The effect of the preceding masking noise on monaural and binaural release from masking

Hyojin Kim[1]*, Bastian Epp[1]

[1]Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark

*hykim@dtu.dk

ABSTRACT

The auditory system uses various signal properties to separate a target signal from background noise. When a target tone is preceded by a noise, the threshold for target detection can be increased or decreased depending on the type of a preceding masker. The effect of the preceding masker on the following sound can be interpreted as stream formation. The effect of stream formation is assumed to be either the result of adaptation at a low-level or high-level auditory processing. In an attempt to disentangle these, we investigated the time constant of the underlying process of adaptation by varying the length of the preceding masker. We designed stimuli consisting of the preceding masker and the following masked tone. Each stimulus induces various stream formation, affecting following target detection or masking release. Target tone was presented in comodulated masking noise and with interaural phase difference (IPD), inducing comodulation masking release (CMR) and binaural masking level difference (BMLD), respectively. We measured CMR and BMLD when the length of preceding maskers varied from 0 (no preceding masker) to 500 ms. We postulate that if the adaptation is dominated by high-level auditory processing, both CMR and BMLD will be affected by an increase in the length of the preceding masker. Results showed that CMR was more affected with longer preceding maskers from 100 ms to 500 ms compared to shorter maskers. On the contrary, the preceding masker did not affect the BMLD. Based on the results, we suggest that the adaptation to a preceding masking sound may arise from a low-level (e.g., cochlear nucleus, CN) rather than the temporal integration by the high-level auditory processing.

1. Introduction

The ability to listen to a particular talker while other people speak simultaneously is one of the most outstanding abilities of the auditory system of humans. The detrimental impact of a hearing loss on this ability is often referred to as the "cocktail party problem" [Cherry, 1953]. One approach to explain this ability is known as "auditory scene analysis," suggesting that our auditory system perceives the acoustic environment as a combination of multiple auditory streams [Bregman, 1994]. Creating auditory streams enables us to segregate a target sound from the remaining sound, which is grouped into a masker. For stream formation, the auditory system makes use of signal properties or auditory cues extracted from the sounds reaching the two ears. This is analogous to the visual system that uses visual cues such as colors, textures, shapes to separate target object from other visual objects [Bregman, 1994]. Hence, a stream can be thought of as a basic unit of sound perception resulting from a perceptual grouping of auditory cues.

Perceptually, frequency components aligned closely in frequency or in time are more likely to be grouped into one stream than distant components [Bregman, 1994]. In a context-dependent case, one of the key principles of stream formation is that sound elements that come from the same sound source are likely to be grouped into the same stream [Moore and Gockel, 2012]. In addition, the information of sound elements can be accumulated in time, affecting following sound perception [Bregman, 1978, Anstis and Saida, 1985]. The temporal effect on the stream formation is referred to in various terms such as "build-up of stream segregation" and "adaptation to auditory streaming" [Moore and Gockel, 2012, Anstis and Saida, 1985].

As our vocal tract vibrates, it modulates all emitted frequency components due to its vibration [Raphael et al., 2007]. Such comodulation is a common property of natural sounds, suggesting that comodulation plays a role as a "grouping cue" for our auditory system [Nelken et al., 1999]. Experimentally, the benefit of comodulation has been shown by using a detection task [Hall et al., 1984, 1990]. The masked threshold of a tone decreases when the noise shows coherent intensity fluctuations across frequency compared to the condition where the tone is masked by a noise masker with uncorrelated intensity fluctuations across frequency. This enhancement in the detection performance can be quantified as a decrease in masked thresholds, often referred to as comodulation masking release (CMR) [Hall et al., 1984]. Based on the concept of stream formation, it can be assumed that comodulation supports the grouping of frequency components of the masker into one stream, thereby facilitating the segregation of the tone from the masker. Secondly, similar to comodulation, spatial information can induce a spatial masking release. When sounds reach the ears from a specific location in space, it induces differences in phase and level between the ears, which also can be beneficial for target detection [van de Par and Kohlrausch, 1999]. The decrease in masked thresholds induced by interaural disparities can be quantified as binaural masking level difference (BMLD) [Jeffress et al., 1956, van de Par and Kohlrausch, 1999]. In the study by Epp and Verhey [2009], they showed the superposition of CMR and BMLD where the overall masking release was close to the sum of CMR and BMLD. To shed light on a possible underlying mechanism, they implemented a conceptual model based on the assumption of serially-aligned processing of CMR and BMLD. However, the exact order of CMR processing and IPD processing could not be separated with their modeling approach.

Studies in both psychoacoustics and physiology have suggested that CMR can be interpreted as the combined effect of "primitive segregation" and "schema-driven segregation" [Bregman, 1994]. Primitive segregation can be understood as the auditory object arising by innate and direct processing of incoming acoustic input. Schema-driven processing can be understood that the auditory system uses stored knowledge of sound input on long time scales to form the auditory object [Moore and Gockel, 2012]. In physiological studies, neural correlates of CMR have been found at various stages along the auditory pathway. The earliest neuronal encoding was found in the cochlear nucleus (CN) as an increased neuronal response to a tone in comodulated noise compared to ones presented in uncorrelated noise [Pressnitzer et al., 2001, Neuert et al., 2004]. This enhanced neuronal response progressively sharpens along the auditory pathway passing the inferior colliculus (IC), medial geniculate body (MGB), and auditory cortex [Nelken et al., 1999, Las et al., 2005]. This suggests that CMR is induced by bottom-up processing (primitive segregation). Other studies propose, however, that CMR is the effect of high-order auditory processing (schema-driven segregation). When a tone in a comodulated masker is preceded and followed by another masker (temporal fringe), CMR decreases or increases depending on the type of temporal fringe maskers [Grose et al., 2009]. This implies that CMR can be affected by the stream formed before and after the target is presented. Dau et al. [2005] found that both pre-cursors and post-cursors could eliminate CMR. In their later study, they suggested that for the effect of pre-cursors, it is hard to exclude the effect of peripheral neural adaptation. Furthermore, the neural correlates of the influence of preceding maskers on CMR (e.g., reduced/increased CMR depending on temporal contexts) have also been found in A1 [Sollini and

Chadderton, 2016]. This study showed that a preceding stream is formed as the auditory system adapts to the sound. Nevertheless, it is still unknown whether neural representation at A1 level is the relayed neural encoding from the CN and IC (bottom-up), or there is cortical feedback to the sub-cortical processing of CMR (top-down) or an additional CMR processing occurs at A1 (high-level auditory processing). In the case of BMLD induced by IPD, physiological evidence suggests that IPD information is processed at the IC, which is located after the CN [Shackleton et al., 2005, 2003, Zohar et al., 2011]. Contrary to CMR, little has been found regarding the stream formation effect on BMLD.

The neural basis of the effect of the preceding masker on CMR and BMLD is not clarified yet. In this study, we try to contribute to provide more insights on this problem. By varying the duration of preceding maskers to disentangle the adaptation effect on masking release at a low-level and at high-level auditory processing. The underlying assumption is that the time constants inherent in auditory processing increase along the ascending auditory pathway. We investigated the effect of adaptation on CMR, and whether the same adaptation effect can be found in BMLD. Our hypothesis was that if the adaptation is the result of higher-order processing at the system level, this will affect both CMR and BMLD. In the first experiment, we measured masked thresholds by varying the length of the preceding maskers to test the effect of the preceding masker on CMR. In the second experiment, we measured masked thresholds in the presence of the binaural cue (IPD of π) in the same conditions as the first experiment. Results showed that only CMR was affected by the increased exposure time to preceding maskers while BMLD showed no adaptation effect, suggesting the low-level processing for the adaptation effect.

2. Methods

2.1. Stimuli

We measured CMR and BMLD with six different lengths of preceding maskers (0 ms, 20 ms, 100 ms, 200 ms, 300 ms, and 500 ms; see Figure 1). Each masker consisted of five narrow noise bands with 20 Hz bandwidth. The center band (CB) was centered at 700 Hz and the remaining four masker bands (flanking bands, FBs) were centered at 460, 580, 820, and 940 Hz. The spectral distance between the masker bands was chosen to maximize CMR based on Grose et al. [2009]. The tone was centered at 700 Hz, along with the CB. For two conditions with no preceding masker, two types of maskers were used (0 ms): the masker with uncorrelated intensity fluctuations across frequency (R), and with coherent intensity fluctuations across frequency (C). To investigate the effect of preceding maskers, we used four different stimuli consist of a preceding masker (20 - 500 ms) and following masked tone (200 ms): the reference condition with uncorrelated masker for both the preceding and following maskers (RR), and three maskers consisting of a comodulated masker preceded by three different maskers: uncorrelated masker (RC), comodulated masker (CC), and a masker with comodulated flanking-band (FC). The preceding masker was overlapped with the masker presented with the tone using a 20 ms raised-cosine off- and onset ramp with 50% overlap. The overall level of the masker was set to 60 dB SPL. To induce BMLD, we presented the tone with an interaural phase difference (IPD) of π.

2.2. Protocol

After the training session, each listener performed three threshold measurements for all conditions. For each measurement, conditions were presented in randomized order. The thresholds were estimated by averaging three trials. An additional measurement was done if the thresholds from the last three measurements had high variance (SD > 3 dB). During the threshold measurement, the listeners were seated in a double-walled, soundproof booth with ER-2 headphones. We used an adaptive, three-interval, three-alternative forced-choice procedure (3-AFC) with a one-up, two-down rule to estimate the 70.7% of the psychometric function [Ewert, 2013], Levitt [1971]. Three sound intervals were presented with a pause of 500 ms in between. Two intervals contained only the maskers, while one interval contained the target tone with maskers. The listeners' task was to choose the interval with a target tone by pressing the corresponding number key (1, 2, 3) on the keyboard. Whenever the listener pressed the keyboard, visual feedback was provided, indicating whether the answer was "WRONG" or "CORRECT". The target tone's start level was set to 75 dB. Depending on the answer, the tone level was adjusted with an initial step size of 8 dB. The step size was halved after each lower reversal until it reached the minimum step size of 1 dB. The signal level at a minimum step size of 1dB was measured six times, and the mean of those measurements was used as the estimated threshold.

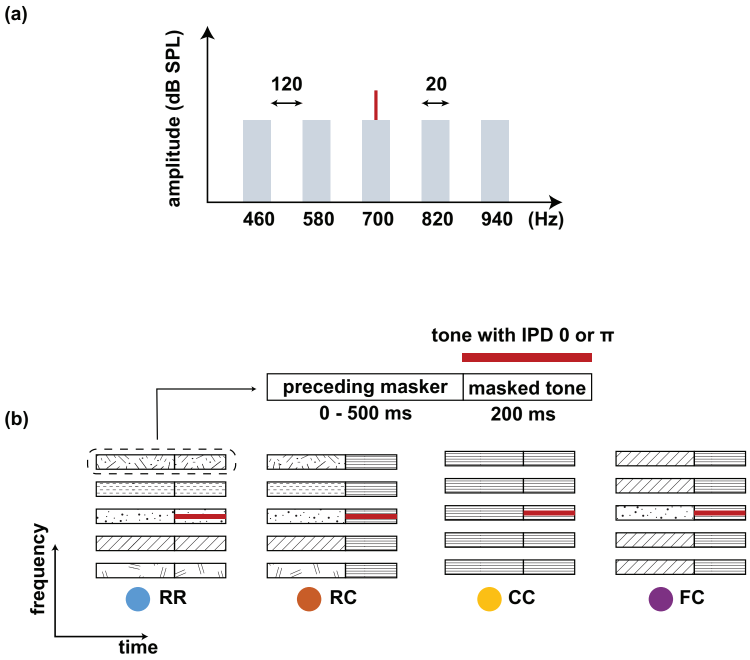

Figure 1: (a) Spectra of the stimulus. A target tone (700 Hz) was presented with a masking noise consisting of five narrow-band maskers: One signal centered band (SCB) and four flanking bands (FBs). The bandwidth of each masker band was 20 Hz, and the frequency spacing was 120 Hz. The overall level of the noise was set to 60 dB SPL. (b) Schematic spectrograms of the stimulus conditions. Each stimulus consisted of a preceding masker (0 - 500 ms) and a masked tone (200 ms). Four types of maskers were used: RR, RC, CC, and FC. The RR was used as the reference condition with uncorrelated masker bands. In the other three conditions, the maskers consisted of a comodulated masker preceded by three different maskers: uncorrelated masker (RC), comodulated masker (CC), and the masker with comodulated flanking-bands (FC). The thick line represents a tone that was presented with an IPD of 0 or π.

2.3. Listeners

We recruited eleven normal-hearing listeners after the initial hearing screening. None of them reported any history of hearing impairment and had pure-tone hearing thresholds within 15 dB HL for the standard audiometric frequencies from 125 to 4000 Hz. All participants provided informed consent, and all experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391).

3. Results

3.1. The threshold measurements

We measured masked thresholds in diotic and dichotic conditions. Figure 2(left panels, a-c) show the mean thresholds across all listeners. Masked thresholds in diotic conditions are plotted with solid lines and masked thresholds in dichotic conditions are plotted with dotted lines. Each plot shows the mean threshold values from 20 ms to 500 ms of preceding maskers for four masker types: RR, RC, CC, and FC. As a reference, we had two masker types without preceding maskers (0 ms): R and C. Thresholds in the RR condition were almost constant as the length of preceding masker increased. The introduction of an IPD reduced the masked threshold by approximately 15 dB. With 20 ms duration of the preceding maskers, thresholds in all three conditions (CC, RC, and FC) were lower than the thresholds in the RR condition in both diotic and dichotic conditions. As the duration of preceding maskers increased, all conditions showed changes in thresholds. Thresholds in the CC conditions showed a slight decrease as the duration of preceding maskers increased. On the contrary, thresholds in the RC and FC conditions showed increased values with longer preceding maskers. In the FC conditions, thresholds became higher than the RR conditions in both diotic and dichotic conditions.

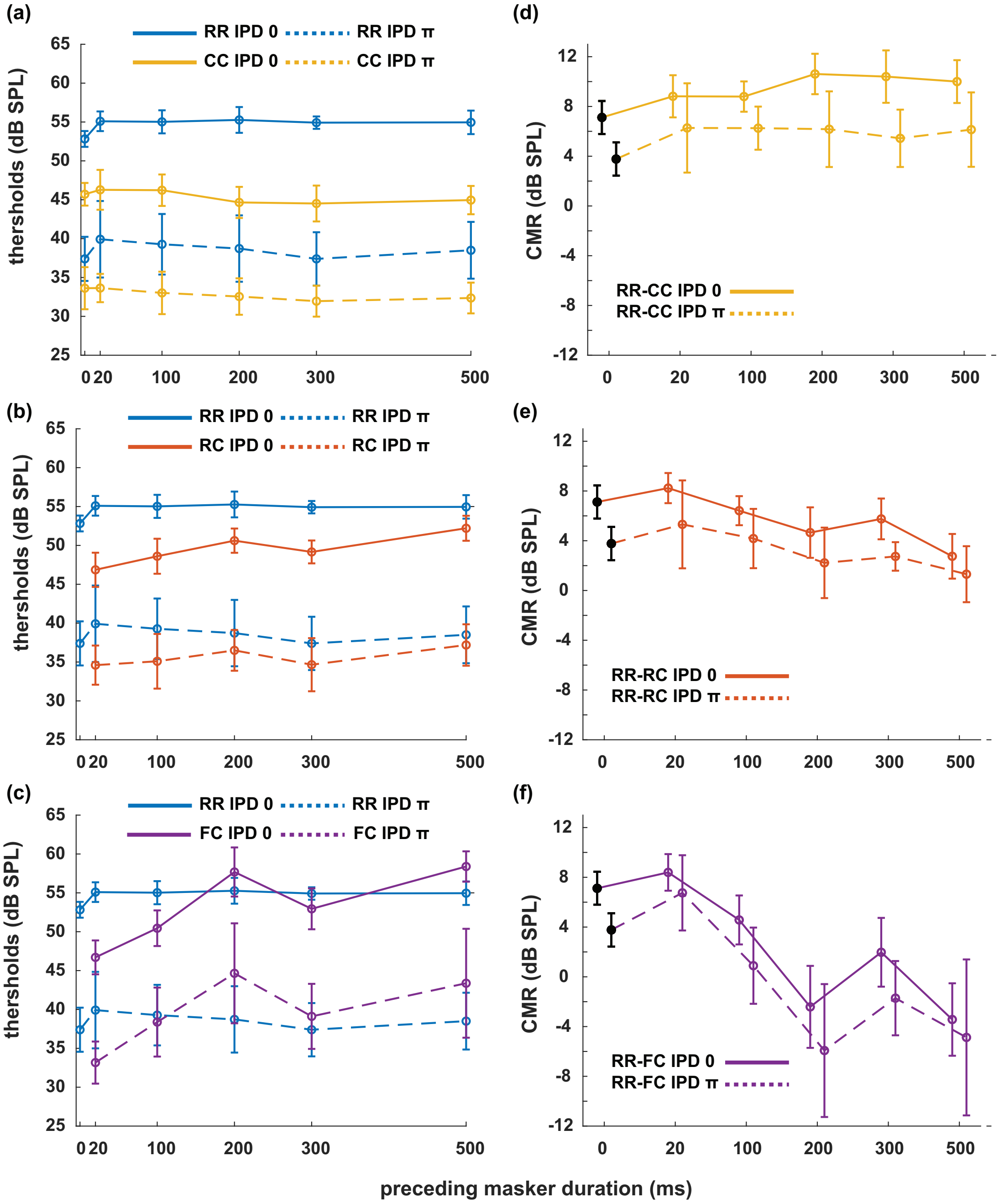

Figure 2: Mean masked thresholds (left column) and CMR values (right column) for each masker condition with the RR masker as a reference. Data from all listeners were averaged. Data are plotted for each masker type with colors (RR - blue, RC - orange, CC - yellow, FC - purple). For diotic conditions with an IPD of 0, masked thresholds are plotted with solid lines. For dichotic conditions with IPD of π, masked thresholds are plotted with dotted lines. Error bars indicate plus-minus one standard deviation.

3.2. Comodulation masking release

We calculated CMR for diotic conditions for three conditions (CC, RC, and FC) by subtracting their thresholds from the thresholds of the RR condition. The resulting CMR is shown in Figure 2 (right panels, d-f). Each plot shows changes in CMR with varying the duration of preceding maskers (see Table 1 for details). As a reference, we calculated CMR for the condition with no preceding masker by subtracting the masked threshold

	CMR	CC	CC_π	RC	RC_π	FC	FC_π
duration	0 ms	7.12	3.77	7.12	3.77	7.12	3.77
	20 ms	8.82	6.27	8.23	5.32	8.39	6.75
	100 ms	8.79	6.26	6.42	4.18	4.57	0.89
	200 ms	10.61	6.17	4.66	2.22	-2.42	-5.93
	300 ms	10.40	5.44	5.76	2.74	1.97	-1.72
	500 ms	10.00	6.14	2.75	1.31	-3.43	-4.87

Table 1: Summary of mean CMR values in all conditions. * CMR values for no preceding masker (0 ms) were calculated by subtracting the threshold of C from that of R.

	BMLD	RR	RC	CC	FC
duration	0 ms	15.43	12.08	12.08	12.08
	20 ms	15.18	12.26	12.63	13.53
	100 ms	15.75	13.51	13.21	12.07
	200 ms	16.54	14.11	12.11	13.03
	300 ms	17.52	14.51	12.56	13.83
	500 mx	16.46	15.02	12.59	15.02

Table 2: Summary of mean BMLD values in all conditions. * BMLD values for no preceding masker (0 ms) were calculated by subtracting the threshold of C from that of R.

of C from the that of R condition. The reference condition is plotted with black symbols. CMR in the diotic condition was estimated as 7.1 dB while CMR was 3.8 dB in the dichotic condition. When the preceding masker was short (20 ms), all conditions showed a slight increase in CMR by around 1.5 dB in diotic conditions and by 2.4 dB in dichotic conditions. With the increased length of the preceding maskers, CMR values in the CC conditions showed a slow increase while CMR values in the RC showed a slight decrease. FC conditions showed a larger decrease compared to the other conditions. In general, CMR values were lower in dichotic conditions compared to diotic conditions.

We tested the significance of the changes in CMR when the duration of the preceding masker was increased from 20 ms to 500 ms (see the table in Supplement section **??**). In the CC condition with IPD of 0, CMR was not significantly changed. However, the CMR was significantly increased from 7.12 dB to 10 dB (one-way ANOVA: F(5,54) = 6.44, $p < 0.001$) compared to the condition without preceding masker. In the RC condition with IPD of 0, CMR was significantly decreased from 8.23 dB to 2.75 dB (one-way ANOVA: F(4,45) = 16.17, $p < 0.001$). In the FC condition with IPD of 0, CMR was significantly decreased from 8.39 dB to -3.43 dB (one-way ANOVA: F(4,45) = 36.47, $p < 0.001$). For dichotic conditions with IPD of π, there was no significant change in CMR in the CC condition. In the RC condition, CMR was decreased 5.32 dB dB to 1.31 dB (one-way ANOVA: F(4,45) = 3.9, p = 0.008). In the FC condition, CMR significantly decreased from 6.75 dB to -4.87 dB (F(4,45) = 13.57, $p < 0.001$).

3.3. Binaural masking level difference

We calculated BMLD for four conditions (RR, RC, CC, and FC) by subtracting the masked threshold of dichotic conditions from those of the corresponding diotic conditions. Figure 3 shows the BMLD for each condition (RR, RC, CC, and FC). Unlike CMR, BMLD did not show a significance change (one-way ANOVA) with varying the duration of preceding maskers (RR condition (F(5,54)=0.64, p=0.67), CC condition (F(5,54)=0.37, p=0.86), RC condition (F(4,45)=1.43, p=0.24, and FC condition (F(4,45)=0.54, p=0.70)). For the FC condition, we observed that the variance of BMLD measures increased from 2 dB to 6.7 dB, while the variances of the other conditions were between 2 to 3.8 dB.

4. Discussion

In this study, we investigated the time constant of the effect of preceding masker on CMR and BMLD, and whether the effect of preceding masker is the result of high-level auditory processing at the system level. We measured CMR by varying the length of the preceding maskers. By adding short preceding maskers (20 ms), all types of preceding maskers showed increased CMR. This supports the idea of bottom-up processing where the neural encoding of CMR is assumed to be an instantaneous process with the wideband inhibition

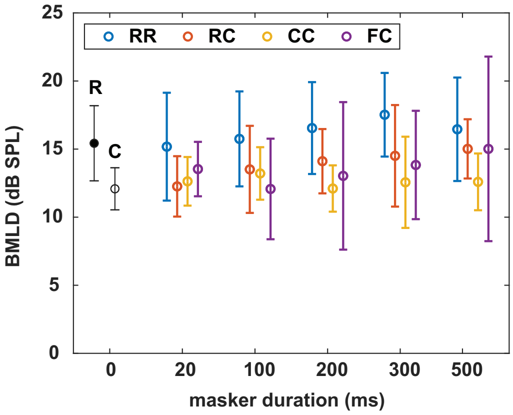

Figure 3: Mean BMLD for each masker condition. Data from all listeners were averaged. Data are plotted for each masker type with colors (RR - blue, RC - orange, CC - yellow, FC - purple). As a reference, BMLD in the R and C conditions are shown in black.

mechanism at the cochlear nucleus (CN) level [Epp and Verhey, 2009, Pressnitzer et al., 2001, Neuert et al., 2004]. In the CC condition, the effect of preceding maskers was small compared to other conditions (RC and FC). Contrary to the CC condition, in the RC and FC condition, "build-up of stream segregation" [Moore and Gockel, 2012] was more prominent as the CMR decreases with longer exposure to the preceding maskers. This indicates that the preceding modulation patterns disrupts the following comodulation cue for the target detection. This is in line with the hypothesis of the adaptation at the high-level auditory processing. However, this is at odds with the finding that BMLD was hardly affected by preceding maskers. That is, the grouping of frequencies in the preceding masker only has influence on following frequency grouping by comodulation but not by IPD. This supports the idea that low-level auditory processing is more involved in the effect of preceding maskers than high-level auditory processing.

In the visual system, the adaptation to light contrast occurs at the retinal level (equivalent to the cochlea), ranging from 0.1 to 17 seconds Baccus and Meister [2002]. These time constants can be realized by different retinal cells such as bipolar cells, amacrine cells, and ganglion cells [Kohn, 2007]. Similarly, in the CN, various types of neurons exist, and these neurons are connected to each other forming neural circuits such as feed-forward excitation/inhibition, feedback inhibition, and mutual excitation [Oertel et al., 2011, Manis and Campagnola, 2018, Ngodup et al., 2020]. Possible neuronal correlates of the slow adaptation effect may exist between the CN level and the cortical level (top-down influence). We speculate that the possible mechanism of the adaptation to preceding maskers may occur at a low-level of auditory processing (e.g., the CN level) rather than at a high level of auditory processing (e.g., cortical level). However, further physiological evidence is needed to support current speculation.

Furthermore, for the FC condition, the effect of preceding maskers was larger compared to other conditions. At the same time, CMR and BMLD values were highly variable depending on the listeners. Some listeners showed high CMR and BMLD, while others showed low CMR and BMLD. With linear regression, we conducted an additional analysis of the relationship between CMR and BMLD for the RC, CC, and FC conditions (Figure 4 - 6). The FC dichotic conditions showed a strong positive correlation between CMR and BMLD compared to other conditions. We speculate that the IPD cue facilitates target tone segregation from the noise. This may induce the grouping of FBs and the CB into one stream, thereby enhancing CMR.

We also observed non-linearity in combination of CMR and BMLD. In dichotic conditions, CMR measures were lower compared to diotic conditions. Interestingly, unlike the FC condition, the CC and RC conditions showed negative correlation in diotic conditions. This is in line with previous studies [Schooneveldt and Moore, 1989, Hall et al., 1990, Ernst and Verhey, 2006]. A possible neural basis has been proposed as a contralateral projection of a wideband inhibitor cell [Verhey et al., 2003, Ernst and Verhey, 2006, Ingham et al., 2006]. In addition, the CC condition did not show any changes in CMR with increased duration of the preceding masker. These may be due to the physiological limit of the auditory system in which a maximum amount of masking release is reached.

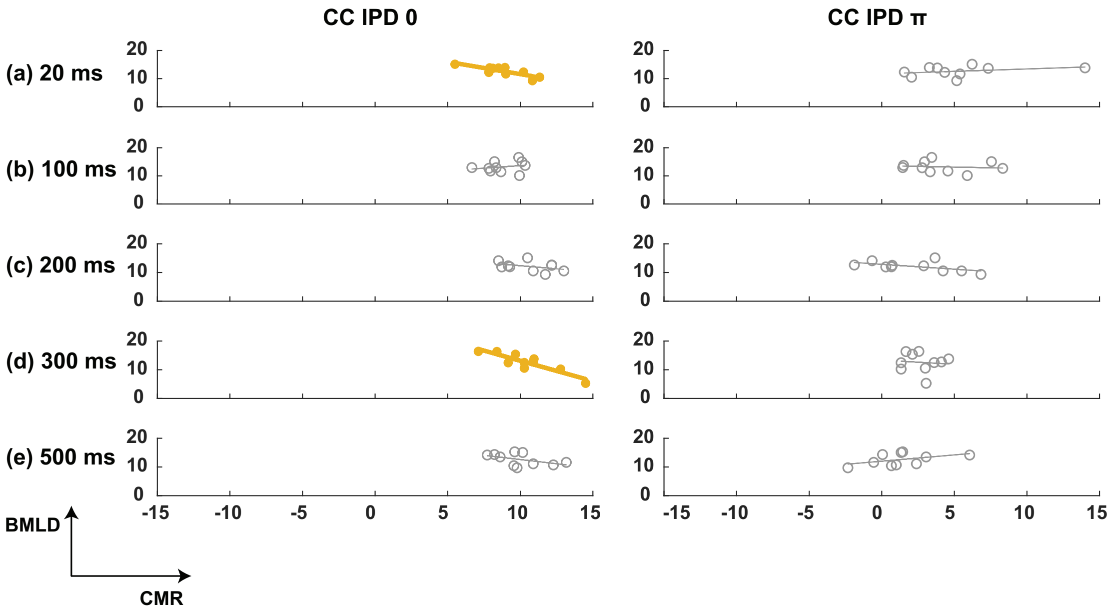

Figure 4: Linear regression analysis between CMR and BMLD for CC conditions. Only the conditions with significant p-values were plotted with color. With 20 ms of preceding masker and 300 ms of preceding masker, CC conditions with no IPD cue showed negative correlation between CMR and BMLD (see Table 3).

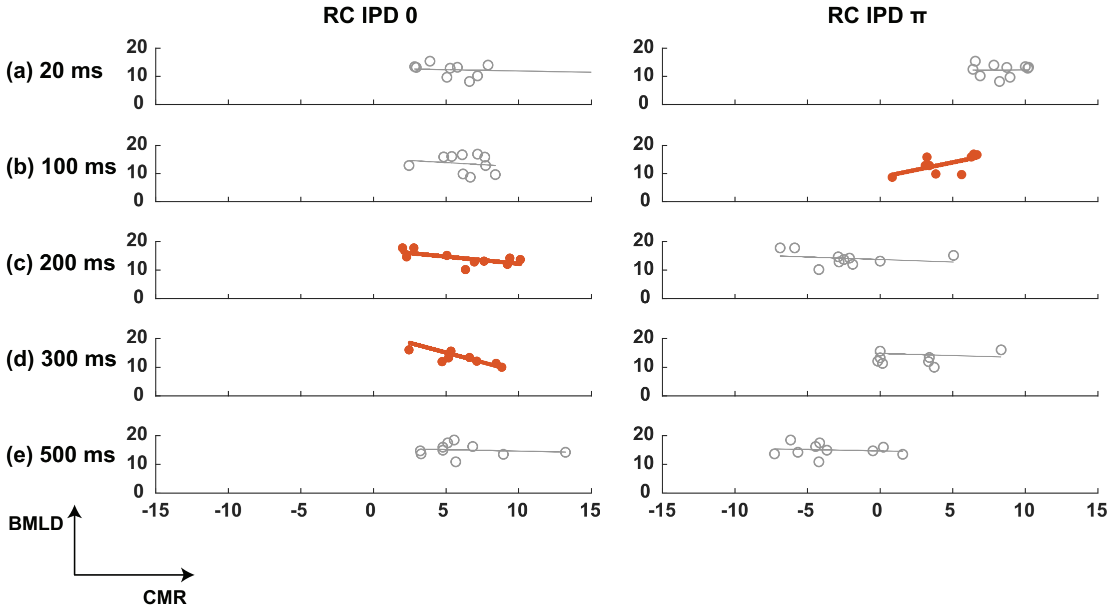

Figure 5: Linear regression analysis between CMR and BMLD for RC conditions. Only the conditions with significant p-values were plotted with color. With 200 ms of preceding masker and 300 ms of preceding masker, RC conditions with no IPD cue showed negative correlation between CMR and BMLD. With IPD of π, RC condition with 100 ms of preceding masker showed positive correlation between CMR and BMLD (see Table 3).

5. Conclusion

The present study investigated the effect of the preceding masker on CMR and BMLD by varying the length of the preceding masker. The effect of the preceding masker on the following sound has been suggested to

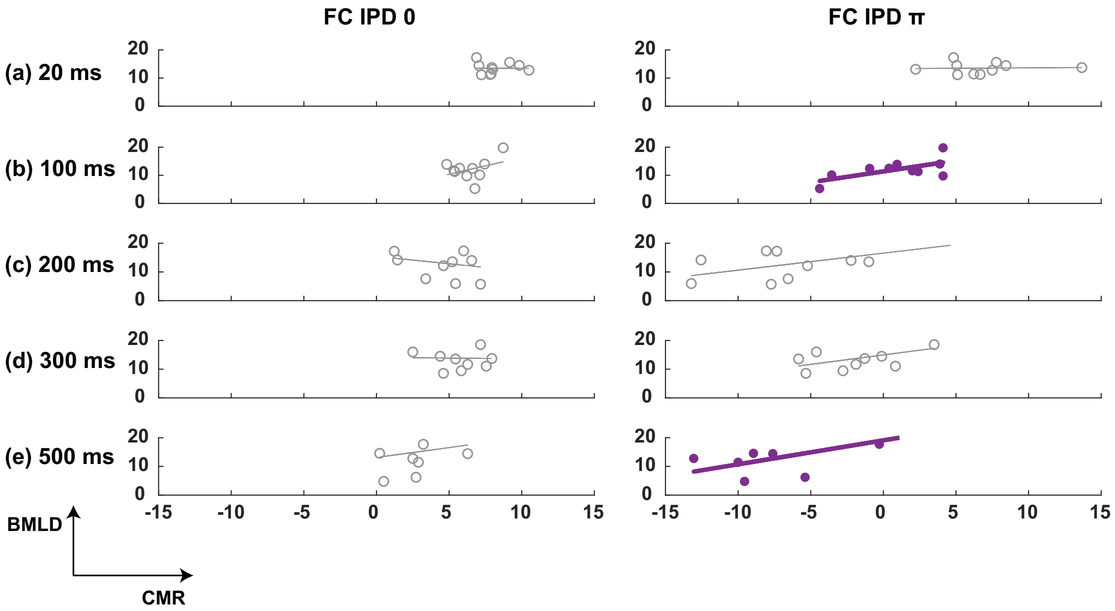

Figure 6: Linear regression analysis between CMR and BMLD for FC conditions. Only the conditions with significant p-values were plotted with color. With IPD of π, RC condition with 100 ms and 500 ms of preceding masker showed positive correlation between CMR and BMLD (see Table 3).

$y = a * x + b$

	a	*b*	*p-values*		*a*	*b*	*p-values*
CC	-0.87	20.34	0.003*	CC_π	0.17	11.72	0.337
	0.37	9.98	0.518		-0.10	13.64	0.727
	-0.43	16.69	0.236		-0.34	12.86	0.090
	-1.42	27.28	0.001*		-0.29	13.36	0.784
	-0.59	18.46	0.156		0.44	12.02	0.168
RC	-0.09	12.84	0.678	RC_π	0.01	12.14	0.978
	-0.30	15.38	0.655		1.06	8.67	0.041*
	-0.50	17.19	0.046*		-0.18	13.68	0.492
	-1.35	21.86	0.003*		-0.14	14.79	0.768
	-0.10	15.64	0.700		-0.09	14.69	0.730
FC	-0.02	13.71	0.972	FC_π	0.03	13.31	0.892
	1.18	4.48	0.286		0.77	11.38	0.045*
	-0.51	15.43	0.594		0.59	16.55	0.076
	-0.05	14.15	0.951		0.66	14.96	0.147
	0.69	13.13	0.615		0.83	19.08	0.009*

Table 3: Linear regression results summary. Here, *a* indicates CMR and *b* indicates BMLD.

be either the result of the adaptation at peripheral level or system level. Our hypothesis was that if CMR and BMLD change with increased duration of the preceding masker, the effect of preceding masker could be due to higher-level processing. From the data acquired in this study, we showed that the time constant of the effect of preceding masker can range up to 500 ms on CMR. However, the preceding masker did not affect BMLD. This may indicate that the effect of preceding masker may be the result of neural processing at the low-level rather than that of higher-order processing at the system level (e.g., temporal integration).

Acknowledgments

We would like to thank Viktorija Ratkute for support with the data collection.

6. Supplement

Table 4: The significant differences in CMR between conditions $p - values$ with varying the length of the preceding masker.

Condition	IPD	Duration (ms)	0	20	100	200	300	500
RR-CC	**IPD 0**	**0**	-	0.205	0.217	$<$ **0.001**	$<$ **0.001**	**0.003**
		20	-0.205	-	0.999	0.159	0.274	0.594
		100	0.217	0.999	-	0.149	0.260	0.574
		200	$<$ **0.001**	0.159	0.149	-	0.999	0.960
		300	$<$ **0.001**	0.274	0.260	0.999	-	0.993
		500	**0.003**	0.594	0.574	0.960	0.993	-
	IPD π	**0**	-	0.285	0.292	0.329	0.713	0.344
		20	0.285	-	0.999	0.999	0.979	0.999
		100	0.292	0.999	-	0.999	0.981	0.999
		200	0.329	0.999	0.999	-	0.988	0.999
		300	0.713	0.979	0.981	0.988	-	0.990
		500	0.344	0.999	0.999	0.999	0.990	-
RR-RC	**IPD 0**	**0**	-	-	-	-	-	-
		20	-	-	0.101	$<$ **0.001**	**0.010**	$<$ **0.001**
		100	-	0.101	-	0.119	0.887	$<$ **0.001**
		200	-	$<$ **0.001**	0.119	-	0.546	0.076
		300	-	**0.010**	0.887	0.546	-	**0.001**
		500	-	$<$ **0.001**	$<$ **0.001**	0.076	**0.001**	-
	IPD π	**0**	-	-	-	-	-	-
		20	-	-	0.855	0.068	0.177	**0.008**
		100	-	0.855	-	0.436	0.716	0.106
		200	-	0.068	0.436	-	0.991	0.929
		300	-	0.177	0.716	0.991	-	0.721
		500	-	**0.008**	0.106	0.929	0.721	-
RR-FC	**IPD 0**	**0**	-	-	-	-	-	-
		20	-	-	**0.014**	$<$ **0.001**	$<$ **0.001**	$<$ **0.001**
		100	-	**0.014**	-	$<$ **0.001**	0.176	$<$ **0.001**
		200	-	$<$ **0.001**	$<$ **0.001**	-	**0.003**	0.901
		300	-	$<$ **0.001**	0.176	**0.003**	-	$<$ **0.001**
		500	-	$<$ **0.001**	$<$ **0.001**	0.901	$<$ **0.001**	-
	IPD π	**0**	-	-	-	-	-	-
		20	-	-	**0.033**	$<$ **0.001**	$<$ **0.001**	$<$ **0.001**
		100	-	**0.033**	-	**0.009**	0.669	**0.038**
		200	-	$<$ **0.001**	**0.009**	-	0.214	0.982
		300	-	$<$ **0.001**	0.669	0.214	-	0.495
		500	-	$<$ **0.001**	**0.038**	0.982	0.495	-

References

S. Anstis and S. Saida. Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception and Performance*, 11(3):257, 1985.

S. A. Baccus and M. Meister. Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5):909–919, 2002.

A. S. Bregman. Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance*, 4(3):380, 1978.

A. S. Bregman. *Auditory scene analysis: The perceptual organization of sound*. MIT press, 1994.

E. C. Cherry. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5):975–979, 1953.

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

B. Epp and J. L. Verhey. Superposition of masking releases. *Journal of computational neuroscience*, 26(3): 393–407, 2009.

S. M. Ernst and J. L. Verhey. Role of suppression and retro-cochlear processes in comodulation masking release. *The Journal of the Acoustical Society of America*, 120(6):3843–3852, 2006.

S. D. Ewert. Afc—a modular framework for running psychoacoustic experiments and computational perception models. In *Proceedings of the international conference on acoustics AIA-DAGA*, pages 1326–1329, 2013.

J. H. Grose, E. Buss, and J. W. Hall III. Within-and across-channel factors in the multiband comodulation masking release paradigm. *The Journal of the Acoustical Society of America*, 125(1):282–293, 2009.

J. W. Hall, M. P. Haggard, and M. A. Fernandes. Detection in noise by spectro-temporal pattern analysis. *The Journal of the Acoustical Society of America*, 76(1):50–56, 1984.

J. W. Hall, J. H. Grose, and M. P. Haggard. Effects of flanking band proximity, number, and modulation pattern on comodulation masking release. *The Journal of the Acoustical Society of America*, 87(1):269–283, 1990.

N. J. Ingham, S. Bleeck, and I. M. Winter. Contralateral inhibitory and excitatory frequency response maps in the mammalian cochlear nucleus. *European Journal of Neuroscience*, 24(9):2515–2529, 2006.

L. A. Jeffress, H. C. Blodgett, T. T. Sandel, and C. L. Wood III. Masking of tonal signals. *The Journal of the Acoustical Society of America*, 28(3):416–426, 1956.

A. Kohn. Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of neurophysiology*, 97 (5):3155–3164, 2007.

L. Las, E. A. Stern, and I. Nelken. Representation of tone in fluctuating maskers in the ascending auditory system. *Journal of Neuroscience*, 25(6):1503–1513, 2005.

H. Levitt. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical society of America*, 49(2B):467–477, 1971.

P. B. Manis and L. Campagnola. A biophysical modelling platform of the cochlear nucleus and other auditory circuits: from channels to networks. *Hearing research*, 360:76–91, 2018.

B. C. Moore and H. E. Gockel. Properties of auditory stream formation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1591):919–931, 2012.

I. Nelken, Y. Rotman, and O. B. Yosef. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*, 397(6715):154–157, 1999.

V. Neuert, J. L. Verhey, and I. M. Winter. Responses of dorsal cochlear nucleus neurons to signals in the presence of modulated maskers. *Journal of Neuroscience*, 24(25):5789–5797, 2004.

T. Ngodup, G. E. Romero, and L. O. Trussell. Identification of an inhibitory neuron subtype, the l-stellate cell of the cochlear nucleus. *Elife*, 9:e54350, 2020.

D. Oertel, S. Wright, X.-J. Cao, M. Ferragamo, and R. Bal. The multiple functions of t stellate/multipolar/chopper cells in the ventral cochlear nucleus. *Hearing research*, 276(1-2):61–69, 2011.

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

L. J. Raphael, G. J. Borden, and K. S. Harris. *Speech science primer: Physiology, acoustics, and perception of speech*. Lippincott Williams & Wilkins, 2007.

G. P. Schooneveldt and B. C. Moore. Comodulation masking release for various monaural and binaural combinations of the signal, on-frequency, and flanking bands. *The Journal of the Acoustical Society of America*, 85(1):262–272, 1989.

T. M. Shackleton, B. C. Skottun, R. H. Arnott, and A. R. Palmer. Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of guinea pigs. *Journal of Neuroscience*, 23(2): 716–724, 2003.

T. M. Shackleton, R. H. Arnott, and A. R. Palmer. Sensitivity to interaural correlation of single neurons in the inferior colliculusof guinea pigs. *Journal of the Association for Research in Otolaryngology*, 6(3):244–259, 2005.

J. Sollini and P. Chadderton. Comodulation enhances signal detection via priming of auditory cortical circuits. *Journal of Neuroscience*, 36(49):12299–12311, 2016.

S. van de Par and A. Kohlrausch. Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *The Journal of the Acoustical Society of America*, 106(4): 1940–1947, 1999.

J. L. Verhey, D. Pressnitzer, and I. M. Winter. The psychophysics and physiology of comodulation masking release. *Experimental Brain Research*, 153(4):405–417, 2003.

O. Zohar, T. M. Shackleton, I. Nelken, A. R. Palmer, and M. Shamir. First spike latency code for interaural phase difference discrimination in the guinea pig inferior colliculus. *Journal of Neuroscience*, 31(25): 9192–9204, 2011.