ARCHIVES OF ACOUSTICS
25, 4, 423–431 (2000)

THE TWO FACES OF AUDITORY MASKING

U. JORASZ

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

B.C.J. MOORE

Department of Experimental Psychology
University of Cambridge
(Cambridge, United Kingdom)

The experiments are concerned with the comodulation masking release (CMR) and modulation discrimination interference (MDI) that can occur when modulated carriers are added to a target modulated sound at frequencies remote from the target frequency. The results are discussed in terms of the factors influencing the both effects.

1. Introduction

Fletcher [2] suggested that the peripheral auditory system behaves as if it contained a bank of bandpass filters, with continuously overlapping passbands (now called "auditory filters"). Fletcher thought — and recent data are consistent with this point of view — that the basilar membrane inside the cochlea provided the basis for the auditory filters. When an observer is trying to detect a sinusoidal signal of a given frequency in a noise background, it has often been assumed that performance is based on the output of the single auditory filter that gives the highest signal-to-noise ratio. The centre frequency of this filter is usually the same as or close to the signal frequency. Threshold is assumed to correspond to a constant signal-to-noise ratio.

This model works very well in many situations [9] but it clearly fails in others. In some cases, the outputs of auditory filters tuned away from the signal frequency can be used to enhance signal detection, as in the phenomenon of comodulation masking release (CMR) [4]. CMR occurs when the task is to detect a signal centred in a narrow-band masker (the on-frequency band) that is amplitude modulated in some way. The addition of other components to the masker, remote from the signal frequency, can enhance signal detection, provided the extra components have a similar pattern of modulation to the on-frequency band. These extra masker components are sometimes called "flankers".

In the majority of experiments demonstrating CMR, the on-frequency band and flankers have been modulated at the same depth. It is usually assumed that the flankers improve detection of the signal because the listener can make use of the disparity in modulation pattern at the outputs of different auditory filters; when the signal is absent, the pattern is similar for all filters, but when the signal is present the modulation pattern differs for filters tuned close to the signal frequency and filters tuned close to the flanker centre frequencies.

In other cases, the outputs of auditory filters tuned away from the signal frequency degrade signal detection. This degradation seems to happen mainly when the task of the observer is to discriminate changes in modulation depth of the signal or to detect a change in the modulation pattern of the signal [12]. In this case, the signal carrier frequency is often called the "target". The ability to discriminate/detect these changes is adversely affected by the presence of other modulated sounds (also called flankers, but sometimes also called "interferers"), even when those sounds have centre frequencies well away from that of the target. We use the acronym MDI (modulation detection/discrimination interference) to describe this effect. The greatest MDI occurs when the target and "interfering" sounds are modulated at similar rates [7, 13]. In this respect, MDI resembles CMR, except that the remote components enhance detection in CMR and degrade it in MDI. MDI often occurs when the modulation depth of the target is small, but the modulation depth of the interferers is large.

In many ways, the conditions in which CMR and MDI occur are only a little different. Both CMR and MDI involve modulated stimuli. In the case of CMR, the signal is added to a modulated sound, and results in a change in the modulation pattern of that sound. The flankers make it easier to detect the change in modulation pattern. In the case of MDI, the signal is also a change in modulation of a target sound. But now, the flankers make it harder to hear the change.

A point of special interest of our experiments was to examine the factors that determine whether the flankers produce MDI or CMR. One difference between conditions giving CMR and those giving MDI concerns the relative modulation depths of the target and flankers; a second difference concerns the ability of subjects to detect a decrease or an increase in modulation depth. MOORE *et al.* [6] showed that stimuli that were otherwise very similar could give CMR when the task of subjects was to detect a decrease in modulation depth, and MDI when the task was to detect an increase. In the experiments reported here, we manipulated the relative modulation depth of the target and flankers and compared the ability to detect a decrease and an increase in modulation depth.

2. Conditions of the experiments

2.1. Stimuli

Thresholds were measured for detecting either an increase or a decrease in modulation depth of a sinusoidally amplitude modulated (SAM) target signal, with that signal presented either alone, or with flankers. The following expression describes a SAM sound:

$$A(t) = A_o[1 + m\cos(2\pi f_m t + \theta)]\cos(2\pi f_c t + \phi),$$

where A_o is the amplitude envelope of the unmodulated carrier, t is time, f_c is the carrier frequency, f_m is the modulator frequency, m is the modulation depth, and θ and ϕ are terms determining the starting phases of the modulator and carrier, respectively. In our experiments, the value of θ for the target sound had a random value from one stimulus to the next; thus the phase of the modulation was random relative to the onset of the stimulus.

The target was a 1000-Hz carrier SAM at a 10-Hz rate with a modulation index m of 0.5. The flankers, when present, consisted of two carriers which were not harmonically related to the target. Their modulation depths were 0.0, 0.25, 0.5, or 0.75. The two carriers were always modulated at the same rate as each other, and they were modulated at the same rate as the target, and in-phase with the target. In one set of conditions, they were centred at 230 and at 3300 Hz. In a second set they were centred at 550 and 1550 Hz. Both sets were chosen to be well outside the passband of the auditory filter centred at the target frequency.

All carriers were presented at a level of 60 dB SPL. To prevent changes in intensity providing a cue for discrimination of modulation depth, carriers which were SAM were "intensity compensated" to eliminate changes in overall intensity with changes in modulation depth [10]. The time pattern of the sounds is illustrated in Fig. 1. The relatively long stimulus duration was chosen for two reasons: so that several cycles of modulation would occur during each stimulus; and to reduce the possible role of perceptual grouping caused by synchronous gating of the target and flankers. In one stimulus the 1000-Hz carrier was modulated at the reference modulation depth (0.5). In the other, the modulation depth was greater, or, in a separate set of trials, less. The order of the two stimuli was random.

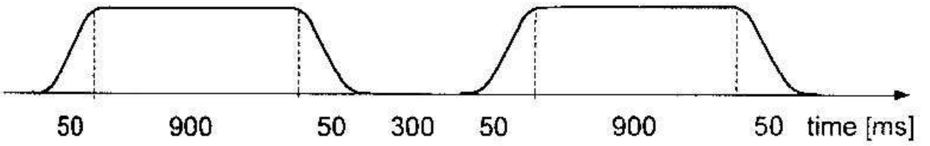


Fig. 1. Time pattern of the sounds. Each stimulus had a 900-ms steady-state portion and 50-ms raised-cosine rise/fall ramps. On each trial, two stimuli were presented, separated by a 300-ms silent interval.

All stimuli were generated using a Masscomp 5400 computer equipped with 16-bit digital-to-analogue converters. The sampling rate was 10 kHz, and stimuli were low-pass filtered at 4 kHz (−3 dB point) using a Fern Electronics EF16 filter with an attenuation rate of 100 dB/oct. The stimuli for each trial were calculated on line in the inter-trial interval.

Subjects were tested individually or in pairs in separate double-walled sound attenuating chambers. Stimuli were delivered via a manual attenuator to one earpiece of a Sennheiser HD414 headset.

2.2. Procedure

An adaptive 2AFC procedure was used to estimate the 79.4% correct point on the psychometric function. A run always started with a large change in modulation depth in the signal interval. After three successive correct responses the change in modulation depth was reduced while after each incorrect response it was increased. The modulation index was not allowed to be greater than 1.0 or less than 0. Initially, the step size for the change in modulation index was 5 dB in units of $20\log(m)$. After four reversals, the step size was decreased to 2 dB and eight further reversals were obtained. The mean value of $20\log(m)$ at last eight reversals was used to estimate the change in modulation index corresponding to threshold. At least four estimates were obtained for each condition, and the threshold was calculated as the geometric mean of the four. When the standard deviation of the log values exceeded 0.2, at least one further estimate was obtained and all estimates were averaged. The standard deviation of the log values was typically between 0.05 and 0.15. The standard error of the log values was never greater than 0.1 and was typically about 0.05. Correct-answer feedback was given after each trial by means of lights on the response box.

2.3. Subjects

Three subjects with a normal hearing at all audiometric frequencies were used. One was author UJ. The others were paid for their services. During initial training, performance in the reference condition (target sound alone) stabilised quite rapidly. Thresholds continued to decrease for some time in the conditions with flankers. All subjects found these conditions very difficult at first. They were given at least 15 hours of practice prior to collection of the data reported here.

3. Factors determining whether CMR or MDI occurs

3.1. The task of the subjects

The solid triangles in Fig. 2 (the results shown as open triangles will be discussed later (Subsec. 3.2)) show results for one subject (S3) from experiment 1, where all sounds were gated synchronously. For all subjects thresholds were consistent and very similar. Thresholds are expressed as the change in modulation index at threshold, divided by the reference modulation index (0.5), and are plotted as a function of the modulation depth of the flankers. The left-most points show thresholds in the reference condition (no flankers). These thresholds are consistent with those in the literature [8, 3, 11].

The unmodulated flankers raised thresholds somewhat. This effect may be partly caused by the synchronous gating of the target and flankers, a point we will return to later. However, modulated flankers produced more MDI, and the amount of MDI increased progressively as the modulation depth of the flankers was increased.

The pattern of results was similar for the two types of flankers (550 and 1550 or 230 and 3300 Hz), but the amount of MDI was slightly greater for the former. This could indicate that MDI is affected by the proximity in frequency of the target and flankers.

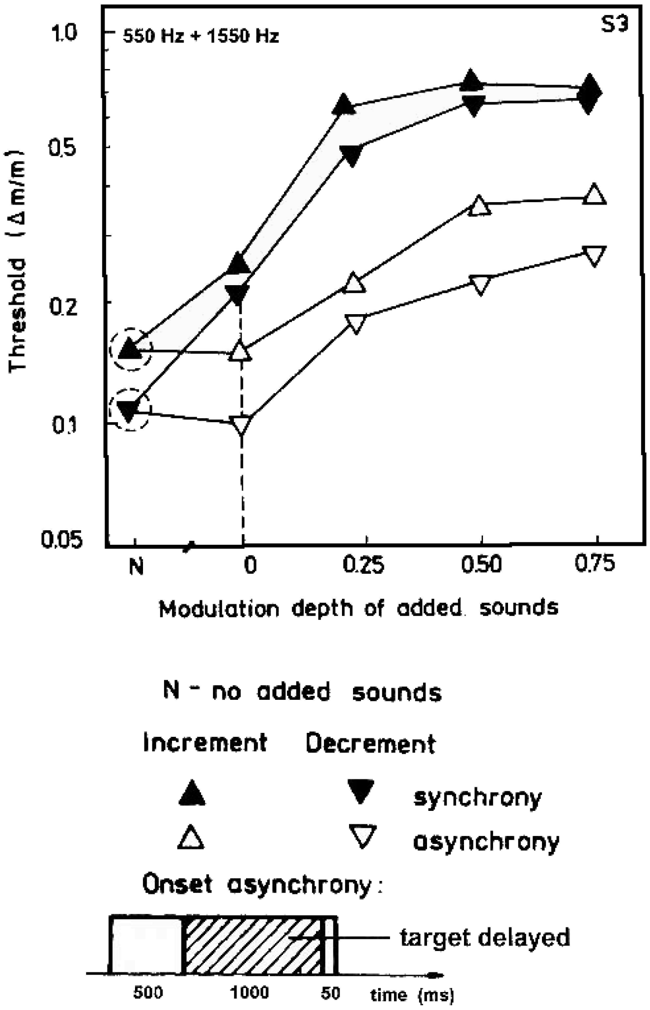


Fig. 2. The solid triangles show results from the first part of the experiment, where all sounds were gated synchronously. Thresholds are shown for detecting an increment in modulation depth (up-pointing triangles) or decrement in modulation depth (down-pointing triangles). Thresholds are expressed as a proportion of the modulation depth of the standard sound ($m = 0.5$) and are plotted as a function of the modulation index of the flankers. The open triangles show results when the onset of the target was delayed relative to the onset of the flankers.

However, the greater effect for the 550- and 1550-Hz sounds might reflect a small degree of interaction of the target and flankers in the peripheral auditory system.

Thresholds for detecting decrements in modulation depth were lower than those for detecting increments, but the difference was small. It may be asked why these data showed no sign of CMR, even in though there was a potential detection cue in at least one condition, where the flankers were modulated with the same depth, 0.5, as the reference sound)? We will try to provide the answer when considering our later experiments.

3.2. Onset asynchrony

Onset asynchrony is known to be a powerful factor in perceptual grouping. When a new sound is introduced after another sound has been on for some time, the new sound seems to "pop out" and to be perceived as a separate sound. It has been suggested that perceptual grouping processes play a role in both CMR and MDI [12]. One principle that operates in perceptual grouping is "common fate" [1]. Elements of a sound that change in the same way tend to be grouped and heard as a single stream. Two applications of this principle are relevant here: Elements that start and stop together tend to be perceived as a single sound; and elements that are modulated in similar way tend to be perceived as a single sound.

If MDI depends on perceptual grouping of the target and flankers, delaying the target sound relative to the flankers should result in a marked decrease in MDI. We also thought that the introduction of an onset asynchrony would increase the chances of obtaining CMR in the conditions of the experiment where the target and flankers had the same modulation depth as the target sound, and where the task was to detect a decrease in modulation depth.

The stimuli were essentially the same as those of the first part of experiment, except for the difference in time pattern (see the bottom of Fig. 2). The flankers were gated on 500 ms before and gated off 50 ms after the 1000-ms target sound. The subjects and procedure were the same as before. The standard error of the log values of the threshold estimates was never greater than 0.1 and was typically about 0.05.

The results are shown as the open triangles in Fig. 2. Now the unmodulated flankers did not produce MDI. This suggests that the MDI found with these sounds previously did not result from the interactions of the sounds in the peripheral auditory system. Rather, it may have been a consequence of perceptual grouping of the target and flankers caused by their synchronous gating. Onset asynchrony reduced the amount of MDI for all subjects. This fact suggests that perceptual grouping does play a role in MDI. However, even though the target sound was easily heard out, there was still no evidence for CMR.

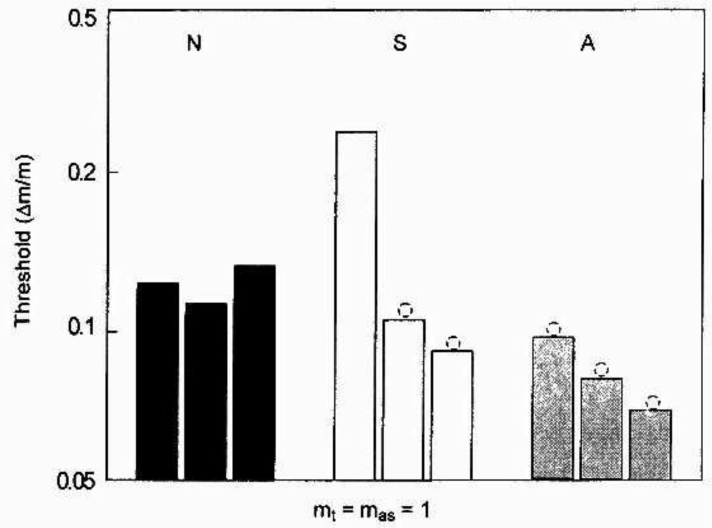
3.3. The reference conditions

We consider now another difference between situations giving rise to CMR and those giving rise to MDI. This difference is related to the difficulty of signal detection in the condition with the on-frequency sound alone (the reference condition). Usually, large CMRs occur when there is a high degree of stimulus uncertainty in the reference condition, or when thresholds are high for some other reason. In contrast, in experiments demonstrating MDI, thresholds in the reference condition tend to be quite low. Thus, our failure to obtain CMR could have been a consequence of our use of too small a modulation depth in the reference condition, which gave rise to low thresholds. Hence, we decided to examine the effect of using greater modulation depths.

The following modulation depths of the target/flanking sounds were used: 0.75 and 0.75; 1.0 and 1.0; 0.75 and 1.0. Since the reference modulation indexes were high, it was not generally possible to measure thresholds for detecting increments in modulation depth. Hence, thresholds were measured only for the detection of decrements in modulation depth. Thresholds were measured under three conditions: the reference condition (N – no flankers); with the target and flankers gated together (S – synchronous condition); and with the onset of the target delayed by 500 ms relative to the onset of the flankers (A). In all cases, the signal duration was 1000 ms.

To increase the chances of obtaining CMR, the two sets of flankers used earlier were combined, giving a set of flankers consisting of carriers at 230, 550, 1550, and 3300 Hz. The magnitude of CMR has been shown to increase with increasing numbers of flankers (for review, see [5]).

The results for the greatest modulation depth ($m = 1$) are shown in Fig. 3. The rest of results is very similar. In each triad of bars, the left-most bar shows results for subject 1, the middle bar for subject 2, and the right bar for subject 3. For all three subjects, the thresholds tended to be lower in all conditions when the flankers were modulated with $m = 1$ than when they were modulated with $m = 0.75$; increasing the modulation depth of the flankers produced a masking release. This contrasts with the results for first experiment, where thresholds tended to increase with increasing modulation depth of the flankers (Fig. 2), i.e. increasing the modulation depth of the flankers produced more MDI.



N — no added sounds, S — added sounds synchronous,
A — target delayed, ◌ — CMR,
m_t — modulation depth of target, m_{as} — modulation depth of added sounds.

Fig. 3. Results of the experiment when the modulation depth was 1.0 for both the target and the flankers. In each group of three bars, the left bar shows results for subject 1, the middle bar for subject 2, and the right bar for subject 3.

In the introduction, it was pointed out that, in most experiments demonstrating CMR, the on-frequency and flanking bands have been modulated at the same depth. Our results suggest, however, that this is not critical; rather, the masking release can be greater when the flanking bands have a greater modulation depth than the on-frequency sound.

It seems likely that our results reflect a balance between mechanisms producing MDI and those producing CMR. The higher modulation index of the target sound — the harder reference condition — created conditions more conducive to obtaining CMR. Now, the tendency for a greater modulation depth of the flankers to produce greater MDI was more than offset by the tendency for CMR to increase with increasing modulation index of the flankers. The balance between mechanisms producing MDI and those CMR varied across subjects. It seems that, under the conditions of this experiment, the most important factor affecting the balance between mechanisms producing MDI and those producing CMR was the onset asynchrony of the target sound.

4. Conclusions

Adding modulated carriers (flankers) at a frequency remote from the target frequency, sometimes impaired discrimination of modulation depth of the target (MDI) and sometimes improved it (CMR). When the modulation depth of the target sound was 0.5, the flankers always gave MDI rather than CMR. The amount of MDI generally increased progressively with increasing modulation depth of the flankers. The amount of MDI was markedly reduced, but not usually eliminated, by gating the target sound on after the flankers. This supports the suggestion that perceptual grouping mechanisms play a role in MDI; the asynchronous onsets of the target and flankers made the target appear to "pop out", making it easier to discriminate changes in the target sound.

CMR occurs when both the target and flankers have a high modulation depth, and the signal is a decrease in modulation depth. It seems likely that a large amount of CMR only occurs when performance in the reference condition (target sound alone) is poor.

Acknowledgements

Preparation of this paper was supported by KBN. Grant 1HO1FO8012.

References

[1] A.S. Bregman, *Auditory scene analysis: The perceptual organization of sound*, MIT, Cambridge, MA 1990.

[2] H. Fletcher, *Auditory patterns*, Rev. Mod. Phys., **12**, 47–65 (1940).

[3] D.W. Grantham and S.P. Bacon, *Detection of increments and decrements in modulation depth of SAM noise*, J. Acoust. Soc. Am. Suppl. 1, **84**, S140 (1988).

[4] J.W. Hall, M.P. Haggard and M.A. Fernandes, *Detection in noise by spectrotemporal pattern analysis*, J. Acoust. Soc. Am., **76**, 50–56 (1984).

[5] B.C.J. Moore, *Comodulation masking release: Spectro-temporal pattern analysis in hearing*, Brit. J. Audiol., **24**, 131–137 (1990).

[6] B.C.J. Moore, B.R. Glasberg and G.P. Schooneveldt, *Across-channel masking and comodulation masking release*, J. Acoust. Soc. Am., **87**, 1683–1694 (1990).

[7] B.C.J. Moore, B.R. Glasberg, T. Gaunt and T. Child, *Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals*, Q.J. Exp. Psychol., **43A**, 327–347 (1991).

[8] E. Ozimek and A. Sęk, *AM difference limens for noise bands*, Acustica, **66**, 153–160 (1988).

[9] 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.

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

[11] G.H. Wakefield and N.F. Viemeister, *Discrimination of modulation depth of SAM noise*, J. Acoust. Soc. Am., **88**, 1367–1373.

[12] W.A. Yost and S. Sheft, *Across-critical-band processing of amplitude-modulated tones*, J. Acoust. Soc. Am., **85**, 848–857 (1989).

[13] W.A. Yost, S. Sheft and J. Opie, *Modulation interference in detection and discrimination of amplitude modulation*, J. Acoust. Soc. Am., **86**, 2138–2147 (1989).