Detection and Discrimination of Masked Tones in Masking Release Conditions

Hyojin Kim[1], Viktorija Ratkute[1], Bastian Epp[1]

[1] *Auditory Physics Group, Hearing Systems Section, Depeartment of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark*

Email: hykim@dtu.dk

Abstract

Comodulated masking noise and binaural cues can enhance one's performance in detecting a masked tone. The difference in detection performance between the absence and the presence of these cues is commonly referred to as "masking release". However, few studies have investigated sound perception at supra-threshold levels, which is more relevant for communication, in masking release conditions. Here, we used intensity just-noticeable difference (JND) to quantify intensity discrimination performance at supra-threshold levels in masking release conditions. We employed eight masking release conditions with combinations of comodulation (comodulation masking release, CMR) and interaural phase disparity (binaural masking level difference, BMLD), paired with various types of preceding maskers. Results for the detection of a tone in the different masking conditions showed that the preceding masker either enhance or diminish CMR while it did not affect BMLD. From the intensity JND experiment results, we found that the discrimination performance was hardly affected by the masking release conditions. The intensity JND was equal at the same physical intensity of the tone regardless of the masked threshold in that condition. At supra-threshold levels, we suggest that the physical intensity of the tone is encoded in addition to the enhanced neural representation at thresholds as an internal signal-to-noise ratio.

Introduction

Masking release is a perceptual phenomenon that describes enhanced detection performance of a masked signal. Masking release induced by comodulation is commonly referred to as "comodulation masking release (CMR)" [1], and the masking release induced by binaural cues to as "binaural masking level difference (BMLD)" [2]. Previous studies showed that the masking release in conditions with interaural phase differences (IPD) and comodulation cues present is equal to the sum of CMR and BMLD [3]. While comodulation and binaural cues have shown significant improvement in detection performance up to 10 dB or more [4][5], less is known about their benefit and relevance for listening at supra-threshold levels. Several studies have investigated the salience of a masked tone at supra-threshold levels in the absence and presence of masking release. They used categorical loudness scaling and continuous scaling of the perceived salience [6][7]. In categorical scaling, the supra-threshold perception in comodulated noise was similar to that in uncomodulated noise. Data obtained by continuous scaling showed individual variability in the ratings, presumably because some listeners confused the loudness of the tone with the salience. The limitation of those methods is that those measures strongly depend on the subjective criterion used by the listener. Here, instead of asking listeners to estimate how clear they hear the tone (salience), we estimated the minimum level of increase in tone intensity to make listeners perceive the an increase in perceived intensity (intensity just-noticeable difference, JND). Previous study showed that the loudness JND (ΔL) is proportional to $L^{1/p}$, where L is the loudness and p is a piece-wise intensity-independent constant ranging from 1 to 2 [8]. In this study, we investigated whether the intensity JND can be a measure of the salience of the tone in masking release conditions. We hypothesized that with the same increase of tone level, the condition with lower JNDs would have higher salience than the condition with higher JNDs.

Methods

Listeners Seven listeners with normal-hearing participated in the experiments. 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. One listener was tested with 20 dB at 125 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).

Stimuli Eight masking release conditions were used (Figure 1). The stimulus consisted of a masker constructed of five noise bands and a tone as a target signal. One noise band was spectrally centred at the tone frequency (signal centred band, SCB). The other noise bands were placed spectrally above and below the tone frequency (flanking bands, FBs). Each masker band had a bandwidth of 20 Hz and a level of 60 dB SPL. The target tone was centered at 700 Hz. The FBs were centered at 460, 580, 820, and 940 Hz to maximize the CMR [9].

Temporally, as shown in Figure 2, the stimulus was constructed of two elements: 1) a 500ms interval with a "preceding masker", 2) a 200ms interval with the masker, and an added tonal signal ("signal interval"). We used four masker conditions: the reference condition with maskers with uncorrelated intensity fluctuations across frequency for the "preceding masker" and the "signal interval" (RR). In the three other conditions, the comodulated noise in the "signal interval" was preceded by one of three masker types: i) a masker with uncorrelated intensity fluctuations across frequency (RC), ii) a comodulated masker (CC), iii) a masker where only the FBs were comodulated (FC). In each of these four masker conditions,

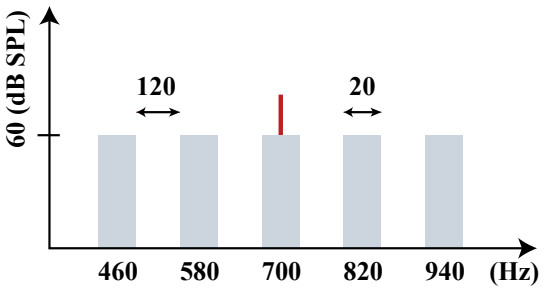

Figure 1: Spectra of the stimulus. A target tone (700 Hz) was presented with a masking noise consisting of five narrow-band maskers: One signal centered band (SCB) and four flanking bands (FBs). The bandwidth of each masker band was 20 Hz and the frequency spacing was 120 Hz. The overall level of the noise was set to 60 dB SPL.

the tone was presented with an IPD of 0 or π.

Apparatus During the experiment, the listeners were seated in a double-walled, soundproof booth with ER-2 headphones (Etymotic). 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 [10][11]. All stimuli were generated in MATLAB with a sampling rate of 44100 Hz and a 16-bit resolution, converted from digital to analogue (RME Frieface UCX). The headphones were calibrated and equalized at the signal frequency of the tone.

Masked thresholds We measured the masked detection thresholds individually for eight masking release conditions. Each listener performed four threshold measurements for all conditions. For each measurement, conditions were in randomized order. Two intervals contained the masker only, and the third randomly chosen interval contained the masker and the tone. The listeners' task was to select the interval with the target tone in noise. Visual feedback was provided, indicating whether the answer was "WRONG" or "CORRECT". The target tone's initial level was set to 75 dB and 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 the last six reversals were used as the estimated threshold. The first measurement was discarded, and the rest of the measurements were used for the threshold estimate. An additional measurement was acquired if the thresholds from the last three measurements showed a high variance > 3 dB.

Intensity just-noticeable differences (JNDs) We measured the intensity JNDs individually at the threshold and at five supra-threshold levels for all eight conditions. The intensity of the tone was individually adjusted for each listener to match levels of +0 dB (threshold), +5dB, +10dB, +15 dB, +20 dB, and +25 dB above the individual listeners' threshold. Three intervals were presented in random order. Two intervals contained the tone at the same fixed supra-threshold level, and the third with the tone at higher levels than the other two intervals. The listeners' task was to choose one interval with the highest intensity of the 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 start level of the tone with the higher level was set to 75 dB, and the tone level was adjusted with the 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. Each listener performed three trials for all conditions. If the supra-threshold level exceeded 80 dB, the intensity JND measure was skipped. The mean of three measurements was used to estimate the intensity JND. We calculated the intensity JND by subtracting the supra-threshold level from the minimum level of discriminable tone.

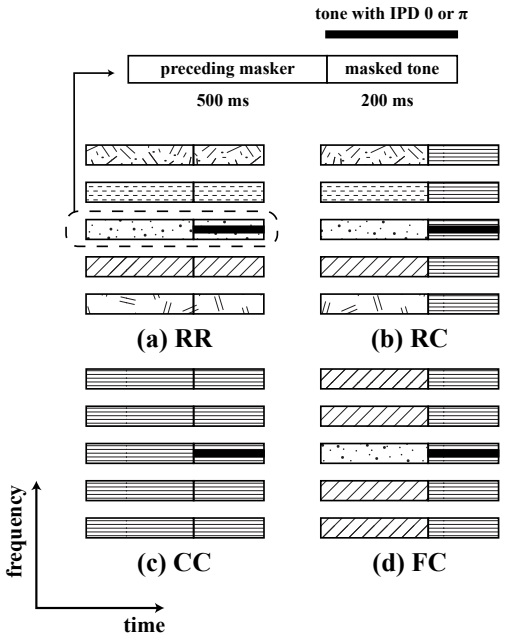

Figure 2: Schematic spectrograms of the stimulus conditions. Each stimulus consist of a preceding masker (500 ms) and masked tone (200 ms). Four types of maskers were used: RR, RC, CC, and FC. The RR was used as the reference condition with uncorrelated masker bands. In the other three conditions, the 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 black line represents a tone which was presented with an IPD of 0 or π.

Results

Figure 3(a) shows the threshold measurements in eight masking release conditions. The observed mean threshold across all the participants for the RR IPD0 condition was 54.6 dB. The RC IPD0 condition had a mean threshold of 52 dB. In the CC IPD0 condition, the threshold

was 44.6 dB. In the FC IPD0 condition, the mean threshold was found to be 58.6 dB. With the IPD of π, the RR IPD π had a mean threshold of 39 dB, and the RC IPD π condition had a mean threshold of 38.2 dB. In the CC IPD π, the mean threshold was 32.8 dB, and that of the FC IPD π condition was 45.6 dB.

$$CMR_{masker} = threshold_{RR} - threshold_{masker} \quad (1)$$

Each condition's CMR was calculated by subtracting the threshold of RR condition from its threshold (eq. (1)). In the diotic conditions, CMR_{RC} was 2.6 dB, CMR_{CC} was 10 dB and CMR_{FC} was -4 dB. In dichotic conditions, CMR_{RC} was 0.8 dB, CMR_{CC} was 6.2 dB and CMR_{FC} was -6.6 dB.

$$BMLD_{masker} = threshold_{IPD0} - threshold_{IPD\pi} \quad (2)$$

Each conditions's BMLD was calculated by subtracting the threshold of dichotic condition (eq. (2)). $BMLD_{RR}$ was 15.6 dB, $BMLD_{RC}$ was 13.8 dB, $BMLD_{CC}$ was 11.8 dB, and $BMLD_{FC}$ was 13 dB.

Figure 3(b) shows the averaged intensity JND at six supra-threshold levels: + 0 dB, + 5 dB, + 10 dB, + 15 dB, + 20 dB and + 25 dB. In Figure 3(c), the intensity JND measures of all conditions and listeners are plotted with scatter, and fitted with the power function and plotted with the line. The intensity JNDs decreased with increasing level of the target tone in all masking release conditions. In Figure 4, the intensity JND measure for each condition was plotted separately, starting from the threshold (+ 0 dB) to + 25 dB. JND values were more scattered below 60 dB compared to the JNDs at above 60 dB. Overall, the condition with lower detection thresholds (e.g., CC IPD π) showed higher JNDs compared to the one with higher detection thresholds (e.g., RR IPD 0). This indicates that enhancement in detection is not related to the supra-threshold levels. Rather, the degree of enhancement in the salience depends on the level of the target tone.

Discussion

We found that the combination of comodulation, IPD, and the preceding masker can induce different amount of masking release. We designed our stimuli based on a previous study with both preceding and following maskers [9]. Although our stimuli had no following masker after the offset of the target tone, the thresholds measures were in line with their results. This may suggest that the preceding masker plays a role in inducing the streaming effect more than following maskers. Furthermore, BMLD was not affected by the preceding masker, indicating a robustness of BMLD over this streaming cue.

The results from the intensity JND experiment show that the intensity JNDs depend on the physical intensity of the target tone rather than the level above masked

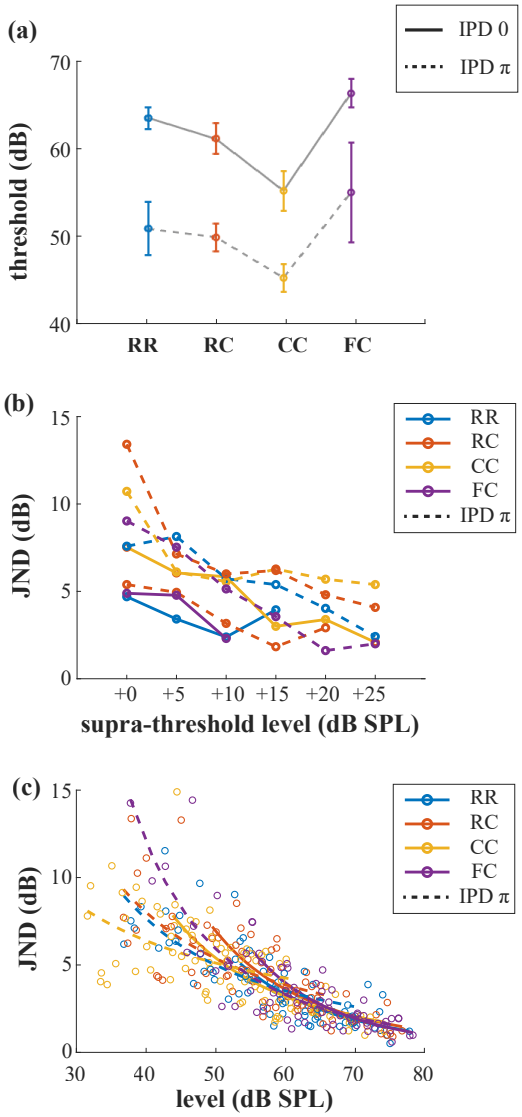

Figure 3: The plots of threshold measurements and the intensity JNDs. (a) The plot of the threshold measurements from eight masking release conditions. Error bars indicate plus-minus one standard deviation. (b) The plot of the averaged intensity JND at six supra-threshold levels: + 0 dB, + 5 dB, + 10 dB, + 15 dB, + 20 dB and + 25 dB. (c) The plot of the intensity JNDs as a function of the level. Individual data is plotted with scatter. The data for each condition is fitted with the power function and plotted with the line. Blue corresponds to RR condition, orange to RC, yellow to CC, and purple to FC condition. Solid line represents the data of IPD 0 and dotted line represents the data of IPD π.

threshold (supra-threshold levels). At supra-threshold levels, we suggest that the physical intensity of the tone is encoded in addition to the enhanced neural representation at thresholds as an internal signal-to-noise ratio (iSNR). At the level of the cochlea nucleus (CN), neural response to the tonal component in comodulated noise was enhanced compared to in random noise [12]. Based on this finding, we assumed that the iSNR is identical at threshold in all conditions. If the intensity-encoding

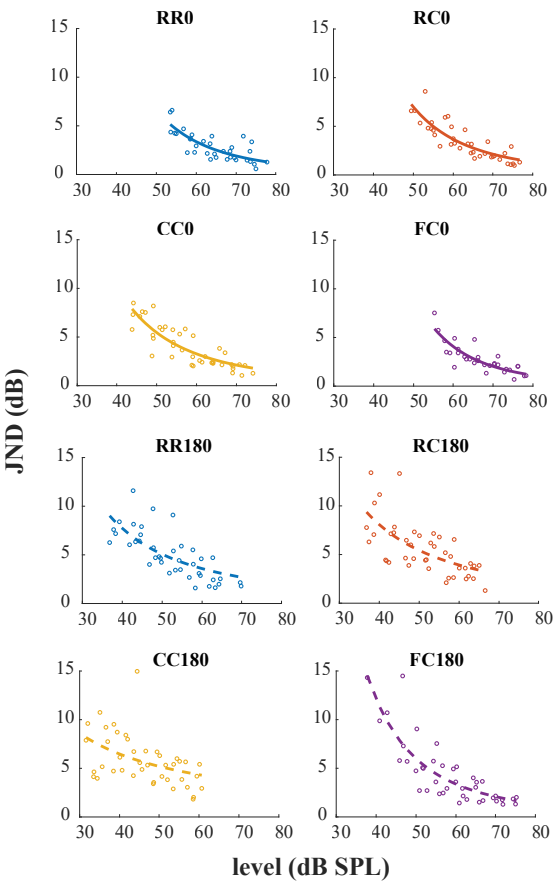

Figure 4: The plots of the intensity JND measurements for each masking release conditions, starting from the threshold (+ 0 dB) to + 25 dB. Four plots above are the intensity JNDs with no IPD. Four plots below are the intensity JNDs with IPD of π.

is encoded by the spike rate, the iSNR will increase as the spike rate increases. In the same manner, we also assumed that the intensity JND can be encoded with spike rate differences. Behaviorally, the intensity JNDs decrease as the level increase. This means that the changes in spike rates should get steeper as level increases. At CN, small cells showed such responses to tone in noise [13].

Conclusion

In this study, we investigated the detection and discrimination of masked tones in masking release conditions. Auditory cues such as comodualtion and IPD, and the preceding masker, could enhance the detection performance. On the other hand, at supra-threshold levels, the discrimination performance was highly dependent on the level of the tone. Regardless of the masking release conditions, the intensity JND measures were correlated with the intensity of the tone.

References

[1] Hall, Joseph W., Mark P. Haggard, and Mariano A. Fernandes.:Detection in noise by spectroâ€temporal pattern analysis. The Journal of the Acoustical Society of America 76.1 (1984): 50-56.

[2] Jeffress, Lloyd A., et al.: Masking of tonal signals. The Journal of the Acoustical Society of America 28.3 (1956): 416-426.

[3] Epp, Bastian, and Jesko L. Verhey.: Superposition of masking releases. Journal of computational neuroscience 26.3 (2009): 393-407.

[4] van de Par, Steven, and Armin 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 (1999): 1940-1947.

[5] Verhey, Jesko L., Daniel Pressnitzer, and Ian M. Winter.:The psychophysics and physiology of comodulation masking release. Experimental Brain Research 153.4 (2003): 405-417.

[6] Verhey, Jesko L., and Wiebke Heeren.:Categorical scaling of partial loudness in a condition of masking release. The Journal of the Acoustical Society of America 138.2 (2015): 904-915.

[7] Egger, Katharina, Torsten Dau, and Bastian Epp.:Supra-threshold perception and neural representation of tones presented in noise in conditions of masking release. Plos one 14.10 (2019): e0222804.

[8] Allen, Jont B., and Stephen T. Neely.:Modeling the relation between the intensity just-noticeable difference and loudness for pure tones and wideband noise. The journal of the acoustical society of America 102.6 (1997): 3628-3646.

[9] Grose, John H., Emily Buss, and Joseph 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 (2009): 282-293.

[10] Ewert, Stephan D.:AFCâ€"A modular framework for running psychoacoustic experiments and computational perception models. Proceedings of the international conference on acoustics AIA-DAGA. 2013.

[11] Levitt, H. C. C. H.:Transformed up-down methods in psychoacoustics. The Journal of the Acoustical society of America 49.2B (1971): 467-477.

[12] Neuert, Veronika, Jesko L. Verhey, and Ian M. Winter.:Responses of dorsal cochlear nucleus neurons to signals in the presence of modulated maskers. Journal of Neuroscience 24.25 (2004): 5789-5797.

[13] Hockley, Adam, Calvin Wu, and Susan E. Shore.: Cochlear nucleus small cells use olivocochlear collaterals to encode sounds in noise. bioRxiv (2021).