Downloaded from orbit.dtu.dk on: Dec 10, 2024

DTU Library

Physiological correlates of masking release

Kim, Hyojin; Epp, Bastian

Published in:
Proceedings of the International Symposium on Auditory and Audiological Research

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Kim, H., & Epp, B. (2020). Physiological correlates of masking release. In *Proceedings of the International Symposium on Auditory and Audiological Research: Auditory Learning in Biological and Artificial Systems* (Vol. 7, pp. 453-460). The Danavox Jubilee Foundation.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Physiological correlates of masking release

HYOJIN KIM[1,*] AND BASTIAN EPP[1]

[1] *Hearing Systems Section, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Masking release is one example of auditory object segregation where the masked threshold of a target sound decreases in the presence of beneficial cues. Two such cues are comodulation and interaural phase disparity (IPD) underlying the phenomena of comodulation masking release (CMR) and binaural masking level difference (BMLD) respectively. While the effect of these cues have been shown in behavioral studies, little is known about the underlying physiological mechanisms of masking release. In this study, we postulated an ”internal signal-to-noise ratio (iSNR)” that reflects neuronal representation of a masked tone. As the proxy for iSNR, we investigated the applicability of late auditory evoked potentials (LAEPs). We added an onset asynchrony cue with comodulation and IPD cues. Results showed that onset asynchrony had a negative effect on CMR while it did not affect BMLD. The P2 component of the vertex LAEPs was suggested to be an objective measure of iSNR. This will provide us information about whether temporal contexts affect the neuronal representation of CMR and BMLD at the level of the auditory cortex.

INTRODUCTION

Our auditory system has the remarkable ability of sound object segregation. In a simple case, where S is a tone and M is a masker, the task for our auditory system can be defined as the detection of a masked-tone by separating the tone (S) from the masker (M). According to the power-spectrum model of masking (e.g. Fletcher, 1940), the detection threshold is correlated to a certain constant signal-to-noise ratio ($k = S/M$) that is based on the physical intensity of the stimulus. This model cannot explain a masking release where the detection threshold decreases by adding cues to the stimulus with identical power spectra. To account for the effect of beneficial cues without changes in the power spectrum, this model can be reformulated in terms of an internal signal-to-noise ratio (iSNR) at the cortex level. As a masked tone ($S+M$) is transmitted through the auditory system, it will have a neuronal representation of $S_i + M_i$ at the cortex level.

$$S+M \rightarrow S_i+M_i \rightarrow \textit{behavioral measures}$$

Assuming that there exists a mapping between internal representations ($S_i + M_i$) and behavioral measures (e.g. audibility measure, masked thresholds), it is possible to

*Corresponding author: hykim@dtu.dk

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 7: Auditory Learning in Biological and Artificial Systems, August 2019, Nyborg, Denmark. Edited by A. Kressner, J. Regev, J. C.-Dalsgaard, L. Tranebjærg, S. Santurette, and T. Dau. The Danavox Jubilee Foundation, 2019. ©The Authors. ISSN: 2596-5522.

predict the behavioral outcome from the internal representation. To achieve this, an objective measure of iSNR is required. Masking release has been investigated on various levels. At the single-cell level, correlates of masking release induced by comodulation (comodulation masking release, CMR) was found in the cochlear nucleus (CN) (Pressnitzer *et al.*, 2001). The neural responses indicated the suppression of a comodulated masker and the enhancement of a tone at the first neural stage after the cochlea. Masking release induced by interaural phase difference (binaural masking level difference, BMLD) was suggested to be processed at the inferior colliculus (IC) (Jiang *et al.*, 1997) that is located upstream of the CN. At the cortical level, Epp et al. (2013) investigated the neural representation of a masked tone with comodulation and IPD cues using EEG. This result suggested that the neuronal representations of these cues are combined at the level of the auditory cortex, supporting the idea of bottom up processing and a superposition of masking release. They found that the late auditory evoked potential P2 can be an objective measure of the audibility of the stimulus (iSNR). The amplitude of P2 was correlated with the individual level above masked threshold rather than to the physical signal-to-noise ratio (SNR) of the stimulus (Epp *et al.*, 2013). Contrary to this idea, CMR was found to be reduced by the streaming effect, which is assumed to be a higher-level process (Dau *et al.*, 2005; Grose *et al.*, 2009). This has only been shown by psychoacoustical experiments measuring masked thresholds. Only few studies investigated with physiological experiments (e.g. the mismatch negativity (MMN)), however, no study has found a neural correlate of the streaming effect on masking release (Verhey *et al.*, 2012). In addition, there is no study regarding the effect of streaming on BMLD. Hence, it remains unclear how streaming affects combined CMR and BMLD (the streaming effect on masking release). We postulate that the streaming effect is related to temporal information processed at the level of the CN. As the IPD cue is likely processed at the level of the IC, we hypothesize that BMLD will not be affected by streaming (Figure 1). In this study, we investigate whether the streaming effect is: i) a result of bottom up processing; ii) merely additional neuronal processing after summation of neuronal representation of CMR and BMLD.

As an extension of the study by Epp et al. (2013), we added onset asynchrony as a grouping cue to induce the streaming effect. We first measured masked thresholds for each condition, $TH_m([condition])$. The level of the tone was set to the same level above the individual masked threshold (e.g. 15dB + $TH_m([condition])$) with a fixed level of the noise. If case i) is true, we hypothesize that the P2 amplitude will be the same for all conditions where the level of the tone was set to the same level above masked threshold in each condition (Figure 2a). The same P2 amplitude indicates the same iSNR for all conditions despite different SNRs of the stimulus for each condition. If the second case ii) is true, however, the P2 amplitude measured at masked thresholds will be higher as the detrimental effect of the streaming effect would require a higher level of neuronal representation of CMR to achieve the same audibility (Figure 2b).

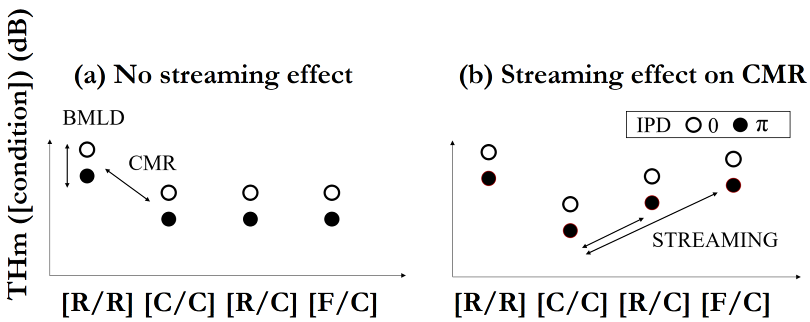

Fig. 1: Hypotheses about psychoacoustic experiment results. (a) CMR and BMLD are independent of the temporal context of the masker bands. (b) CMR are reduced as flanking bands (FBs) and signal-centred bands (SCB) are grouped separately by the onset asynchrony cue (the streaming effect) while there is no streaming effect on BMLD

METHODS

The same stimuli were used for the behavioural and the electrophysiological experiment (Figure 3). Masker conditions were designed based on Grose *et al.* (2009). For each condition, different first masker (Masker 1) and second masker (Masker 2) intervals were used. Four different masker conditions were used (Figure 4): (a) [R/R]: both masker intervals had uncorrelated envelope fluctuations; (b) [C/C]: both masker intervals had comodulated envelopes; (c) [R/C]: the first masker interval had uncorrelated envelope fluctuations and the second masker interval had comodulated envelope fluctuations; (d) [F/C]: the first masker interval had comodulated flanking bands (FBs) and the second masker interval had comodulated FBs and signal-centred bands (SCB). The masker consisted of five narrow-band noises with a width of 20 Hz, centered at 460 Hz, 580 Hz, 820 Hz, 940 Hz (FBs) and at 700 Hz (SCB). The bandwidth and center frequency of each noise band were chosen to maximize CMR (Grose *et al.*, 2009). The total duration of the signal was 700 ms including 20 ms on- and offset ramps. The first masker interval was gated on for 500 ms followed by a second masker of duration 200 ms. The target tone had a frequency of 700 Hz and was presented with the second masker interval. The target tone had an interaural phase difference of either 0 or π. The stimuli were digitally generated with a sampling frequency of 48 kHz.

The stimuli were presented using ER-2 headphones. For the psychoacoustical experiment, a modular framwork for running psychoacoustic experiments and computational perception models (AFC) software package for MATLAB was used (Ewert, 2013). An adaptive and three-alternative forced choice procedure was used with a one-up, two-down rule (Levitt, 1971). The listener was asked to choose the interval with the tone. During the EEG experiment, the stimuli were presented while participants were

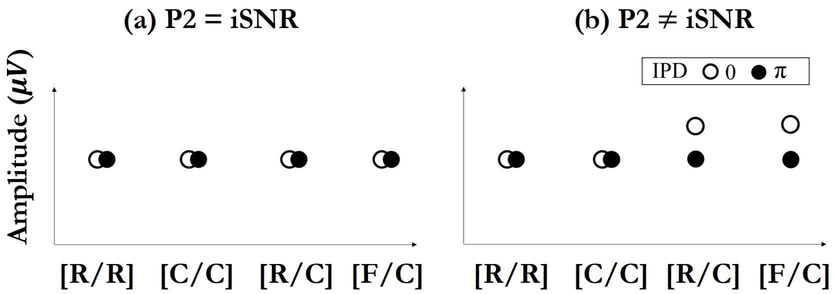

Fig. 2: Hypotheses about electrophysiological experiment results. (a) The P2 component of the LAEP reflects iSNR. (b) The P2 component of the LAEP cannot reflect iSNR and only reflect the summed neuronal representation of CMR and BMLD.

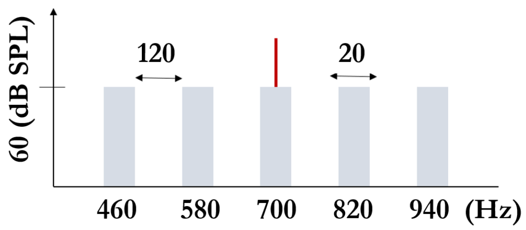

Fig. 3: Spectra of the stimulus. A target tone (700 Hz) was presented with a set of narrow-band masker bands: One signal centered band (SCB) and four flanking bands (FBs). Thresholds were measured individually and used to adjust the levels for the EEG experiment to set equal levels above masked threshold.

watching a silent movie. Late auditory evoked potentials (LAEPs) were measured using a 144 channel EEG amplifier (g.Tec HiAmp research) with active electrodes. A conductive gel was used to reduce the impedance of electrodes. Electrodes with an impedance higher than 10 kΩ were excluded from the analysis. The reference electrode was placed close to the mastoid (P8) and the region of interest was the central position (Cz). The data analysis was performed using FieldTrip (Oostenveld *et al.*, 2011). The EEG data were partitioned into epochs from -300 to 1200 ms relative to the onset of the masker. Each epoch was low pass filtered with a cut-off frequency of 20 Hz. Detrending, base line correction and weighted averaging (Riedel *et al.*, 2001) were applied to increase the signal-to-noise ratio. Trials containing signals exceeding 100 μV in any channel were rejected as artifacts.

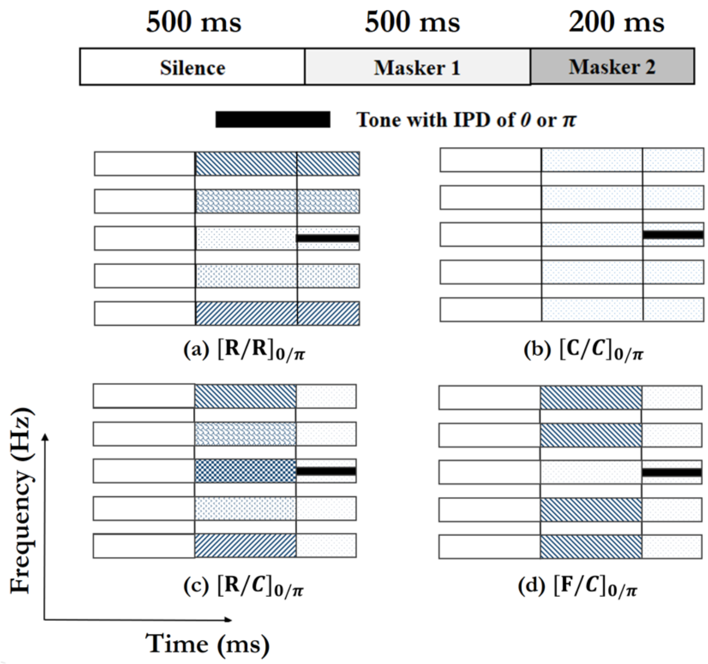

Fig. 4: Schematic spectrograms of the stimulus. Each block represents a noise band. The thick black line represents a tone. Tone is presented either with IPD of 0 or π. RAN: All noise bands have random envelope fluctuations. COM: All noise bands have the same envelope fluctuations (comodulated). FCOM: Only the flanking bands are comodulated.

RESULTS AND DISCUSSION

Comodulation masking release, CMR

$$CMR_* = THm([R/R]) - THm([*/C]) \quad \text{(Eq. 1)}$$

Here, * stands for one of three masker types (RAN, FCOM, COM).

Figure 5 shows the results of the psychoacoustic experiment. In the diotic condition (Figure 5, circle), CMR_C was observed for all listeners. The CMR_R and CMR_F are smaller than CMR_C. In the streaming conditions ([R/C], [FC/C]), we postulate that the auditory system grouped masker bands into separate objects due to their uncorrelated intensity fluctuations during the first masker (Masker 1). Reduced CMR indicates that the comodulation cue is not beneficial when masker bands were separated before the comodulation cue is provided.

Binaural masking level difference, BMLD

$$BMLD_{masker} = THm([condition]_0) - THm([condition])_\pi \quad \text{(Eq. 2)}$$

When an IPD of π was introduced, BMLD was observed for all conditions (Figure 5, cross). BMLD was almost constant, except for the [F/C] condition. For participant A - C, $BMLD_{F/C}$ was larger than in the other conditions. This might suggest that there might be individual differences in the effects of temporal contexts to BMLD. Therefore, it remains to be seen which pattern of results is dominant over a larger cohort of listeners.

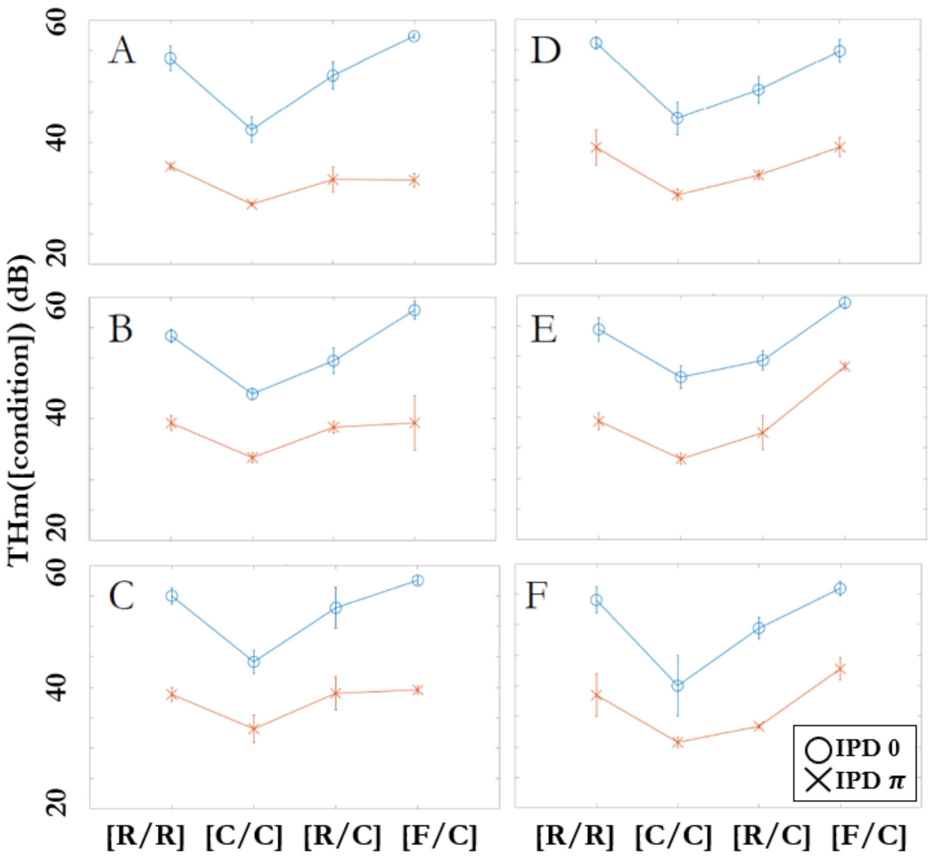

Fig. 5: Masked threshold for the different stimulus conditions for diotic (circle) and dichotic (cross) presentation of the tone. Thresholds measurements were repeated three times.

Onset response

To check for the presence of an onset response evoked by the onset of the different masker intervals, event-related potentials (ERPs) for the conditions were analyzed. Figure 6 and 7 show the grand average ERPs for the conditions with two identical masker intervals ([R/R] and [C/C]) and two different masker intervals ([R/C] and

[F/C]), respectively. Both masker groups evoked a typical onset response at the start of the first masker, caused by the change in stimulus energy. The onset response decayed after about 400 ms, which shows a suitable scaling of the masker-only time interval before the onset of the tone. No such onset was found at the transition from the first to the second masker interval (vertical line). These results also confirm the suitability of the masker design to study the response to the tone independent of the masker onset response in the first masker interval.

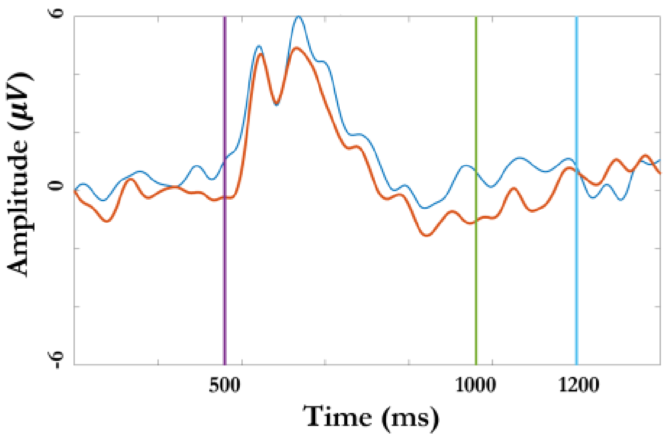

Fig. 6: Average ERPs of [R/R] (thin line) and [C/C] (thick line) conditions. Masker 1 onset (500 ms), Masker 2 onset (1000 ms) and offset (1200 ms).

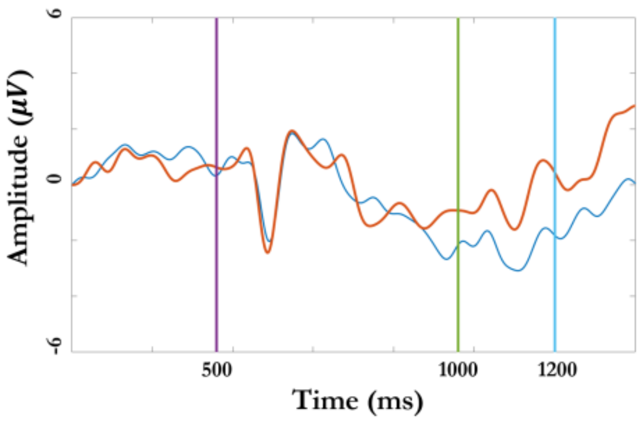

Fig. 7: Average ERPs of [R/C] (thin line) and [F/C] (thick line) conditions. Masker 1 onset (500 ms), Masker 2 onset (1000 ms) and offset (1200 ms).

CONCLUSION

In this study, we designed the stimuli to investigate the streaming effect on CMR and BMLD. The effect of streaming on CMR was observed as shown in previous studies. There was no streaming effect on BMLD in all conditions except [F/C] condition where there were individual differences. To avoid overlapping between evoked potentials induced by Masker 1 and Masker 2, Masker 1 needs to be 500 ms in length. The preliminary data confirms the applicability of the design. Additional data will provide more conclusive results for the effect of streaming on BMLD and correlations between psychophysics and electro-physiology.

REFERENCES

Dau, T., Ewert, S. D., and Oxenham, A. J. (**2005**), "Effects of concurrent and sequential streaming in comodulation masking release," Auditory signal processing (Springer), 334–342.

Epp, B., Yasin, I., and Verhey, J. L. (**2013**), "Objective measures of binaural masking level differences and comodulation masking release based on late auditory evoked potentials," Hear. Res., **306**, 21–28.

Ewert, S. D. (**2013**), "AFC—A modular framework for running psychoacoustic experiments and computational perception models," Proc. AIA-DAGA, 1326–1329.

Fletcher, H. (**1940**), "Auditory patterns," Rev. Mod. Phys., **12**(1), 47.

Grose, J. H., Buss, E., and Hall III, J. W. (**2009**), "Within-and across-channel factors in the multiband comodulation masking release paradigm," J. Acoust. Soc. Am., **125**(1), 282–293.

Jiang, D., McAlpine, D., and Palmer, A. R. (**1997**), "Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions," J Neurophysiol., **77**(6), 3085–3106.

Levitt, H. (**1971**), "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am., **49**(2B), 467–477.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (**2011**), "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," Comput. Intel. Neuroscience., **2011**, 1–9.

Pressnitzer, D., Meddis, R., Delahaye, R., and Winter, I. M. (**2001**), "Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus," J. Neurosci., **21**(16), 6377–6386.

Riedel, H., Granzow, M., and Kollmeier, B. (**2001**), "Single-sweep-based methods to improve the quality of auditory brain stem responses Part II: Averaging methods," Zeitschrift fur Audiologie, **40**(2), 62–85.

Verhey, J. L., Ernst, S. M., and Yasin, I. (**2012**), "Effects of sequential streaming on auditory masking using psychoacoustics and auditory evoked potentials," Hear. Res., **285**(1-2), 77–85.