ICA 2019 AACHEN

PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Spatial unmasking of circular moving sound sources in the free field

Norbert KOLOTZEK[1]; Bernhard U. SEEBER

Audio Information Processing, Technical University of Munich, Germany

ABSTRACT

This study investigates the unmasking of a horizontally moving tonal source in noise under free-field presentation. Thresholds of static and circularly moving 500 Hz sine tones presented in octave-band noise from the front were measured with four normal hearing listeners with an adaptive forced choice method. Six velocities were tested from 10°/s to 200°/s. Stimuli were of fixed duration, started in the front and moved to an azimuthal end position, determined by the overall probe tone duration in relation to their speed. The moving sound sources were generated with 17th order Ambisonics and played via 36 horizontally arranged loudspeakers in an anechoic chamber.

Preliminary results suggest that for the six velocities tested, spatial unmasking increased with increasing velocity up to 10 dB for 200°/s. However, unmasking remained substantially lower than for static sources, at the respective most lateral position reached in the movement trajectory at the end of the overall probe tone duration. This is consistent with observations in the literature on binaural sluggishness. Even for a small velocity of 50°/s, and therefore relatively slow changes in interaural phase, spatial unmasking decreased by 4 dB. This velocity is comparable to that of a walking person in 2 meters distance. It seems that in the free field, also slow changes in the interaural cues can lead to a decrease in binaural unmasking. This finding is important for extending binaural models to more realistic, dynamic listening situations in the free field.

Keywords: Moving Sound Sources, Spatial Unmasking, Binaural Hearing

1. INTRODUCTION

In our everyday life, many sound sources are in motion. In many realistic hearing situations, sound sources surround us that move with different velocities. Such sound sources are, for example, cars driving with relatively high velocity, or also pedestrians walking with a speed of up to 2 m/s. Besides the signal of interest, environmental noise sources might be present which can mask the target signal and increase detection thresholds. Because we receive two ear signals, phase differences between the left and right ear signals can be used to enhance the detection threshold in noisy situations (1, 2, 3). This difference in detection threshold is known as Binaural Masking Level Difference (BMLD).

Motion of a sound source, and therefore continuously changing sound source position, leads to time varying changes in the interaural level and interaural phase differences (ILD & IPD). Changes in the interaural cues can influence the BMLD (4, 5). Furthermore, it is known from the literature that the auditory system behaves sluggish on fast changes of interaural cues which can result in decreased detection thresholds (5, 6).

Several studies have measured BMLDs for interaural cues changing sinusoidally using stimuli presented over headphones (2, 3, 4). However, neither a sinusoidal stimulus variation nor the headphones condition represent a natural hearing situation that we deal with in everyday life. Normally, we perceive our environment through our own ears. In the free field, subjects listen with their individual head related transfer function (HRTF) which represents a more natural hearing situation. However, in the free-field, ILDs and IPDs co-vary in their natural fashion, which, to our knowledge, has not been considered in former BMLD measurements. Because ILD and IPD cannot be investigated separately from each other in a free field experiment, additionally a headphone

[1] norbert.kolotzek@tum.de

experiment was carried out with manipulated binaural recordings of the free field stimuli to separate ILD and IPD cues. This allows us to examine the individual contribution of interaural level and phase differences to binaural unmasking in the free field.

2. FREE FIELD EXPERIMENT

2.1 Stimuli

In this study, a 500 Hz sine tone was used to measure detection threshold in a static octave-band noise centered around 500 Hz. The probe tone had an effective duration of 300 milliseconds (ms) and additional 30 ms Gaussian shaped rise and fall times. Effective duration was measured after reaching 67.5% of the signal's amplitude. This results in an overall probe tone duration of 400 ms. The noise signal, with a duration of 600 ms, was played from a single loudspeaker at 0° in front of the listener, leading approximately to a binaural correlated noise (N_0 condition). To validate the binaural correlation of the noise signal, it was recorded with an artificial head at the listener's position. The measured binaural correlation coefficient was therefore 0.98. The level of the noise masker was set to 60 dB SPL at the listener's head center.

The stimuli were presented in a Simulated Open Field Environment (SOFE) (7) in an anechoic chamber with 36 horizontally arranged loudspeakers (Dynaudio BM6A mkII, Dynaudio, Skanderborg, Denmark) with a 10° spacing. The loudspeakers are mounted on a custom made 4.8 m x 4.8 m squared holding frame in a height of 1.4 m. The loudspeaker at 0°, in front of the listener, has a distance of 2.4 m to the listener's position. Equalization filters were used for the playback to compensate for the loudspeakers' frequency response and the time-of-arrival difference.

The moving sine tone was generated with two-dimensional 17th order Ambisonics with mode matching decoding, implemented with Matlab (Math Works, Inc, Nantick, US-MA). Positions of the motion trajectories and the resulting stimuli were calculated sample wise to ensure the correct phase information during the whole signal. Seven different constant sound source velocities were tested: 0°/s (as a reference, static at 0°), 10°/s, 30°/s, 50°/s, 100°/s, 150°/s and 200°/s. The starting position of the probe tone was always 0° in front of the listener and it moved clockwise at constant distance, i.e. on an arc. Additionally and for comparison, static stimuli were generated with 17^{th} order Ambisonics at the absolute lateral end position of the signal, corresponding to the azimuthal position reached with a given velocity at the end of the overall probe tone duration.

2.2 Subjects

Until now, four voluntary participants (two male) finished the experiment. Participant's age ranged from 23 to 27 (mean: 25.25, std.: 2.06). All participants had normal hearing thresholds with a hearing loss less than 15 dB up to 8 kHz as assessed with a clinical audiometer (Madsen Astera2, GN Otometrics A/S, Taastrup, Denmark). All subjects gave written consent and no one was payed for participating in the experiment. The study was approved by the ethics committee of the TUM, 65/18S.

2.3 Experimental procedure

All participants sat in a completely darkened anechoic chamber in the middle of the loudspeaker array. The detection threshold of the sine stimulus in noise was measured with a three-interval three-alternative-forced-choice method (3I-3AFC) with a three-down/one-up adaptive staircase procedure. The noise intervals were played back from a single loudspeaker at 0° with a fixed level of 60 dB SPL at the listener's position. Listeners' task was to indicate in which of the three noise intervals they perceived the sine tone by pressing the corresponding number on a keyboard on their knees. The initial level of the probe tone was set to 65 dB SPL. The initial step size was 5 dB, after the first two reversals the step size was decreased to 2 dB, after the next two reversals the step size was set to the final step size of 1 dB. The last 12 reversals at the final step size were used to calculate the threshold. The next randomly selected test condition of sound source velocity or position of the static sound source was started after the previous condition was finished, i.e. stimuli of each adaptive track were blocked.

2.4 Results

The measured thresholds of the first four subjects for moving as well as for a static sound sources is shown in Figure 1. In both conditions, a decrease in the threshold with increasing velocity or azimuthal location indicates a spatial unmasking of the sine tone in binaurally correlated noise. This difference compared to the reference condition at 0° is up to 14 dB in the static condition and up to

11 dB for moving sound sources for the highest tested velocity of 200°/s and thus lateral position. Even though the preliminary data show a spatial unmasking for both conditions, there is a clear difference between thresholds for moving and static sound sources presented in the free field. For static sound sources, detection threshold is lower than for moving sound sources especially for velocities larger than 30°/s. This indicates an effect of sluggishness of the auditory system, which was also reported in headphone studies (5, 6). It is noteworthy that even for a relatively small velocity of 50°/s, which corresponds to a walking speed in a distance of two meters, and therefore relatively slow changes in interaural phase difference, difference between static and moving condition is up to 6 dB.

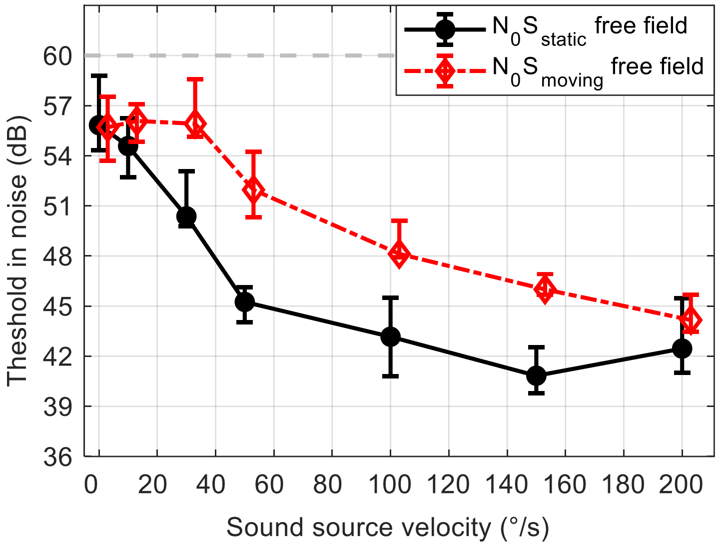

Figure 1 – Medians with upper and lower quartiles of detection thresholds in octave-band noise for a static stimulus (black solid line with filled circles) and moving stimulus (red, dash-dotted line with diamonds) at the most lateral end position of the trajectory.

3. HEADPHONE EXPERIMENT

3.1 Stimuli

To investigate the contribution of interaural cue differences to the detection threshold in the previous free field experiment, we manipulated the stimuli such that either ILD or IPD information of the free field condition is preserved dynamically in the stimulus while the other interaural cue is set to zero. Therefore, the moving as well as the static sine stimuli from the first experiment were recorded with the HMS II.3 artificial head with an anatomically formed pinna according to ITU-T P.57 (HMS II, Head acoustics GmbH, Herzogenrath, Germany). The artificial head was placed in the same loudspeaker array used in the previous experiment at the listeners' position. To extract the interaural level difference from the recorded binaural signals, the envelope of both ear signals was extracted and multiplied with an in-phase sine tone carrier with a frequency of 500 Hz. This stimulus preserves the ILD information (only ILD) of the previous free field experiment combined with a diotic carrier tone which effectively points to the front.

To extract the interaural phase difference of the free field stimuli, the binaural recordings were divided by the signal's envelope separately for the left and the right ear signal. This results in a stimulus with only the IPD information as in the free field experiment (only IPD), without any interaural level difference. The amplitude as well as the onset of the only IPD stimuli was corrected to ensure the correct level and identical onset slope.

The binaural correlated noise masker, presented from 0° in the free field experiment, was also recorded with the artificial head and was not manipulated. This signal generation ensures a separation of ILD and IPD information but preserves the same interaural information as in the previous experiment.

For the second experiment, only five velocity conditions were used. 10°/s and 150°/s were excluded to shorten the experiment since no unusual effects was seen in Exp. 1 for these velocities.

3.2 Subjects

All four subjects from the first were also included in the second experiment.

3.3 Experimental procedure

Experimental procedure was the same as in the experiment before except that the binaurally recorded and manipulated stimuli were presented via headphones and participants sat in a sound-attenuated booth. They again heard three successive noise intervals and had to indicate the interval in with they perceived the probe tone. An adaptive three-down/one-up staircase procedure was used and the threshold was calculated from the last 12 reversals. Six different conditions were tested: a reference condition with non-manipulated binaural recordings containing both interaural cues as in the free field, the only ILD and the only IPD condition each for both static and moving sine tone stimuli. As in experiment 1, the next test condition of sound source velocity or position of the static sound source and binaural condition (binaural recording, only ILD or only IPD) was started after the previous condition was finished.

3.4 Results

The measured thresholds for the second experiment with manipulated binaural cues as well as for the non-manipulated reference binaural recordings are shown in Figure 2. For the non-manipulated binaural recordings spatial unmasking is up to 12 dB for static and 10 dB for a moving sound source in the 200°/s velocity condition.

In the only ILD condition, spatial unmasking is substantially lower compared to the non-manipulated binaural recordings, especially for velocities higher than 50°/s. The results show that for the highest velocity condition tested, spatial unmasking is 3 dB for a moving stimulus and 6 dB for the corresponding static sound source when only ILD information is preserved in the stimulus.

The data for the only IPD condition seem to match the thresholds for the non-manipulated binaural recordings in most conditions, thus the only ILD condition is not fully additive with the IPD thresholds. It seems that detection thresholds of a 500 Hz sine tone in noise presented in the free field are mainly driven by the IPD information of the binaural signals. This is in contrast to findings for spatial unmasking of speech signals. Previous studies assume that better-ear-listening, mainly caused by the ILDs, and binaural unmasking caused by IPDs is fully additive to the overall spatial benefit on a speech signal (8, 9).

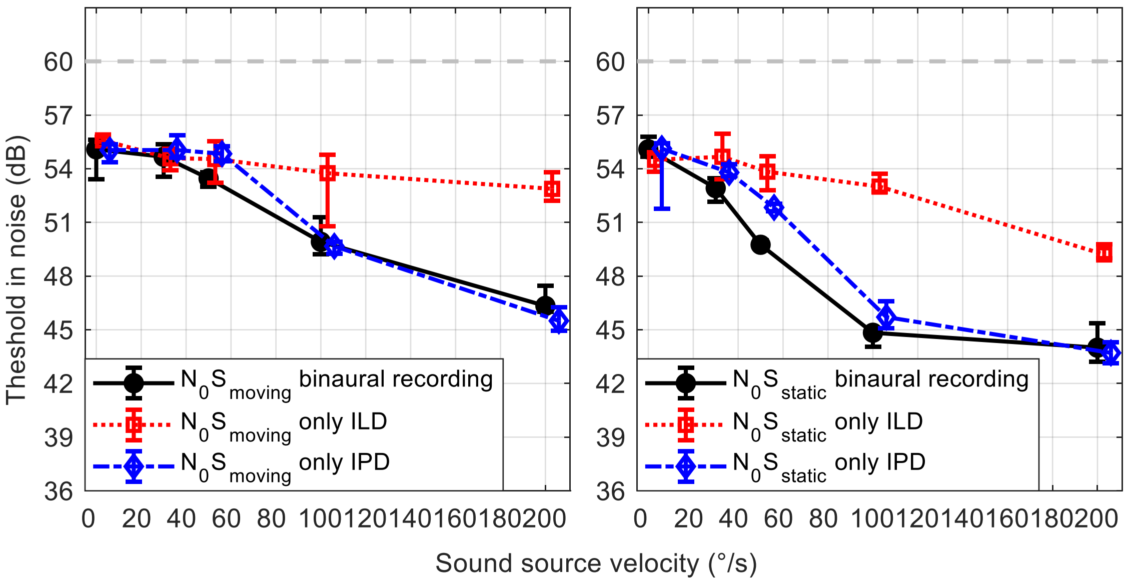

Figure 2 – Median with upper and lower quartiles of the measured threshold in octave-band noise for a moving stimulus (left pannel) and static stimulus (right panel) for the binaural recording (black solid lines with filled circles), only ILD cues (red dotted lines with squares) and only IPD cues (blue dashed-dotted lines with diamonds). All stimuli were generated from binaural recordings of the free field stimuli and played back via headphones.

4. CONCLUSIONS

This study investigates the binaural unmasking of moving or static sine tones in noise presented in free field. Additionally, this study investigates the contribution of interaural phase and level differences to the binaural unmasking.

Results from this preliminary study show that for a moving sound source, detection threshold is increased compared to a static sound source located at the absolute end position of the movement trajectory. This is in accordance with findings on the sluggish behavior of the auditory system to fast changes in interaural phase differences. Noteworthy is here that even for slow velocities binaural unmasking decreases by 6 dB for moving sound sources.

Results of the second experiment show that binaural unmasking is mainly driven by changes in the interaural phase difference. The interaural level difference does not additionally contribute to binaural unmasking nor does it hinder it, even though lower detection thresholds were also observed in the only ILD condition, indicating spatial unmasking solely based on ILDs.

These findings suggest that both IPD and ILD information contribute to binaural unmasking in the free field but are not fully additive as it is assumed for speech signals.

This may not be taken into account in current auditory models and is a step in extending such models to dynamic scenes presented in the free field.

ACKNOWLEDGEMENTS

This study was funded by DFG SFB 1330 project C5 and the rtSOFE system by BMBF 01 GQ 1004B.

REFERENCES

1. Licklider JCR. The Influence of Interaural Phase Relations upon Masking of Speech by White Noise. J Acoust Soc Am. 1948; 20(2):150-159.
2. van de Par S, Kohlrausch A. A new approach to comparing binaural masking level differences at low and high frequencies. J Acoust Soc Am. 1997; 101(3):1671-1680.
3. van de Par S, Kohlrausch A. Dependence of binaural masking level difference on center frequency, masker bandwidth, and interaural parameters. J Acoust Soc Am. 1999; 106(4):1940-1947.
4. Breebaart J, van de Par S, Kohlrausch A. The contribution of static and dynamic ITDs and IIDs to binaural detection. J Acoust Soc Am. 1999; 106(2):979-992.
5. Grantham DW, Wightman FL. Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. J Acoust Soc Am. 1979; 65(6):1509-1517.
6. Holube I, Kinkel M, Kollmeier B. Binaural and monaural auditory filter bandwidths and time constraints in probe tone detection experiments. J Acoust Soc Am. 1998; 104(4):2412-2425.
7. Seeber, BU, Kerber S., and Hafter ER. A System to Simulate and Reproduce Audio-Visual Environments for Spatial Hearing Research. Hear Res. 2010; 260:1-10.
8. Culling JF, Mansell ER. Speech intelligibility among modulated and spatially distributed noise sources. J Acoust Soc Am. 2012; 133(4):2254-2261.
9. Jelfs S, Culling JF, Lavandiere M. Revision and validation of a binaural model for speech intelligibility in noise. Hearing Research. 2011; 93(1-2):96-104.