ARCHIVES OF ACOUSTICS
31, 4, 543–549 (2006)

Technical Notes

MODULATION MASKING FOR RECURRENT LOW-NOISE NOISE MASKER

D. J. KUTZNER

Adam Mickiewicz University
Institute of Acoustics
Umultowska 85, 60-614 Poznań, Poland
e-mail: konsbol@wp.pl

(*received June 15, 2006; accepted September 30, 2006*)

The main purpose of these investigations was to examine modulation masking phenomenon for recurrent low-noise noise masker. Such masker is characterized by three parameters, namely: repetition frequency, f_{rep}, centre frequency, f_o, and bandwidth, b. The parameter f_{rep} is not reflected in the signal power spectrum and is related to the autocorrelation period. The parameters f_o and b describe spectral properties of the interfering signal, i.e. localization and concentration of its power in the frequency domain. In order to separate possible effects of the masker temporal repetition and its spectral parameters, modulation masking measurements were carried out for $f_o = 64$ Hz, $b = 16$ Hz and $f_{\mathrm{rep}} = 1$ s^{-1} (without repetition), 4 s^{-1} and 8 s^{-1} and probe signal of frequencies $f_p = 1, 2, 4, 6, 8, 12, 16, 32, 52$ and 64 Hz. The masker *rms* modulation depth was 30%; carrier signal was a 4-kHz sinusoid. The main conclusion is that modulation masking patterns are determined by the spectral properties of the masker.

Key words: amplitude modulation, signal envelope, modulation masking, modulation filterbank concept, perceptual grouping, low-noise noise, wavelet transformation.

1. Introduction

The main purpose of this study was to provide further information about how the auditory system analyses and processes the sound envelope in a presence of the so-called masking modulator. Results of many experiments related to this topic suggest that auditory perception of amplitude modulation in masking conditions is, to some extent, comparable to the perception of acoustic stimuli in the audible frequency domain. It has turned out that modulation masking patterns show similar shape to those obtained in the audio-frequency domain, i.e. they reveal clearly noticeable local maximum falling at the frequency of a masker. In other words, modulation masking magnitude decreases as the spectral separation between masked and masking modulation signal increases in the envelope rate domain. Such regularity was found in experiments in which sinusoidal and noiseband modulation maskers were used [2, 8]. Some kind of auditory tuning in the modulation frequency domain was found in the measurements concerning psychophysical tuning curves in the modulation rate domain [13]. Also results of experiments related to detection of modulator phase spectrum [12], amplitude modulation rate discrimination [7] and detection of asynchrony in the modulation rate domain [6] can be easily interpreted on basis of this idea.

Since the results of the mentioned experiments reveal similar relationships as the previous measurements carried out in the audio-frequency domain, it has been suggested that there are two stages of auditory filtering in the auditory system. The fist one is related to selectivity of the basilar membrane and is reflected in an activity of the auditory filters. The second stage, namely modulation filterbank, MFB, is assumed to function at the higher stages of the auditory pathway. The MFB is thought to analyse the sound envelope in a given auditory channel by means of decomposition of its spectral structure. The standard MFB concept presupposes that the modulation filter bank is composed of a set of linear, bandpass, overlapping filters tuned to different modulation rates [2, 8].

The most recent version of the model [5] argues that the hypothetical modulation filters, if they existed, should reveal frequency characteristics of two local maxima. The first one is related to a characteristic frequency of the filter, CF_{mod}, while the second falls at the frequency of 0 Hz and is "responsible" for transmission of the envelope *DC* component. Since an impulse response of such filter is a function of non-negative values, this approach to the auditory frequency selectivity in the modulation rate domain has been called as non-negative-impulse-response, NNIR, modulation filters concept.

It was found that calculation of variance (the so-called variance excitation pattern) of the NNIR filter output signal accounts for many aspects of modulation perception against an interfering modulation signal without predicting negative values of the envelope. Both versions of MFB model argue that the modulation masking effectiveness is determined by a spectral separation of masked and masking modulation and masking patterns are determined by the power spectrum of masking modulator.

An alternative approach interprets modulation masking phenomenon in terms of perceptual grouping [10]. In this case, it is assumed that the less spectral separation between masked and masking modulation in the modulation domain is, the two modulators become more and more similar and the auditory system tends to perceive these amplitude modulations as one object (common fate principle) [3]. Therefore, the results of these experiments suggest that much more difficult detection of a probe signal in a presence of a masker modulator is a consequence of some temporal similarity of the masking and the probe modulation. The aim of this paper was to provide information about possible mechanisms underling perception of the sound envelope in masking conditions, i.e. to answer the question: is modulation masking determined by spectral or temporal properties (or both) of masker modulator?

The main inspiration of the present study was experiment carried out by DAU [2] in which modulation masking pattern was determined for harmonic tone complex acting as a masking modulation. The frequency components of the multi-tone modulation masker were: 90, 120, 150, 180 and 210 Hz. Therefore, masking waveform had a repetition period equal to 30 s^{-1}, but its power spectrum did not reveal local maximum at frequency of 30 Hz. It turned out that modulation masking magnitude was the largest for the probe frequencies about 90–120 Hz and decreased as the probe rate was decreased. A local maximum was not observed for probe modulation frequency of 30 Hz; therefore the modulation masking seems to be determined by a spectral separation between the probe modulation and the masker modulation in the envelope rate domain. In the present measurements this paradigm was expanded with noise modulation masker.

2. Stimuli

A general formula describing acoustic signals used in the present measurements is given by an expression:

$$y(t) = (1 + x_{\text{mod}}(t)) \sin(2\pi f_c t + \phi_c), \tag{1}$$

where f_c and ϕ_c are carrier signal frequency and its initial phase, respectively; $x_{\text{mod}}(t)$ is a modulation signal given by the following formula:

$$x_{\text{mod}}(t) = m_p \cos(2\pi f_p t + \phi_p) + m_m\, n_r(t). \tag{2}$$

The first part of superposition (2) represents a probe modulation signal, while the second part is a masking modulation; m_p, f_p and ϕ_p denote modulation depth, frequency and initial phase of a probe signal, respectively; m_m is a modulation depth of the masker, $n_r(t)$. The masking signal was generated by means of copying some randomly selected, short-term realisation of low-noise-noise, LNN [11], in the time domain. The LNN was chosen as a modulation masker due to its low intrinsic amplitude fluctuations that are crucial in avoiding an over-modulation effect. The LNN was produced according to an iterative procedure described by KOHLRAUSH [4]. Notice, that it is highly probable that such recurrent signal might reveal some discontinuities falling at "sticking" of consecutive LNN realisation. These broadband discontinuities were eliminated by the following procedure: 1) discrete time wavelet transformation [1], *dwt*, (wavelet *db2*, number of decomposition levels: 12) was computed, 2) wavelet components correlated with the discontinuities were attenuated, 3) the final recurrent LNN signal without discontinuities was synthesised by means of invert discrete time wavelet transformation [1], *idwt*.

Figure 1 presents characteristics (waveform, autocorrelation, power spectrum density) of LNN (left panels) and the same representations of recurrent LNN masker composed by temporal copying of a randomly selected 250-ms realisation of the LNN signal and reduction of discontinuities (right panels). As can be seen, both signals have similar power spectra, however the autocorrelation of the recurrent LNN masker is characterized by repetitiveness, i.e. in this case $f_{\text{rep}} = 4\ s^{-1}$. In the present experiments centre frequency, f_0, and bandwidth, b, of the modulation masker were 64 Hz and 16 Hz, respectively, while the probe signal rate f_p was varied: 1, 2, 4, 6, 8, 12, 16, 32, 52 and

64 Hz (therefore the spectral separation between the masking modulation and the probe modulation was changed). Modulation masking patterns were determined for the following masker periods $f_{\rm rep}$: 1 s^{-1} (without repetition), 4 s^{-1} and 8 s^{-1}.

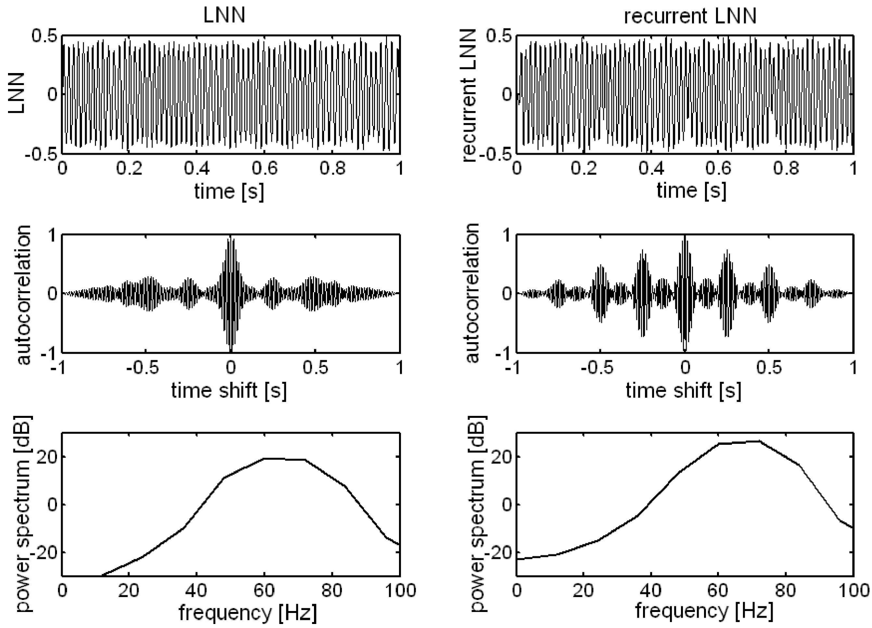

Fig. 1. Characteristics of initial (left panels) and recurrent LNN masker $f_{\rm rep} = 4\ s^{-1}$ (right panels): temporal waveforms (upper panels), autocorrelations (middle panels) and power spectra (bottom panels).

All the spectral components of the signal (1) were processed within one critical band. The carrier frequency was $f_c = 4$ kHz, the overall stimuli sound pressure level was normalized to 70 dB SPL.

3. Measurement method and apparatus

A standard psychoacoustical equipment was used in the investigations. It was composed of a PC-computer, the 24-bit signal processor RP2 and the headphone amplifier HB7 (*Tucker-Davies Technologies*). The processor generated the signals according to the experiment-devoted software (*Matlab 6.5, MathWorks*) at a sampling rate of 48288 Hz. The stimuli were presented monaurally via Sennhesiser HD 580 headphones. The measurements were carried out in an acoustically insulated booth. A three-interval, two-alternative, forced-choice (3I, 2AFC) paradigm with adaptive procedure (*1-up, 3-down*) [9] corresponding to 79.4% correct responses point, was employed to determine unmasked and masked modulation detection thresholds. Three stimuli ("intervals") were presented to a subject in a random order. In unmasked condition one of them was modulated by the probe ("signal interval"), while in masked measurements each of them was modulated by the masker and one of them was modulated additionally by the probe ("signal interval"). The subject's task was to indicate the signal interval. Listeners were informed whether their answer was correct or wrong (response feedback). The probe *rms* modulation depth was varied with respect to a subject's response: it was increased after one incorrect answer (*1-up*) and decreased after three succeeding correct answers (*3-down*). The initial step was 3 dB (in terms of $20 \log m_{p\,\mathrm{rms}}$) and was decreased to 1 dB after first four turnpoints. 12 turnpoints were determined during the session, whereas the modulation threshold was computed as a geometric mean of last 8 turnpoints. The final threshold (unmasked and masked) value was a mean of data gathered for three separate experimental runs.

So as to determine modulation masking patterns, two independent sessions were performed. In the first one, unmasked modulation thresholds (in an absence of the masker, $m_{m\,\mathrm{rms}} = 0\%$) were determined for all probe modulation rates f_p. In the second session the same measurements were done, but in a presence of the masker (masked modulation thresholds, $m_{m\,\mathrm{rms}} = 30\%$) and for the masker repetition equals to 1 s^{-1} (without recurrence), 4 s^{-1} and 8 s^{-1}, respectively. Three subjects with clinically normal hearing took part in the measurements. Mean age (standard deviation) of the subjects was 24.7 (1.2).

4. Results and discussion

The obtained threshold values were subjected to four-way analysis of variance, ANOVA, with respect to the following factors: probe frequency, masker presentation, masker repetition and subject. It has turned out that subject was statistically insignificant $\{F(2, 359) = 1.27, p < 0.28\}$, while the other factors were highly statistically significant: probe modulation frequency $\{F(9, 359) = 16.75, p < 0.001\}$, masker presentation $\{F(1, 359) = 929.48, p < 0.001\}$ and repetition $\{F(2, 359) = 90.66, p < 0.001\}$. Moreover, a significant interaction between masker presentation and probe modulation frequency was found $\{F(9, 359) = 90.66, p < 0.001\}$, which means that difference between masked and unmasked thresholds, i.e. masking effectiveness, depends strongly on the spectral separation between the probe signal and the masker. Figure 2 depicts modulation masking patterns, i.e. difference between masked and unmasked modulation detection threshold, for LNN masker repetitions: 8 s^{-1} (squares), 4 s^{-1} (triangles) and 1 s^{-1} (circles). The main conclusions that can be drawn from inspection of Fig. 2 are as follows: 1) modulation masking magnitude decreases as the spectral separation between the probe modulation and the masking modulation increases, 2) local maxima are not observed for probe rates equal to respective masker repetition.

Thus, the obtained results suggest that modulation masking phenomenon is determined entirely by the power spectrum of the modulation masker and is not related to the masker temporal properties. Nevertheless, the masking values obtained for $f_{\mathrm{rep}} = 1\ \mathrm{s}^{-1}$ are slightly lower that those for $f_{\mathrm{rep}} = 4\ \mathrm{s}^{-1}$ and 8 s^{-1}. This is related to some differences (being by-products of wavelet decomposition and synthesis) between power spectra of the non-recurrent ($f_{\mathrm{rep}} = 1\ \mathrm{s}^{-1}$) and the recurrent ($f_{\mathrm{rep}} > 1\ \mathrm{s}^{-1}$) LNN mod-

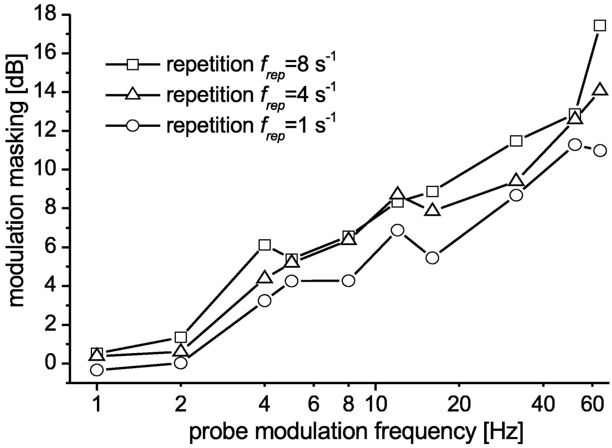

Fig. 2. Modulation masking patterns for recurrent LNN modulation masker of the following repetitions: 8 (squares), 4 (triangles) and 1 s^{-1} (circles). Respective points present data averaged across the subjects.

ulation masker. As can be seen from Fig. 1, the power spectrum for $f_{\mathrm{rep}} = 1\ s^{-1}$ (left bottom panel) is slightly downshifted with respect to the power spectrum for recurrent LNN masker, i.e. $f_{\mathrm{rep}} > 1\ s^{-1}$ (right bottom panel). This difference is directly reflected in the discrepancies between masking patterns for the non-recurrent and the recurrent LNN maskers.

5. Conclusions

The obtained results are in line with the outcome of the previous investigation concerning spectral and temporal masker properties [2] and other experiments related to modulation masking phenomenon [8, 13]. Since the determined relationship between masking effectiveness and masker signal frequency is analogous to that found in the audible frequency domain [10], the outcomes of the present investigation support an idea that there are some channels tuned to different envelope rates at the higher stages of the auditory pathway.

Acknowledgment

This work was supported by State Committee for Scientific Researches (KBN) 4T11E 01425 (2003–2006).

References

[1] BIAŁASIEWICZ J. T., *Falki i aproksymacje*, WNT, Warszawa 2004.

[2] DAU T., *Modelling auditory processing of amplitude modulation*, University of Oldenburg, Oldenburg 1996.

[3] JORASZ U., *Selektywność układu słuchowego*, Wydawnictwo UAM, Poznań 1999.

[4] KOHLRAUSCH A., FASSEL R., VAN DER HEIJDEN M., KORTEKAAS R., VAN DE PAR S., OXENHAM A., *Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations*, Acustica united with Acta-Acustica, **83**, 659–669 (1997).

[5] KUTZNER D., *Modelling the sound envelope auditory processing using the non-negative-impulse-response modulation filters concept. I. Initial simulations*, Archives of Acoustics, **31**, 1, 29–45 (2006).

[6] KUTZNER D., LEMAŃSKA J., SĘK A. P., *Detection of asynchronicity in the modulation rate domain*, Archives of Acoustics, **30**, 3, 291–307 (2005).

[7] LEMAŃSKA J., SĘK A., SKRODZKA E., *Discrimination of the amplitude modulation rate*, Archives of Acoustics, **27**, 1, 3–22 (2002).

[8] LEMAŃSKA J., SĘK A. P., RYBICKA W., *Masking in the amplitude modulation rate domain*, Archives of Acoustics, **28**, 3, 151–159 (2003).

[9] LEVITT H., *Transformed up-down methods in psychoacoustics*, J. Acoust. Soc. Am., **49**, 467–477 (1971).

[10] MOORE B. C. J., *An introduction to the psychology of hearing*, 5th Ed., Academic Press, London 2003.

[11] PUMPLIN J., *Low-noise noise*, J Acoust. Soc. Am., **78**, 100–104 (1985).

[12] SĘK A., MOORE B. C. J., *Testing the concept of a modulation filter bank: The audibility of component modulation and detection of phase change in three-component modulators*, J. Acoust. Soc. Am., **113**, 2803–2811 (2003).

[13] SĘK A. P., SKRODZKA E. B., *Tuning in the amplitude modulation rate domain*, Archives of Acoustics, **23**, 4, 491–498 (1998).