ARCHIVES OF ACOUSTICS
26, 3, 215–222 (2001)

CONTRIBUTION OF CMR AND MDI TO THE PICTURE OF THE HEARING SYSTEM SELECTIVITY

U. JORASZ

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

I hear all sounds running together, combined,
fused or following,
Sounds of the city and sounds out of the city,
Sounds of the day and night

Walt Whitman, *Song of myself*

The general picture of the hearing system selectivity should take into account the involving of the peripheral and central mechanisms. Everything what is known about CMR and MDI can contribute to the formation of auditory scene and take the important place in this picture.

1. Introduction

The fundamental characteristic of the auditory system is its action as a frequency analyser. This action is revealed by our ability to "hear out" the individual partials, or harmonics, of complex sounds. More generally, our ability to hear one sound in the presence of other sounds depends crucially on frequency resolution also known as frequency selectivity; I will use the latter term.

It seems likely that frequency selectivity depends to a large extent on the filtering that takes place in the cochlea. Frequency selectivity is most often quantified by masking which may be regarded as reflecting the limits of frequency analysis; by measuring when one sound is just masked by another, it is possible to characterize the frequency analysis capabilities of the auditory system.

The problem facing the auditory system is in fact much more complicated. In many everyday situations our auditory system is presented with an acoustic waveform made up from a mixture of sounds originating from a variety of sources. The role of the system is to interpret this complex waveform as sound-producing events.

The problem of interpreting sound in terms of separate events is closely related to the visual problem of interpreting, in terms of three-dimensional objects. This is the case that assigning the different frequency components in the sound to the appropriate sources is called the separation of auditory objects or the identification of sound sources. Our ability to do this is greater than we might expect from simple studies of masking hence the mechanism involved in this action should take into consideration not only what is happen in the cochlea, on the basilar membrane. This mechanism is called "perceptual grouping" or "stream formation" and — unlike filtering on the basilar membrane which is peripheral one — perceptual grouping belongs to central or cognitive processes in hearing. Perceptual grouping may be a very useful tool to construct the "auditory scene" and build the most general picture of selectivity of the hearing system. This form of selectivity has to be connected with both peripheral and central processes in hearing.

Now our task is to find out effects which could carry information about this high degree of selectivity of auditory system. I suggest that everything what is known about comodulation masking release (CMR) and modulation detection/discrimination interference (MDI) can contribute to the formation of auditory objects. CMR and MDI because of their nature play the role of bridging peripheral and central auditory processes [7] and *ipso facta* take the important place in the general picture of the hearing system selectivity.

2. Frequency selectivity   peripheral considerations

Frequency selectivity refers to the ability of the auditory system to separate or resolve the components in a complex sound. Usually we extend this terminology and say about the detection of single sound in a noise background. It seems likely that frequency selectivity is determined to a large extent on the level of cochlea.

Fletcher [2] suggested that the peripheral auditory system (cochlea and the neurones of the auditory nerve connected to inner hair cells and outer hair cells) behaves as if it contained a bank of bandpass filters (now called auditory filters), with continuously overlapping passbands. Recent data are consistent with Fletcher's point of view that basilar membrane inside cochlea provided the basis for the auditory filters [14]. 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 (within-channel mechanism) that gives the highest signal-to-noise ratio. The centre frequency of this filter is usually the same as or close to (off-frequency listening) the signal frequency. Threshold is assumed to correspond to a constant signal-to-noise ratio.

If an aspect of auditory perception can be explained entirely by consideration of processes occurring within one frequency channel, then that aspect might reflect mainly peripheral processing.

Over the past ten years this traditional model of masking clearly fails. Instead, performance appears to depend on the pattern of outputs across different auditory filters [4, 16].

3. Identi cation of the sound sources. Auditory grouping central processes

A basic problem faced by audition is not as simple as detection single sound in a noise background. The problem is in dealing with mixtures of sounds. The sensory components that arise from distinct environmental events have to be segregated into separate perceptual representations. The process of doing this is often described as "perceptual grouping" or "stream formation". These representations, which Bregman [1] called streams, provide basis of description that connect sensory features so that the right combinations can be useful for recognizing the environmental events or — saying differently — identification of the sources of sound. A very spectacular example is cocktail party effect. To recognize the unique timbre of the friend's voice we have to isolate the frequency components that are responsible for it from others that are present at the same time. An error of the hearing system, a wrong choice of frequency components would change the perceived timbre of the voice. The fact that we can usually identify the timbre implies that we regularly choose the right components.

The elements of the sound are grouped across-frequency and across-time to form percepts — or auditory objects — of coherent streams each with its own loudness, pitch, timbre and spatial location. This is named the building of "the auditory scene". Auditory object((1)) is perceptual representation of acoustic event. The constructing of the auditory scene and object identification is a form of selectivity of the hearing system, a form which needs more operations and additional principles than the plain frequency selectivity.

Our ability to identify the sound source is better than we might expect from simple studies of masking taking to consideration only peripheral processes. Auditory grouping is in a sense analogous to the filtering on the basilar membrane but in terms of perception. Grouping "puts in motion" special perceptive principles. Our decisions can only be made after some degree of perceptual grouping has occurred; we do not have access to the raw "sensory data" [9].

The principle which is known as a powerful factor in perceptual grouping is onset asynchrony. 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. Although the perceptual organization of hearing is rather functional than physiological in nature the onset asynchrony is supported by physiology; immediately after the onset of a sound, the firing rate of the auditory nerve increases then decreases rapidly [11]. This is a form of adaptation that may highlight change and the onset of new sounds.

It is worth to notice that grouping is not an "all-or-none" process. The rules which govern the auditory perception are not completely independent of one another, and not one of them always works perfectly. There exists some kind of economy of perception. The auditory system chooses this principle (or principles) which will be the best, more "economical", in particular condition.

---

(1) Why the term "auditory object/stream" is better than the word "sound"? First of all the representations of acoustic events can be multifold in a way that the word "sound" does not suggest. An example: the singer and piano together form a perceptual entity: the "performance" that is distinct from other sounds that are occuring [1].

4. The role of CMR and MDI on the construction of the picture of the hearing system selectivity

So far we have seen that the construction of the general picture of the hearing system selectivity needs the contribution of peripheral and central processes in hearing. The phenomena exploration of which could be of interest in creation this picture should be instrumental in bridging the both aspects of hearing: peripheral and central.

There is above-mentioned if an aspect of auditory perception can be explained entirely by consideration of processes occurring within one frequency channel, then that aspect might reflect mainly peripheral processing. However, if an aspect of auditory perception can only be explained by processes that involve comparing or combining information across frequency channels, then those processes must occur relatively centrally, at a level higher than the auditory nerve.

The phenomena known under acronyms CMR (comodulation masking release) and MDI (modulation detection/discrimination interference) may be very useful in illustrating these aspects of peripheral and central processing in hearing.

In the phenomenon of comodulation masking release (CMR) [4], the outputs of auditory filters tuned away from the signal frequency can be used to enhance signal detection. CMR occurs when the task is to detect a signal centred in a narrow-band masker that is amplitude modulated in some way. The addition of other components to the masker (the on-frequency band), remote from the signal frequency, can enhance signal detection, provided the extra components (sometimes called "flankers"([2])) have a similar pattern of modulation to the on-frequency band. CMR is usually assumed to reflect a relatively central across-channel process.

In the phenomenon of modulation detection/discrimination interference (MDI), 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 [16]. The ability to discriminate/detect these changes is adversely affected by the presence of other modulated sounds (also called flankers), even when those sounds have centre frequencies well away from that of the target. Again, MDI is usually assumed to reflect a relatively central across-channel process.

In many ways, the conditions in which CMR and MDI occur are only a little different. MDI resembles CMR, except that the remote components enhance detection in CMR and degrade it in MDI. Jorasz and Moore [8] suggest that the both effects are the two faces of the same phenomenon: the auditory masking. It is very important and "comfortable" because we can illustrate the nature of the both phenomena showing results of measurements for one of them (CMR or MDI).

The necessity of involving of central processes is the elementary feature of CMR and MDI, their "to be or not to be". But if the paradigm used to measure the detection of the signal had previously been associated with highly successful within-channel explanation that was based upon peripheral auditory processes we could expect that peripheral,

---

([2]) In reference to vision perception the role of "optical flankers" have been assigned — in a sense — to additional lights.

within-channel mechanisms contribute to CMR and MDI [7]. Additionally MOORE and JORASZ [12] demonstrated that perceptual grouping — closely connected with the identification of the sound sources — play a role in the amount of MDI.

Figure 1 (data from experiment of MOORE and JORASZ [12]; see also [8]) shows that the amount of MDI was markedly reduced by gating the target on after the flankers. It reflects the role of onset asynchrony known as a very important factor in perceptual grouping. It seems that perceptual grouping does play a role in MDI. All subjects reported that the onset asynchrony made it much easier to "hear out" the signal (a decrement in modulation depth). We can see that thresholds for detecting the target sound are lower when the target is delayed.

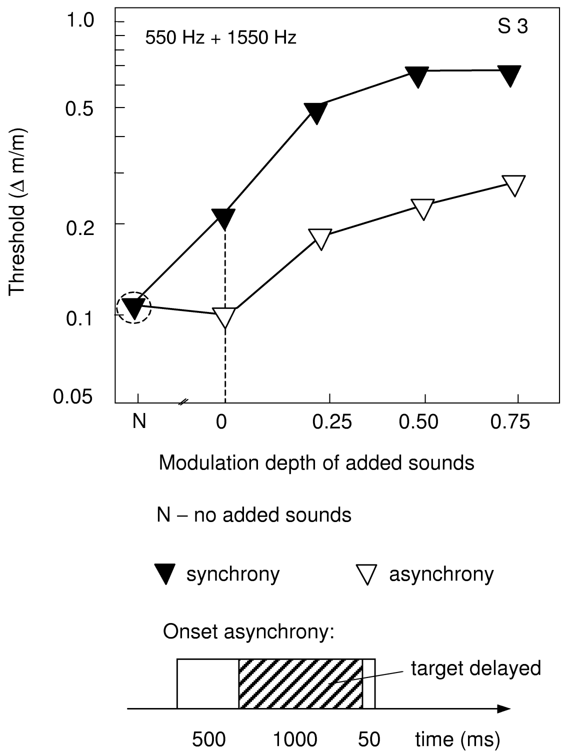

Fig. 1. Thresholds for detecting a decrement in modulation depth. Thresholds are expressed as a proportion of the modulation index of the standard sound ($m = 0.5$) and are plotted as a function of the modulation depth of the added sounds.

Consider now the results shown in Fig. 2 (data from JORASZ and MOORE [7]). There is a tendency for the amount of MDI to be greater for flankers centred above (conditions 4–6) the target frequency (1000 Hz) than for flankers centred below the target frequency (conditions 1–3). This is consistent with a role for within-channel processes. The peripheral processes involve the flankers either masking part of the excitation pattern of the target or introducing extra modulation into part of the excitation pattern evoked by

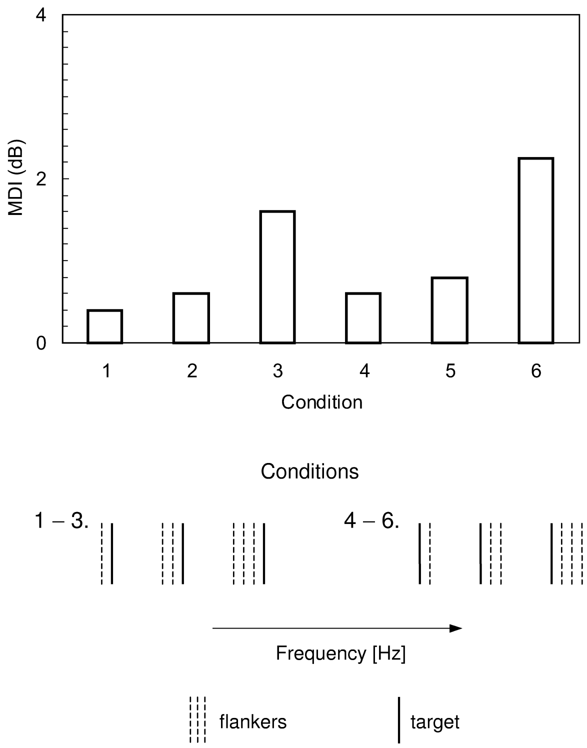

Fig. 2. The amount of MDI. The cartoons show schematically the spectra of the target ($f = 1000$ Hz) and the flankers (416, 572, 765, 1301, 1670, and 2127 Hz) and their presentation in each forced-choice trial for particular condition.

the target. This might be particularly important for flankers higher in frequency than the target, since changes in excitation level of the target are effectively magnified on the high-frequency side of the excitation pattern [17].

5. Conclusions

Figure 3 illustrates the conclusions of our speculation about general picture of the hearing system selectivity. Our ability to hear out one sound in the presence of other sounds (left side of the graph) or — more generally — our dealing with mixtures of sounds effecting the identification of sound sources (right side of the graph) is the basic problem facing the auditory system.

The full picture of the hearing system selectivity should take into consideration not only the filtering on the basilar membrane inside the cochlea, i.e. peripheral mechanism, but also the perceptual mechanism called perceptual grouping, i.e.central mechanism, involved in this action of the hearing system.

Exploration of CMR and MDI, the effects which demonstrate the two faces of auditory masking, could take an important place in the picture of the hearing system selectivity.

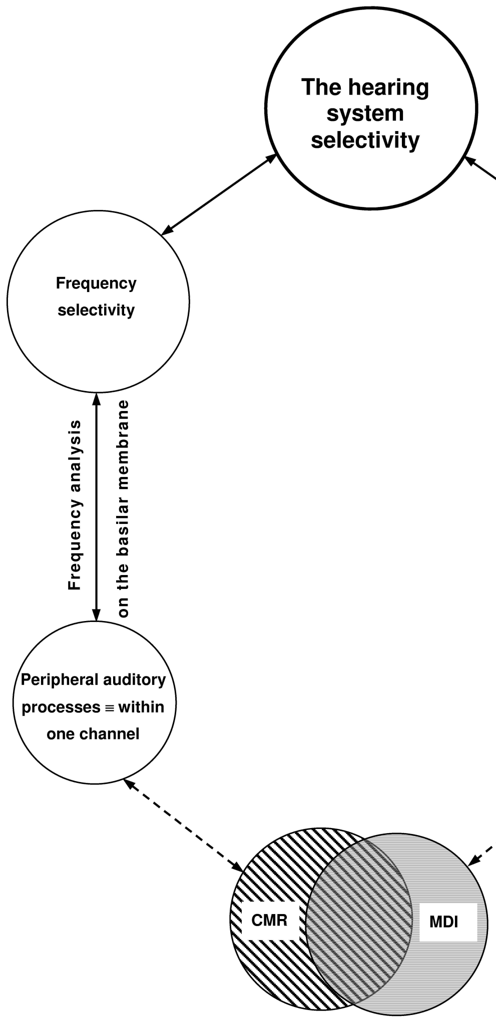

Fig. 3. The general picture of the hearing system selectivity.

The relative importance of within-channel and across-channel processes in CMR and MDI may shed light on the peripheral versus central nature of these processes contribution of whom is irremissible for understanding the selectivity of the hearing system. The overlapping of the graphs — CMR and MDI — reflects a delicate balance between mechanisms producing MDI and those producing CMR [12, 8].

Ackowledgements

Preparation of this paper was partially supported by KBN. Grant 1H01F 08012. I wish to thank the anonymous reviewer for his friendly constructive remarks.

References

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

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

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

[4] 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).

[5] U. Jorasz, *Selektywność układu słuchowego*, Wydawnictwo Naukowe UAM, Poznań 1999.

[6] U. Jorasz, *Wykłady z psychoakustyki*, Wydawnictwo Naukowe UAM, Poznań 1998.

[7] U. Jorasz and B.C.J. Moore, *Peripheral and central processes in hearing*, Archives of Acoustics, **25**, 4, 433–439 (2000).

[8] U. Jorasz and B.C.J. Moore, *The two faces of auditory masking*, ibid. 423–431 (2000).

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

[10] B.C.J. Moore, *An introduction to the psychology of hearing* (4th ed.), London 1997.

[11] B.C.J. Moore [Ed.], *Hearing*, London 1995.

[12] B.C.J. Moore and U. Jorasz, *Detection of changes in modulation depth of a target sound in the presence of other modulated sounds*, J. Acoust. Soc. Am., **91**, 1051–1061 (1992).

[13] B.C.J. Moore and U. Jorasz, *Modulation discrimination interference and comodulation masking release as a function of the number and spectral placement of narrow-band noise modulators*, J. Acoust. Soc. Am., **100**, 2373–2381 (1996).

[14] D. Patterson and B.C.J. Moore, *Auditory filters and excitation patterns as representations of frequency resolution*, in B.C.J. Moore (ed.), Frequency Selectivity in Hearing, Academic, London 1986.

[15] J.L. Verhey, T. Dau and B. Kollmeier, *Within-channel cues in comodulation masking release (CMR): Experiments and model predictions using a modulation-filterbank model*, J. Acoust. Soc. Am., **106**, 2733–2745 (1999).

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

[17] E. Zwicker, *Masking and psychological excitation as consequences of the ear's frequency analysis*, [in:] Frequency Analysis and Periodicity Detection in Hearing, R. Plomp, G.F. Smoorenburg [Eds.], Sijthoff, Leiden 1970.