ARCHIVES OF ACOUSTICS
23, 3, 379–390 (1998)

PERCEPTION OF MIXED MODULATION FOR SINGLE COMPONENTS IN HARMONIC COMPLEX FOR HIGH MODULATING FREQUENCIES

M.J. KIN and A.B. DOBRUCKI

Institute of Telecommunications and Acoustics
Wrocław University of Technology
(50-370 Wrocław, Wybrzeże Wyspiańskiego 27)

Results and problems of perception of mixed modulation in a harmonic multitone are discussed. The experimental research was done for higher modulating frequencies (70 and 200 Hz) and a 5-harmonics complex with fundamental frequency of 256 Hz. The results indicated that for a modulating frequency of 70 Hz, the perception of frequency modulation in the presence of the threshold values of the amplitude modulation is different from the perception of amplitude modulation in the presence of the threshold values of the frequency modulation. This means clearly that there are two mechanisms of the perception for these kinds of modulation.

1. Introduction

Problems of perception of simultaneous amplitude and frequency modulation have been widely discussed [3, 4, 6, 14, 15]. For low modulating frequencies, the perception is based on the changes in pitch and loudness and the auditory system follows the temporal structure of the sound. For higher modulating frequencies, when the spectrum covers a wider frequency range, the perception of AM, FM and MM is based mostly on the spectrum of the modulated sound, i.e. on the low and high sidebands produced by modulation process, and it could be said that, in this case, the auditory system analyses the spectral structure of the sound. This means that we perceive a steady "rough" sound or two tones of different pitches depending on the frequency of modulation and the carrier. VOGEL [21, 22] presented a model, derived partially from the TERHARDT'S model [19], in which the partial roughness is evaluated within each critical band on the basis of the fluctuation in excitation. FASTL [5] suggested that for AM-modulated broad-band noise the differences in the level are evaluated using the masking period patterns produced by AM broad-band noise differing in the modulation frequency, degree of modulation, and level. ZWICKER [25] compared the sensations produced by the AM and FM octave-band noise. The result indicated a correlation between the sensations produced by these two kinds of noise which suggests that the roughness for both AM and FM is perceived by the same mechanism. In other ZWICKER'S works [23, 24, 26] it was found that the just-noticeable modulation index for AM and the just-noticeable

frequency deviation for the FM tones, changed in the same manner of change as the function of the modulating frequency when other parameters were constant. This means that the threshold of roughness behaves similarly for AM and FM tones. TERHARDT [18, 20] compared the roughness of AM tones with that produced by FM tones at a constant modulation index and concluded that the mechanism producing this effect was similar for the two kinds of modulation.

ZWICKER [27] and MAIWALD [12, 13] proposed a functional scheme in which the frequency and intensity differences are assumed to be detected by one single mechanism rather than by two independent mechanisms. CONINX [3, 4] found an independent detection of the pitch and loudness differences on the basis of his experiment in which the detection of combined differences in frequency and intensity was investigated. HARTMANN and HNATH [7] determined the influence of each component of the AM, FM and MM signal spectrum on the modulation threshold values. Most important is the relation between the MM threshold and the ratio of the frequency and amplitude modulation indices for coincident and opposed phases between the amplitude and frequency modulating signals. OZIMEK and SĘK'S research [15] shows that there exists a fairly complex perception mechanism for MM signals which depends on the kind and frequency of the modulating signal. For the modulating frequency in the "spectral" region, there are probably two independent mechanisms that may operate either separately or in combination; the component whose frequency is lower than the carrier frequency of the signal determines the perception of simultaneous amplitude and frequency changes.

From the musical point of view, two aspects of the changes in the spectra should be taken into consideration: firstly, when two sidebands have frequencies in harmonic order with the carrier frequency and they assemble in some kind of a harmonic multitone; secondly, when those sidebands are not in harmonicity that causes some kind of dissonance. The second aspect was widely discussed by BREGMAN [2] on the basis of perceptual grouping. According to this theory, listeners could "remove" some nonharmonic partials from the consonance complex creating a sound image based on two separate perceptual streams: one with the products of modulation, and another one – the harmonic signal with the carrier.

The main aim of this paper is to find how the human ear can percept a mixed modulation in the roughness and "spectral" regions for the single components at the presence of other partials of the harmonic complex. Another interesting problem was the interaction between these two kinds of modulation.

2. Stimuli and procedure

When a pure tone with amplitude A and frequency f_c:

$$x(t) = A \cos 2\pi f_c t, \tag{1}$$

is processed by frequency and amplitude modulation using another pure tone with amplitude M and frequency F_m:

$$y(t) = M \cos 2\pi F_m t, \tag{2}$$

the result is given by the following formula:

$$X(t) = A(1 + m\cos 2\pi F_m t)\cos(2\pi f_c t + \beta \sin 2\pi F_m t), \tag{3}$$

where: $m = M/A$ – AM modulation index, $\beta = \Delta f / F_m$ – FM modulation index, and Δf – frequency deviation.

Formula (3) expresses the spectral structure of the mixed-modulated tone containing three components:

— a carrier of frequency f_c and amplitude $A \cdot J_0(\beta)$,

— a low sideband of frequency $f_c - F_m$ and amplitude $A\left|J_1(\beta)\cdot\left(1-\frac{m}{\beta}\right)\right|$,

— a high sideband of frequency $f_c + F_m$ and amplitude $A\left|J_1(\beta)\cdot\left(1+\frac{m}{\beta}\right)\right|$,

(for $\beta \ll 1$ only), where $J_0(\beta)$ and $J_1(\beta)$ are the Bessel functions of zero and first order of the argument β, respectively.

When only one of the partials is modulated, the spectrum contains other harmonics with their own amplitudes, modulated harmonic and two sidebands which result from the modulation process.

Test signals were generated by an IBM PC computer with a 16-bit sound board at a sampling rate of 16 000 points per second. The phases of each of the generated components and the modulation signal were equal to zero. Both the duration of the signal and the pause between two stimuli was 1 s. The stimulus presentation, timing and response recording were controlled by the computer. In the first interval, subjects listened to the AM or FM modulated single components of the investigated complex at the threshold values of m or β obtained in the experiment, presented partially previously [9]. The second interval always contained the same modulated partial with constant threshold values of m or β and a co-existing FM or AM modulation, respectively. This sequence was used because it could help the listeners to decide whether the sounds in those two intervals were the same or different and listeners did not have to care which one contained AM, FM, and MM-modulated partials. During this experiment, the AM and FM modulating signals were in-phase, i.e. the maximum in amplitude coincided with the maximum in frequency. An adaptive PEST method [16, 17] with the "yes – no" subjects' task was used. The modulation index values were decreased after two correct responses and increased by an apprioprate step after one incorrect response. Initially, the step size was 6 dB, but it was reduced to 0.5 dB after the first four reversals. A response feedback was not provided. The threshold value of the AM index, in the presence of threshold values β of the FM presented alone, was denoted m_β, and similarly, the threshold value of the FM index in the presence of threshold values m of the AM presented alone, was denoted by β_M. The values of m_β and β_M were defined as an average reversal level occuring during 10 trials starting with the fourth reversal. Ten threshold estimates were obtained for each observer under every condition.

In the experiment, the complex consisting of 5 components of 256 Hz as the fundamental frequency was presented via an Audiostatic ES 100 electrostatic loudspeaker and a Pioneer A 400 X amplifier with the level according to the isophone of 60 phones. Five male subjects with normal, good hearing participated in the experiment. All of them

were experienced in psychoacoustic tests and were paid for their service. The subjects' task was to answer the question: is the sound in the second interval the same as in the first one or not?

3. Results

The results of the experiment are presented in Figs. 1 and 2. They show the threshold values of m_β and β_M obtained for a mixed modulation in comparison with the values of m and β obtained under the same listening conditions but for the modulation of amplitude or frequency only. The standard deviations for these results did not exceed 10% of the obtained values. It can be seen that the MM thresholds are different from those

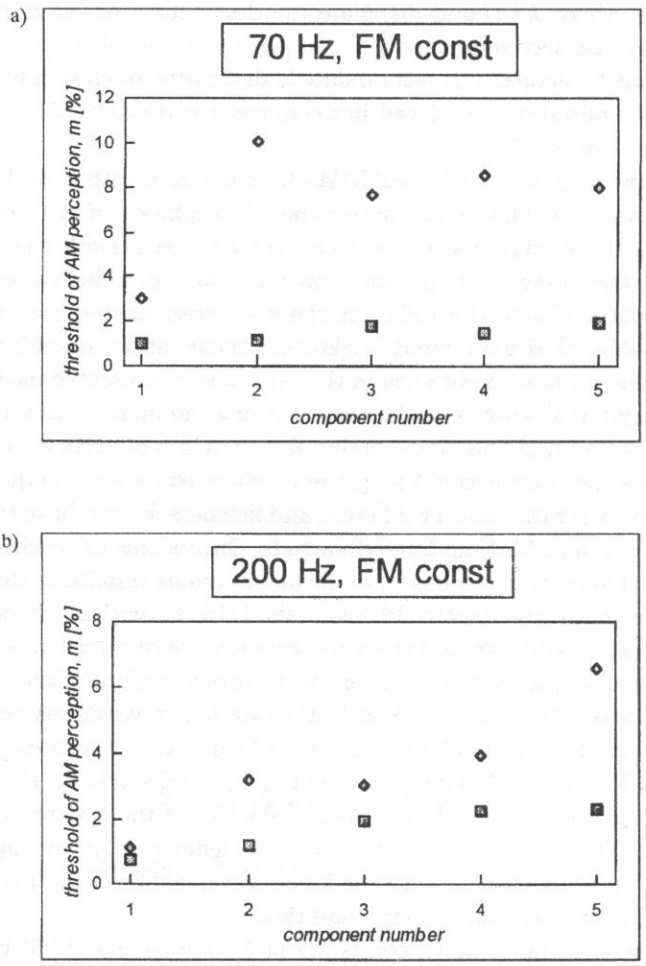

Fig. 1. Threshold values of the perception of amplitude modulation for the modulating frequencies: a) 70 Hz, and b) 200 Hz. ■ – amplitude modulation exposed separately, ♦ – amplitude modulation exposed with FM at its threshold values.

of AM and FM. The amplitude modulation in the presence of the threshold values of the FM perception has higher threshold values than the amplitude modulation without simultaneous fast changes in the frequency (Fig. 1). This takes place for both the modulating frequencies, i.e. for 70 and 200 Hz. The perception of the frequency modulation in the presence of the threshold values of the AM perception depends on the modulating frequency (Fig. 2). For modulating frequency of 200 Hz, the situation is similar to the AM perception in the presence of the FM threshold (Fig. 1), while for modulating frequency of 70 Hz the situation is completely different, i.e. the thresholds of the perception of frequency changes in the presence of the threshold values of AM are higher than in the case when only FM exists. Another interesting fact, which has been found in this experiment, is that for the 70 Hz-amplitude modulation in the presence of frequency

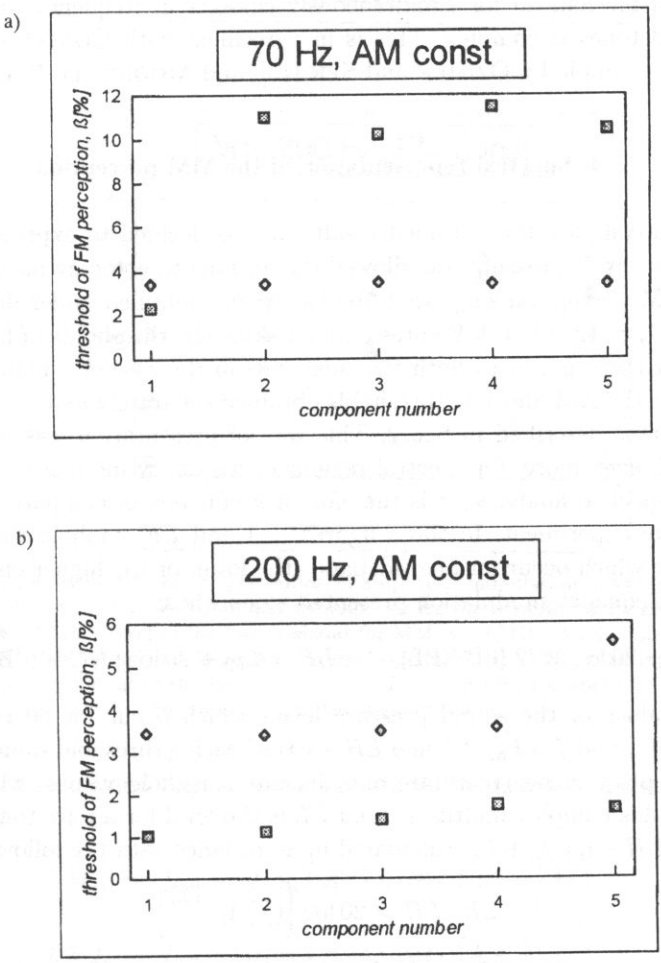

Fig. 2. Threshold values of the perception of frequency modulation for the modulating frequencies: a) 70 Hz, and b) 200 Hz. ■ – frequency modulation exposed separately, ♦ – frequency modulation exposed with AM at its threshold values.

modulation the threshold values of m are approximately the same for the partials 2 to 5, while for the fundamental component the threshold of AM is three times lower. For a modulating frequency of 200 Hz, the threshold values of m increase for higher components however this increase is greater than for the case when only the AM perception was measured.

For the 70 Hz-frequency modulation in the presence of an amplitude modulation, the perceived values of β are smaller than the thresholds of perception when only FM occured. These values are equal for all the investigated partials. For the modulating frequency of 200 Hz, the threshold values of the FM perception are higher for a mixed modulation in comparison to these when only FM exists alone.

The above mentioned facts suggest that for modulating frequencies equal to 70 and 200 Hz, an interaction between the modulations of amplitude and frequency exists, and the perceptual mechanism for simultaneously changes in frequency and amplitude of the modulated tones is complex. This is in agreement with CONINX suggestion [3, 4] confirmed, for example by OZIMEK and SĘK [15], and MOORE and SĘK [14].

4. Spectral representation of the MM perception

In order to compare the obtained results, it was decided to express them as levels of sound. This way of presentation allowed the authors to compare all the results referring to the MM perception suggested previously but obtained under different listening conditions [3, 7, 8, 12, 14, 15]. Figures 3 and 4 show the thresholds of hearing for tones that frequency corresponds to both the sidebands in the case of a multitone. They are compared with the AM and FM thresholds obtained separately and the MM thresholds for the conditions described in Sec. 2. This way of presentation was used because for roughness and, even more, for spectral regions of frequency modulation the human ear behaves as a spectral analyzer; it is therefore a useful way of comparing all the results obtained in the experiment. In those figures, LA and LF symbolize the levels of the sound pressure which occur for one sideband (the lower or the higher one) as a result of amplitude or frequency modulation presented alone; thus:

$$LA = L_{60} + 20\log M/2 \text{ [dB SPL]}, \qquad LF = L_{60} + 20\log|J_1(\beta)| \text{ [dB SPL]}. \tag{4}$$

L_{60} are the values of the sound pressure levels which occur for 60 isophone for the frequency $f_c - F_m$ and $f_c + F_m$. LL and LH are the hearing threshold values resulting only from masking phenomena (these are not absolute threshold values) which "normally" could exist in this complex multitone, and LL is the level for a pure tone at frequencies $f_c - F_m$, and LH – for $f_c + F_m$ calculated in accordance with the following formula [1].

$$LL,\ LH = 20\log\left[\left(\sum A_i^{\alpha}\right)^{1/\alpha}\right], \tag{5}$$

where A_i is the amplitude of the masked threshold produced by i-th component of a 5-harmonic complex and $\alpha = 0.8$ is a typical value for the frequency domain when more than one masker is present at the same time [10, 11].

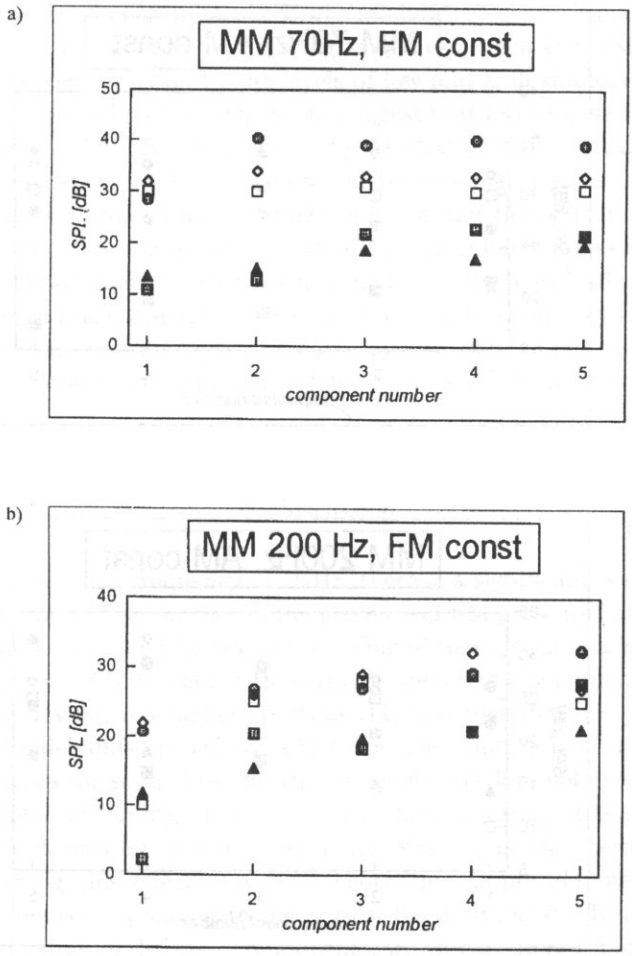

Fig. 3. Spectral representation of the perception of amplitude modulation in the presence of the threshold values of frequency modulation and without FM for the modulating frequencies: a) 70 Hz, and b) 200 Hz. ■ – LML – level of the low sideband for MM, • – LMH – level of the high sideband for MM, ▲ – LA – level of sidebands for AM exposed separately, □ – LL – masking threshold for the frequency corresponding to the low sideband, ◊ – LH – masking threshold for the frequency corresponding to the high sideband.

The levels of the low and high sidebands produced by mixed modulation were calculated as follows:

$$\begin{aligned} LML &= L_{60} + 20\log\left(A \cdot \left|J_1(\beta) \cdot \left(1 - \frac{m}{\beta}\right)\right|\right), \\ LML &= L_{60} + 20\log\left(A \cdot \left|J_1(\beta) \cdot \left(1 + \frac{m}{\beta}\right)\right|\right). \end{aligned} \tag{6}$$

It can be seen that the perception of mixed modulation changes in a different manner for the two cases, i.e. when m = const or when β = const. In the case when both

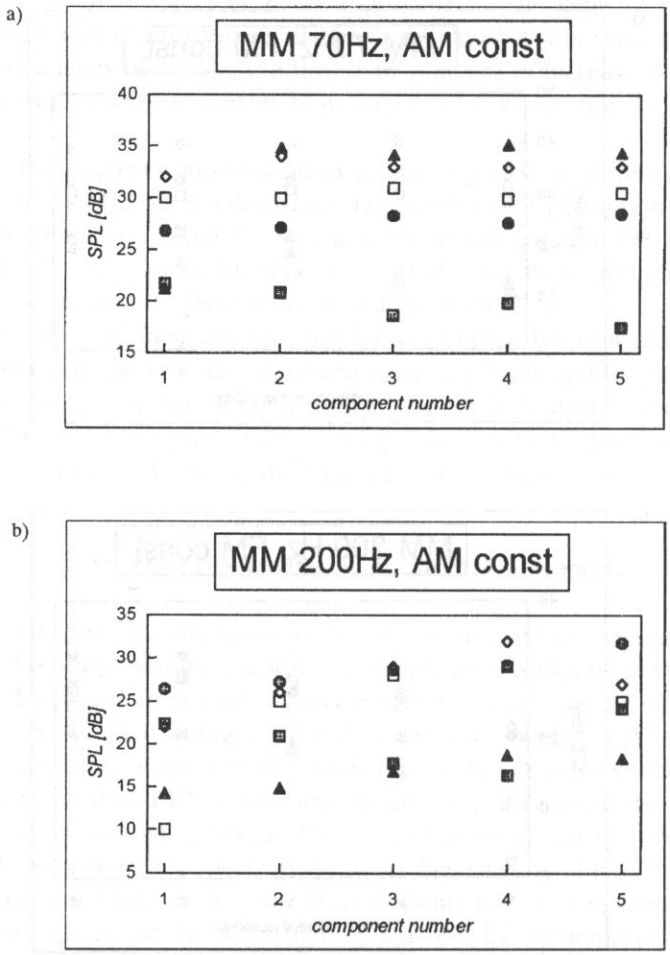

Fig. 4. Spectral representation of the perception of frequency modulation in the presence of the threshold values of amplitude modulation and without AM for the modulating frequencies: a) 70 Hz and b) 200 Hz. ■ – LML – level of the low sideband for MM, • – LMH – level of the high sideband for MM, ▲ – LF – level of sidebands for FM exposed separately, □ – LL – masking threshold for the frequency corresponding to the low sideband, ◊ – LH – masking threshold for the frequency corresponding to the high sideband.

kinds of modulation (AM and FM) are applied to a single component in the harmonic complex and frequency modulation is presented at the threshold levels (Fig. 3), the levels of the low sideband, resulting from mixed modulation, are lower than the thresholds of hearing at the corresponding frequencies when all the 5-component complex is on. Previous results [3, 5, 7, 15] indicated rather the opposite – the perception of AM, FM and MM is based on the lower sideband and the higher one is masked by the carrier. However, those results were obtained in experiments without masking and under different hearing conditions different from those in our experiment. Our results show that this kind of perception is based rather on the higher sidebands the levels of which

exceed the thresholds of hearing (for a 70-Hz modulator) or are equal to them (for a 200 Hz modulation). For a 70 Hz modulating frequency, when the AM is presented at the threshold of perception (Fig. 4), the levels of low and high sidebands are below the threshold of hearing for all partials which can suggest that in this case the summation of partially roughnesses takes place as proposed by TERHARDT [20] and VOGEL [21, 22]. For the modulating frequency of 200 Hz, the low sideband is higher than the corresponding hearing threshold (resulting from masking) only for the fundamental harmonic, and equal for the 5-th one, but for the basic partial this low sideband does not play any role because its level is lower than the absolute hearing threshold. This leads to the conclusion that the low sideband resulting from mixed modulation does not play an important role in the MM perception; this perception is based mostly on the high sideband caused by the modulation process or on both the sidebands as a result of summation of partial roughnesses.

5. Discussion

For the the modulating frequency of 70 Hz, there is a significant decrease in the AM threshold for the 2-nd to 5-th partials in the presence of FM presented at the threshold of detection. This may be caused by two factors. One of them is the masking phenomenon according to which some components appearing in the middle of a multitone are masked stronger than the lowest and highest partials [11] and therefore it is necessary to use a higher level of sidebands caused by AM to produce an audible "roughness effect". Moreover, in all cases the sound level for the low sidebands does not exceed the threshold of hearing calculated according to equation (5), which can suggest that the "roughness effect" is based not only on the low sidebands existing in the spectrum. The main argument supporting this effect may be a short-time amplitude spectrum evaluated with a time window corresponding to the critical-band filters [20]. In the case like this, two subthreshold values of different components, which can cause separate roughnesses, should be added in one critical band to make an audible sensation. This fact is confirmed for all partials in the experiment, if it can be assumed that both the lower and higher sidebands are in one critical band. Secondly, in this case there is no specific pitch which could help the listeners to recognize the MM process as a dissonance effect and for all the partials a disturbing specific sound appears consisting of two additional components in the spectrum which are nonharmonic to the other ones. Despite the levels of the sidebands, we can still say that a partial roughness summation mechanism may exist when the 70 Hz sinusoid is used as an amplitude modulator. The third aspect, which should be noticed here, is that the change of the excitation place appearing for the 70 Hz-modulation may play an important role in the MM-perception, especially when the modulating signal contains only a clear sinusoid.

For the higher modulating frequency (200 Hz), the perceptual grouping mechanism [2] plays the main role in the AM perception in the presence of FM for all the modulated harmonics. For all partials, the threshold values of the modulation index increase for the higher partials. For the first four harmonics, the level of the low sideband is lower

than the hearing threshold which suggests that a summation mechanism of subthreshold values may exist even for partials from different critical bands except for the 5-th partial where the low sideband is higher than the hearing threshold for its frequency. For high sidebands, the levels are equal to the hearing threshold values obtained on the basis of the masking phenomena.

For a modulating frequency of 70 Hz, and the FM perception in the presence of AM presented with the threshold of perception, the levels of both the sidebands do not exceed the hearing threshold values; this is the main difference between the AM and FM perceptions found in this experiment. This means that when the frequency deviation is smaller than the critical bandwidth, the human ear can also perceive temporal changes in the sound level of the processed partials. Finally, it can be concluded that the perception of "roughness" is based on both the sound spectra with sidebands produced by the modulation process and on fast changes in the frequency of the modulated component. However, a boundary which would clearly separate these two mechanisms has not been defined precisely for MM-modulated tones.

For a modulating frequency of 200 Hz, the spectral model of hearing plays the decisive role in the FM perception in the case of AM presented simultaneously. In the case of modulated partials of the investigated harmonic complex, all the rules applying to the masking phenomena have been confirmed by the obtained results, however, for higher components the high sideband produced by mixed modulation plays the most important role in the MM perception. Another phenomenon that can help listeners in the MM detection is the perceptual grouping mechanism. According to this, for some of the modulated partials MM can be perceived as an additional pitch occuring in multitones which is usually out of tune of the fundamental frequency of the complex; this causes two independent perceptual streams [2]. The summation of the next (or previous) partial and the higher (or lower) sideband, respectively, takes place in one critical band, which suggests that a specific pitch for that partial could be changed and, from the musical point of view, this could be considered as a simple way to detect these simultaneous fast changes in the amplitude and frequency of the modulated partials.

6. Conclusions

The results of this experiment have confirmed all the rules applying to the masking process in the investigated multitones. On the basis of the sound spectra of investigated complexes, it should be noticed that the higher sidebands are more important for MM-perception than the lower ones. However, it can be said that two underthreshold values of the lower and higher sideband levels may be added in some cases, even if the products of the modulation process are spread widely in the frequency domain. Another important rule refers to the summation of excitation: in most cases, for 70 Hz-modulation, such a summation does exist allowing listeners to detect fast changes in the amplitudes and frequencies of a single components in the harmonic complex.

This work was supported by the Polish Committee of Scientific Research, grant no. 7 T07 B 034 08.

References

[1] J.G. BEERENDS and J.A. STEMERDINK, *A perceptual audio quality measure based on a psychoacoustic sound representation*, J. Audio Eng. Soc., **40**, pp. 963–978 (1992).

[2] A.S. BREGMAN, *Auditory Scene Analysis – The Perceptual Organization of Sound*, MIT Press, Cambridge, Mass 1990.

[3] F. CONINX, *The Detection of Combined Differences in Frequency and Intensity*, Acustica, **39**, pp. 137–150 (1978).

[4] F. CONINX, *The Perception of Combined Frequency and Amplitude Modulations with Clearly Audible Modulation Depths*, Acustica, **39**, pp. 151–154 (1978).

[5] H. FASTL, *Roughness and temporal masking patterns of sinusoidally amplitude modulated broadband noise*, [In:] Psychophysics and psychology of hearing, E.F. EVANS and J.P. WILSON [Eds.], Academic Press, London 1977.

[6] L.L. FETH, *Combinations of amplitude and frequency differences in auditory discrimination*, Acustica, **26**, pp. 67–77 (1972).

[7] W.M. HARTMANN and G.M. HNATH, *Detection of mixed modulation*, Acustica, **50**, pp. 297–312 (1982).

[8] S. KEMP, *Roughness of frequency-modulated tones*, Acustica, **50**, pp. 126–133 (1982).

[9] M.J. KIN and A.B. DOBRUCKI, *Roughness effect as a result of amplitude or frequency modulation of single components in harmonic complex*, Proc. of 102nd AES Convention, Munich, Germany, preprint 4481 (L2) (1997).

[10] R.A. LUFTI, *Additivity of simultaneous masking*, J. Acoust. Soc. Am., **73**, pp. 262–267 (1983).

[11] R.A. LUFTI, *A power-law transformation predicting masking by sounds with complex spectra*, J. Acoust. Soc. Am., **77**, pp. 2128–2136 (1985).

[12] D. MAIWALD, *Ein Funktionsschema des Gehörs zur Beschreibung der Erkennbarkeit klein Frequenz und Amplitudenänderungen*, Acustica, **18**, pp. 81–92 (1967).

[13] D. MAIWALD, *Die Berechnung von Modulationsschwelten mit Hilfe eines Funktionsschemas*, Acustica, **18**, pp. 193–207 (1967).

[14] B.C.J. MOORE and A. SĘK, *Detection of combined frequency and amplitude modulation*, J. Acoust. Soc. Am., **92**, pp. 3119–3130 (1992).

[15] E. OZIMEK and A. SĘK, *Perception of amplitude and frequency modulated signals (mixed modulation)*, J. Acoust. Soc. Am., **87**, pp. 1598–1603 (1987).

[16] M.M. TAYLOR, *PEST: Efficient estimates on probability functions*, J. Acoust. Soc. Am., **41**, pp. 1224–1235 (1967).

[17] M.M. TAYLOR, *PEST reduces bias in forced choice psychophysics*, J. Acoust. Soc. Am., **74**, pp. 375–387 (1983).

[18] E. TERHARDT, *Über akustische Rauhigkeit und Schwankungsstrke*, Acustica, **20**, p. 215 (1968).

[19] E. TERHARDT, *Schallfluctuation und Rauhigkeitsempfinden: Akustik und Schwingungstechnik* VDI-Verlag, Düsseldorf 1971.

[20] E. TERHARDT, *On the perception of periodic sound fluctuation (Roughness)*, Acustica, **30**, pp. 201–213 (1974).

[21] A. VOGEL, *Roughness and its relation to the time-pattern of psychoacoustical excitation*, [In:] Facts and Models in Hearing, E. ZWICKER and E. TERHARDT [Eds.], Springer, Berlin – Heidelberg – New York 1974.

[22] A. VOGEL, *Ein gemeinsames Funktionsschema zur Beschreibung der Lautheit und der Rauhigkeit*, Biol. Cybernetics, **18**, pp. 31–37 (1975).

[23] E. ZWICKER, *Die Grenzen der Hrbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones*, Acustica, **2**, pp. 125–133 (1952).

[24] E. ZWICKER, *Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs*, Acustica, **6**, pp. 356–381 (1956).

[25] E. ZWICKER, *Direct comparisons between the sensation produced by frequency modulation and amplitude modulation*, J. Acoust. Soc. Am., **34**, pp. 1425–1430 (1962).

[26] E. ZWICKER and R. FELDTKELLER, *Das Ohr als Nachrichtenempfanger*, Hirzel Verlag, Stuttgart 1967.

[27] E. ZWICKER, *Masking and psychological excitation as consequences os the ear's frequency analysis* [In:] Frequency analysis and periodicity detection in hearing, Driebergen Symposium 1969, R. PLOMP and G. SMOORENBURG [Eds.], Sythof, Leiden 1970.