PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Modeling binaural detection of a Gaussian noise target in the presence of a lead/lag masker

Jonas BRAASCH[(1)], M. Torben PASTORE[(2)], Yi ZHOU[(3)]

[(1)]Rensselaer Polytechnic Institute, USA, braasj@rpi.edu

[(2)]Arizona State University, USA, torben.pastore@asu.edu

[(3)]Arizona State University, USA, yizhou@asu.edu

Abstract

Recently, we presented the results of a binaural masked detection experiment in which a noise target was temporally embedded within a lead/lag noise masker pair [4]. The results show that the inter-stimulus interval (ISI) between the masker and its reflection changed the detection threshold significantly. For low ISIs of 2-ms, the average masked detection was -11 dB, but for a greater ISI of 20 ms, the threshold was much higher (-2 dB). In the experiment, masked detection thresholds did not depend on whether the masker lead was on the same side as the target (with the lag on the contralateral side) or the other way round.

Keywords: Binaural Hearing, Binaural Masking Level Difference (BMLD), Precedence Effect

1 INTRODUCTION

In reverberant environments, sound source localization is often dominated by the directional cues of the wavefronts that propagate directly from the sound source and precede the arrival of reflections off nearby surfaces. That is, despite the presence of reflections that arrive along many different trajectories, listeners' perceived location of sound sources often corresponds closely with the actual location of the sound source itself. This is called the precedence effect (PE). Two competing theories exist to explain the precedence effect. One assumes that the auditory system at least partly "removes" the lag stimulus from the incoming signal and then estimates the position of the sound source based on the remaining signal (Research Hypothesis I – signal removal). This could occur as the result of neural processes not specific to the PE (e.g., [8]) or mechanisms specifically purposed for the PE (e.g., [2]). Another theory assumes that the auditory system "disregards" the localization cues of the lag signal and estimates the location of the sound source based on the directional cues of the lead (Research Hypothesis II – cue disregard) [9]. The present study reported here used a masked detection experimental paradigm to investigate which of the two research hypotheses is correct. For this purpose, the detectability of a target signal embedded in a lead/lag-based masker was tested. If the lagging stimulus is removed (Research Hypothesis I), the test participants would have clear access to the lead signal on its own and the results would then be similar to those where the same target is embedded in a masker with no lag. If, instead, the lag stimulus remained during auditory processing but the directional information of the lag was minimized (Research Hypothesis II), the results for both test conditions should differ substantially because the lag would continue to act as an additional masker.

2 PREVIOUS PSYCHOACOUSTICS STUDY

In this section, the results of a binaural detection threshold experiment from [4] are described that will be analyzed in the next section. Masked detection thresholds were obtained using an adaptive staircase method after Levitt [?] based on a three-alternative, forced-choice paradigm. In this experiment, Gaussian noise signals were used for both the target and the masker. All signals were bandpass filtered to restrict the frequency range of the signal to 200–1000 Hz using 128-coefficients FIR filters. The masker duration was 500 ms with 20-ms $\cos^2$ on and off ramps. The target duration was 200 ms with 5-ms $\cos^2$ on- and offset ramps. In all cases, the

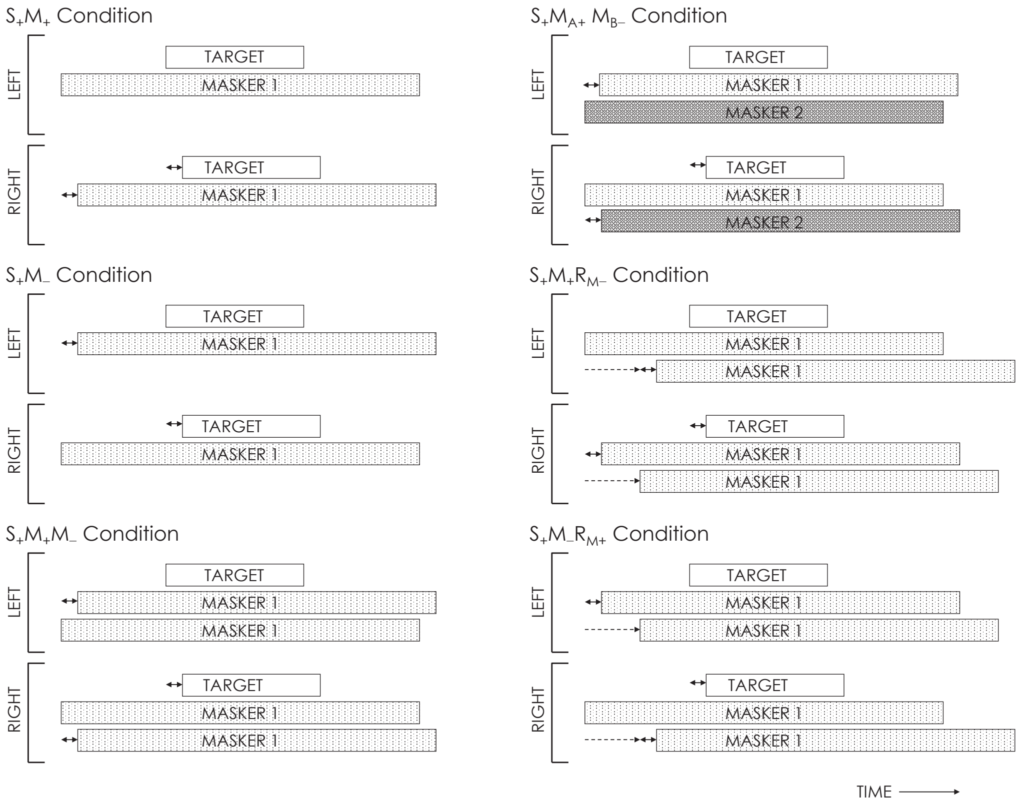

Figure 1. Schematic for the six stimulus configurations.

target was always presented 200 ms after the masker onset. A masker-independent noise sample was used to compute the target. The target signals were lateralized using an interaural time difference (ITD) of 300 μs. Six different spatial conditions were tested for the masker – see also Figure 1:

1. Ipsilateral Masker Condition (S_+M_+): The masker was spatially placed at the side ipsilateral to the target using an ITD of $+300$ μs (dichotic masker).
2. Contralateral Masker Condition (S_+M_-): the masker was spatially placed at the opposite side using an ITD of -300 μs.
3. Summing Localization Masker Condition, diotic masker: ($S_+M_+M_-$): using a masker with two identical components with simultaneous onsets. The first component was spatially placed at the ipsilateral target side using an ITD of $+300$ μs ms, the second presented contralaterally with an ITD of -300 μs.
4. Summed Uncorrelated (dichotic) Masker Condition ($S_+M_{A+}M_{B-}$): a masker pair as in Condition 3, but using two summed uncorrelated pairs.
5. Ipsilateral Masker with Contralateral Reflection Condition ($S_+M_+R_{m-}$): the masker was spatially placed at the opposite side using an ITD of -300 μs (dichotic masker) with an additional reflection from the same side -300 μs, presented after an ISI of 2, 5, 10, or 20 ms.
6. Contralateral Masker w/ Ipsilateral Reflection Condition ($S_+M_-R_{m+}$): the masker was spatially placed at the opposite side using an ITD of -300 μs with an additional reflection from the opposite side $+300$ μs,

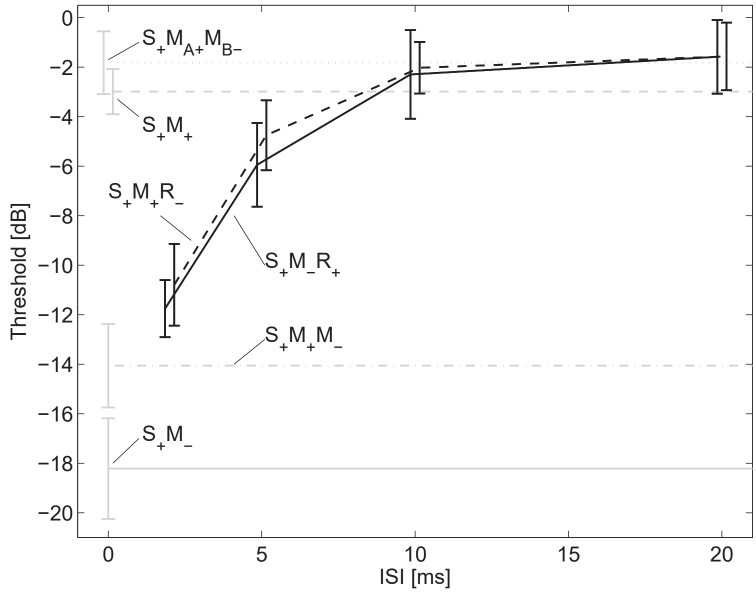

Figure 2. Results of the detection experiment showing the detection threshold for a Gaussian noise-burst target in the presence of a lead-lag pair masker as a function of the interstimulus-interval (ISI). The solid black curve shows the results for the contralateral masker lead. The dashed curve depicts the results for the ipsilateral masker lead. The gray curves show the baseline conditions. Each data point shows the average over all five listeners, the error bars depict the standard deviation – from [4].

presented after an ISI of 2, 5, 10, or 20 ms.

Figure 2 shows the results for all six conditions of the experiment. For the S_+M_+ condition, the target and the masker have the same spatial properties, the masked detection threshold was measured at -3 dB (gray dashed line). For the S_+M_- condition, where the masker was presented from the side opposite to the target, the masked threshold was found to be -18 dB (gray solid line), indicating a spatial release from masking of 15 dB. Similarly, for the $S_+M_+M_-$ diotic masker condition, the masked detection was measured at -14 dB (gray dashed-dotted line). In this case, the spatial release from masking was 11 dB. The highest masking threshold at -1.8 dB was found for the uncorrelated masker ($S_+M_{A+}M_{B-}$).

For conditions where the masker was a lead/lag pair ($S_+M_+R_{m+}$ and $S_+M_+R_{m-}$), performance was nearly the same regardless of whether the lead masker ITD favored the same or opposite side to the target. For both conditions, the masking thresholds approach the threshold for the diotic masking condition at low ISI values and approach the threshold of the uncorrelated masking condition at high ISI values. Also, variability is very consistent across all the different conditions, with values between 1 and 2 dB.

When the target and masker were presented from the same direction (300 μs ITD), our baseline condition, the threshold was -3 dB. This value is in general agreement with data by [1], who measured a -5 dB threshold for a 200-ms broadband noise target masked by conditions, 500-ms Gaussian noise masker (500-ms duration with its onset preceding the target onset by 200 ms). The 2-dB difference between both studies can be easily explained; the study reported here used band-pass filtered signals instead of the broadband signals used by Braasch [1], so less information was available to the test participants for making a decision than in [1]. Good et al. [7] previously showed that binaural masking level difference (BMLD), a measure of spatial release from masking, strongly depends on the frequency range of sounds, in their case; low, mid, and high, using free-field signals, so the inclusion of more high-frequency energy in the stimuli of [1] could be expected to make a

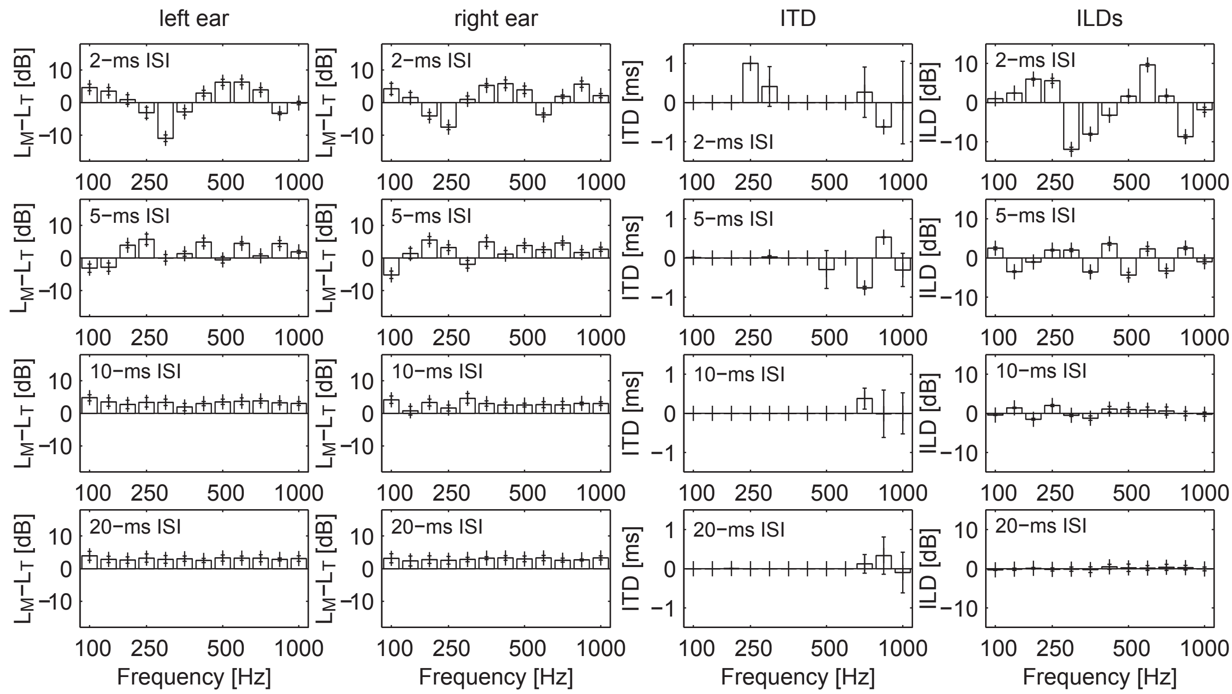

Figure 3. Monaural spectral changes resulting from interference between the leading and lagging masker components for different ISI values for an adjusted 0-dB target-to-masker ratio. The actual masker levels are shown for different auditory bands as the difference to the target level. L_M denotes the level of the masker, and L_T denotes the level of the target. A gammatone filterbank was used to simulate the auditory filters. Each bar shows the mean result for 10 simulation runs using different noise samples. Error bars denote ± 1 standard deviation.

difference in results as well.

In addition, we found:

1. The effect clearly depends on the ISI, for long ISI we see the effect of the ipsilateral masker plus a sight threshold elevation for an additional contralateral noise.
2. For short ISIs both ipsilateral and the contralateral masker components interfere and cannot be perceptually separated, leading to similar results as a diotic masker
3. Over the whole tested ISI range it does not make any significant difference if the ipsi or the contralateral masker component comes first, meaning there is no evidence for the PE effect playing a role here.

3 BINAURAL CUE ANALYSIS

Binaural and monaural cues were analyzed within auditory bands to understand how the direct sound masker and the reflection interfere at different ISIs. The leftmost column of Fig. 3 shows the computed level differences between the masker and the target for the left ear, therewith, if both the masker and the target have the same level before the reflection is applied. The second column from the left depicts the computed level differences between the masker and the target for the right ear. At an ISI of 2 ms, the reflection cancels out the signal at a frequency of 250 Hz, because 2 ms is half the period of a 250-Hz sine tone, shifting this frequency component out of phase. Since both the direct (lead) masker signal and its reflection (lag) have ITD cues pointing in

	2	3	4	5	6	7	8	9	10	11	12
1. S_+M_+	0.000	0.000	0.139	0.000	0.042	0.119	0.091	0.000	0.012	0.191	0.021
2. S_+M_-	-	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
3. $S_+M_+M_-$	-	-	0.000	0.001	0.000	0.000	0.000	0.016	0.000	0.000	0.000
4. $S_+M_{A+}M_{B-}$	-	-	-	0.000	0.016	0.754	0.684	0.000	0.001	0.545	0.714
5. $S_+M_+R_-$ [2]	-	-	-	-	0.001	0.000	0.000	0.099	0.001	0.000	0.000
6. $S_+M_+R_-$ [5]	-	-	-	-	-	0.045	0.028	0.000	0.094	0.036	0.017
7. $S_+M_+R_-$ [10]	-	-	-	-	-	-	0.211	0.000	0.013	0.710	0.392
8. $S_+M_+R_-$ [20]	-	-	-	-	-	-	-	0.000	0.009	0.368	0.974
9. $S_+M_-R_+$ [2]	-	-	-	-	-	-	-	-	0.001	0.000	0.000
10. $S_+M_-R_+$ [5]	-	-	-	-	-	-	-	-	-	0.006	0.004
11. $S_+M_-R_+$ [10]	-	-	-	-	-	-	-	-	-	-	0.036
12. $S_+M_-R_+$ [20]	-	-	-	-	-	-	-	-	-	-	-

Figure 4. All possible pairwise comparisons with a Student's t-test. Individual conditions are listed in the leftmost column. For conditions which included a lead/lag masker pair, the delay between lead and lag (in milliseconds) is indicated within square brackets. All other columns show the comparisons, with the listed number of the comparison condition indicated at the top of each column. For example, the p-value for the comparison of the the 5-ms ISI $S_+M_+R_-$ condition with the same condition but with a 10-ms ISI is shown at the intersection of the 6th row and the 7th column.

opposite directions, the actual delay of the reflection is 2.3 ms in the left ear and 1.7 ms in the right ear. This explains why the minimum level difference is observed at two different frequency bands for the left and right ear signals. In both cases, the minimum level difference is slightly less than -10 dB. At this frequency, the masker signal is about 10 dB below the target signal if the nominal target-to-masker level difference is 0 dB. At an ISI values of 5-ms, the cancellation frequency declines to 100 Hz. While this frequency is below the lower cut-off frequency of the stimuli, residual energy still exists because of the roll-off characteristics of the high-pass finite-impulse-response filter (Matlab function fir1.m with 128 filter coefficients). At the right ear, the lowest masker-to-target level difference is found in the 100-Hz band with a value of -6 dB. At the 10-ms and 20-ms ISI values, the masker level no longer varies much from the average level difference of 3 dB in the relevant frequency range. The 3-dB value corresponds to the level increase found for adding two equally intense uncorrelated noise sources.

Further, it was expected that the threshold for the condition with the additional uncorrelated noise masker ($S_+M_{A+}M_{B-}$) of approximately -2 dB, would be even higher than the baseline condition. The additional masker component only affected the masked detection threshold slightly, because its spatial properties differ from the congruent spatial cues of the first masker component and the target. The BMLD between the S_+M_- and S_+M_+ conditions was 15 dB, which corresponds to the general BMLD for the M_0S_π versus M_mS_m condition cited in the literature [11] (Table 12.2).

Listener performance for the conditions with lead-lag pair maskers ($S_+M_+R_{m-}$ and $S_+M_-R_{m+}$) was highly dependent on the inter-stimulus interval (ISI) between the lead and the lag masker components. However, it made hardly any difference whether the lead masker was presented from the same or opposite side to the target. In general, the thresholds rose monotonically with the ISI. When the delay between the first and second masker was 2 ms ISI, the results were similar to those of the diotic masker condition ($S_+M_+M_-$). In fact, the diotic masker condition can be seen as a special case of the lead/lag pair masker with an ISI of 0 ms where lead and lag arrive simultaneously. The measured thresholds for stimuli with lead/lag masker pairs were similar to those found by [6] who tested a lead/lag pair masker (noise or talker) for a fixed ISI setting (4 ms). For the noise masker the authors found no BMLD improvement, but for the talker interference, the BMLD was in the order of 6 dB for most conditions. The authors mainly attribute the improvement to informational masking effects. Similar to our findings, [6] found that which side lead and which side lagged also made essentially no difference to the results for both noise and speech.

In this context, it is noteworthy that the echo threshold, and therefore the ISI dependency on the lag's impact, is highly signal dependent as previous studies have shown – e.g., see [5]. However, more important than the actual values is the general trend that the perceptual influence of the lag on detecting the target highly depends on the ISI. For low ISIs, the lead-lag pair demonstrates spatial properties that perceptually differ from the spatial cues of the target, while for high ISIs, the lead-lag pair behaves like an uncorrelated noise pair. For high values of ISI, one of the underlying signals has the same spatial properties as the target, causing a relatively high masked detection threshold.

For the low ISI conditions, we assume that the signals interfere with each other and produce new spatial cues that can be used by the listener to spatially separate the target from the masker. Studies, such as [3] and [10] have shown that for small ISIs two interfering ITD-based signals that are identical except for a delay can produce very high Interaural Level Differences (ILD) due to interference effects. In both cases (high and low ISIs), the temporal order of the arrival of lead and lag does not matter.

4 CONCLUSIONS

In a sound localization tasks, the auditory system largely disregards the information from the later arriving signals if lead and lag are identical and the delay between both signals is within a few milliseconds. For detection tasks, however, the temporal order does not matter and the hypothesis that that auditory system completely removes the lag from the superposed signal (Research Hypothesis I) can be ruled out. These considerations, together with the results presented here, suggest that the PE does not need to be taken into account to explain our data. The interference between the direct sound and the reflection for small ISIs leads to energy reductions in isolated frequency bands. The magnitude of these reductions is in the same order as the threshold decrease measured from the 20-ms ISI condition. For small ISI's, the interference effects also create large shifts in binaural cues in isolated auditory bands.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant Nos. BCS-1539276 and NSF BCS-1539376, and the Cognitive and Immersive Systems Laboratory (CISL).

REFERENCES

[1] J. Braasch. *Auditory Localization and Detection in Multiple-Sound Source Scenarios*. VDI Verlag, Duesseldorf, 2002.

[2] J. Braasch. A precedence effect model to simulate localization dominance using an adaptive, stimulus parameter-based inhibition process. *J. Acoust. Soc. Am.*, 134(1):420–435, 2013.

[3] J. Braasch, J. Blauert, and T. Djelani. The precedence effect for noise bursts of different bandwidths. I. Psychoacoustical data. *Acoust. Sci. & Tech.*, 24:233–241, 2003.

[4] J. Braasch and M. T. Pastore. Binaural detection of a gaussian noise target in the presence of a lead/lag masker. *J. Acoust. Soc. Am.*, 143(1):EL49–EL54, 2018.

[5] J. Braasch, H. V. Savitala, and J. Blauert. Modeling the precedence effect for percussive sounds with different attack transients. In *19th International Congress on Acoustic*, pages PPA–05–001–IP, Madrid, Spain, Sept. 2-7 2007.

[6] R. L. Freyman, K. S. Helfer, D. D. McCall, and R. K. Clifton. The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America*, 106(6):3578–3588, 1999.

[7] M. D. Good, R. H. Gilkey, and J. M. Ball. *Binaural and Spatial Hearing in Real and Virtual Environment*, compendium The Relation Between Detection in Noise and Localization in Noise in the Free Field, pages 1108–1117. Lawrence Erlbaum Associates, 1997.

[8] K. Hartung and C. Trahiotis. Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. *J. Acoust. Soc. Am.*, 110:1505–1513, 2001.

[9] W. Lindemann. Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization of stationary signals. *J. Acoust. Soc. Am.*, 80:1608–1622, 1986.

[10] M. T. Pastore and J. Braasch. The precedence effect with increased lag level. *J. Acoust. Soc. Am.*, 138(4):2079–2089, 2015.

[11] W. A. Yost. *Fundamentals of Hearing: An Introduction*. Academic Press, San Diego, 4th edition edition, 2000.