Across-channel processes in auditory masking

Brian C. J. Moore

*Department of Experimental Psychology, University of Cambridge,
Downing Street, Cambridge CB2 3EB, England*

The peripheral auditory system is often modelled as containing an array of bandpass filters (called the auditory filters), each tuned to a different centre frequency. When a subject tries to detect, or discriminate changes in, a narrowband signal in noise, it has traditionally been assumed that performance is based on the output of the single auditory filter which gives the highest signal-to-masker ratio. However, recent experiments demonstrate that outputs from other filters, tuned to frequencies remote from the signal frequency, can both enhance and degrade signal detection and discrimination. The enhancement takes place especially when the envelope of the masker fluctuates over time, and when the fluctuations are correlated across different frequency bands. This phenomenon is called comodulation masking release (CMR). In other situations, frequency components remote from the frequency of a signal may impair the discrimination of changes in that signal, especially changes in the modulation of the signal. This has been called modulation detection interference or modulation discrimination interference (MDI). Both CMR and MDI may depend partly on basic processes of auditory grouping that are involved in assigning elements of complex sounds to perceptual streams. However, other processes are probably involved.

Keywords: Masking, Masking release, Comodulation, Modulation discrimination, Across-channel processes

PACS number: 43. 66. Dc, 43. 66. Ba, 43. 66. Mk, 43. 66. Lj

1. INTRODUCTION

It is widely assumed that the peripheral auditory system contains a bank of overlapping bandpass filters (the auditory filters) which are approximately linear.¹⁾ The shape of the auditory filter at a given centre frequency can be thought of as a weighting function which is applied to the power spectrum of a sound to determine the effective magnitude of the output of the filter. The auditory filters seem to play an important role in many aspects of auditory perception and discrimination, including masking, loudness perception, timbre perception and pitch perception.²⁻⁴⁾

The detection and discrimination of narrowband signals has traditionally been explained by considering just one auditory filter at a time. For example, when an observer is trying to detect a sinusoidal signal of a given frequency 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, and threshold is assumed to correspond to a constant signal-to-masker ratio.¹⁾

Although this model works well in many situations⁵⁾ 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 profile analysis⁶⁾ or comodulation masking release (CMR)⁷⁾; see below for details. 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 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.^{8,9)} Hence, this phenomenon has been called modulation detection interference⁸⁾ or modulation discrimination inter-

ference¹⁰⁾ (MDI).

In this paper, each of these situations is reviewed in turn. There follows a discussion of whether CMR and MDI can be accounted for in terms of mechanisms of perceptual grouping. Finally, other possible factors affecting CMR and MDI are considered.

2. COMODULATION MASKING RELEASE

2.1 Initial Demonstrations of Comodulation Masking Release

Hall *et al.*⁷⁾ provided one of the first demonstrations that across-filter comparisons could enhance the detection of a sinusoidal signal in a fluctuating noise masker. The crucial feature for achieving this enhancement was that the fluctuations should be correlated across different frequency bands. In one of their experiments, the threshold for a 1-kHz 400-ms signal was measured as a function of the bandwidth of a continuous noise masker, keeping the spectrum level constant. The masker was centred at 1 kHz. They used two types of masker. One was a random (Gaussian) noise; this has irregular fluctuations in amplitude, and the fluctuations are independent in different frequency regions. The other was a random noise which was modulated in amplitude at an irregular, low rate. The modulation resulted in fluctuations in the amplitude of the noise which were the same in different spectral regions. This across-frequency correlation was called 'comodulation' by Hall *et al.* Figure 1 shows the results of this experiment.

For the random noise (denoted by R), the signal threshold increased as the masker bandwidth increased up to about 100–200 Hz, and then remained constant. This is exactly as expected from the traditional model of masking.^{1,5)} The auditory filter at this centre frequency has a bandwidth of about 130 Hz.¹¹⁾ Hence, for noise bandwidths up to about 130 Hz, increasing the bandwidth results in more noise passing through the filter. However, increasing the bandwidth beyond 130 Hz does not increase the amount of noise passing through the filter, so threshold does not increase. The pattern for the modulated noise (denoted by M) is quite different. For noise bandwidths greater than 100 Hz, the signal threshold decreases as the bandwidth increases. This indicates that subjects can compare the outputs of different auditory filters to enhance

Fig. 1 Thresholds for a sinusoidal signal plotted as a function of the bandwidth of a noise masker centred at the signal frequency. The noise was either random (R) or modulated (M). From Hall *et al.*⁷⁾ by permission of the authors and publisher.

signal detection. The fact that the decrease in threshold with increasing bandwidth only occurs with the modulated noise indicates that fluctuations in the masker are critical and that the fluctuations need to be correlated across frequency bands. Hence, this phenomenon has been called 'comodulation masking release' (CMR).

A second way in which CMR can be demonstrated is by using narrow bands of noise, which have inherent relatively slow amplitude fluctuations. One band, the on-frequency band, is centred at the signal frequency. One or more other bands, the flanking bands, are centred away from the signal frequency. If the flanking bands are uncorrelated with the on-frequency band, then when they are added to the on-frequency band they either have no effect on signal threshold, or they increase the signal threshold slightly. However, when the envelopes of the flanking bands are correlated with that of the on-frequency band (*i.e.* they have the same pattern of fluctuation over time), the flanking bands can produce a release from masking, a CMR.^{7,12-14)} The release from masking can occur even if the signal and on-frequency band are presented to one ear and the flanking bands are presented to the other ear, although in this case the magnitude of the release is usually a little smaller.¹³⁻¹⁵⁾ The magnitude of the CMR is usually defined as the difference in signal threshold for the condition with the uncor-

related flanking bands and the condition with the correlated flanking bands. Sometimes it is defined as the difference in thresholds for the condition with the on-frequency band alone and the condition with the correlated flanking bands.

2.2 The Role of Within-Channel Cues

Schooneveldt and Moore^{14,16)} and Moore *et al.*¹⁷⁾ have shown that modulation of a masker can produce a release from masking even when the masker's bandwidth is less than the auditory-filter bandwidth (for example, when both on-frequency and flanking bands fall within the bandwidth of a single auditory filter). This release from masking cannot arise from comparisons of the outputs of different auditory filters, and hence it does not represent a CMR. Rather it results from a cue or cues available in the output of a single auditory filter. Schooneveldt and Moore called such cues 'within-channel cues.' One example of such a cue is a change in the pattern of envelope modulation which occurs when the signal is added to the masker; the envelope fluctuates less and the minima in the envelope tend to be less deep when the signal is present.¹⁶⁾ This cue appears to be used in band-widening experiments such as that illustrated in Fig. 1, but it can only be used when the signal duration is greater than about 100 ms.¹⁶⁾ Of course, this within-channel cue does not account for the decrease in thresholds with increasing bandwidth beyond 130 Hz. Its primary effect seems to be to reduce thresholds in modulated noise for all bandwidths.

A second possible within-channel cue is connected with the detailed temporal waveform at the outputs of auditory filters tuned close to the signal frequency. This waveform is reflected in the patterns of phase locking in neurones tuned to that frequency¹⁸⁾; the nerve spikes tend to occur at a specific phase of the stimulating waveform. When the masker envelope is close to a maximum, the neurones are phase locked primarily to the masker. When the masker envelope is close to a minimum the neurones phase lock at least partly to the signal. This change in phase locking is a cue which may be used in experiments where an on-frequency band and a flanking band are used as maskers. When the two bands are close in frequency, they interact at the output of the auditory filter to give a waveform with regular envelope fluctuations, and with a centre frequency corresponding to the mean frequency of the two

bands. Thus, the pattern of phase locking fluctuates as the masker envelope fluctuates. Schooneveldt and Moore¹⁴⁾ presented evidence that this cue was most effective for centre frequencies of 4,000 Hz and below when the flanking band was close in frequency to the on-frequency band.

These within-channel cues make it easy to overestimate the magnitude of CMR. However, the cues do not appear to be important for brief signals in band-widening experiments, or when the on-frequency and flanking bands are widely separated in frequency. They also do not occur when the on-frequency and flanking bands are presented to opposite ears. Hence, results for these conditions can be used to estimate the magnitude of the 'true' CMR.

2.3 Factors Influencing the Magnitude of CMR

CMR measured in band-widening experiments (as in Fig. 1) occurs over a wide range of signal frequencies (500 to 4,000 Hz), and does not vary greatly with signal frequency.^{7,16,19)} CMR is largest when the modulation of the masker is at a low rate, and when the masker covers a wide frequency range.^{16,19)} For signal durations of 100 ms or less, for which within-channel cues probably have little effect, CMR can be as large as 11 dB.¹⁶⁾

When CMR is measured using an on-frequency band and a single flanking band, it generally falls in the range 1–6 dB when the flanking band is distant in frequency from the on-frequency band^{7,12–14)}; this probably reflects a true CMR. The release from masking can be as large as 14 dB when the flanking band is close in frequency to the on-frequency band, but in this case within-channel cues probably influence the results.¹⁴⁾ The 'true' CMR appears to vary little with centre frequency or with flanking-band frequency, although it may be somewhat greater for flanking band frequencies close to the signal frequency.²⁰⁾ CMR measured with a flanking band presented in the opposite ear to the signal-plus-masker also varies little with centre frequency or flanking-band frequency, and is typically 2–6 dB.¹⁴⁾ Again, however, there is a trend for CMRs to be larger when the flanking bands are centred close to the signal frequency.²⁰⁾ Thus, across-filter comparisons can be made over a wide frequency range, but they are made most effectively for filters that are reasonably closely spaced in centre frequency.

CMR measured with an on-frequency band and a flanking band tends to increase as the width of the bands of noise is decreased. This is probably a consequence of the fact that the rate of envelope fluctuations decreases as the bandwidth decreases; slow fluctuations lead to large CMRs.^{14,16,21,22)} CMR also increases if more than one flanking band is used. Indeed, if a large number of bands is used, the CMR can be as large as 15 dB.^{17,21,23-25)}

3. MODULATION DISCRIMINATION INTERFERENCE

I turn now to situations where the detection or discrimination of a signal is *impaired* by the presence of frequency components remote from the signal frequency. Usually, this happens when the task is either to detect modulation of the signal, or to detect a change in depth of modulation of the signal. Yost and Sheft⁸⁾ showed that the threshold for detecting sinusoidal amplitude modulation (AM) of a sinusoidal carrier was increased in the presence of another carrier, amplitude modulated at the same rate, even when the second carrier was remote in frequency from the first. They called this modulation detection interference (MDI). They showed that MDI did not occur if the second carrier was unmodulated. Wilson, Hall and Grose²⁶⁾ have demonstrated a similar phenomenon for the detection of frequency modulation (FM). They showed that the threshold for detecting sinusoidal FM of a given carrier frequency could be elevated in the presence of a second carrier which was also frequency modulated. The effect showed some tuning for carrier frequency, being greater when the carrier frequencies of the target and interfering sounds were close, but it persisted even when the carriers were widely separated in frequency (by more than one octave). The effect also showed tuning for modulation rate, being greatest when the modulation rates of the target and interfering sounds were close.

Moore *et al.*⁹⁾ carried out a series of experiments measuring the threshold of a brief 700-Hz signal as a function of its temporal position relative to the modulation cycle of a 700-Hz masker 100% amplitude modulated at a 10-Hz rate. When extra components, also 100% amplitude modulated at 10 Hz, were added to the masker, signal detection was often impaired, even when all of the added components were remote in frequency from 700 Hz.

The effect was usually greater when the added components were modulated in phase with the 700-Hz component than when they were modulated with a different phase. Moore *et al.* called this effect across-channel masking (ACM), since it did not seem to be explicable in terms of the output of the single auditory filter centred at the signal frequency.

It seems likely that the across-channel masking found by Moore *et al.*⁹⁾ is related to the MDI described by Yost and Sheft.⁸⁾ For the signal and maskers used by Moore *et al.*, the main cue for signal detection was probably the change in the pattern of envelope modulation produced by adding the signal to the masker. The flanking modulated components appear to interfere with the ability to detect the change in modulation pattern (an effect opposite to CMR).

Moore *et al.*²⁷⁾ determined how thresholds for detecting an increase in modulation depth (sinusoidal AM or FM) of a 1,000-Hz carrier frequency (the target) were affected by modulation of carriers (interference) with frequencies of 230 Hz and 3,300 Hz. The target sound was presented twice for a duration of 1,000 ms, with 300-ms inter-stimulus interval, and the subject was required to indicate the interval in which the sound was more modulated. The interference was gated with the target sound. The target sound was always modulated at a 10-Hz rate; the modulation index of the standard sound was 0.25 for AM and 2 for FM. An example of their results is shown in Fig. 2. They found that modulation increment thresholds were increased (worsened) when the remote carriers were modulated. This MDI effect was greatest when the target and interference were modulated at similar rates, but the effect was broadly tuned for modulation rate. When the target and interfering sounds were both modulated at 10 Hz, there was no significant effect of the relative phase of modulation of the target and interfering sounds. Moore *et al.* also found cross-modal interference effects; the detection of an increment in AM depth was adversely affected by FM interference, and vice versa. Moore *et al.* suggested that this could be explained by the transformation of FM to AM in the peripheral auditory system.

In summary, the detection of modulation (either AM or FM) or the detection of changes in modulation depth in a target sound can be impaired by modulated sounds remote from the target frequency.

Fig. 2 Results for one subject from Moore *et al.*²⁷⁾ Thresholds for detecting a change in modulation depth are plotted as a function of the modulation rate of the interfering sound (log scale). The box in each panel indicates the signal type and interference type. Squares show thresholds in the absence of any interfering sound and triangles show thresholds with the interfering sound unmodulated. Diamonds show thresholds for interference modulated at a 10-Hz rate, the same as the target, averaged across the four values of modulator phase. Error bars show $\pm$ one standard error.

The effects are greatest when the target and interference are modulated at the same rate, but they persist for quite large differences in modulation rate between the target and the interference.

4. THE ROLE OF PERCEPTUAL GROUPING IN CMR AND MDI

In everyday life we often listen to several sound sources simultaneously. The auditory system is faced with the task of deciding which 'elements' of the complex mixture arise from one source, and which from another. The process of doing this is often 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.^{4,28)} 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.

Consider first CMR. When the masker is a single narrow band of noise centred at the signal frequency, a sinusoidal signal added to the masker is not usually heard as a separate auditory 'object' or stream. Instead, the signal is detected as a

slight change in loudness, or a subtle change in the pattern of fluctuation of the stimulus.^{16,20)} However, when comodulated flanking bands are added to the masker, they tend to be grouped with the on-frequency band to form a coherent percept. When the signal is added it appears to 'pop out'; a pure tone is heard at the signal frequency. Moore and Shailer²⁵⁾ reported that this can even happen when the on-frequency band is presented to one ear, and multiple flanking bands are presented to the other. If the spectrum level of the flanking bands is close to or above that of the on-frequency band, then the entire masker may appear to come from the ear receiving the flanking bands. When a sinusoidal signal is added to the on-frequency band, it appears as a separate auditory event in the ear opposite to the ear receiving the flanking bands. In both these examples, the crucial feature seems to be that comodulation of the on-frequency band and flanking bands causes all bands to form a coherent perceptual stream. The signal causes a change in the modulation pattern in a local frequency region (*i.e.* at the output of the auditory filter centred at the signal frequency), and this disruption of the across-frequency comodulation causes the signal to 'pop out.'

Perhaps the most direct evidence for a role of perceptual grouping processes in CMR comes from some experiments of Hall and Grose.³⁰⁾ They measured CMR in conditions where some flanking bands were comodulated with the on-frequency band, but other bands (called 'deviant' bands) were not. These deviant bands tended to reduce the magnitude of CMR, especially when they were centred close to the on-frequency band. In one condition of their experiment, CMR was measured for a 1,000-Hz signal and a masker composed of seven comodulated 20-Hz wide bands, spaced from 400 to 1,600 Hz (with one at 1,000 Hz). The CMR for this condition, relative to the threshold for the on-frequency band alone, was about 15 dB. In a second condition, two deviant bands were added at 900 and 1,100 Hz. These bands reduced the CMR to only 3–4 dB. They then investigated the effect of adding more deviant bands, interleaved in frequency with the comodulated bands. When the deviant bands shared a common envelope (codeviant bands) the extra bands resulted in increased CMR; the codeviant bands distant from the signal frequency reduced the disruptive effect of the code-

viant bands close to the signal frequency. However, when the extra deviant bands all had independent envelopes (multideviant bands), no such beneficial effect was observed. The benefit of adding extra codeviant bands was also found when they were added in the opposite ear to the remainder of the stimulus. These results suggest a strong role for perceptual grouping. The added codeviant bands presumably formed a perceptual group with the codeviant bands close to the signal frequency, making all of the codeviant bands sound like a single stream. This helped to reduce their effect in disrupting CMR. As pointed out by Hall and Grose,³⁰⁾ these results indicate that "some processes contributing to CMR occur after, or receive feedback from, stages of processing that affect auditory grouping."

I turn now to the possible role of perceptual grouping in MDI. Yost and Sheft⁸⁾ suggested that the MDI found by them might be a consequence of perceptual grouping; the common AM of the target and interfering sounds might make them fuse perceptually,^{31,32)} making it difficult to 'hear out' the modulation of the target sound. A similar explanation might apply to the other experiments demonstrating MDI, as described earlier. Certainly, the impression of subjects in these experiments is that the target sound is difficult to 'hear out' in the presence of the modulated interfering sounds. 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 interference are modulated at different rates.^{26,27)}

Recently, in collaboration with Michael Shailer and Urszula Joras, I have been investigating the effects on MDI of gating the target sound on after the interfering sounds. 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 (for reviews, see Moore⁴⁾ and Bregman²⁸⁾). Thus, if MDI depends on perceptual grouping of the

Fig. 3 Thresholds for detecting an increase in modulation index from a standard value of 0.25 for a 1,000-Hz carrier modulated at a 10-Hz rate. The threshold without any interfering sound is indicated by the dashed horizontal line. The interfering sounds were centred at 230 and 3,300 Hz. Symbols show thresholds plotted as a function of the onset asynchrony between the interference and the target sound.

target and interference, delaying the target sound relative to the interfering sounds should result in a marked decrease in MDI. In one task, the subjects had to detect an increase in depth of 10-Hz AM of a 1,000-Hz carrier, relative to a baseline modulation index of 0.25. The target sound was presented twice for a duration of 400 ms, with 300-ms inter-stimulus interval, and the subject was required to indicate the interval in which the sound was more modulated. An example of some results obtained by Michael Shailer and myself is shown in Fig. 3.

There were four conditions. In one, the target sound was presented alone. The modulation increment threshold in this case is indicated by the dashed line in Fig. 3. In the other conditions, interfering sounds were added at centre frequencies of 230 Hz and 3,300 Hz. The interfering sounds were either gated synchronously with the target (0-ms asynchrony), or were gated on 25, 50 or 100 ms before the target. In one of these conditions (circles), the interfering sounds were unmodulated. In another (up-pointing triangles) they were mod-

ulated in phase with the target with modulation index 0.5. In the last condition (down-pointing triangles), they were modulated in antiphase with the target.

Even when the interference was unmodulated, it produced some MDI; the amount of MDI declined as the onset asynchrony increased. The MDI in this case presumably occurred because gating the target and interference together tended to promote their perceptual fusion. MDI was markedly greater when the interference was modulated. Again, the MDI declined with increasing onset asynchrony, but, even at the largest asynchrony (100 ms) a substantial amount of MDI remained.

Overall, these results suggest that perceptual grouping does play a role in MDI. The MDI for both modulated and unmodulated interference declined with increasing onset asynchrony. Similar results have been reported by Moore and Jorasz¹⁰⁾ for AM modulation depth discrimination and by Hall and Grose⁸³⁾ for the detection of AM and FM. Hall and Grose also found that they could reduce or even abolish MDI by presenting the interference continuously, rather than gated with the target. However, in our experiments, significant MDI remained even at large onset asynchronies, suggesting that some factor other than perceptual grouping plays a role.

Finally, it is worth noting that, if we were to interpret both CMR and MDI as arising from perceptual grouping processes, we would come to quite different conclusions about the precision with which grouping mechanisms operate. Results from CMR experiments suggest that the auditory system is very sensitive to disparities in modulation pattern across bands; even a small disparity appears sufficient to cause the signal to 'pop out' and to enhance signal detection. In contrast, to explain MDI in terms of grouping, we would have to say that sounds with quite different modulation patterns tend to be fused and heard as a single sound.

5. FACTORS DETERMINING WHETHER CMR OCCURS OR MDI OCCURS

In many ways, the situations where CMR occurs and where MDI occurs are very similar. Both CMR and MDI involve modulated stimuli. In the case of CMR, the signal is added to a modulated sound (the on-frequency band), and results in a

change in the modulation pattern of that sound. The flanking bands 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. However, the flanking (remote) components make it harder rather than easier to detect the change. What then, determines whether CMR occurs or MDI occurs? The following sections describe some relevant factors.

5.1 The Relative Modulation Depth of the On-Frequency and Flanking Bands

In the great majority of experiments demonstrating CMR, the flanking bands have been modulated in the same way, and with the same modulation depth, as the on-frequency band. Thus, the addition of the signal to the masker has resulted in a *disparity* in modulation depth across bands, providing a potential detection cue. In contrast, in experiments demonstrating MDI, the flanking bands have usually had a greater modulation depth than the on-frequency band. Thus, even in the non-signal interval there was a disparity in modulation depth across bands. However, this difference between CMR and MDI does not appear to be critical. Grose and Hall³⁴⁾ showed that when the on-frequency band was a sinusoid modulated with a modulation index of 0.63, substantial CMRs could be obtained when the flanking bands were modulated with a modulation index of 1.0. Thus, CMR *can* be obtained when the flanking bands are modulated more than the on-frequency band. Also, Moore and Jorasz¹⁰⁾ have shown that MDI for AM increment detection can occur when the on-frequency band and the flanking bands are modulated at the same depth. Thus, the relative modulation depths of the on-frequency band and flanking bands do not seem to be critical in determining whether CMR or MDI will occur.

5.2 The Difficulty of Signal Detection in the Reference Condition

A second difference between CMR and MDI is related to the difficulty of signal detection in the condition with the on-frequency band 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. For example, when the masker is a Gaussian noise with a very narrow band-

width, the main cue for signal detection appears to be the change in level produced by adding the signal to the masker.^{16,29)} However, this is a poor cue, because narrowband noise fluctuates markedly in level from moment to moment. Under these conditions, the signal level required for threshold has to be roughly equal to the masker level; the change in overall level at threshold is then about 3 dB. Co-modulated flanking bands provide an indication of the inherent fluctuations in the masker, thus reducing uncertainty, and giving lower thresholds. Hence, a large CMR occurs.

In contrast, in experiments demonstrating MDI, thresholds in the reference condition tend to be quite low. For example, in Fig. 3, a change in modulation index from 0.25 to 0.286 was detectable. This corresponds to a change in level in the 'valleys' of the target of only 0.42 dB and a change in peak-to-valley ratio from 4.44 dB to 5.51 dB. In general then, one might argue that when performance in the reference condition is poor, possibly because of stimulus uncertainty, adding comodulated flanking bands tends to give a CMR. When performance in the reference condition is good, adding flanking bands either has no effect or produces MDI.

5.3 The Nature of the Signal to be Detected

In experiments demonstrating CMR, the addition of the signal to the masker usually results in an effective decrease in the modulation depth of the on-frequency band. If the signal is a relatively long duration sinusoid then it partially 'fills in' the valleys in a fluctuating noise masker.¹⁶⁾ Sometimes the signal consists of one or more brief tone pips presented in the valleys of an AM masker.^{9,34)} In either case, the result is that the level in the valleys increases. Several researchers have suggested that CMR may depend strongly on the detection of changes in the valleys ('dip listening').^{9,34-36)} In contrast, experiments demonstrating MDI have usually required subjects to detect an *increase* in modulation depth, resulting in a decrease in level in the valleys. Moore *et al.*⁹⁾ showed that stimuli that were otherwise very similar could give CMR when the task was to detect an increase in level in the valleys, and MDI when the task was to detect a decrease. Moore and Jorasz¹⁰⁾ studied the ability to detect changes in modulation depth of a target sound in the presence of added sounds modulated at the same rate. They found that MDI could occur

when the task of the subject was to detect either an increase in modulation depth or a decrease in modulation depth. However, they were only able to demonstrate CMR when the task was to detect a decrease in modulation depth.

5.4 The Need for a Change in Modulation Pattern

It may be the case that large CMRs only occur when there is a difference in modulation *pattern* across bands, rather than simply a difference in modulation depth.⁹⁾ In experiments where large CMRs are observed, the addition of the signal usually changes the modulation pattern of the on-frequency band, for example transforming a local minimum to a local maximum or plateau. Most experiments on MDI have involved only a change in modulation depth, not a change in modulation pattern. Moore and Jorasz¹⁰⁾ and Fantini and Moore³⁷⁾ have shown that it is *possible* to obtain CMR when the only cue is an across-band disparity in modulation depth, but the effect tends to be small, and it only occurs for large modulation depths of the target sound.

5.5 The Synchrony or Asynchrony of Signal and Masker

A final difference between CMR and MDI concerns the timing of the signal and masker. Most CMR experiments have involved continuous maskers, with the signal gated on during a certain observation interval. This tends to promote perceptual segregation of the signal and masker. Although it is possible to obtain CMR when the signal and masker are gated synchronously, the CMR tends to be rather small.¹⁰⁾ On the other hand, MDI is largest when the signal and masker are gated synchronously, which tends to promote perceptual fusion of the signal and masker. And, as described above, MDI becomes small or non-existent if the masker is gated on before the signal or if the masker is continuous.

In summary, large CMRs tend to occur when the signal threshold is high in the reference condition (on-frequency band alone), when the signal produces an increase in level in the valleys of the on-frequency band, when the addition of the signal to the on-frequency band results in a change in modulation *pattern* and when the signal and masker are not gated synchronously. When all these conditions are satisfied, the addition of comodulated flanking

bands to the on-frequency band results in a large masking release. When only one of these conditions is satisfied, a small CMR may occur. When none of these conditions is satisfied, the addition of flanking bands either produces no effect or produces MDI.

6. OTHER MECHANISMS INVOLVED IN CMR AND MDI

The discussion so far has concentrated on explanations of CMR and MDI in terms of perceptual grouping mechanisms. However, it seems likely that other factors are involved. This section considers a few of these other factors.

6.1 Mechanisms Involved in CMR

As described earlier, a common modulation pattern across channels (at the outputs of different auditory filters) may be used as a basis for the perceptual grouping of different frequency components in a complex sound. Conversely, when one channel has a different modulation pattern from the others, this may be used for perceptual segregation, making it easier to 'hear out' the signal in that channel. This can probably explain many cases of CMR. However, there are some clear examples of cases that cannot be explained in this way. For example, Hall and Grose^{36,38)} used an on-frequency band and one or more flanking bands as maskers, but the signal was a band of noise identical to the on-frequency band. Thus, the addition of the signal to the on-frequency band merely resulted in a change in overall level, without changing the modulation pattern. They found a significant CMR; detection of the signal was better in the presence of one or more comodulated flanking bands than in the presence of non-comodulated bands. This indicates that a disparity in modulation pattern or depth across channels is not necessary to obtain a CMR. In this case, the flanking bands probably improved signal detection because subjects were able to judge the level of the on-frequency band (or that band plus the signal) *relative* to the level(s) of the flanking band(s). Essentially, one might regard this task as a special case of profile analysis involving dynamically varying stimuli.³⁹⁾

A second example of CMR that is difficult to explain in terms of perceptual grouping comes from some experiments by Hall *et al.*²³⁾ and by van den Brink *et al.*⁴⁰⁾ They found that CMR could be obtained when the signal itself was complex and

covered the same spectral region as the masker. For example, Hall *et al.* obtained a CMR when the masker was composed of three noise bands and the signal was composed of three sinusoids, each centred in one of the bands. In this case, when the signal was present, the modulation pattern would have been different for each of the three bands. Thus, there would have been no effective cue to allow perceptual separation of the signal and masker. However, a cue based on modulation patterns was available. When the masker alone was present, the comodulated bands would have produced a similar pattern of modulation at the outputs of all the auditory filters. When the (multi-component) signal was present, there would have been disparities in modulation pattern across bands. These disparities could provide a cue to the presence of the signal even though they might not allow perceptual segregation of the signal from the masker.

It seems likely that CMR can arise in several different ways, and that no single model will account for all of the results. While many cases of CMR can be explained in terms of perceptual grouping, other cases require different explanations.

6.2 Mechanisms Involved in MDI

One explanation of MDI, not mentioned so far, is that it reflects the operation of 'channels' specialized for detecting and analysing modulation. There is both physiological evidence⁴¹⁻⁴³ and psychophysical evidence⁴⁴⁻⁴⁷ for such channels. Yost *et al.*⁴⁸ 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.

Recently, Hall and Grose⁴⁹ have described some results consistent with the idea that modulation is detected in channels responsive to a range of centre frequencies. 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

Fig. 4 An example of the results of Shailer and Moore (unpublished). The subject heard a 1,000-Hz sinusoid amplitude modulated at a 10-Hz rate, followed by that same sound together with two other carriers ("maskers"), one at 230 Hz and the other at 3,300 Hz. The maskers were modulated at either 10 Hz, 4 Hz or 25 Hz. The subject was asked to rate how clearly the modulation of the 1,000 Hz carrier could be heard in the second sound. The duration of the sounds was varied from 200 to 1,600 ms, but on any given trial the first sound (single carrier) and the second sound (three carriers) always had the same duration. The pentagons indicate the results obtained when only the 1,000-Hz carrier was present in the second sound.

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.

A very different experiment, conducted by myself and Michael Shailer (unpublished results), also suggests that subjects can have difficulty assigning modulation to particular components. Subjects were presented with a 1-kHz tone amplitude modulated at a 10-Hz rate, followed by that same tone together with some interfering sounds. They were asked to listen to the modulation of the 1-kHz tone alone, and to rate (on a scale from 1 to 7) how clearly they could hear the modulation of that tone

in the complex mixture.

An example of the results is given in Fig. 4. When the interfering sounds were modulated at a 10-Hz rate (a condition giving maximum MDI), clarity ratings were higher than when the interfering sounds were modulated at 4 Hz or 25 Hz (conditions giving less MDI). Thus there was a discrepancy between the MDI results and the clarity ratings. Subjects could make an overall judgement of how clearly they could hear 10-Hz modulation, but it was hard to judge the modulation of the 1-kHz component alone. It seems that, under some conditions, it is hard to assign a particular modulation to a particular carrier frequency, consistent with the idea that the hypothetical modulation channels respond to a wide range of carrier frequencies. Exactly what these conditions are remains to be determined.

7. CONCLUDING REMARKS

The results reviewed above are not consistent with the traditional idea that stimuli exciting different auditory filters (or critical bands) are analysed independently from each other. Components separated by many critical bands from a target sound can either enhance or degrade the detection and discrimination of that target, depending on the stimulus conditions. The case of MDI is especially interesting, since it indicates that people are not able to attend to the output of a single auditory filter even if that would be the best way of performing the task required of them. The results from experiments on both CMR and MDI suggest that decisions can only be made *after* some degree of perceptual grouping has occurred; we do not have access to the raw "sensory data."

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council (U.K.). I thank Brain Glasberg, Michael Shailer, Michael Stone and Joe Hall for helpful discussions.

REFERENCES

- 1) H. Fletcher, "Auditory patterns," *Rev. Mod. Phys.* **12**, 47-65 (1940).
- 2) R. Plomp, *Aspects of Tone Sensation* (Academic Press, London, 1976).
- 3) B. C. J. Moore, *Frequency Selectivity in Hearing* (Academic Press, London, 1986).
- 4) B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 3rd ed. (Academic Press, London, 1989).
- 5) 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 Press, London, 1986), pp. 123-177.
- 6) D. M. Green, *Profile Analysis* (Oxford University Press, Oxford, 1988).
- 7) J. W. Hall, M. P. Haggard, and M. A. Fernandes, "Detection in noise by spectro-temporal pattern analysis," *J. Acoust. Soc. Am.* **76**, 50-56 (1984).
- 8) W. A. Yost and S. Sheft, "Across-critical-band processing of amplitude-modulated tones," *J. Acoust. Soc. Am.* **85**, 848-857 (1989).
- 9) 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).
- 10) B. C. J. Moore and U. Joras, "Detection of changes in modulation depth of a target sound in the presence of other modulated sounds," *J. Acoust. Soc. Am.* (in press) (1992).
- 11) B. C. J. Moore and B. R. Glasberg, "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750-753 (1983).
- 12) D. McFadden, "Comodulation masking release: Effects of varying the level, duration, and time delay of the cue band," *J. Acoust. Soc. Am.* **80**, 1658-1667 (1986).
- 13) M. F. Cohen and E. D. Schubert, "Influence of place synchrony on detection of a sinusoid," *J. Acoust. Soc. Am.* **81**, 452-458 (1987).
- 14) G. P. Schooneveldt and B. C. J. Moore, "Comodulation masking release (CMR): effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band," *J. Acoust. Soc. Am.* **82**, 1944-1956 (1987).
- 15) J. W. Hall, M. P. Haggard, and A. D. G. Harvey, "Release from masking through ipsilateral and contralateral comodulation of a flanking band," *J. Acoust. Soc. Am.* **76**, S76 (1984).
- 16) G. P. Schooneveldt and B. C. J. Moore, "Comodulation masking release (CMR) as a function of masker bandwidth, modular bandwidth and signal duration," *J. Acoust. Soc. Am.* **85**, 273-281 (1989).
- 17) B. C. J. Moore, M. J. Shailer, J. W. Hall, and G. P. Schooneveldt, "Comodulation masking release in subjects with unilateral and bilateral cochlear hearing impairment," *J. Acoust. Soc. Am.* (submitted) (1992).
- 18) D. J. Anderson, "Quantitative model for the effects of stimulus frequency upon synchronization of auditory discharges," *J. Acoust. Soc. Am.* **54**, 361-364 (1973).
- 19) M. P. Haggard, J. W. Hall, and J. H. Grose, "Comodulation masking release as a function of band-

- width and test frequency," *J. Acoust. Soc. Am.* **88**, 113–118 (1990).
- 20) J. W. Hall, J. H. Grose, and M. P. Haggard, "Effects of flanking band proximity, number, and modulation pattern on comodulation masking release," *J. Acoust. Soc. Am.* **87**, 269–283 (1990).
- 21) J. W. Hall, "Experiments on comodulation masking release," in *Auditory Processing of Complex Sounds*, W. A. Yost and C. S. Watson, Eds. (Lawrence Erlbaum, Hillsdale, N.J., 1987), pp. 57–66.
- 22) R. P. Carlyon, S. Buus, and M. Florentine, "Comodulation masking release for three types of modulator as a function of modulation rate," *Hear. Res.* **42**, 37–46 (1989).
- 23) 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).
- 24) G. P. Schooneveldt and B. C. J. Moore, "Comodulation masking release (CMR) for various monaural and binaural combinations of the signal, on-frequency and flanking bands," *J. Acoust. Soc. Am.* **85**, 262–272 (1989).
- 25) B. C. J. Moore and M. J. Shailer, "Comodulation masking release as a function of level," *J. Acoust. Soc. Am.* **90**, 829–835 (1991).
- 26) A. S. Wilson, J. W. Hall, and J. H. Grose, "Detection of frequency modulation (FM) in the presence of a second FM tone," *J. Acoust. Soc. Am.* **88**, 1333–1338 (1990).
- 27) 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).
- 28) A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound* (Bradford Books, MIT Press, Cambridge, Mass., 1990).
- 29) B. C. J. Moore, J. W. Hall, J. H. Grose, and G. P. Schooneveldt, "Some factors affecting the magnitude of comodulation masking release," *J. Acoust. Soc. Am.* **88**, 1694–1702 (1990).
- 30) J. W. Hall and J. H. Grose, "Comodulation masking release and auditory grouping," *J. Acoust. Soc. Am.* **88**, 119–125 (1990).
- 31) S. McAdams, "Spectral fusion, spectral parsing and the formation of the auditory image," Ph. D. Thesis, Stanford (1984).
- 32) A. S. Bregman, R. Levitan, and C. Liao, "Fusion of auditory components: Effects of the frequency of amplitude modulation," *Percept. Psychophys.* **47**, 68–73 (1990).
- 33) J. W. Hall and J. H. Grose, "Consequences of onset asynchrony for modulation detection interference," *J. Acoust. Soc. Am.* **88**, S145 (1990).
- 34) J. H. Grose and J. W. Hall, "Comodulation masking release using SAM tonal complex maskers: Effects of modulation depth and signal position," *J. Acoust. Soc. Am.* **85**, 1276–1284 (1989).
- 35) S. Buus, "Release from masking caused by envelope fluctuations," *J. Acoust. Soc. Am.* **78**, 1958–1965 (1985).
- 36) J. W. Hall and J. H. Grose, "Relative contributions of envelope maxima and minima to comodulation masking release," *Q. J. Exp. Psychol.* **43A**, 349–372 (1991).
- 37) D. A. Fantini and B. C. J. Moore, "Comodulation masking release (CMR) and profile analysis: The effect of varying modulation depth," in *9th Int. Symp. on Hearing: Auditory Physiology and Perception*, Y. Cazals, L. Demany, and K. Horner, Eds. (Pergamon Press, Oxford, 1991) (in press).
- 38) J. W. Hall and J. H. Grose, "Comodulation masking release: Evidence for multiple cues," *J. Acoust. Soc. Am.* **84**, 1669–1675 (1988).
- 39) D. A. Fantini and B. C. J. Moore, "Across-band processing of dynamically varying signals," in *The Psychophysics of Speech Perception, II*, M. E. H. Schouten, Ed. (Nijhoff, Dordrecht, The Netherlands, 1992) (in press).
- 40) W. A. C. van den Brink, T. Houtgast, and G. F. Smoorenburg, "Signal detection in temporally modulated and spectrally shaped maskers," *J. Acoust. Soc. Am.* (submitted) (1992).
- 41) R. H. Kay, "Hearing of modulation in sounds," *Physiol. Rev.* **62**, 894–975 (1982).
- 42) A. Rees and A. R. Møller, "Responses of neurons in the inferior colliculus of the rat to AM and FM tones," *Hear. Res.* **10**, 301–310 (1983).
- 43) C. E. Schreiner and J. V. Urbas, "Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF)," *Hear. Res.* **21**, 227–241 (1986).
- 44) R. H. Kay and D. R. Mathews, "On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency modulated tones," *J. Physiol.* **225**, 657–667 (1972).
- 45) B. W. Tansley and J. B. Suffield, "Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation," *J. Acoust. Soc. Am.* **74**, 765–775 (1983).
- 46) T. Houtgast, "Frequency selectivity in amplitude-modulation detection," *J. Acoust. Soc. Am.* **85**, 1676–1680 (1989).
- 47) S. P. Bacon and D. W. Grantham, "Modulation masking: Effects of modulation frequency, depth and phase," *J. Acoust. Soc. Am.* **85**, 2575–2580 (1989).
- 48) 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).
- 49) J. W. Hall and J. H. Grose, "Some effects of auditory grouping factors on modulation detection interference (MDI)," *J. Acoust. Soc. Am.* (in press) (1991).

B. C. J. MOORE: ACROSS-CHANNEL PROCESSES IN MASKING

Brian Moore was born in London, England on February 10, 1946. He received his B. A. in Natural Sciences in 1968 and his Ph.D. in Psychoacoustics in 1971, both from the University of Cambridge, England. He held posts as Lecturer in Psychology at Reading University and Cambridge

University before taking up his present position as Reader in Auditory Perception at the University of Cambridge. He has also been a Visiting Professor at Brooklyn College, the City University of New York, and

the University of California at Berkeley. His research interests are: the perception of sound; mechanisms of normal hearing and hearing impairments; relationship of auditory abilities to speech perception; design of signal processing hearing aids for sensorineural hearing loss; electrical stimulation as a means of restoring hearing to the totally deaf; design and specification of high-fidelity sound-reproducing equipment. He is a Fellow of the Acoustical Society of America and a member of the Experimental Psychology Society (U.K.), the British Society of Audiology, the Audio Engineering Society and the Acoustical Society of Japan.