PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Modulation Discrimination Interference and Auditory Grouping

Brian C. J. Moore and Michael J. Shailer

Phil. Trans. R. Soc. Lond. B 1992 **336**, 339-346
doi: 10.1098/rstb.1992.0067

Modulation discrimination interference and auditory grouping

BRIAN C. J. MOORE AND MICHAEL J. SHAILER

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K.

SUMMARY

The detection of a change in the modulation pattern of a (target) carrier frequency, f_c (for example a change in the depth of amplitude or frequency modulation, AM or FM) can be adversely affected by the presence of other modulated sounds (maskers) at frequencies remote from f_c, an effect called modulation discrimination interference (MDI). MDI cannot be explained in terms of interaction of the sounds in the peripheral auditory system. It may result partly from a tendency for sounds which are modulated in a similar way to be perceptually 'grouped', i.e. heard as a single sound. To test this idea, MDI for the detection of a change in AM depth was measured as a function of stimulus variables known to affect perceptual grouping, namely overall duration and onset and offset asynchrony between the masking and target sounds. In parallel experiments, subjects were presented with a series of pairs of sounds, the target alone and the target with maskers, and were asked to rate how clearly the modulation of the target could be heard in the complex mixture. The results suggest that two factors contribute to MDI. One factor is difficulty in hearing a pitch corresponding to the target frequency. This factor appears to be strongly affected by perceptual grouping. Its effects can be reduced or abolished by asynchronous gating of the target and masker. The second factor is a specific difficulty in hearing the modulation of the target, or in distinguishing that modulation from the modulation of other sounds that are present. This factor has effects even under conditions promoting perceptual segregation of the target and masker.

1. INTRODUCTION

It is widely assumed that the peripheral auditory system contains a bank overlapping bandpass filters (the auditory filters) (Fletcher 1940). When an observer is trying to detect a narrowband signal in the presence of a masking sound, it is usually assumed that performance is based on the output of the single auditory filter which gives the highest signal-to-masker ratio (Fletcher 1940). Although this model works well in many situations (Patterson & Moore 1986) it clearly fails in others. In the situations considered in this paper, the outputs of auditory filters tuned away from the signal frequency degrade signal detection. It appears that sometimes subjects cannot attend to a single auditory filter even though it would pay them to do so. This degradation seems to happen mainly when the masker and signal are modulated in some way and the task of the subject is to detect a change in the modulation of the signal (Yost & Sheft 1989; Yost *et al.* 1989; Moore *et al.* 1990, 1991; Wilson *et al.* 1990). Hence, this phenomenon has been called modulation detection interference (Yost & Sheft 1989) or modulation discrimination interference (Moore & Jorasz 1992) (MDI).

As an example of MDI, we will briefly describe some experiments reported by Moore *et al.* (1991), as the experiments described in this paper use very similar stimuli. Moore *et al.* determined how thresholds for detecting an increase in modulation depth (sinusoidal amplitude modulation, AM, or frequency modulation, FM) of a 1000 Hz carrier frequency (the target) were affected by the presence of modulated carriers (maskers) with frequencies of 230 Hz and 3300 Hz. The carrier frequencies of the maskers were sufficiently far from the target frequency that they would have produced a negligible output from the auditory filter centred at the target frequency. The target was presented twice for a duration of 1000 ms, with 300 ms inter-stimulus interval, and the subject was required to indicate the interval in which the sound was more modulated. The maskers were gated with the target. They found that modulation increment thresholds were increased (worsened) in the presence of the maskers. This MDI effect was greatest when the target and maskers were modulated at similar rates, but the effect was broadly tuned for modulation rate.

2. THE POSSIBLE ROLE OF PERCEPTUAL GROUPING IN MDI

In everyday life we often listen to several sound sources simultaneously. The auditory system has to decide which 'elements' of the complex mixture arise from one source, and which from another. The process of doing this is described as 'perceptual grouping' or 'stream formation'; the elements of the sound are grouped across-frequency and across time to form percepts of coherent streams each with its own loudness, pitch, timbre and location (Moore 1989; Bregman 1990). One principle that operates in perceptual grouping is 'common fate'; 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 a similar way tend to be perceived as a single sound.

Yost & Sheft (1989) suggested that MDI might be a consequence of perceptual grouping. In the case of AM sounds, the common AM of the target and maskers might make them fuse perceptually (McAdams 1984; Bregman *et al.* 1990), making it difficult to 'hear out' the modulation of the target sound. This explanation is consistent with the subjective impression of subjects in these experiments. However, certain aspects of the results on MDI are difficult to reconcile with an explanation in terms of perceptual grouping. One would expect that widely spaced frequency components would only be grouped perceptually if their modulation pattern was very similar. Grouping would not be expected, for example, if the components were modulated at different rates. Yet, it is possible to obtain large amounts of MDI under conditions where the target and the maskers are modulated at different rates (Yost *et al.* 1989; Wilson *et al.* 1990; Moore *et al.* 1991).

The present experiments were intended to clarify the role of perceptual grouping in MDI. The first experiment was similar to that of Moore *et al.* (1991). Thresholds for detecting a change in AM depth of a target sound were measured with and without maskers whose carrier frequencies were remote from the target frequency. However, in this experiment we manipulated a second factor that is known to affect perceptual grouping, namely the overall duration of the stimuli. Sounds that are gated on and off synchronously for a brief duration tend to fuse perceptually and be heard as a single sound, whereas at longer durations more than one sound may be perceived (Moore *et al.* 1986). If perceptual grouping is responsible for MDI, then at short durations the MDI produced by the synchronous gating of the target and maskers would be expected to combine with the MDI produced by modulation of the maskers to produce a greater overall effect. The second experiment used the same subjects and similar stimuli, but involved a rating task intended to provide a direct indication of the extent to which grouping of the target with the masker made it difficult to hear the modulation of the target sound. Because we wish to compare the results of the two experiments, the method for each will be described first.

3. METHOD

(*a*) *Experiment 1*

Thresholds were measured for detecting an increase in modulation depth of an AM target. The target was presented either alone, or with various additional sounds.

(*i*) *Stimuli*

The target was a 1000 Hz carrier amplitude modulated at a 10 Hz rate. The phase of the modulation was random relative to the onset of the stimulus. The masker, when present, consisted of two carriers, chosen to be non-harmonically related to the target. The two carriers were always modulated at the same depth as each other, and at the same rate. Their modulation index, m, was either 0 (no modulation) or 0.5. They were centred at 230 Hz and at 3300 Hz, the same as used by Moore *et al.* (1991). When m was 0.5, the modulation rate of the masker was either 4, 10 or 25 Hz. When the rate was 10 Hz, the masker was modulated either in phase with the target or in antiphase. Two conditions were also included where the masker consisted of a single carrier frequency, either 230 Hz or 3300 Hz, modulated at 10 Hz in phase with the target.

All carriers were presented at a level of 60 dB SPL. The root-mean-square pressure was held constant regardless of modulation depth. Thus, intensity changes in the target could not be used as a cue for detecting changes in modulation depth.

On each trial, two stimuli were presented, separated by a silent interval of 300 ms. Each stimulus had 50 ms raised-cosine rise and fall ramps. The steady state duration was either 100, 300 or 1500 ms, giving overall durations of 200, 400 or 1600 ms. In one stimulus the 1000 Hz carrier was modulated with index $m = 0.25$. In the other, m was greater. The order of the two stimuli was random.

All stimuli were generated using a Masscomp 5400 computer equipped with 16-bit digital-to-analog converters. The sampling rate was 10 kHz, and stimuli were lowpass filtered at 4 kHz (−3 dB) using a Fern Electronics EF16 filter with an attenuation rate of 100 dB per octave. Subjects were tested individually in a double-walled sound attenuating chamber. Stimuli were delivered via a manual attenuator (Hatfield 2125) to one earpiece of a Sennheiser HD414 headset.

(*ii*) *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 m in the signal interval. After three successive correct responses the change in m was reduced whereas after each incorrect response it was increased. The value of m was not allowed to be greater than 1.0. Initially, the step size for the change in m 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 the last eight reversals was used to estimate the change in m ($\Delta\ m$) 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.

During initial training, performance with the target sound alone stabilized quite rapidly, but thresholds continued to decrease for some time in the conditions with maskers; all subjects found these conditions very difficult at first, and they sometimes had difficulty scoring above chance. All subjects were given at least 15 h practice before collection of the data reported here. However, even after this time there was some evidence of systematic improvements in certain conditions, especially those involving two modulated maskers. Subjects were given several additional hours of practice in these conditions, until their performance appeared to be stable.

(*iii*) *Subjects*

Four subjects with normal hearing at all audiometric frequencies were used. One was author M.S. The others were paid for their services.

(*b*) *Experiment 2*

Subjects were presented with the target (a sinusoid amplitude modulated at a 10 Hz rate with $m = 0.25$, and with a carrier frequency close to 1000 Hz) followed by the target together with some maskers. They were asked to listen to the modulation of the target alone, and to rate (on a scale from 1 to 7) how clearly they could hear the modulation of the target in the complex mixture. To prevent subjects remembering the pitch of the target across trials, its carrier frequency was randomly varied ($\pm 10\%$, rectangular distribution) from trial to trial. The maskers were the same as used in experiment 1. The first and second sounds always had the same overall duration of either 200, 400 or 1600 ms, and the interval between the two stimuli in a trial was 300 ms. To control for possible range effects, all stimuli (types of masker and durations) were presented in quasi-random order within a single large block of trials. The only constraint was that each stimulus should be presented five times during the first half of a block and five times during the second half. Each block took about 15–20 min. Three blocks were run for each subject. The first two ratings of each stimulus in a block tended to be rather variable, and also to be inconsistent with ratings obtained later on in that block, so they were discarded; data presented are based on 24 ratings per stimulus.

4. RESULTS

Some individual differences were apparent in the results, but the overall pattern was similar across subjects. For simplicity, we present only the mean results. The upper panels of figure 1 show the results for experiment 1, while the lower panels show the corresponding results for experiment 2. Open hexagons show results without any masker. Error bars indicate $\pm$ one standard error of the mean across subjects; they thus give an idea of the degree of individual variability. The likely error of the data points is somewhat less than indicated by the error bars. The rating data are plotted with values decreasing up the ordinate to make it easier to see similarities and differences between the results of the two experiments; one would expect large amounts of MDI to be associated with low clarity ratings.

Consider first the results shown in the left-hand panels. MDI is usually described as a difficulty in hearing modulation at one frequency when modulation is also present at another frequency. Consistent with this, most previous researchers have not found interference effects for unmodulated maskers at relatively long durations (Yost & Sheft 1989; Moore *et al.* 1991). However, our results show some interference for the unmodulated maskers. The effect varied somewhat with duration; at 1600 ms, thresholds increased (relative to those with no masker) by a factor of 1.8, whereas at 200 ms they increased by a factor of 2.3. This is broadly consistent with an explanation for the interference effect in terms of perceptual grouping. These findings may indicate a need to broaden the definition of MDI to include interference in modulation discrimination caused by unmodulated maskers.

Modulated maskers produced more MDI. At the longest duration, maskers modulated in phase with the target produced more MDI than maskers modulated in antiphase. However, at the two shorter durations the phase effect was very small. Using similar stimuli, Moore *et al.* (1991) found no significant effect of modulator phase for a duration of 1000 ms, although one of their subjects did show some evidence of a phase effect.

As expected, clarity ratings (bottom left panel) were highest when no maskers were present. The ratings decreased when maskers were added. The effect of duration was small. Ratings were not generally lower when the maskers were modulated than when they were unmodulated. Thus, there is a discrepancy between the results of the two experiments; modulated maskers produced the greatest MDI, but they did not have the greatest effect on subjective clarity of the modulation.

The middle panels illustrate the effect of the modulation rate of the maskers, which always had carriers of 230 and 3300 Hz. Data for the 10 Hz rate are averaged over the in phase and antiphase conditions. Moore *et al.* (1990) found that MDI measured using similar stimuli was tuned for modulation rate, peaking at 10 Hz, but the tuning was very broad. The data in figure 1 do not show any clear tuning; MDI did not vary greatly with modulation rate of the maskers. This may reflect the fact that the range of modulation rates (4–25 Hz) was smaller than used by Moore *et al.* (1991) (2–50 Hz). Clarity ratings were somewhat lower when the maskers were modulated at 4 Hz or 25 Hz than when they were modulated at 10 Hz. Thus, once again, there is a discrepancy between the results of the two experiments.

The right-hand panels show the effect of the number of masker carriers. There is a clear trend for two carriers to produce more MDI than one. In this case, there is a reasonable correspondence between the

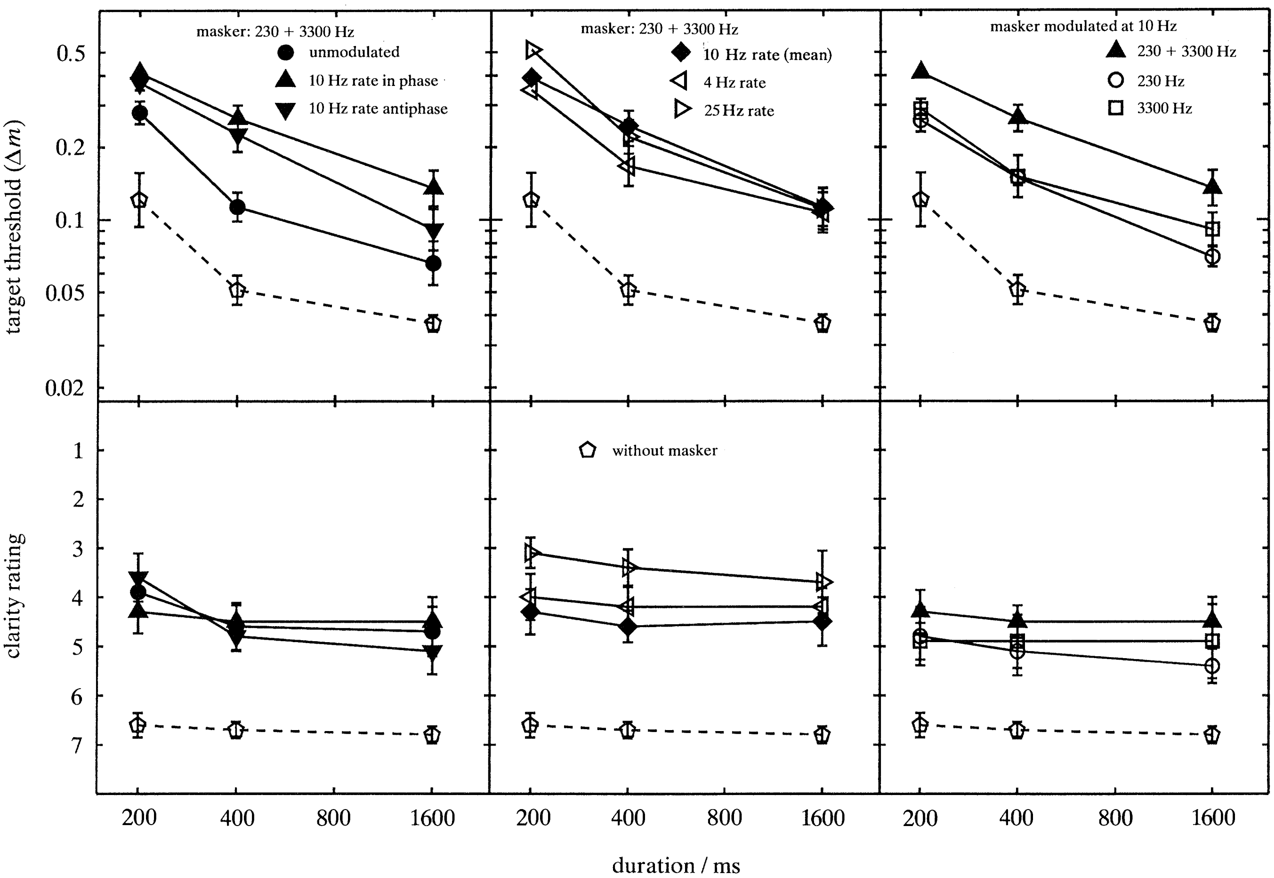

Figure 1. Results of experiment 1 (upper panels) and experiment 2 (lower panels). Results are means across four subjects. Error bars indicate ± one standard error of these means. Error bars are not shown where they would be smaller than the symbol used. The masker type is indicated in the figure.

clarity ratings and the thresholds. Clarity ratings were lower for two interfering carriers than for one.

In all conditions of experiment 1, thresholds declined with increasing duration. However, the question of whether MDI declines with increasing duration is somewhat unclear; it depends on how 'amount' of MDI is defined. If the thresholds are expressed relative to the thresholds obtained without any maskers, there is no clear evidence of a change in MDI with duration. If perceptual grouping were entirely responsible for the MDI produced by modulated maskers, then at short durations the MDI produced by the synchronous gating of the target and maskers would have been expected to combine with the MDI produced by modulation of the maskers to produce a greater overall effect. This did not happen.

In summary, the results of experiment 1 suggest that some factor other than perceptual grouping plays a role in MDI. This conclusion is supported by the finding that maskers giving the largest amount of MDI did not always produce the lowest clarity ratings.

5. EXPERIMENTS 3 AND 4

In experiments 3 and 4 we investigated the effect of another factor known to influence perceptual grouping, namely temporal asynchrony between the target sound and the masker. Asynchrony has been shown to have a powerful tendency to produce perceptual segregation (Rasch 1978; Darwin 1981; Darwin & Sutherland 1984; Roberts & Moore 1991). Thus, we expected that the asynchrony might counteract the effects of perceptual grouping produced by the common modulation of the target and maskers, reducing MDI. Experiment 3 was similar to experiment 1; thresholds were measured for detecting a change in modulation depth of the target. Experiment 4 was a rating experiment, like experiment 2, but using stimuli similar to those used in experiment 3.

(*a*) *Conditions*

The target always had an overall duration of 400 ms. In one set of conditions, the masker was gated on before the target, with an onset asynchrony of 0, 25, 50, 100 or 200 ms. The masker and target ended simultaneously. In a second set of conditions, the target and masker started simultaneously, but the masker continued after the target for 0, 25, 50, 100 or 200 ms. The masker was either absent, unmodulated, or modulated at a 10 Hz rate either in phase with the target or in antiphase. The masker consisted of two carriers, one at 230 Hz and the other at 3300 Hz. Other aspects of the simuli were the same as for experiments 1 and 2.

(*b*) *Procedure and subjects*

The procedure was essentially the same as for experiments 1 and 2. In experiment 4, two separate blocks of trials were conducted, one using conditions where the masker was gated on before the target, and the other using conditions where the masker was gated off after the target. Data presented are based on a minimum of 20 ratings per stimulus per subject. Three of the subjects from experiments 1 and 2 were used, including author M.S.

6. RESULTS

The results were broadly similar across subjects, so only mean results will be presented. The upper panels of figure 2 show data for experiment 3, and the lower panels show data for experiment 4. Both when the masker was gated on before the target (top left panel) and when it was gated off after the end of the target (top right panel), MDI decreased with increasing asynchrony, but the effect was somewhat greater for onset than for offset asynchronies. For the unmodulated masker, the amount of MDI was almost absent at the largest asynchrony used. This is consistent with the idea that the MDI produced by the unmodulated masker in experiment 1 occurred because the synchronous gating of the target and masker caused them to be grouped perceptually. The asynchrony in experiment 3 apparently was sufficient to create perceptual segregation of the target and masker, thus eliminating the MDI.

The amount of MDI also declined with increasing asynchrony for the modulated maskers. However, even at the longest asynchrony, MDI still occurred. The results of Hall & Grose (1991) and of Moore & Jorasz (1992) are consistent with this. Moore & Jorasz found that an onset asynchrony of 500 ms between the masker and target was not sufficient to abolish MDI when the masking sounds were modulated at the same rate as the target. Previous work examining the effects of asynchrony on perceptual grouping (Rasch 1978; Darwin 1981; Darwin & Sutherland 1984; Bregman 1990; Roberts & Moore 1991) indicates that an asynchrony of 200 ms is usually sufficient to produce perceptual segregation. Thus, our finding that MDI persisted at an asynchrony of 200 ms suggests that factors other than auditory grouping play a role in MDI.

The results of experiment 4 (bottom panels) show a very different pattern. Surprisingly, there was no clear change in clarity ratings with either onset or offset asynchrony. If anything, there was a slight trend for clarity ratings to decrease with increasing onset asynchrony. There is another discrepancy between the threshold data and the rating data: the unmodulated masker produced less MDI but gave lower clarity ratings than the modulated maskers.

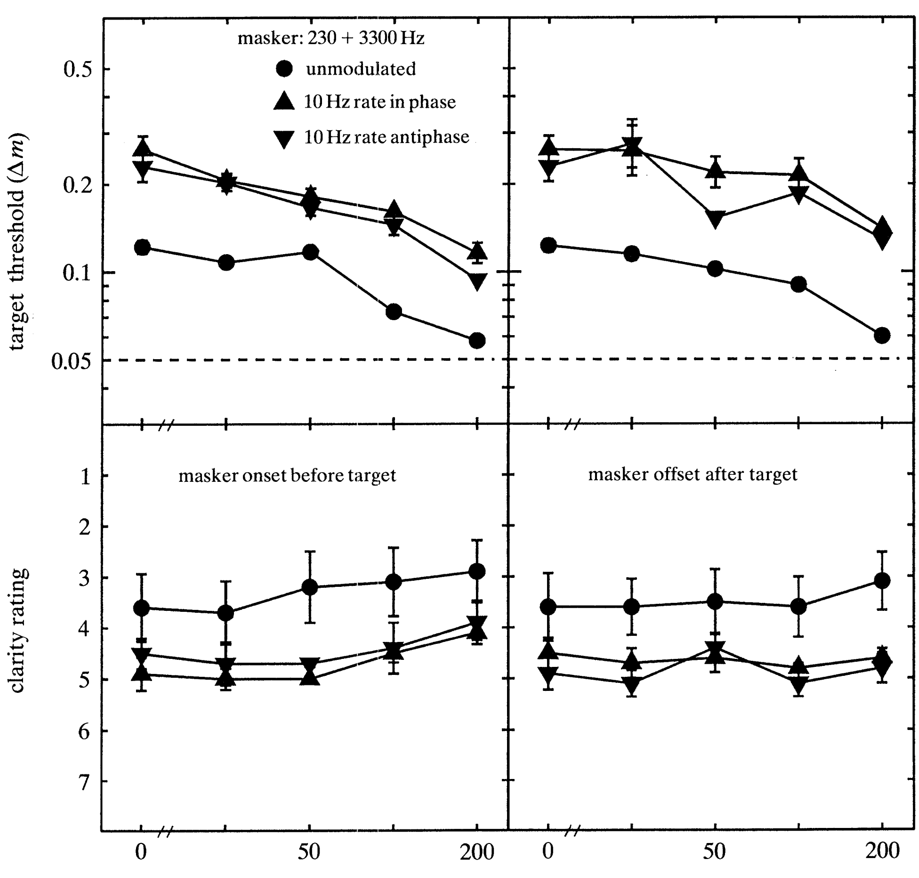

Figure 2. Results of experiment 3 (upper panels) and experiment 4 (lower panels). The dashed line indicates the threshold obtained without any masker. Otherwise as figure 1.

7. GENERAL DISCUSSION

Some aspects of the results are consistent with the idea that perceptual grouping plays a role in MDI. In experiment 1, the synchronous gating of the target and masker was sufficient to produce some MDI even when the masker was unmodulated. Experiment 3 showed that this MDI was largely abolished by gating the masker on 200 ms before the target or off 200 ms after the target, conditions designed to produce perceptual segregation of the masker and target.

On the other hand, several aspects of the results suggest that perceptual grouping is not the only factor involved. Firstly, experiment 1 showed that the MDI produced by modulated maskers was very broadly tuned for modulation rate. As pointed out by Moore *et al.* (1991), one would not expect perceptual grouping of components whose modulation frequencies were very different. Secondly, there were several discrepancies between the results of the masking experiments and of the rating experiments. The conditions giving the largest MDI did not always give the lowest clarity ratings. Finally, the MDI produced by modulated maskers was not abolished by an onset or offset asynchrony of 200 ms, even though that asynchrony was probably large enough to produce clear perceptual separation of the target and masker.

In interpreting the role of perceptual grouping in MDI, a distinction should be made between the ability to hear a pitch corresponding to the target frequency, and the ability to hear the modulation of the target frequency. It seems clear that the pitch of a given component can be harder to hear when that component is grouped perceptually with other sounds (Bregman & Pinker 1978; Moore *et al.* 1986; Bregman 1990; Roberts & Bregman 1991). Thus MDI might be a consequence of the failure to hear out the target frequency. If the target frequency could not be heard, then it would be difficult to judge the modulation of the target separately from the modulation of the other sounds present.

When we started the rating experiments, we had the above reasoning in mind. Hence, initially we asked the subjects to rate how clearly they could hear the target frequency in the complex mixture of sounds. To our surprise, the ratings were uniformly very high (all sixes and sevens), regardless of the masker type. It appeared that subjects had no difficulty in hearing out the target frequency. Hence, the instructions were modified; the rating data presented were obtained by asking subjects to rate how clearly they heard the modulation of the target in the complex mixture.

Given that the target frequency could be clearly heard in the rating experiment, even when the target and masker were gated synchronously, the results of experiment 4 become less surprising. The onset and offset asynchrony presumably had little effect on the clarity ratings because the target frequency was already clearly audible. This, however, creates a further puzzle; why did MDI change with asynchrony when the clarity ratings did not?

A possible explanation is that the stimulus sequence used in the rating experiments – target tone followed by target plus maskers – served to focus attention on the frequency region of the target, making it easier to hear out the target frequency. To check on this, experiment 2 was partially re-run, but with the sequence of stimuli reversed: target plus maskers followed by target alone. This had very little effect on the pattern of the results. Also, subjects reported that the target frequency was still very easy to hear in all conditions. It appears that simply hearing the target on its own, either before or after the complex mixture, makes it easier to hear out the target from the complex. Of course, in the masking experiments the target was never heard alone immediately before or after presentation with a masker. Thus, subjects may have had trouble hearing out the target frequency in the masking experiments, but not in the rating experiments.

To check on this possibility, some of the conditions of experiment 1 were re-run, but with a 'cue' tone presented before each forced-choice trial. The cue tone had the same duration, level and frequency as the target, but it was unmodulated. The interval between the cue tone and the first test stimulus was 300 ms. Subjects were told that the cue tone had the same frequency as the target tone whose modulation depth they were required to discriminate. Subjects reported that the cue tone did make it easier to hear out the target frequency from the complex mixture. Furthermore, thresholds with the cue tone were lower than those obtained without the cue tone. This suggests that part of the MDI observed in experiments 1 and 3 may have been caused by difficulty in hearing out the frequency of the target tone. However, the decrease in MDI produced by the cue tone was relatively small. For example, for 230 and 3300 Hz maskers modulated in phase with the target at a 10 Hz rate, the mean value of delta *m* for a duration of 400 ms was 0.26 without the cue tone and 0.23 with the cue tone. The value of delta *m* without any masker was 0.051, so it is clear that substantial MDI occurred even with the cue tone. The only exception was with the unmodulated masker, where thresholds with the cue tone were only slightly higher than those obtained without any masker.

It appears, then, that MDI arises in two ways: (i) from difficulty in hearing a pitch corresponding to the frequency of the target; (ii) from difficulty in hearing the modulation of the target, or distinguishing that modulation from the modulation of other components present in the sound. We will denote the former 'carrier-specific' MDI and the latter 'modulation-specific' MDI. Carrier-specific MDI can occur both for unmodulated and for modulated maskers, but it is reduced or abolished by conditions designed to promote perceptual segregation of the target and maskers, such as those using onset and offset asynchronies or using a cue tone. Modulation-specific MDI occurs mainly for modulated maskers, and it persists under conditions designed to promote perceptual segregation of the target and maskers. It remains somewhat unclear whether carrier-specific MDI results partly from perceptual grouping caused by common or similar modulation; our results suggest that common gating is a more powerful factor.

Some indication of the possible nature of modulation-specific MDI is provided by an experiment of Hall & Grose (1991). Their subjects were presented with 1 and 2 kHz tones, gated synchronously. On each trial, one of the tones, selected at random, was amplitude modulated and the other was unmodulated. The task of the subjects was to identify which tone was modulated. Remarkably, most subjects performed rather poorly at this task. Even when the modulation was clearly audible, it was hard to say whether the higher or lower tone was modulated. It appears that subjects sometimes have difficulty assigning modulation to its appropriate carrier frequency.

In our experiments 2 and 4, subjects may have been able to listen for the overall clarity of 10 Hz modulation, but they may have had difficulty telling which carrier had that modulation. When all carriers were modulated at 10 Hz, the 10 Hz modulation would obviously have been easy to hear. When the maskers were modulated at 4 Hz or 25 Hz, the overall impression of hearing 10 Hz modulation definitely decreased, and clarity ratings were lower. When the maskers were unmodulated, this also reduced the overall impression of hearing 10 Hz modulation, again giving lower clarity ratings (in experiment 4).

The mechanisms underlying modulation-specific MDI remain unclear. One possibility is that it reflects the operation of 'channels' specialised for detecting and analysing modulation. There is both physiological evidence (Kay 1982; Rees & Moller 1983; Schreiner & Urbas 1986) and psychophysical evidence (Kay & Mathews 1972; Tansley & Suffield 1983; Bacon & Grantham 1989; Houtgast 1989) for such channels. Yost *et al.* (1989) suggested that MDI might arise in the following way. The stimulus is first processed by an array of auditory filters. The envelope at the output of each filter is extracted. When modulation is present, channels are excited that are tuned for modulation rate. All filters responding with the same modulation rate excite the same channel, regardless of the filter centre frequency. Thus, modulation at one centre frequency can adversely affect the detection and discrimination of modulation at other centre frequencies. Hall & Grose (1991) pointed out that such channels could provide an explanation for their finding that subjects have difficulty assigning modulation to its appropriate carrier frequency. They could also account for the general form of the rating data obtained with modulated maskers in experiment 2.

The function of the modulation channels remains unclear. Yost *et al.* (1989) suggested that they could provide a means for perceptual grouping of components with common modulation patterns. However, our data suggest that modulation-specific MDI persists under conditions where the target is perceptually segregated from the masker. In addition, the tuning of MDI in the modulation domain appears to be too broad for it to be useful for perceptual grouping.

It should be emphasized that such modulation channels can only explain part of the MDI observed in our experiments, namely modulation-specific MDI. For example, they do not account for the MDI produced by unmodulated maskers; such maskers should not excite the modulation channels. Also it is difficult to explain the effects of onset and offset asynchrony in terms of such channels. To do this it would be necessary to assume that the channels are not 'hard wired' but operate after perceptual grouping processes.

In summary, it appears that MDI may have more than one cause. One component of MDI, carrier-specific MDI, is caused by difficulty in hearing a pitch corresponding to the target frequency. Carrier-specific MDI appears to be strongly affected by perceptual grouping. Its effects can be reduced or abolished by asynchronous gating of the target and masker or by presenting a cue tone at the target frequency. The second component, modulation-specific MDI, is caused by a specific difficulty in hearing the modulation of the target, or in distinguishing that modulation from the modulation of other sounds that are present, or in assigning modulation to the appropriate carrier frequency. Modulation-specific MDI occurs even under conditions promoting perceptual segregation of the target and masker. It may reflect the operation of channels tuned for detecting modulation but very broadly tuned for carrier frequency.

This work was supported by the Medical Research Council (U.K.). We thank Brian Glasberg for writing the computer programs to run these experiments, and Tom Baer, Bob Carlyon, Joe Hall, Brian Glasberg and Michael Stone for helpful comments on an earlier version of this paper.

REFERENCES

Bacon, S.P. & Grantham, D. W. 1989 Modulation masking: effects of modulation frequency, depth and phase. *J. acoust. Soc. Am.* **85**, 2575–2580.

Bregman, A.S. 1990 *Auditory scene analysis: the perceptual organisation of sound*. Cambridge, Massachusetts: Bradford Books, MIT Press.

Bregman, A.S., Levitan, R. & Liao, C. 1990 Fusion of auditory components: effects of the frequency of amplitude modulation. *Percept. Psychophys.* **47**, 68–73.

Bregman, A.S. & Pinker, S. 1978 Auditory streaming and the building of timbre. *Can. J. Psychol.* **32**, 19–31.

Darwin, C.J. 1981 Perceptual grouping of speech components differing in fundamental frequency and onset time. *Q. Jl exp. Psychol.* **33A**, 185–287.

Darwin, C.J. & Sutherland, N.S. 1984 Grouping frequency components of vowels: when is a harmonic not a harmonic? *Q. Jl exp. Psychol.* **36A**, 193–208.

Fletcher, H. 1940 Auditory patterns. *Rev. Mod. Phys.* **12**, 47–65.

Hall, J.W. & Grose, J.H. 1991 Some effects of auditory grouping factors on modulation detection interference (MDI). *J. acoust. Soc. Am.* **90**, 3028–3035.

Houtgast, T. 1989 Frequency selectivity in amplitude-modulation detection. *J. acoust. Soc. Am.* **85**, 1676–1680.

Kay, R.H. 1982 Hearing of modulation in sounds. *Physiol. Rev.* **62**, 894–975.

Kay, R.H. & Mathews, D.R. 1972 On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency modulated tones. *J. Physiol., Lond.* **225**, 657–667.

McAdams S. 1984 Spectral fusion, spectral parsing and the formation of the auditory image. Ph.D. thesis, University of Stanford.

Moore, B.C.J. 1989 *An introduction to the psychology of hearing*, 3rd edn. London: Academic Press.

Moore, B.C.J., Glasberg, B.R., Gaunt, T. & Child, T. 1991 Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals. *Q. Jl exp. Psychol.* **43A**, 327–347.

Moore, B.C.J., Glasberg, B.R. & Peters, R.W. 1986 Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *J. acoust. Soc. Am.* **80**, 479–483.

Moore, B.C.J., Glasberg, B.R. & Schooneveldt, G.P. 1990 Across-channel masking and comodulation masking release. *J. acoust. Soc. Am.* **87**, 1683–1694.

Moore, B.C.J. & Jorasz, U. 1992 Detection of changes in modulation depth of a target sound in the presence of other modulated sounds. *J. acoust. Soc. Am.* **91**, 1051–1061.

Patterson, R.D. & Moore, B.C.J. 1986 Auditory filters and excitation patterns as representations of frequency resolutions. In *Frequency selectivity in hearing* (ed. B. C. J. Moore), pp. 123–177. London: Academic Press.

Rasch, R.A. 1978 The perception of simultaneous notes such as in polyphonic music. *Acustica* **40**, 21–33.

Rees, A. & Moller, A.R. 1983 Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear. Res.* **10**, 301–310.

Roberts, B. & Moore, B.C.J. 1991 The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels under conditions of asynchrony. *J. acoust. Soc. Am.* **89**, 2922–2932.

Roberts, B.R. & Bregman, A. 1991 Effects of the pattern of spectral spacing on the perceptual fusion of harmonics. *J. acoust. Soc. Am.* **90**, 3050–3060.

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

Tansley, B.W. & Suffield, J.B. 1983 Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation. *J. acoust. Soc. Am.* **74**, 765–775.

Wilson, A.S., Hall, J.W. & Grose, J.H. 1990 Detection of frequency modulation (FM) in the presence of a second FM tone. *J. acoust. Soc. Am.* **88**, 1333–1338.

Yost, W.A. & Sheft, S. 1989 Across-critical-band processing of amplitude-modulated tones. *J. acoust. Soc. Am.* **85**, 848–857.

Yost, W.A., Sheft, S. & Opie, J. 1989 Modulation interference in detection and discrimination of amplitude modulation. *J. acoust. Soc. Am.* **86**, 2138–2147.