

# Interaction of Object Binding Cues in Binaural Masking Pattern Experiments

Jesko L.Verhey, Björn Lübken and Steven van de Par

**Abstract** Object binding cues such as binaural and across-frequency modulation cues are likely to be used by the auditory system to separate sounds from different sources in complex auditory scenes. The present study investigates the interaction of these cues in a binaural masking pattern paradigm where a sinusoidal target is masked by a narrowband noise. It was hypothesised that beating between signal and masker may contribute to signal detection when signal and masker do not spectrally overlap but that this cue could not be used in combination with interaural cues. To test this hypothesis an additional sinusoidal interferer was added to the noise masker with a lower frequency than the noise whereas the target had a higher frequency than the noise. Thresholds increase when the interferer is added. This effect is largest when the spectral interferer-masker and masker-target distances are equal. The result supports the hypothesis that modulation cues contribute to signal detection in the classical masking paradigm and that these are analysed with modulation bandpass filters. A monaural model including an across-frequency modulation process is presented that account for this effect. Interestingly, the interferer also affects dichotic thresholds indicating that modulation cues also play a role in binaural processing.

**Keywords** Binaural processing · Binaural masking level difference · Beating · Modulation processing · Across-frequency process · Modulation detection interference · Off-frequency masking · Object binding cues · Modulation filterbank · Auditory modelling

---

J. L.Verhey (✉) · B. Lübken

Department of Experimental Audiology, Otto von Guericke University Magdeburg,  
Leipziger Straße 44, 39120 Magdeburg, Germany  
e-mail: jesko.verhey@med.ovgu.de

S. van de Par

Acoustics group, Cluster of Excellence “Hearing4All”, Carl von Ossietzky Universität  
Oldenburg, 26111 Oldenburg, Germany  
e-mail: steven.van.de.par@uni-oldenburg.de

© The Author(s) 2016

P. van Dijk et al. (eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, Advances in Experimental Medicine and Biology 894,  
DOI 10.1007/978-3-319-25474-6\_26

## 1 Introduction

In complex acoustical environments, the auditory system makes use of various cues to group sound components from one particular sound source. These cues, sometimes referred to as object binding cues, are often associated with characteristics of the sound such as coherent level fluctuations in different frequency regions (as, e.g., observed in speech), or the location of the sources in space. Differences in location result, among others, in interaural differences. Psychoacousticians developed experiments to study how the human auditory system makes use of the information. For example, the ability to use interaural disparities is shown in experiments on binaural masking level differences (BMLD). The BMLD is a well investigated phenomenon that signifies the reduction in masked thresholds when a tonal signal and masker have disparate interaural cues. For a diotic narrowband masker, a tonal target can be detected at a 25 dB lower level when the interaural target phase is changed from 0 to  $\pi$  (e.g. van de Par and Kohlrausch 1999). This large BMLD can only be observed when masker and target are spectrally overlapping. Several studies have shown that the BMLD reduces to rather small values for off-frequency targets (e.g., van de Par et al. 2012).

This lack of a BMLD could be interpreted as a lack of effective binaural processing in off-frequency conditions. Alternatively, it could result from monaural processing which is much better in off-frequency as compared to on-frequency conditions. The narrowband noise masker has an envelope spectrum that predominantly has components in the range from 0 Hz to the masker bandwidth. Thus, the presence of an on-frequency tonal target signal does not substantially change this envelope spectrum. When, however, the tonal signal is presented off-frequency, new, higher-frequency spectral modulation components are introduced due to the beating between tone and masker that are only masked to a small degree by the inherent masker fluctuations. This modulation cue provides for an additional lowering of off-frequency monaural detection thresholds in excess of the peripheral filtering effect. The modulation frequency selectivity of the auditory system shown in previous studies (e.g., Dau et al. 1997) indicates that the inherent modulations can be dissociated from those due to beating. The contribution of modulation cues in off-frequency masking has been demonstrated by Nitschmann and Verhey (2012) using a modulation filterbank model.

In a previous study (van de Par et al. 2012), this monaural explanation was tested by attempting to reduce the effectiveness of the modulation cues by creating a Modulation Detection Interference (MDI) effect. To this end a tonal interferer was provided at the opposite spectral side of the masker compared to where the target was, but with the same spectral distance between target and masker as between masker and interferer. This extra interferer created beating frequencies that were equal to those that should be created by the presence of the target tone. Although these beating frequencies are then present in different auditory channels, the across-frequency nature of MDI (Yost et al. 1989) implies that they should still interfere with the modulation cues used for detecting the target tone. The level of the interfering tone

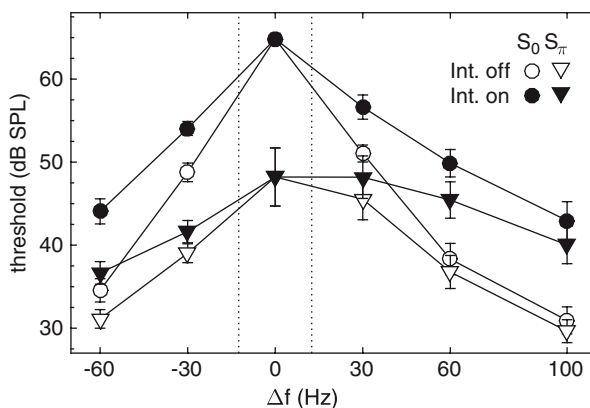
was chosen to be 6 dB lower than the masker level to ensure that no energetic masking should occur. Figure 1 shows the data of van de Par et al. (2012) as masking patterns with (filled symbols) and without (open symbols) interferer for the diotic (circles) and the dichotic (triangles) conditions. As hypothesized the presence of the interferer increased the diotic thresholds substantially and had less effect on the dichotic thresholds.

In the present contribution, we investigated two aspects of the interference effect described above. First it is known that MDI exhibits some degree of frequency selectivity (Yost et al. 1989). By altering the frequency offset between the interferer tone and the narrowband masker, the interference on the fixed sinusoidal signal should change, giving the largest interference effect when the frequency differences between the target and interferer are the same relative to the masker centre frequency. Secondly, we will investigate to what extent the interference effect can be modelled.

## 2 Methods

### 2.1 Psychoacoustic Experiment

Three masking conditions were considered. The masker of the first masking condition was a 25-Hz wide bandpass-filtered Gaussian noise with an overall level of 65 dB SPL generated using a brick-wall filter in the frequency domain. The arithmetic centre frequency  $f_c$  was 700 Hz. Apart from this noise-only condition, two other conditions were used, where an additional pure tone (interferer) was presented



**Fig. 1** Masked thresholds as a function of the spectral distance between target signal and masker centre frequency. Open symbols indicate data for the classical masking pattern experiment with a narrowband diotic noise masker and a sinusoidal signal. Filled symbols indicate data of a modified masking pattern experiment with an additional tonal interferer. Circles denote thresholds for the diotic condition, triangles those of the dichotic condition. The data were taken from our previous study on the effect of an interferer on masking patterns (van de Par et al. 2012)

at 670 Hz (i.e., 30 Hz below  $f_c$ ) or 600 Hz (i.e., 100 Hz below  $f_c$ ). All masker and interferer components were presented diotically. The target was a pure tone that was either presented diotically ( $S_0$ ) or interaurally out of phase ( $S_\pi$ ). The target frequency was either 730 Hz (30 Hz above  $f_c$ ) or 800 Hz (100 Hz above  $f_c$ ). The duration of masker, interferer and target was 500 ms including 50-ms raised cosine-ramps at onset and offset.

Masked thresholds were measured using an adaptive three-interval three-alternative forced-choice procedure. In each of the three 500-ms stimulus intervals of a trial the masker was presented. Stimulus intervals were separated by 500 ms of silence. In one randomly chosen interval of a trial, the target signal was added. The subject had to indicate which interval contained the target. Feedback about correctness of the answer was provided. A two-down one-up adaptive staircase procedure was used. The initial step size of 6 dB was reduced to 3 dB after the first upper reversal and to 1 dB after the second upper reversal. With this minimum step size, the run continued for another eight reversals of which the mean was used as threshold estimate. Four threshold estimates were obtained and averaged leading to the final threshold estimate for this signal and masking condition. The conditions were run in random order.

The stimuli were generated digitally with a sampling rate of 44100 Hz, converted to analogue signals using a 32 bit D/A converter (RME Fireface UC) and presented via headphones (Sennheiser HD650) within a sound insulated booth. Eight experienced subjects (Five male and three female) aged from 21 to 30 years participated. All subjects had pure tone thresholds of 10 dB HL or lower for the standard audiogram frequencies.

## 2.2 Simulations

For the simulations a model was developed based on the model used by Nitschmann and Verhey (2012) to predict their masking patterns. The monaural front end was (i) a first-order band-pass filter with cut-off frequencies of 0.5 and 5.3 kHz to simulate the filtering of the outer and middle ear, (ii) a bank with fourth-order gammatone filters (centre frequencies 313–1055 Hz) with a bandwidth of one equivalent rectangular bandwidth (ERB) and a spectral distance of adjacent filters of one ERB, (iii) white noise with a level of 7 dB SPL added to the output of each filter to simulate thresholds in quiet, (iv) half-wave rectification and a first order low-pass filter with a cut-off frequency of 1 kHz and (v) five consecutive adaptation loops to model adaptation and compression effects in the auditory system. In the monaural pathway the output of the adaptation loops was analysed by a modulation filterbank, where the highest centre frequency of the modulation filters was restricted to a quarter of the centre frequency of the auditory filter. To simulate across-channel (AC) processes, two different versions of the monaural part of the model were used.

The first version is based on the approach of Piechowiak et al. (2007). In this approach a weighted sum of modulation filters at the output of off-frequency auditory filters was subtracted from the corresponding modulation filters at the output

of the on-frequency filter. In contrast to Piechowiak et al. (2007), a multi-auditory-channel version of this approach was used. Since the weighting is essentially based on the energy in the auditory filters it will be referred to as the energy-based AC model in the following.

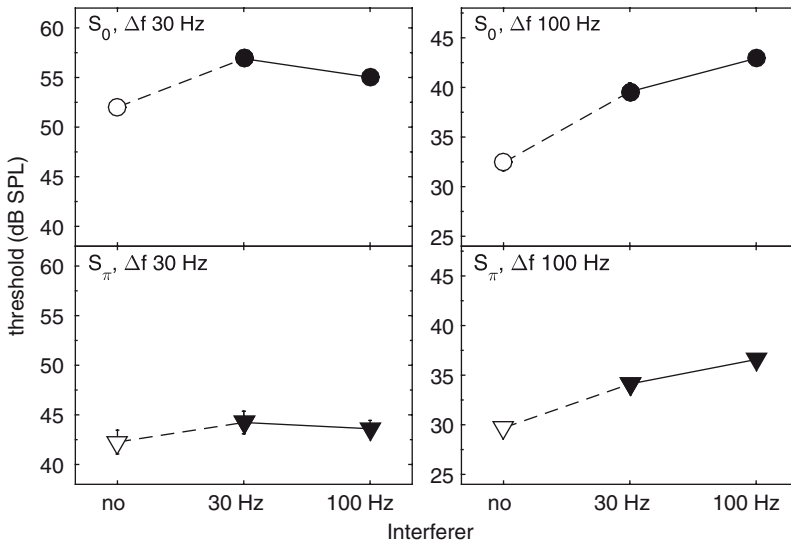
The second version of across-channel processing is based on the assumption that modulation in off-frequency auditory filters channels hampers the processing of the modulation filters tuned to the corresponding modulation frequency in the on-frequency auditory channel. This detrimental effect is realised by assuming that modulation frequencies adjacent to the modulation created by the masker-interferer distance could not be processed. This was implemented in an extreme form by setting the output of the modulation filters tuned to frequencies adjacent to the interferer-masker distance induced modulation frequency to zero. This approach is referred to as the modulation-based AC model.

The binaural pathway of the model is the same in both models. It is realised as an equalisation-cancellation process followed by a low-pass filtering with a first-order low-pass filter with a cut-off frequency of 8 Hz. The final stage is an optimal detector, which calculates the cross correlation between a temporal representation of the signal at a supra-threshold level (template) with the respective actual activity pattern (Dau et al. 1997). For the simulations the same adaptive procedure was used as for the psychoacoustic experiment with the model as an artificial observer. Simulated thresholds are calculated as the mean of the estimates from 24 runs.

### 3 Results

Figure 2 shows average thresholds and standard errors of the eight subjects participating in the psychoacoustic experiment. Each panel shows thresholds of one signal as indicated in the top of the panel. Thresholds are shown for the three interferer conditions: “no interferer”, “interferer 30 Hz below  $f_c$ ” and “interferer 100 Hz below  $f_c$ ”. For the diotic 730-Hz target (top left panel), thresholds increase by 4.5 dB when an interferer was added at a distance of 30 Hz. For the spectral distance interferer-masker of 100 Hz, the threshold was 1.5 dB lower than for the spectral distance interferer-masker of 30 Hz. For the diotic 800 Hz target (top right panel), adding an interferer had a higher impact on threshold than for the lower signal frequency, as shown in Fig. 1. Threshold was highest when the spectral distance masker-target and interferer-masker threshold was the same. It was 4.5 dB lower when the interferer was positioned 30 Hz instead of 100 Hz below the masker. A similar trend was observed for the dichotic 800-Hz target (bottom right panel) although the difference between the highest threshold for an interferer at 100 Hz below the masker and the lowest threshold without interferer was only 8 dB whereas it was 11.5 dB for the diotic 800-Hz target. For the dichotic 730-Hz target (bottom left panel) thresholds were about the same for the two conditions with interferer. The threshold without interferer was 1.5 dB lower than thresholds with interferer.

Figure 3 shows model predictions. Since the binaural pathway is the same in both models, predictions of the two models do not differ for the dichotic target sig-

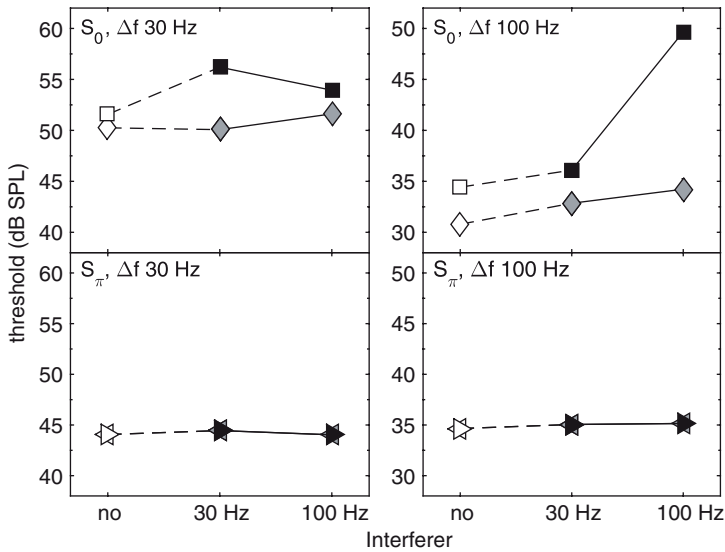


**Fig. 2** Measured thresholds for three different interferer conditions: without interferer (no) and with interferers at 30 or 100 Hz *below* the masker. Each panel shows the average thresholds and standard errors of the eight subjects for one target signal. The signal frequency was either 730 Hz (masker-target  $\Delta f = 30$  Hz, *left* panels) or 800 Hz ( $\Delta f = 100$  Hz, *right* panels). The target was either presented diotically (*top* panels) or dichotically (*bottom* panels)

nals (bottom panels). Both models predict no effect of the interferer on threshold, in contrast to the data. For the diotic 730-Hz target (top left panel), the energy-based AC model predicted no change in threshold when the interferer was added. In contrast, the modulation-based AC model predicted an increase of about 5 dB when the interferer was added 30 Hz below the masker compared to the condition without interferer. For an interferer 100 Hz below the masker, threshold was about 2 dB lower than for the interferer 30 Hz below the masker. Thus, only the modulation-based AC model predicts the effect of the interferer on thresholds for the diotic 730-Hz target. For diotic 800-Hz target, both models predicted highest thresholds for the interferer 100 Hz below the masker and lowest for the condition without interferer. The energy-based AC model predicted an effect of 4.5 dB, i.e., considerably smaller than the 11.5 dB observed in the data. The modulation-based model predicted an effect of 14 dB. For the interferer 30 Hz below the masker, the predicted threshold was similar to the thresholds without interferer, whereas the difference between these two thresholds was 7 dB in the measured data.

## 4 Discussion

The behavioural results of the present study support the hypothesis that modulation cues affect detection of a diotic target in masking pattern experiments and that an additional interfering tone with the same distance below (or above) the masker as



**Fig. 3** Same as Fig. 2 but now showing simulated data. *Black squares* and *rightward-pointing triangles* indicate predictions of the modulation-based AC model. *Grey diamonds* and *leftward-pointing triangles* show the predictions of the energy-based AC model

the spectral distance of the signal spectrally above (or below) the masker makes the beating cue between signal and masker less usable for detection. The present study shows that the same spectral distance masker-target and interferer-masker increases thresholds more than an interferer at a different spectral distance. This result points towards a modulation spectral selectivity as previously observed in MDI experiments.

The present study tested an extended (energy-based AC) model of the modulation filterbank model originally designed to predict CMR by Piechowiak et al. (2007). Although this model predicted the trend for the 800-Hz target the effect was too small. For the 730-Hz target, no effect was predicted, in contrast to the data. Thus this energy-based AC model was unable to predict the effect of the interferer on the diotic thresholds.

The second model assumed that modulation cues with frequencies adjacent to the beating frequency between interferer and masker could not be used for signal detection. This was realised in the modulation-based AC model by setting the output of the corresponding modulation filters to zero. This approach predicted general trends of the diotic data of the present study. Among others it predicted a larger effect of the interferer for the 800-Hz signal than for the 730-Hz target, in agreement with the data. However, there was a tendency of an overestimation of the effect for the same spectral distance masker-signal and interferer-masker and an underestimation of the effect of the interferer with a spectral distance not equal to that of signal and masker, at least for the 800-Hz target. This indicates that the present approach was too simplistic, i.e., modulation cues may only be reduced and not abolished by the presence of the interferer. In addition, not only modulation frequencies close to the beating frequency may

be affected by the presence of the interferer but also other modulation frequencies. This is also consistent with MDI results where an MDI is also observed for modulation frequencies of the interferer that differ from the target modulation frequency (Yost et al. 1989). A better match between data and simulations could be achieved by assuming weights to the output of the modulation filters that gradually increase as the distance between beating frequency (masker—interferer) increases.

As already shown in van de Par (2012), dichotic thresholds are also affected by the presence of an interferer. The present data for the dichotic 800-Hz target indicate a similar effect of interferer-masker distance as for the diotic target, although less pronounced. Thus, also dichotic thresholds may be influenced by a frequency-specific modulation processing. This was already hypothesized in Nitschmann and Verhey (2012). Binaural modulation frequency selectivity seems to be less accurate than monaural modulation frequency selectivity (Thompson and Dau 2008). Thus, if binaural modulation frequency selectivity affects the dichotic off-frequency masked thresholds the mismatch of beating frequency between masker-target and interferer-masker should have less effect on threshold than for the diotic condition. The present data for the 800-Hz target support this hypothesis.

**Acknowledgments** This work was supported by the Deutsche Forschungsgemeinschaft (DFG, sfb trr 31).

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (<http://creativecommons.org/licenses/by-nc/2.5/>) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

## References

- Dau T, Kollmeier B, Kohlrausch A (1997) Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J Acoust Soc Am* 102(5):2892–2905
- Nitschmann M, Verhey JL (2012) Modulation cues influence binaural masking-level difference in masking-pattern experiments. *J Acoust Soc Am Exp Lett* 131(3):EL223–EL228
- Piechowiak T, Ewert SD, Dau T (2007) Modeling comodulation masking release using an equalization-cancellation mechanism. *J Acoust Soc Am* 121(4):2111–2126
- Thompson ER, Dau T (2008) Binaural processing of modulated interaural level differences. *J Acoust Soc Am* 123(2):1017–1029
- van de Par S, Kohlrausch A (1999) Dependence of binaural masking level differences on center frequency, masker bandwidth and interaural parameters. *J Acoust Soc Am* 106(4):1940–1947
- van de Par S, Luebken B, Verhey JL, Kohlrausch A (2012) Off-frequency BMLD: the role of monaural processing. In: Moore BCJ, Patterson RD, Winter IM, Carlyon RP, Gockel HE (eds) *Basic aspects of hearing: physiology and perception*. Springer, New York, pp 293–301
- Yost WA, Sheft S, Opie J (1989) Modulation interference in detection and discrimination of amplitude modulation. *J Acoust Soc Am* 86:2138–2147