ARCHIVES OF ACOUSTICS
27, 1, 3–21 (2002)

DISCRIMINATION OF THE AMPLITUDE MODULATION RATE

J. LEMAŃSKA, A.P. SĘK and E.B. SKRODZKA

Institute of Acoustics, A. Mickiewicz University
(61-614 Poznań, 85 Umultowska, Poland)

This study examines the amplitude modulation rate discrimination for sinusoidal and noise carriers. It was shown that the discrimination of AM rates is a monotonically growing function of modulation rate. Higher values of the discrimination thresholds were observed for a narrowband carrier. It appears that in the case of a narrowband noise carrier, the spectral range of the noise envelope is similar to that of the modulation rates of the signal (up to 120 Hz). It results in a masking in the modulation rate domain and in a much higher threshold growth than that observed for a wideband noise carrier or a sinusoidal carrier. The results are consistent with the idea of the so-called second stage of filtering acting on the envelope of the acoustic signal. This hypothesis postulates the existence of a so-called modulation filter bank, (MFB), responsible for the frequency selectivity observed in the amplitude modulation rate domain. The existence of the MFB suggests that a certain form of the spectral analysis of any acoustic signal envelope may be performed in the auditory system after initial filtering in the auditory filter bank. A model of the modulation rate discrimination based either on the classical concept of the excitation patterns or on the modulation excitation patterns has not accounted for our experimental data. According to both the models, an increase in the frequency discrimination threshold versus modulation rate should be slower than that measured in the experiment.

1. Introduction

The peripheral auditory system is usually considered in the context of a so-called auditory filter bank which is responsible for the most important feature of the system, i.e. the frequency selectivity [21]. The auditory filter bank model is similar to the concept of the critical bands and assumes that in a tone detection process the subject takes advantage of a single auditory filter with the best signal-to-noise ratio. However, in recent years there has been some evidence that subjects, in majority of real situations, take advantage of more than one auditory filters combining their output signals in a specific way like in the case of the comodulation masking release (CMR) effect [11, 18] or the modulation detection/discrimination interference (MDI) [20].

One of the convincing interpretations of the CMR and MDI effects is based on the so-called perceptual grouping concept. It is assumed that if the changes in a multicomponent broadband signal are similar in all its frequency bands (e.g. the rate of changes is similar in all bands), than such signal forms a single perceptual object. Distinguishing between changes within such an object (e.g. modulation depth discrimination) is more difficult than across different auditory objects [2].

Another way of CMR and MDI interpretation assumes that there is a second stage of filtering in the auditory system. According to this hypothesis, the acoustic signal is first filtered in a set of bandpass auditory filters. Thus, the output signal from the auditory filters describe envelope changes in different frequency bands. Then the output signals are filtered in linear bandpass filters tuned to the frequency of the signal envelope. A set of those hypothetical filters is called a modulation filter bank (MFB). The concept of the MFB implies that the auditory system performs a certain form of spectral analysis with limited resolution on the envelope of the output of each auditory filter. Moreover, it is also assumed that it does not perform a time analysis.

The main argument supporting the MFB concept is the existence of an array of auditory neurones that generate action potentials when there are slow amplitude changes (less than 100 Hz) in the input signal. Neurons tuned to such low characteristic frequencies have been found above the auditory nerve (i.e. in the cochlear nucleus [19, 31] and the inferior colliculus [15, 17, 27, 32]). It is generally assumed that they can be considered as the modulation filter bank. In the MFB concept, similarly to the auditory filter bank model, it is assumed that there is a set of overlapping bandpass linear filters (called modulation filters) tuned to different rates of amplitude changes (i.e. modulation rates). Therefore, it seems that in the modulation filter bank many phenomena, analogous to those found for the auditory filter bank, should be observed, e.g. a frequency selectivity in the modulation rate domain, tuning, masking, ringing and temporal integration.

Although the concept of the second stage of filtering performed in the MFB remains still somehow controversial, evidences consistent with this idea have come from many psychophysical experiments. The concept gives a successful interpretation of some effects which cannot be explained using the classical model of the auditory filter bank only. The most important results concerning the modulation masking and suggesting the existence of a frequency selectivity in the modulation rate domain were reported by HOUTGAST [14] and BACON and GRANDHAM [1]. A similar effect was observed by SĘK and SKRODZKA [36, 37]. Their results are consistent with Dau's idea who assumed that if the MFB does exist, the resulting frequency selectivity in the modulation rate domain is much poorer than the frequency selectivity related to the auditory filter bank, i.e. in the audio frequency domain.

Many arguments supporting the MFB concept were given by DAU [3 – 7]. He analysed the sinusoidal AM detection of a noise carrier as a function of its bandwidth. He found that the inherent fluctuations of the noise carrier played a critical role in the modulation detection. He showed that the modulation detection thresholds for the narrowest carrier (3.14 Hz) were the highest for the lowest modulation rates, while for a 314-Hz wide noise carrier the thresholds were highest for the highest modulation rates used. Dau concluded that the elevation of the AM detection thresholds resulted from a masking effect in the modulation rate domain. The signal, in this case sinusoidal amplitude changes being detected, was masked by inherent noise fluctuations. The above result can be described by saying that the AM detection was the most difficult (required high values of AM depth) when the modulation rate of the signal fell within the spectral range of the noise envelope connected with its inherent fluctuations.

The concept of the MFB allows to model the peripheral auditory system as a set of stages, the MFB is one of the stages [6, 7]. The model could account for the results of many masking modulation experiments as well as for the temporal modulation transfer function (TMTF [38]). A significant progress in the MFB development was done by Everet and Dau [8] who examined the frequency selectivity in the modulation rate domain and suggested values of the main parameters of the modulation filters.

Regardless many experiments supporting the concept of the second stage of filtering, the idea is still controversial because no unique psychophysical evidence of it has been found so far. An assumption about the MFB implies that the modulating signal spectrum, rather than its temporal pattern, is crucial for the signal changes perception. However, many experimental data on the modulation masking can be explained without the MFB concept by taking into account temporal models comparing the similarity between a temporal pattern of the masker and the signal. Thus, it seems that in order to accept the auditory system model containing the MFB, it is necessary to present a variety of arguments and evidences.

2. Aim

One of the most important manifestations of the auditory filters is the so called frequency discrimination describing the ability of the auditory system to detect changes in the signal frequency in course of time. The frequency discrimination can be measured either as frequency modulation detection thresholds or as frequency difference limens [24]. Sęk and Moore [35] have shown that for continuous frequency changes (e.g. for FM signals) and for modulation rates higher than 5 Hz, the peripheral filtering was responsible for the detection of frequency changes. The modulation detection thresholds increasing with increasing carrier frequency, but related to equivalent rectangular bandwidths of the auditory filters, were constant. This suggested that the frequency discrimination mechanism is based on the peripheral filtering exclusively. Frequency discrimination thresholds (difference limens) for two successive sinusoidal tones are monotonically increasing functions of frequency, too [39]. It should be pointed out that the basic mechanism responsible for this type of discrimination is based on phase locking [29], however, the initial signal processing takes place in the auditory filter bank which significantly influences the threshold value. Thus, if the second stage of bandpass filtering, analogous to the first stage of filtering, does exists, it should be possible to observe a frequency discrimination (for modulation rate in this case) for two successive amplitude modulated sinusoidal tones.

In order to analyse the ability of the auditory system to discriminate the frequency of the acoustic signal envelope changes and the factors responsible for this discrimination, it seems worthwhile to repeat the classical frequency discrimination experiment but in the modulation rate domain. If there are modulation filters in the auditory system and if they operate similarly to the auditory filters, one can expect that the modulation rate discrimination thresholds should be monotonic functions of the modulation rate. Furthermore, the ratio of the discrimination thresholds to the modulation rate may enable the estimation of the modulation filter bandwidth and the comparison of the selectivity of those filters with the selectivity of the auditory ones. Therefore, the main aim of the experiments presented was to measure the frequency discrimination thresholds in the amplitude modulation frequency domain, i.e. to measure just noticeable modulation rate differences for two successive bursts of amplitude modulated sinusoidal tones.

Intrinsic fluctuations observed for the noise carrier are an important element of the Dau's model. The model accounts for the modulation perception of noise carriers and bands of noise, but it does not work for sinusoidal carriers. Therefore, in the experiment presented we used a sinusoidal carrier, a band of noise narrower than one critical band and a wide band of noise covering several critical bands.

3. Experiment

AM rate discrimination thresholds were measured using three types of the carrier: a 4-kHz sinusoid, a 450-Hz wide band of noise centred at 4 kHz and a 6-kHz wide band of noise centred at 3.2 kHz. The reason for the high values of the frequency of the sinusoidal carrier and centre the frequency of the noise bands was the fact that the spectral structure of the AM signal does not play a role in the modulation detection. As shown by OZIMEK *et al.* [26], there was a wide range of roughness for the carriers up to 4 kHz (for modulation rates up to about 200 Hz), where AM detection thresholds were roughly independent of the modulation rate. The range covers a band of modulation rates for which the AM signal is not perceived as loudness fluctuations or as a multitone, i.e. based on its spectral structure.

The thresholds were measured using an adaptive two-interval, forced-choice procedure with two-down and one-up stepping rule that estimates a 71% correct point on the psychometric function. Two successive intervals were presented for a subject's preferred ear. Each interval contained a pair of AM signals with the modulation depth $m = 0.5$. In one pair both signals were modulated at a modulation rate f_{mod} (a reference modulation rate), while in the other interval one of the signals was modulated with a modulation rate $f_1 = f_{\mathrm{mod}} + 0.5 \cdot \Delta F_{\mathrm{mod}}$, while the other one was modulated with a rate $f_2 = f_{\mathrm{mod}} - 0.5 \cdot \Delta F_{\mathrm{mod}}$. The modulation rate difference, ΔF_{mod}, of these two signals was changed according to the adaptive procedure: after one incorrect answer it was increased by a certain factor and decreased after two correct answers. The subjects were asked to indicate the interval at which the modulated signals differed in the modulation rate. The reference modulation rate, f_{mod}, was equal to: 10, 20, 35, 50, 75, 100 and 125 Hz. The overall level of the stimuli used in the experiment was 70 dB SPL.

Each run consisted of 12 reversals and the threshold estimate for that run was taken as the arithmetic mean of the last eight reversals. At least three estimates were obtained for each threshold value. The order of the pair and that of the two stimuli in each pair were random. Each signal in the pair had a duration of 500 ms. The time interval between stimuli in the pair was 100 ms and between pairs was 400 ms. The signals were presented monaurally using a Sennheiser HD414 headset in an acoustically isolated chamber. The stimuli were generated using a Tucker-Davis system II. Four subjects without hearing disorders were tested.

4. Results

4.1. AM rate discrimination thresholds

The results of experiments performed are presented in Figs. 1 – 3 for the 4-kHz sinusoidal carrier, the 450-Hz wide carrier centred at 4 kHz and the 6-kHz wide carrier centred at 3.2 kHz respectively. The modulation rate discrimination thresholds vs. reference modulation rate, f_{mod}, for all subjects are plotted in the pictures.

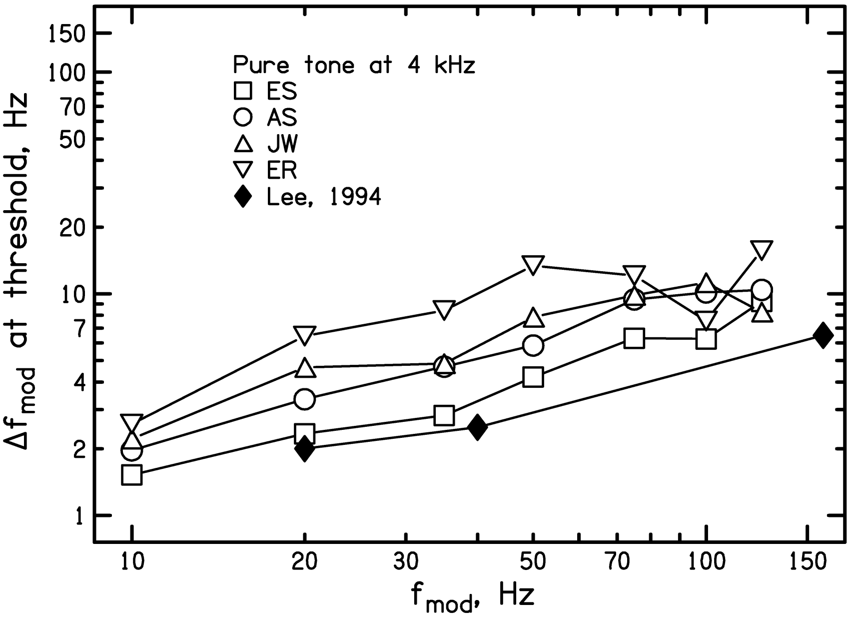

Fig. 1. Amplitude modulation rate discrimination threshold vs. modulation rate for the 4-kHz sinusoidal carrier.

For the sinusoidal carrier the AM rate discrimination thresholds increase with increasing modulation rate from about 2 Hz (for $f_{\mathrm{mod}} = 10$ Hz) to about 10 Hz (for $f_{\mathrm{mod}} = 125$ Hz). The results obtained for subjects taking part in the experiment are qualitatively consistent, although a scatter can be observed. Lee [16] obtained similar results which are also shown in Fig. 1 for comparison.

For the 450-Hz wide carrier the AM rate discrimination thresholds also increase with increasing modulation rate (Fig. 2) as observed in the case of the sinusoidal carrier. However, the threshold values as well as the growing rate of the thresholds as a function

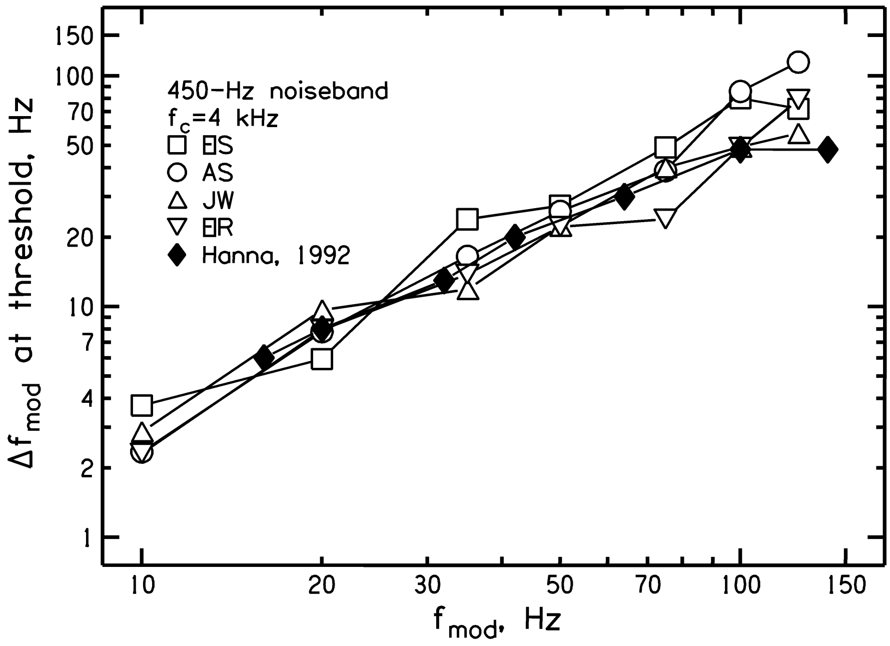

Fig. 2. Amplitude modulation rate discrimination threshold vs. modulation rate for the 450-Hz wide noiseband carrier cantered at 4 kHz.

of the AM reference rate are much higher compared to the sinusoidal carrier; they are about 2–3 Hz (for $f_{\mathrm{mod}} = 10$ Hz) up to about 100 Hz (for $f_{\mathrm{mod}} = 125$ Hz). Significantly higher AM rate discrimination threshold values for the narrow band of the noise carrier can be caused by random changes in the amplitude envelope of this noise which are not observed for the sinusoidal carrier and which are quite different for the wide band of noise. An attempt to interpret the experimental results basing on the analysis of the envelopes of the carrier noise bands can be found in the further text. The results obtained for all the subjects are qualitatively consistent, although there are some differences. HANNA [12] obtained similar results that are shown in Fig. 2 for comparison.

The AM rate discrimination thresholds for the wide noise carrier are an increasing function of the modulation rate, too. The experimental values found for all the subjects are similar and cover a range from about 2 Hz (for $f_{\mathrm{mod}} = 10$ Hz) up to 20 Hz (for $f_{\mathrm{mod}} = 125$ Hz) as can be seen in Fig. 3. These results are very similar to those obtained for the sinusoidal carrier. It is observed that for the wide noise carrier the threshold differences across the subjects are much smaller than for the sinusoidal carrier.

In summary, the AM rate discrimination thresholds for all the carriers are increasing functions of the reference modulation rate. The results obtained for the sinusoidal carrier and the wideband one are similar. The thresholds for the narrowband carrier are significantly higher. The results for both the noise carriers are qualitatively and quantitatively consistent for all tested subjects, however not for all the reference modulation rates used. The differences in the threshold values for the sinusoidal carrier, especially for modulation

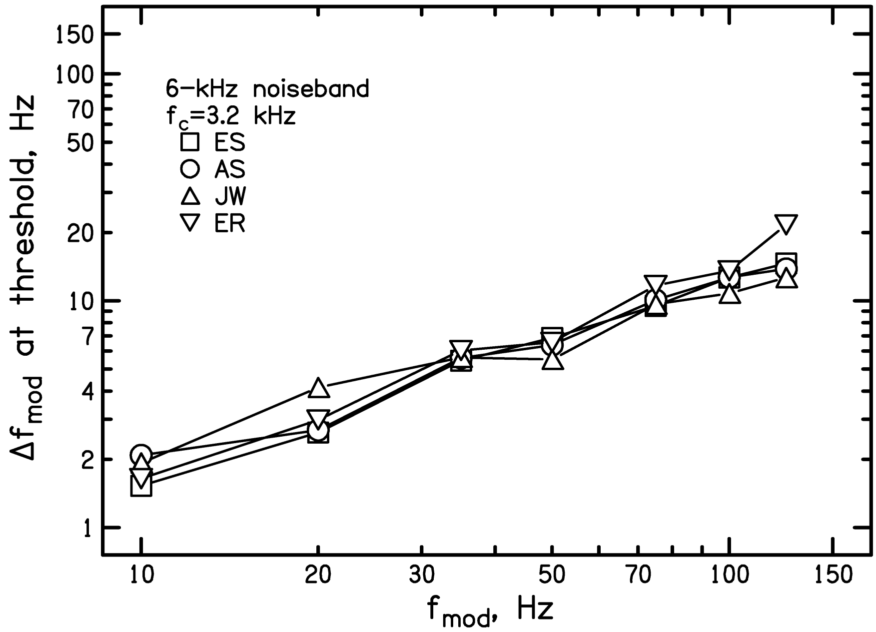

Fig. 3. Amplitude modulation rate discrimination threshold vs. modulation rate for the 6-kHz wide noiseband carrier cantered at 3.2 kHz.

rates higher than 50 Hz, may result from the different subject's sensitivity to AM signals at high modulation rates. It is possible that for the bands of noise carrier their inherent fluctuations as well as the broadband excitation the sensitivity fluctuations for different rates compensate each other resulting in much smaller scatter of the experimental data.

Neglecting the observed scatter it can be stated generally that the AM rate discrimination thresholds are monotonically increasing functions of the reference modulation rate regardless of the type of the carrier. The highest threshold values and a steeper increase with the modulation rate was observed for the narrowband carrier.

To assess the statistical significance of the effects described above the data were subject to an analysis of the variance ANOVA (Genstat 5.1 program package) with the following factors: reference modulation rate and type of the carrier. Taking into account that only for the sinusoidal carrier a statistical scatter for all the subjects was observed, a within-subjects analysis of the variance was performed and the data obtained for all the subjects were treated as repetitions of the same measurement.

As expected, the modulation rate was highly significant [$F(6,\ 18) = 36.33$, $p < 0.001$] what confirmed the earlier observation of a monotonic increase of the AM rate discrimination thresholds with increasing modulation rate for all the carriers used. The type of the carrier signal was highly significant, too [$F(2,\ 6) = 44.61$, $p < 0.001$]. Indeed, on average the thresholds for the sinusoidal carrier and the 6 kHz one are similar, while the thresholds are significantly (by a factor of 10) lower than those for the 450-Hz wide carrier.

The most important result of the analysis is that the interaction between the two factors, i.e. modulation rate and the type of the carrier, was significant too [$F(12, 36) = 20.27$, $p < 0.001$]. The interaction means that the growth of the AM rate discrimination threshold is different for different carriers. This can be easily seen when the data for different carriers are compared.

4.2. Weber's law for the AM rate discrimination

The experimental results were considered in the context of the Weber's law. In our case an analysis of the Weber fraction implies the calculation of the ratio of the modulation rate discrimination threshold to the modulation rate in order to show a relation between the ratio and the modulation rate. Therefore, the results for all the subjects were averaged, or they were treated as repetitions. Further, the Weber fractions were calculated for three carriers used in the experiment and plotted versus modulation rate as shown in Fig. 4. For the sinusoidal carrier, the Weber fraction is constant for modulation rates lower than 50 Hz while for higher rates the fraction slightly decreases. This means that Weber's law is not completely fulfilled. The decrease of the $\Delta F_{\mathrm{mod}}/f_{\mathrm{mod}}$ ratio for modulation rates higher than 50 Hz means that for this range the AM rate discrimination threshold grows slower than for modulation rates below 50 Hz, as seen in Fig. 1. For the narrowband carrier, the Weber fraction slightly increases as a function of the modulation rate. This means that there is also a near miss of Weber's law for the narrowband carrier. The increase of the Weber fraction as a function of the modulation rate implies that the AM rate discrimination threshold rises increasingly faster when the modulation rate increases too (Fig. 2). For the wideband noise carrier the $\Delta F_{\mathrm{mod}}/f_{\mathrm{mod}}$ ratio is constant and independent of the modulation rate within the range of the standard error. This implies that for the 6-kHz wide carrier the Weber's law is fully satisfied and that the modulation rate discrimination threshold is proportional to the modulation rate, as shown in Fig. 3.

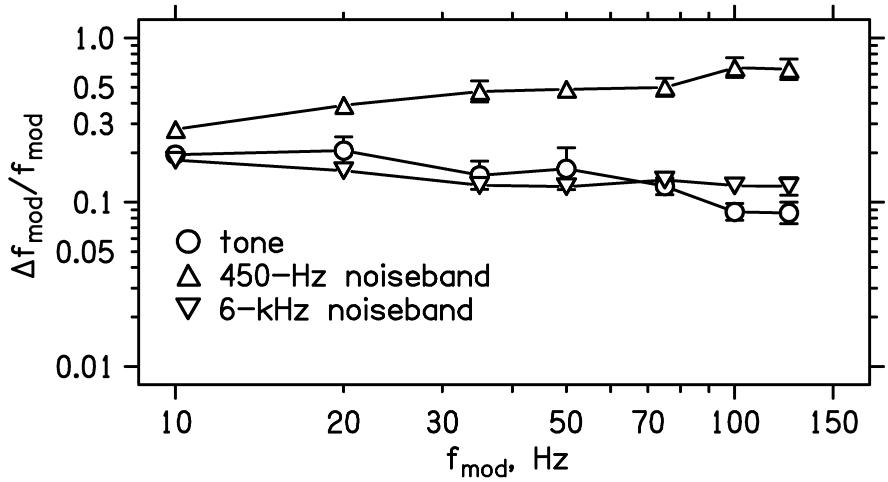

Fig. 4. The Weber fraction for three carriers used in the experiment vs. modulation rate.

The Weber fractions are higher for the narrowband noise carrier than for the sinusoidal and the wideband ones; for the two latter carriers the, Weber fractions are comparable. The above statement supports earlier the conclusion that for the narrowband noise carrier the AM rate discrimination thresholds rise faster than for the other ones; for the tone and the wideband noise carriers the thresholds are similar functions of the modulation rate. It should be pointed out that near miss from the Webers law for the tone and the wideband noise carriers is not significant. Similar discrepancies are observed in the case of the intensity discrimination of some acoustic stimuli.

5. Discussion

5.1. AM rate discrimination vs. excitation pattern models

The most important part of the sound analysis in the auditory system takes place on the peripheral filtering level. Thus, the modulation detection/discrimination is often considered within the frames of the excitation pattern models. The excitation pattern can be defined as an internal activity of the auditory system and is often described as an envelope of the basilar membrane displacement from its equilibrium position being a function of the characteristic frequency of its points. The impulse response of the filters is usually described by the Euler gamma function [28] (see Eq. (5)). However, the frequency characteristics of the auditory filter is often described by the roex (p) function:

$$W(g) = (1 + pg)e^{-pg}, \tag{1}$$

where g is a relative deviation from centre frequency of the filter and p is the steepness of the filter's slopes for the low- and high-frequency sides.

According to ZWICKER's model [40], changes in a stimulus are detected when the change in the excitation pattern at any point (for any characteristic frequency) exceeds a criterion amount, of about 1 dB. The model proposed by FLORENTINE and BUUS [9] is a generalisation of the Zwicker's model. They assumed that the information from the entire excitation pattern is combined in an optimal way instead of the information coming from one point only.

The global detectability d' can be expressed in this case as a sum of the detectabilities observed for the individual filters (critical bands) d'_i:

$$d' = \sqrt{\sum_i d_i'^2}\,. \tag{2}$$

The detectability in the i-th band is proportional to the square of the AM modulation index [13, 20] (the index is proportional to the difference of the extreme levels of the modulated signal for small modulation depths). Hence, the general form of the detectability d' can be expressed as:

$$d' = K\sqrt{\sum_i \Delta L_i^4}\,, \tag{3}$$

where K is a constant.

Basing on the above assumption, MOORE and SĘK [25, 34] suggested the so-called non-optimal excitation pattern model. They assumed that changes in the stimulus are detected basing on the unweighted sum of the detectabilities of all the active auditory filters. Thus, the detectability d' is assumed to be:

$$d' = K \sum_{i=1}^{n} \frac{\Delta L_i^2}{\sqrt{n}}, \tag{4}$$

where n is the number of active filters/channels.

This work the constant modulation depth was used. Therefore, signal levels corresponding to the signal amplitude extreme values were always the same and led to an amplitude changes detection instead of the modulation rate discrimination. The modulation rate change for the constant AM depth is clearly visible in a signal spectral structure which was an input signal to the model. The differences in the excitation patterns evoked by the AM signals on the modulation rate discrimination threshold were analysed and the d' values were calculated according to Eqs. (3) and (4) taking into account active auditory filters only. The calculated detectability d' was an increasing function of the reference signal modulation rate for both the analysed models. The increase was significant because the change in the calculated detectability d' for the sinusoidal carrier and for modulation rate from 10 to 120 Hz was nearly 100 times higher.

If the AM rate differences detection (discrimination) were based exclusively on the excitation pattern changes evoked by the modulated signals with different modulation rates, then the calculated d' should be a more or less constant function of the reference modulation rate. Thus, the model predicts a much higher sensitivity of the auditory system in the modulation rate discrimination task than that observed experimentally. Moreover, if the decision about the modulation rate discrimination were based on differences in the auditory filters activity, the discrimination thresholds would be lower and their increase with modulation rate should be much slower. Therefore, it seems that a different mechanism is responsible for the AM rate discrimination.

5.2. Frequency discrimination vs. modulation rate discrimination

Our results appear to be consistent with those of WIER *et al.* [39]. They found that the frequency discrimination thresholds for tones in the audio frequency domain are an increasing function of frequency. When the logarithm of the threshold is plotted against the square root of frequency, a linear relation is obtained [39]. Similar relation was found for our results: the logarithm of the thresholds was proportional to the square root of reference modulation rate. High values of correlation coefficients were found for each carrier used: $R = 0.96$, 0.98, and 0.98 for the sinusoidal carrier, the narrowband carrier and the broadband carrier, respectively. This qualitative agreement of the analysed data with the frequency discrimination in the audio frequency domain suggests that the frequency discrimination mechanisms in both the domains are similar. Thus, if the frequency discrimination is considered in the context of frequency selectivity based on the peripheral filtering or the auditory filter bank, the AM rate discrimination may be considered as the frequency selectivity in the modulation rate domain. The selectivity may result from the filtering performed in the analogous filter bank working in the signal envelope frequency domain, i.e. in the modulation filter bank. An alternative mechanism may be the phase locking of neurons tuned to modulation rates with signal amplitude envelope maxima.

It is worthwhile to focus the attention on the differences in the slopes of straight lines describing our results and those obtained by WIER *et al.* when the results are plotted in the sqrt(f) and log(ΔF) plane. For our data these slopes are 0.083, 0.175 and 0.106 for the sinusoidal carrier, the narrowband carrier and the broadband carrier, respectively, while for the results of Wier *et al.* the slope is much smaller and no higher than 0.02. If we assume that the frequency discrimination thresholds result from the auditory system selectivity and the modulation rate discrimination thresholds result from the analogous selectivity in the modulation rate domain, it can be stated that in the latter case the selectivity of the modulation filters is much worse than that in the auditory filters.

5.3. AM rate discrimination vs. modulation filter bank

5.3.1. Modulation masking

Our experimental data can be interpreted by virtue of the modulation filter bank concept. If we assume that the MFB works similarly to auditory filters, our results are consistent with this idea. The frequency discrimination thresholds are increasing functions of frequency, as mentioned above [39]. In many models of the auditory system, the increase in the frequency discrimination thresholds is interpreted as an evidence of the auditory filters existence whose bandwidths are increasing functions of their centre frequency. If we assume that there are modulation filters in the auditory system, where signal envelopes of the auditory filter outputs are analysed, the results obtained in the experiment presented suggest that the bandwidths of the modulation filters should be increasing functions of their centre frequency too, consistently with the DAU's assumption [3].

The relative value of the discrimination thresholds for frequency modulation does not exceed 1% in the range of 250 – 4000 Hz [35] and the thresholds expressed as a portion of the equivalent rectangular bandwidth (ERB) of the auditory filters are approximately constant. The average value of the ratio of the AM rate discrimination thresholds to the modulation rate found in our experiment was about 0.35, i.e. it is approximately 100 times higher than the analogous ratio of the frequency discrimination thresholds to frequency in the audio frequency domain. If we assume that the ratios did describe the filters, it could be said that the relative bandwidth of the modulation filters is much broader than the auditory filter bandwidths. The above conclusion is consistent with the DAU's concept [3] and with the measurements of the modulation tuning curves [36]. It is consistent with the studies of SĘK and MOORE [33] who found that the ringing duration in modulation filters for sinusoidal modulation was rather short and that their activity decay in a time shorter than one modulation period. Thus the quality factor, Q, and frequency selectivity for modulation filters should be much poorer compared to those of the auditory filters.

The MFB concept also explains the differences in the discrimination thresholds for the 450-Hz wide carrier, and in the sinusoidal and broadband carriers (see Figs. 1 and 3). An increase in the modulation discrimination thresholds for the narrowband carrier is probably brought about by random changes in the amplitude envelope. The amplitude of the band of noise (its envelope) is random and depends strongly on the noise bandwidth. Thus, if the spectral range of the inherent noise amplitude fluctuations coincided with the spectral range of modulation rates used in the experiment, we would expect some interaction between the inherent noise those and changes evoked by the modulation being detected. Especially, random noise changes (inherent changes) could be mixed up with changes evoked by modulation, or inherent noise fluctuations could mask the modulation. Thus, the modulation rate discrimination should be more difficult under such conditions and the thresholds should be much higher. To be sure that such kind of interaction (masking) is taking place, the envelope changes spectra were determined for unmodulated bands of noise with a centre frequency of 4 kHz and bandwidths 100, 450 and 600 Hz and for a 6-kHz wide band of noise centred at 3.2 kHz. The Hilbert Transform was used in the amplitude envelope calculations. The next step was the calculation of the spectrum of the AC component of the envelope, as shown in Fig. 5. The maximum values of the amplitude changes were normalised to unity to enable a comparison of the dynamics of the amplitude changes in the analysed noise bands. Figure 5 shows that for each noise bandwidth the spectrum of the amplitude envelope changes is constant in a certain frequency range staring from 0 Hz. The range increases with increasing bandwidth and its level changes too. For example, for the 100-Hz, 450 Hz, 600 Hz and 6 kHz wide noise the ranges are about 70 Hz, 300 Hz, 500 Hz and 4.6 kHz, respectively. However, the

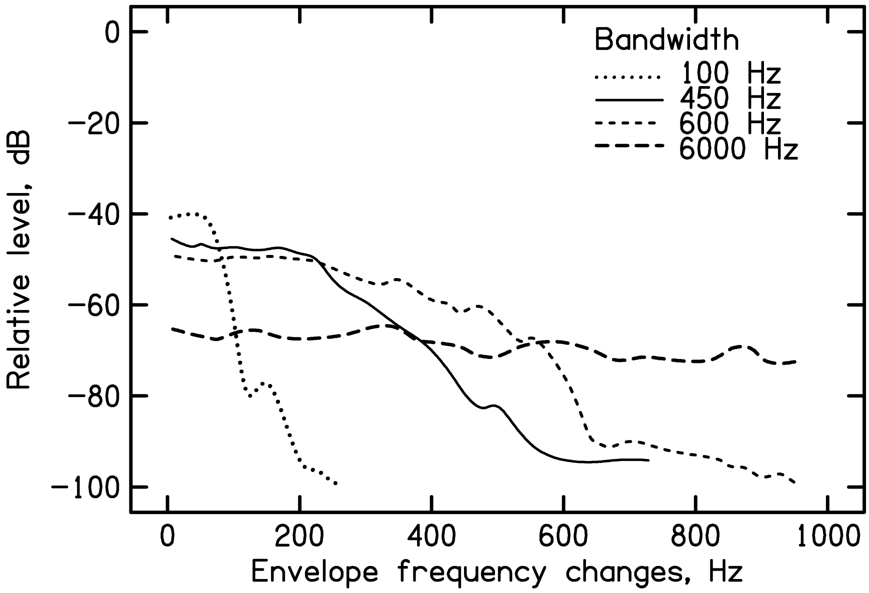

Fig. 5. Spectra of amplitude envelope changes for the noise bandwidths: 100 Hz, 450 Hz, 600 Hz and 6 kHz, respectively.

level of fluctuations for the 450-Hz and the 600-Hz wide noise in the ranges mentioned above was by 20 dB higher than for the 6-kHz wide noise. A considerably high level of amplitude fluctuations for the 450-Hz wide noise compared to that of the 6-kHz wide noise in the frequency range up to 300 Hz suggests that for the narrow band of noise inherent changes in the amplitude envelope may be mixed up with AM-evoked amplitude changes. Presumably, these inherent fluctuations of the 450-Hz wide noise with their high level in the frequency range up to 300 Hz may effectively mask amplitude changes evoked by the sinusoidal modulation. Thus, it is clear that the modulation detection threshold for this carrier is higher than the threshold for the sinusoidal carrier for which no amplitude fluctuation occurs.

For the 6-kHz wide noise, the spectral range of the amplitude envelope changes is very broad; its level is by about 20 dB lower than for the 450-Hz wide noise. Thus, we can assume that the level of the inherent amplitude envelope changes for the 6-kHz wide carrier is low and does not influence the discrimination of the modulation rate or that the influence is much smaller than in the case of the 450-Hz wide carrier. Therefore, the AM rate discrimination thresholds for the 6-kHz wide carrier are similar to thresholds found for the sinusoidal carrier.

The spectral range of the noise envelope fluctuations for the 450-Hz wide noise extends up to 300 Hz, but the modulation rate used in the experiment does not exceed 150 Hz. Thus, we can say that for the 450-Hz wide noise changes evoked by the AM signal at a given modulation rate were masked by the inherent amplitude envelope changes occurring at a frequency close to the modulation rate. The masking observed in our experiment may result from the MFB, similarly to the masking observed in the audio frequency domain which was one of the arguments supporting the auditory filter concept.

5.3.2. Filtering in the modulation filter bank

An alternative interpretation of the pattern of the results can be performed by virtue of the auditory system model containing the MFB. A block diagram of the model is shown in Fig. 6. The first stage of the signal processing (not shown in Fig. 6) in the model is a constant bandpass filter describing the outer- and the middle ear performance [29]. The next stage of the model is a bank of the auditory filters (Gammatone Auditory Filter Bank, GAFB) modelling the signal processing on the basilar membrane. The impulse response of the auditory filters is given by the 5th order Eulers gamma function:

$$\Gamma(t) = t^{n-1} e^{-2\pi t B} \cos(2\pi t f_0), \tag{5}$$

where n is the order of the function, B is the equivalent rectangular bandwidth of the auditory filter with centre frequency f_0. The quality factor, Q, of the auditory filters is about 8 and their bandwidths are described by GLASBERG and MOORE [10] as:

$$B = 24.7(0.00437 f_0 + 1). \tag{6}$$

The next stage of the model is a nonlinear transformation which can be performed by raising signal intensity to the power of 0.3 [24]. This stage mimics the transformation of the basilar membrane motion to compound action potentials in the neurons of the

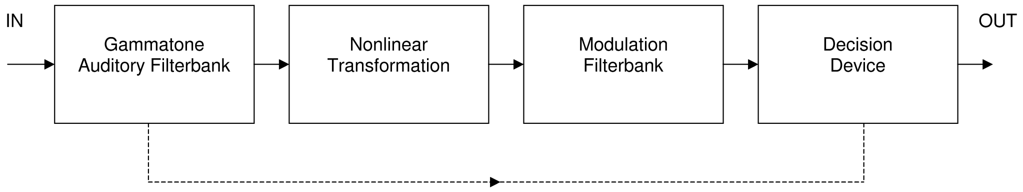

Fig. 6. Block diagram of the auditory system with the modulation filter bank.

auditory nerve. The output of the compressive nonlinearity is fed through the modulation filterbank (MFB).

The impulse response of the filters was described by the Euler's gamma function of the order of 5 and it was assumed, similarly to DAU *et al.* [4, 5], that their quality factor, Q, was 2. In the model presented to each auditory filter a separate MFB was attributed. When a sinusoidal carrier was delivered to the auditory system, only one auditory filter was used, i.e. this one centred at the signal frequency. Many auditory filters spread uniformly in the frequency domain (usually every 10 Hz) were used for the bands of noise. The outputs from the auditory filters were then fed through separate MFBs which worked in the same manner as those for the sinusoidal carrier. The total modulation excitation for given modulation rate consisted of output signals from all the active modulation filters having the same centre frequency. By analogy to the auditory filters, the pattern of the modulation filters activity as a function of their centre frequency was called a modulation excitation pattern. Examples of the patterns for the AM 4-kHz sinusoidal carrier with a modulation rate of 50 Hz and 60 Hz and modulation depth of 50% are shown in Fig. 7. The excitations are normalized according to their maximum values and plotted in a liner scale. The excitation, similarly to the excitation pattern in the audio frequency domain, has a steeper slope at the low-frequency side and a shallow slope at the high-frequency one. However, the dynamic range of the modulation excitation pattern (no higher than 10 dB) is much lower compared to that of the excitation patterns in the audio frequency domain. It is worthy to remember that the dynamic range of the signal envelope for AM with a modulation depth of $m = 0.5$ is about 9.5 dB and it is subsequently decreasing by the non-linear transformation and the temporal integration which smooths the signal amplitude changes. The last stage of the model is a decision device. The nature of the decision device may vary with the task being considered. The decision device transforms outputs from modulation filters or modulation excitation patterns.

The main assumption of our model is that in the AM detection process a subject relies on the modulation filter which is the most active or which has the biggest modulation excitation pattern or which has the best signal-to-noise ratio. It is assumed that for the detection of the AM rate differences (or the modulation rate discrimination measured in our experiment) the subject detects the difference in the modulation excitation patterns evoked by two AM signals with different modulation rates. If the difference, expressed in the model as the RMS values of all modulation filter excitation differences, is equal or bigger than a certain value, the AM rate differences are detected.

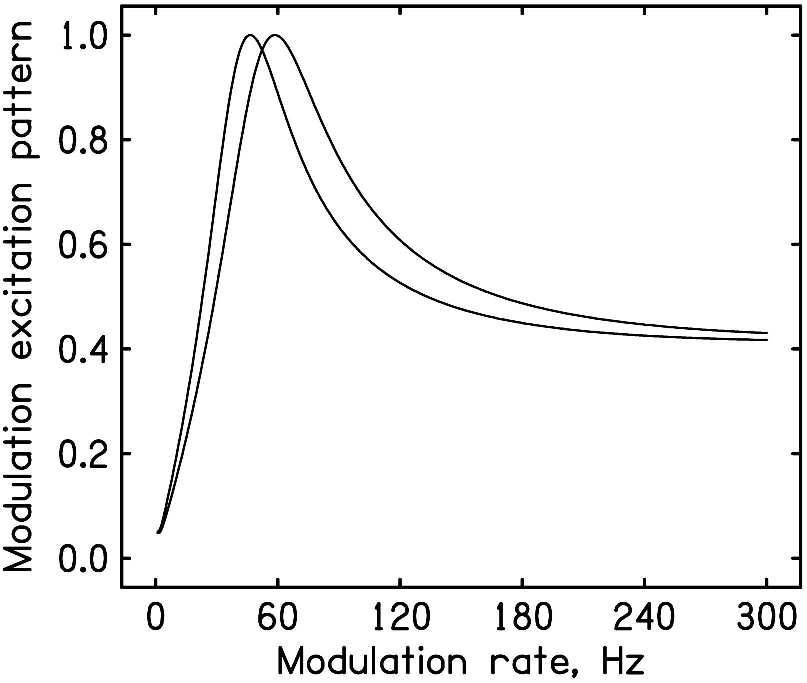

Fig. 7. Normalised modulation excitation patterns describing the activity of neurons tuned to the envelope changes frequency vs. their characteristic frequency (linear units) for the amplitude modulated 4-kHz sinusoidal carrier with modulation rate of 50 Hz and 60 Hz and a modulation depth $m = 0.5$.

The application of this model to the prediction of AM detection thresholds was rather successful because the thresholds calculated for the 4-kHz sinusoidal carrier were independent of the modulation rate in the roughness region, i.e. for modulation rates not higher than 200 Hz.

However, the model application to the AM rate discrimination task was not as successful as it was for the AM detection thresholds: only a qualitative agreement between the model predictions and the experimental data was found. In other words: the AM rate discrimination thresholds were increasing functions of the reference modulation rate, however the increase of the calculated thresholds was smaller than for those measured experimentally. Generally, it can be stated that the predicted ability to discriminate the modulation rate was much better than found in the experiment. Similar results were obtained for different quality factors of the modulation filters (i.e. $Q = 0.5$, 1 and 2) and for different statistics calculated by the decision device (e.g. a maximum value of excitation differences, a difference between the excitation on one side of the pattern and a maximum excitation or a difference in the excitation maxima). Hence, it seems that the decision about the AM rate discrimination is not only supported by the differences in the modulation excitation patterns or by the MFB having the best signal-to-noise ratio. If it were so, the AM rate discrimination thresholds would increase much slower than it was found in our experiment.

It seems reasonable that each model stage influences the subjects decision. Apart from the modulation excitation pattern the most important stage of the model seems to be signal processing on the basilar membrane shown schematically in Fig. 6 by a dashed line. Moreover, it is also reasonable to take into account an internal noise which is present on each level of the auditory system caused by spontaneous activity of the neurons. When the internal noise is arbitrary chosen on each level, or when an arbitrary signal from the output of the auditory filterbank is fed through the decision device, the predictions of the model are much closer to the experimental data.

Our model, as well as many other models, represents a class of models which do not account for all known phenomena and processing rules in the auditory system. The most important phenomenon, strongly related to signal processing in the auditory system and not taken into account, is the phase locking: neural firings tend to occur at a particular phase of the stimulating waveform, so that there is a temporal regularity in the firing pattern in response to a periodic stimulus. The frequency discrimination (in the audio frequency domain) is much better than that resulting from the quality factors of the auditory filters. The phase locking seems to be the mechanism responsible for this observation. A similar situation is observed for the frequency selectivity in the modulation rate domain: the selectivity of the modulation filters is too poor to explain the auditory system ability to resolve two modulation rates. Probably, the phase locking is the crucial factor for detection of the signal envelope frequency changes.

6. Conclusions

The most important conclusions of this paper are as follows:

1. The AM rate discrimination thresholds are monotonically increasing functions of the reference modulation rate. For the sinusoidal carrier and the 6-kHz wide noiseband carrier, the thresholds are similar. For the 450-Hz wide noiseband carrier the thresholds are significantly higher.

2. The relation between the AM rate discrimination thresholds and the modulation rate is qualitatively similar to that between the frequency discrimination thresholds and the frequency in the audio frequency domain. The similarity allows to assume that the frequency discrimination processes in these two domains may be similar to each other.

3. The modulation rate discrimination thresholds obey the Webers law approximately: they are also constant fractions of the modulation rate, as it is in the audio frequency domain. By analogy it can be stated that the mechanisms responsible for the audio frequency discrimination and modulation rate discrimination work in a similar manner.

4. The values of the modulation rate discrimination thresholds for the 450-Hz wide carrier higher than for other carriers may result from amplitude changes masking (interaction) evoked by AM detected by internal changes in the noise band amplitude. This form of masking also confirms the MFB concept.

5. The overall pattern of the results is consistent with the MFB concept in that amplitude envelope changes are processed. The monotonically increasing relation between the modulation rate discrimination thresholds and the modulation rate supports the idea that the bandwidths of the modulation filters increase with the centre frequency.

6. The model which could account for our experimental data should take into account the decision cues coming both from the MFB and from the AFB as well as the phase locking in the audio frequency domain and in the modulation rate domain.

Acknowledgements

This work was supported partly by the Committee for Scientific Research, project no. 7T07B05216, 7T11D 00520 and 8T07B06020. Computer calculations were carried out in Poznań Supercomputer and Networking Centre. We would like to thank an anonymous reviewer for helpful comments on earlier version of this paper.

References

[1] S.P. Bacon and D.W. Grantham, *Modulation masking: effects of modulation frequency, depth and phase*, J. Acoust. Soc. Am., **85**, 2575–2580 (1989).

[2] A.S. Bregman, *Auditory Scene Analysis: The Perceptual Organisation of Sound*, Cambridge, Mass., Bradford Books, MIT Press, 1990.

[3] T. Dau, *Modeling auditory processing of amplitude modulation*, University of Oldenburg, 1996.

[4] T. Dau, B. Kollmeier and A. Kohlrausch, *Modeling auditory processing of amplitude modulation: I. Detection and masking with narrowband carriers*, J. Acoust. Soc. Am., **102**, 2892–2905 (1997).

[5] T. Dau, B. Kollmeier and A. Kohlrausch, *Modeling auditory processing of amplitude modulation: II. Spectral and temporal integration*, J. Acoust. Soc. Am., **102**, 2906–2919 (1997).

[6] T. Dau, B. Kollmeier and A. Kohlrausch, *Modeling modulation perception: modulation low-pass filter or modulation filter bank?*, [in:] Psychoacoustics, Speech and Hearing Aids, B. Kollmeier [Ed.], World Scientific, Singapore 1996.

[7] T. Dau, D. Püschel and A. Kohlrausch, *A quantitative model of the "effective" signal processing in the auditory system.II. Simulations and measurements*, J. Acoust. Soc. Am., **99**, 3623–3631 (1996).

[8] D.S. Evert and T. Dau, *Characterising frequency selectivity for envelope fluctuations*, J. Acoust. Soc. Am., **108**, 1181–1196 (2000).

[9] M. Florentine and S. Buus, *An excitation-pattern model for intensity discrimination*, J. Acoust. Soc. Am., **70**, 1646–1654 (1981).

[10] B.R. Glasberg and B.C.J. Moore, *Derivation of auditory filter shapes from notched-noise data*, Hear. Res., **47**, 103–138 (1990).

[11] J.W. Hall, J.H. Grose and M.P. Haggard, *Comodulation masking release for multi-component signals*, J. Acoust. Soc. Am., **83**, 677–686 (1988).

[12] T.E. Hanna, *Discrimination and identification of modulation rate using a noise carrier*, J. Acoust. Soc. Am., **91**, 2122–2128 (1992).

[13] W.M. Hartmann and M.A. Klein, *Theory of modulation detection for low modulation frequencies*, J. Acoust. Soc. Am., **67**, 935–946 (1980).

[14] T. Houtgast, *Frequency selectivity in amplitude-modulation detection*, J. Acoust. Soc. Am., **85**, 1676–1680 (1989).

[15] R.H. Kay, *Hearing of modulation in sounds*, Physiol. Rev., **62**, 894–975 (1982).

[16] J. Lee, *Amplitude modulation rate discrimination with sinusoidal carriers*, J. Acoust. Soc. Am., **96**, 2140–2147 (1994).

[17] C. Lorenzi, C. Micheyl and F. Berthommier, *Neuronal correlates of perceptual amplitude-modulation detection*, Hear. Res., **90**, 219–227 (1995).

[18] D. McFadden and B.A. Wright, *Temporal decline of masking and comodulation masking release*, J. Acoust. Soc. Am., **92**, 144–156 (1992).

[19] A.R. Moller, *Dynamic properties of primary auditory fibers compared with cells in the cochlear nucleus*, Acta Physiol. Scand., **98**, 157–167 (1976).

[20] B.C.J. Moore, *Across-channel processes in auditory masking*, J. Acoust. Soc. Jpn. (E), **13**, 25–37 (1992).

[21] B.C.J. Moore, *An Introduction to the Psychology of Hearing*, 4th Ed., Academic Press, London 1997.

[22] B.C.J. Moore, *Perceptual Consequences of Cochlear Damage*, Oxford University Press, Oxford 1995.

[23] B.C.J. Moore, *Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids*, Ear Hear., **17**, 133–161 (1996).

[24] B.C.J. Moore, *An Introduction to the Psychology of Hearing* [in Polish], I ed, PWN, Poznań 1999.

[25] B.C.J. Moore and A. Sęk, *Effects of carrier frequency and background noise on the detection of mixed modulation*, J. Acoust. Soc. Am., **96**, 741–751 (1994).

[26] E. Ozimek, J. Konieczny, Y. Suzuki and T. Sone, *Random changes in envelope of AM tones and their detection*, J. Acoust. Soc. Jpn. (E), **19**, 83–91 (1998).

[27] A.R. Palmer, *Neural signal processing*, [in:] Hearing, B.C.J. Moore [Ed.], Academic Press, San Diego 1995.

[28] R.D. Patterson and B.C.J. Moore, *Auditory filters and excitation patterns as representations of frequency resolution*, [in:] Frequency Selectivity in Hearing, B.C.J. Moore [Ed.], Academic, London 1986.

[29] J.O. Pickles, *An Introduction to the Physiology of Hearing*, 2nd Ed., 2nd ed London, Academic Press, 1988.

[30] C.J. Plack and B.C.J. Moore, *Temporal window shape as a function of frequency and level*, J. Acoust. Soc. Am., **87**, 2178–2187 (1990).

[31] A. Rees and A.R. Moller, *Responses of neurons in the inferior colliculus of the rat to AM and FM tones*, Hearing. Res., **10**, 301–310 (1983).

[32] C.E. Schreiner and J.V. Urbas, *Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF)*, Hearing Res., **21**, 227–241 (1986).

[33] A. Sęk and B.C.J. Moore, *Detectability of gaps in the modulation of a 4-kHz carrier*, J. Acoust. Soc. Am., **101**, 3082 (1997).

[34] A. Sęk and B.C.J. Moore, *The detection of modulation at low modulation rates*, [in:] Auditory Perception: Some Principles and Applications, U. Jorasz [Ed.], Adam Mickiewicz University Press, Poznań 1996.

[35] A. Sęk and B.C.J. Moore, *Frequency discrimination as a function of frequency, measured in several ways*, J. Acoust. Soc. Am., **97**, 2479–2486 (1995).

[36] A.P. Sęk and E.B. Skrodzka, *Masking and tuning in the modulation frequency domain*, XIIth Meeting of the International Society of Psychophysics, Poznań, August, 287–291, 1997.

[37] A.P. Sęk and E.B. Skrodzka, *Tuning in the amplitude modulation rate domain*, Archives of Acoustics, **23**, 491–498 (1998).

[38] N.F. Viemeister, *Temporal modulation transfer functions based on modulation thresholds*, J. Acoust. Soc. Am., **66**, 1364–1380 (1979).

[39] C.C. Wier, W. Jesteadt and D.M. Green, *Frequency discrimination as a function of frequency and sensation level*, J. Acoust. Soc. Am., **61**, 178–184 (1977).

[40] E. Zwicker and H. Fastl, *Psychoacoustics — facts and models*, Springer-Verlag, Berlin 1990.