modulation rate domain.

NON-NEGATIVE IMPULSE RESPONSE MODULATION FILTERS

Due to the above-mentioned limitations of usage of the both mentioned models of envelope analysis in the auditory system, a new concept of envelope processing has been developed. This approach argues that if any form of modulation processing took place at the higher stages of the auditory pathway, this process could not be described in terms of band-pass filtration. So as to reflect the auditory frequency selectivity in the modulation domain, a bank of the so-called non-negative impulse response, NNIR, filters have been designed. Fig.2 presents impulse response (the left panel) and frequency characteristics (the right panel) of an exemplary NNIR modulation filter.

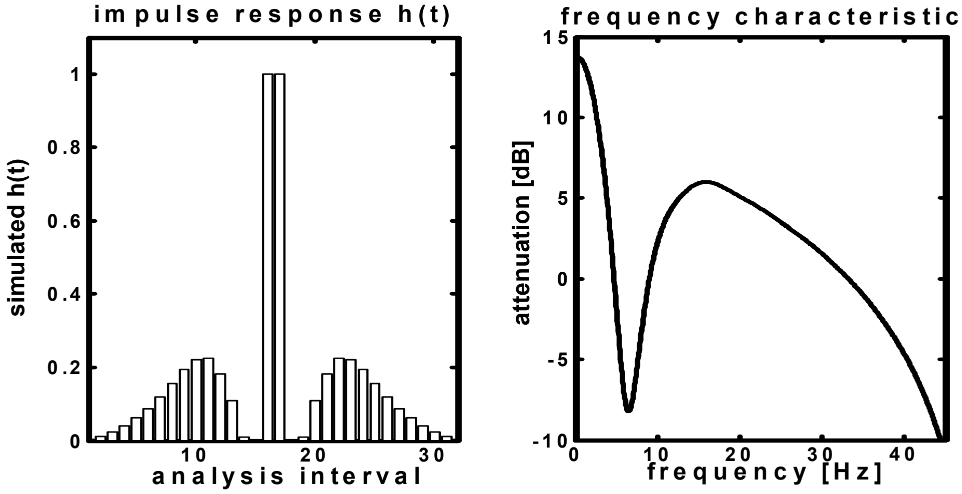

Fig. 2. Impulse response (the left panel) and frequency characteristics (the right panel) of exemplary NNIR modulation of characteristic frequency CF=16 Hz.

As it can bee seen from Fig.2, the frequency characteristics of NNIR filters is characterized by two distinct local maxima, related to characteristic frequency, CF, of the filter and frequency of 0 Hz, respectively. Therefore, such set is selective in some modulation frequency range and, additionally, passes energy of the DC component of an input envelope. Consequently, an output signal is a non-negative function. Thus, such filters could be realised physiologically. So as to reflect band-pass selectivity in the modulation rate domain, an activity of the NNIR filters is described by the so-called variance excitation pattern. For a given input signal, this function presents variance,σ^2, (as $10\log\sigma^2$) of the NNIR filter response as a function of CF. Unlike the traditional MFB model [4], σ^2 is determined instead of rms, thus the DC component is not taken into consideration and the final excitation is a band-pass function. Fig. 3 depicts auditory pathway model of modulation perception used in simulations. Acoustic stimuli are filtered in a set of auditory filters (gamma-tone filters [3]) and processed in a non-linear device simulating a non-linear transformation of the basilar membrane oscillations into spike generation rate [12]. Subsequently, the signals are processed by a low-pass filter (reflecting TMTF shape [2] and simulating a limited temporal resolution of the auditory system). In the next step, the sound envelope in a given auditory channel is passed through the NNIR modulation filters bank and the variance excitation pattern is determined.

Fig. 3. Auditory pathway model of modulation perception and analysis used in the simulations (see the text for details).

EMPIRICAL VERIFICATION OF NNIR MODULATION FILTERS CONCEPT

Experimental equipment and method

This chapter presents results of psychophysical measurements and a comparison of the empirical data and outcomes of model investigations. As frequency selectivity is relatively simple examinable by measurements under masking conditions, two psychoacoustical experiments concerned with modulation masking phenomenon have been carried out. In the first paradigm, the modulation masking patterns for various carriers were determined. The main purpose of the second experiment was to determine modulation masking patterns for stimuli of different durations.
The modulation masking patterns present modulation masking effectiveness for different envelope rates [4],[5]. Like in the audible frequency domain [3], the patterns are determined by means of a subtraction of modulation detection thresholds measured in a presence of the so-called masker modulator (masked thresholds) from absolute detection thresholds (unmasked thresholds).General formula of AM signals used in the investigations was as follows:

$$y(t) = (1 + x_{\mathrm{mod}}(t))n(t) \qquad (1)$$

where: $n(t)$ is a carrier signal; x_{mod} denotes a modulation waveform:

$$x_{\mathrm{mod}}(t) = m_p \cos(2\pi f_p t + \phi_p) + m_m \cos(2\pi f_m t + \phi_m) \qquad (2)$$

where: m_p, f_p and ϕ_p are depth, frequency and phase of a probe modulator and m_m, f_m and ϕ_m are depth, frequency and phase of a masking modulation. A three-interval, alternative forced-choice (3AFC) procedure (1-up, 3-down), corresponding to 79,4% correct responses point, was employed to determine unmasked and masked modulation thresholds [13]. Three stimuli ('intervals') were presented to a subject in a random order. In an unmasked conditions, one of them was modulated by the probe ('signal interval') and m_m=0%, while in a masked measurements each of them was modulated by the masker at modulation depth m_m=25% (rms) and f_m =16 Hz and one of them was modulated additionally by the probe (‘signal interval’). The subject’s task was to indicate the signal interval. The probe modulation depth was varied with respect to a subject’s response: it was increased after one incorrect answer (1-up) and decreased after three successive correct answers (3-down). The initial step was 3 dB (in terms of $20\log m_{rms}$) and was decreased to 1 dB after first four turn-points. 12 turnpoints were determined during the session, whereas the modulation threshold was computed as a geometric mean of last 8 turnpoints. Threshold value was a mean of data gathered for three separate experimental runs. Since the masker rate was constant, the probe rate defined a spectral separation of the masked and the masking envelopes in the modulation domain. The probe rate varied from 1 to 64 Hz in one-octave steps. Final modulation masking patterns were determined as a difference between the masked and unmasked probe detection thresholds functions. Phase shift between the probe and the masker was random. The signals were generated by a TDT (System 3) equipment and were presented monaurally by means of Sennheiser HD 580 headphones. The presentation sound pressure level was 70 dB. During the measurements subjects were seated in an acoustically insulated booth. Three listeners with clinically normal hearing took part the in the measurements.

In the experiment 1, the masking patterns were examined for the various carries *n(t),* namely: sinusoidal, gaussian noises (GN) of the following bandwidths: 50, 100, 400 and 2000 Hz and a 50-Hz wide low-noise-noise (LNN) [14]. All the carriers were rms-normalized and were centred at 4 kHz. The stimuli duration was 1000 ms including 20-ms rise and fall ramps. Each of these carriers is characterized by stochastic amplitude fluctuations that strongly depend on carrier bandwidths and its probability density function. In these measurements an influence of carrier parameters on the modulation masking patterns was examined.

In the experiment 2, the masking function was determined for various stimuli durations 1000, 500, 250 and 125 ms (including 20-ms raise and fall ramps) and for a 4-kHz sinusoidal carrier. Since a filter influence on an input signal depends on the input signal duration, dependency between the stimuli duration and a shape of modulation masking pattern was examined.

Results of measurements and model investigations

Fig. 4 presents the results of the experiment 1, i.e. modulation masking patters for the exemplary carriers: the sinusoidal (the left upper panel), the 2000-Hz wide GN (the right upper panel), the 50-Hz wide GN (the left bottom panel) and 50-Hz wide low-noise-noise (the right bottom panel). A three-way ANOVA indicated that the probe frequency, the carrier type and the subject were statistically significant factors, while two-way ANOVA carried out for each carrier separately revealed that for the 50-Hz wide GN neither the probe rate $\{F(6,62)=0.48, p<0.82\}$ nor the subject factor were statistically significant $\{F(2,62)=1.44, p<0.24\}$. For the sine and the 2000-Hz wide GN carriers modulation masking effectiveness is determined by a spectral separation between the probe and the masker modulators in the envelope rate domain.
The masking magnitude is the largest when the probe rate and the masker rate are the same. On the contrary, there is no local maximum in the function for the 50-Hz wide

GN. This is due to relatively high intrinsic envelope fluctuation; in this case application modulation masker of m_m=25% does not change markedly amplitude fluctuation of the sound, thus significant differences between masked and unmasked thresholds are not observed. This explanation was easily confirmed by using the 50-Hz wide LNN carrier. It has the same power spectrum as the 50-Hz wide GN carrier, while it reveals much less intrinsic envelope fluctuation [14]. Thus applying the modulation masker of m_m=25% results in a band-pass modulation masking pattern (the right bottom panel).

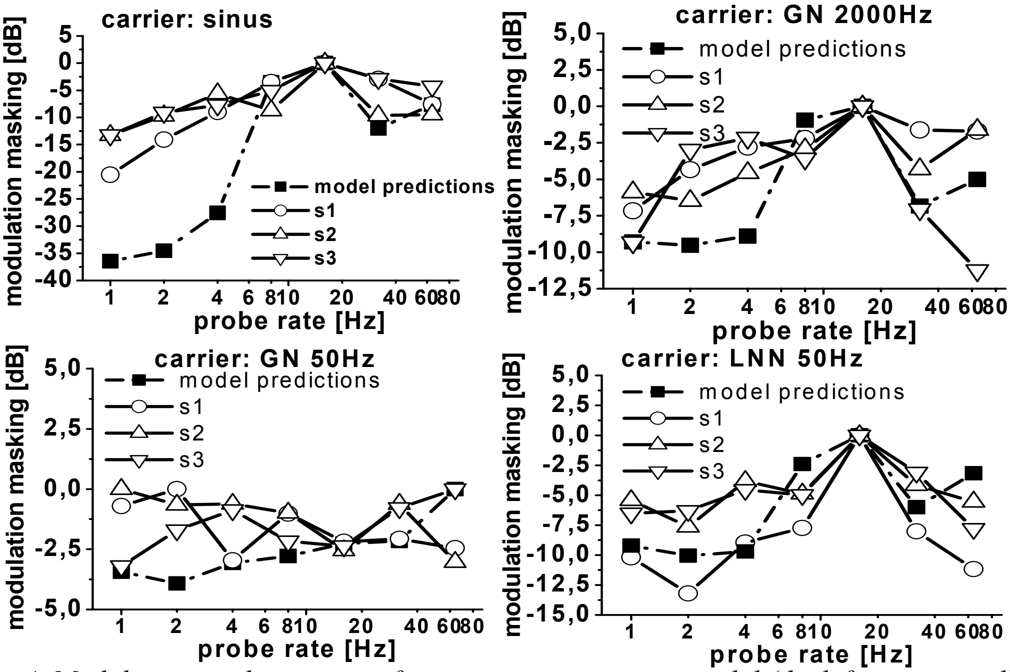

Fig. 4. Modulation masking patters for various carriers: sinusoidal (the left upper panel), 2000-Hz wide GN (the upper right panel), 50-Hz wide GN (the left bottom panel) and 50-Hz wide LNN (the right bottom panel). Open symbols present data for the respective subjects; filled squares depict modulation masking predicted by the NNIR modulation filters model.

Results of the model investigations are comparable to outcomes of empirical measurements (filled squares).

Fig. 5 depicts modulation masking patterns for exemplary stimulus duration: 500 ms (the left panel) and 125 ms (the right panel). A three-way ANOVA indicated that the probe frequency, the stimulus duration and the subject were statistically significant factors. Thus, modulation masking depends on the spectral separation between the probe and the masker in the modulation domain as well as on the stimuli duration. As it can be seen, a dynamic range of the masking functions decreases when the stimuli duration is decreased: 15-20 dB for 1000 ms (the left upper panel of Fig. 4), 8-10 dB for 500 ms (the left panel of Fig. 5), and 5-7 dB for 125 ms (the right panel of Fig. 5). This dependency suggests, indirectly, an activity of some filters tuned to different modulation rates. This is due to 'resultant selectiveness' of a filter depends on an input signal (envelope) duration. This is also reflected in the results of model investigations (filled squares), i.e. the dynamic ranges of simulated patters are comparable to the results of empirical measurements.

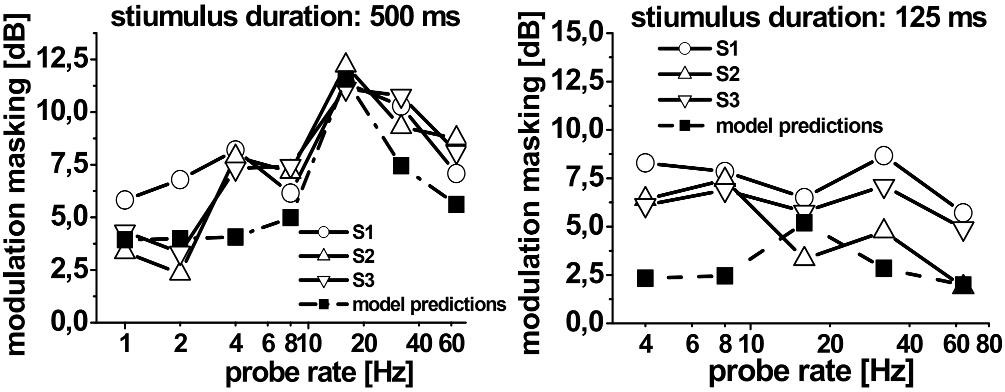

Fig. 5. Modulation masking patterns for two stimulus durations: 500 ms (the left panel) and 125 ms (the right panel). Open symbols present data obtained for the respective subject, while filed squares show the results of the simulations.

CONCLUSIONS

The results of the psychophysical measurements are in line with the prediction of NNIR modulation filters model. In contrast to the standard MFB concept, the new model of the auditory frequency selectivity in the modulation rate domain does not predict negative rates (probabilities) of neural spikes generation (occurrence).

This work was supported by the State Committee for Scientific Research, project number 4T11E01425.

REFERENCES

[1] Chi, T., et al., Spectro-temporal modulation transfer functions and speech intelligibility. JASA **106**: p. 2719-2731.

[2] Viemeister, N.F., Temporal modulation transfer functions based on modulation thresholds. JASA, 1979. **66**: p. 1364-1380.

[3] Moore, B.C.J., An Introduction to the Psychology of Hearing, 5th Ed. 2003, London: A.Press

[4] Dau, T., Modeling auditory processing of amplitude modulation. 1996, Univ. of Oldenburg

[5] Lemańska, J., A. Sęk, and W. Rybicka, Masking in the amplitude modulation rate domain. Archives of Acoustics, 2003. **28**(3): p. 151-159.

[6] Lemańska, J., A. Sęk, and E. Skrodzka, Discrimination of the amplitude modulation rate. Archives of Acoustics, 2002. **27**(1): p. 3-22.

[7] Sęk, A. and B.C.J. Moore, Testing the concept of a modulation filter bank: The audibility of component modulation and detection of phase change in three-component modulators., JASA, 2003. **113**(5): p. 2803-2811.

[8] Kutzner, D., J. Lemańska, and A.P. Sęk, Detection of asynchronicity in the modulation rate domain. Archives of Acoustics, 2005. **30**(3): p. 291-307.

[9] Moller, A.R., Coding of amplitude and frequency modulated sounds in the cochlear nucleus of the rat. Acta Physiological Scandanavica, 1972. **86**: p. 223-238.

[10] Langner, G. and C.E. Schreiner, Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. Journal of Neurophysiology, 1988. **60**: p. 1799-1822.

[11] Jorasz, U., Selectiveness of the auditory system (in Polish). UAM Press, Poznań, 1999.

[12] Bacon, S.P., R.R. Fay, and A.N. Popper, Compression. From Cochlea to Chochlear Implants. Spinger Handbook of Auditory Research, 2004. **17**.

[13] Levitt, H., Transformed up-down methods in psychoacoustics, JASA, 1971. **49**: p. 467-477.

[14] Kohlrausch, A., et al., Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations. Acustica, Acta-Acustica, 1997. **83**: p. 659-669.