Onset importance in a binaural detection task during an arc-shaped movement trajectory

Norbert Kolotzek & Bernhard U. Seeber
Audio Information Processing, Technical University of Munich, 80333 Munich
{norbert.kolotzek, seeber}@tum.de

Introduction

Moving sound sources and dynamic scenes are attracting increasing research interest since they are plentiful in everyday life. We are also surrounded by static sound sources which influence the ability to detect sounds of interest. If a sound source is moving while a static noise masker is present, the detection threshold will change. This is mainly caused by varying interaural cues like level differences (ILDs) and time differences (ITDs). For frequencies below 1500 Hz, ITDs are dominantly used over ILDs by humans to evaluate the location of a sound source [1, 2] and to unmask a signal in binaurally correlated noise, which is known as the binaural masking level difference [3]. Kolotzek & Seeber [4] showed that if a 500 Hz tone is moving from the front to a lateral position in the presence of a static noise masker in the front, thresholds are higher than for a static presentation at the most lateral position reached by the movement trajectory. This is in accordance with previous findings showing less unmasking of signals with time varying interaural cues, as induced by, e.g., a moving sound source in the free field, which is referred to as a sluggish behavior of the auditory system [5, 6].

ITDs as well as ILDs are also used to lateralize a sound source. The literature here shows a higher weighting of these interaural cues at the onset of the stimulus [7].

Stecker & Brown [8] measured thresholds of detecting a changing ITD for short signals in which an ITD was either presented at the onset and varied to 0 µsec or vice versa. Thresholds were lower when the ITD was presented at the onset of the stimulus. They therefore confirmed the onset dominance of binaural cues also for stimuli with time varying interaural differences.

The stimuli used in Kolotzek & Seeber [4] to measure detection thresholds of moving stimuli in noise always started in the front, collocated with the static noise masker. The maximum unmasking was therefore always present at the offset of the movement trajectory. It is of interest whether the onset dominance of binaural cues also extends to binaural detection. In the experiment, this would be seen in lower thresholds if the maximum possible unmasking is at the onset of the stimulus. This paper presents a pilot experiment to investigate the temporal weighting of binaural unmasking in the free field. A 500 Hz tone was used to measure detection thresholds in noise of a moving stimulus. The movement either started at higher angular positions and moved towards lower ones or vice versa. For comparison, also detection thresholds of static sine tones were measured.

Methods

Stimuli

In this study, a 500 Hz sine tone, which was either moving or static, was used as target sound to be detected by the participants in the presence of a static bandpass noise from the front. The target probe tone had either an effective duration of 40 or 300 ms with 5 ms Gaussian shaped rise and fall times. Five equally distributed arc-shaped trajectory segments of 20° were tested in the range from -50° to +50°; positive angles correspond to the right hemisphere. The moving stimuli passed these 20° arc within the effective signal duration mentioned before. The movement for each segment was either clockwise or counterclockwise to present the maximum possible binaural cues either at the onset or at the offset of the signal. Additionally, and for comparison, static stimuli were generated corresponding to the onset and offset azimuth angles of the trajectory segments. Figure 1 illustrates the different segments as well as the static positions. The arrows indicating the direction of movement, whereas blue arrows indicate stimuli with the maximum possible binaural cues at the signal onset and green arrows the maximum cues at the offset of the movement.

All sine tone stimuli were samplewise pre-calculated using 17th-order Ambisonics [9] with *max_r$_E$* decoding [10] and played over the 36 horizontally arranged loudspeakers of the Simulated Open Field Environment [11].

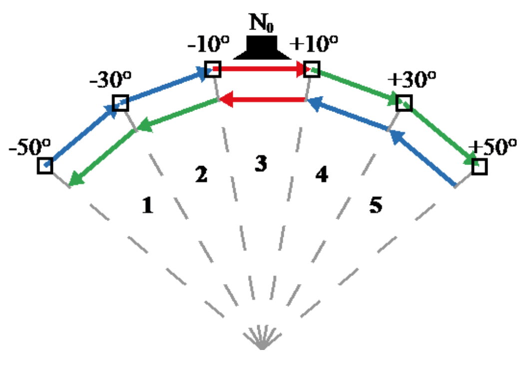

Figure 1: Tested movement trajectory segments of 20° ranging from -50° to +50°. The arrows indicate the direction of movement, with the maximum possible binaural cues at the signal onset (blue) or at the offset of the trajectory (green). The red arrows show the frontal segment where binaural cues have the identical magnitude at on- and offset. The black squares correspond to the static angular positions tested as a reference. The loudspeaker symbol at 0° indicates the position of the noise masker. Each segment is numbered with a digit from 1 to 5.

Uniform exciting noise was used as masker [12]. The noise was band-limited from 250 Hz to 750 Hz. It had an overall duration of 500 ms with 10 ms Gaussian rise and fall times. The noise source had a sound pressure level of 60 dB at the listener's position. The noise was played from a single loudspeaker at 0° leading to highly binaurally correlated noise (almost N_0).

Procedure

The participants sat in the completely darkened anechoic chamber in the center of the loudspeaker array. Detection thresholds of the moving or static sine tone in noise from the front were determined with a three-interval three-alternative-forced-choice method (3I-3AFC) using a two-down/one-up adaptive staircase procedure [13] tracking the 71% point of the psychometric function. Participants listened to three intervals of the anechoic uniform exciting bandpass noise, separated by an inter-stimulus-interval of 200 ms. To one of these intervals the sine tone was added, either static or moving. The listeners' task was to indicate which interval differed from the others by pressing the corresponding number on a keyboard. The overall level of the sine tone was then adjusted depending on a right or wrong answer. The mean of the last eight reversals at the final step size of 1 dB was used to calculate the detection threshold in noise. Before a new random test condition started, the previous one had to be finished (no interleaved tracks). Each subject finished one track for each condition combination of target stimulus duration, trajectory segment or static position and onset condition, resulting in 32 tracks for each subject. Subjects finished the experiment on average in 2,6 hours.

Participants

For this preliminary experiment, four male participants volunteered. Their age ranged from 23 to 30 years (mean: 26,6 yr, std: 3,5). All participants had a self-reported normal hearing threshold. They gave written consent and were not payed for participating in the experiment. The study was approved by the ethics committee of the TUM, 65/18S.

Experimental results

Figure 2 shows medians and quartiles of the measured thresholds averaged across symmetric segments, since the raw data showed symmetric performance in the left and right hemisphere. The following section will first discuss results for the short 40 ms target signal followed by the results of a longer signal duration of 300 ms.

Stimulus duration of 40 ms

The measured thresholds for the short stimulus with static presentation show decreasing detection thresholds with increasing lateral angle, which was expected and is in accordance with previous experiments [3, 4]. If the sound source is moving between ±10° and passing 0° (segment 3), thresholds are around 4 dB higher than for the tested static boundary position. In segments 2/4 as well as in segments 1/5 detection thresholds of the moving stimuli stay in between measured thresholds for the margin angles of the corresponding segments. Note that almost the same detection thresholds are reached irrespective if higher binaural cues occur at the onset or at the offset.

Stimulus duration of 300 ms

With a longer signal duration the measured detection thresholds decrease by about 5 dB. A difference to the short stimulus duration of 40 ms can be observed in segment 3. Here the measured detection thresholds of a static stimulus at ±10° is almost the same as for a moving stimulus starting and ending at ±10°. If the movement trajectory does not have the same onset and offset cues (segment 2/4 as well as 1/5), detection thresholds of the 300 ms stimuli correspond to those measured at the most lateral boundary position of the segments. Interestingly, this trend appears to be independent of whether the maximum possible unmasking is at the onset or at the offset of the movement trajectory. This suggests that if the stimulus duration is long enough, the most unmasked position can be evaluated independently of the temporal appearance of the most lateral position.

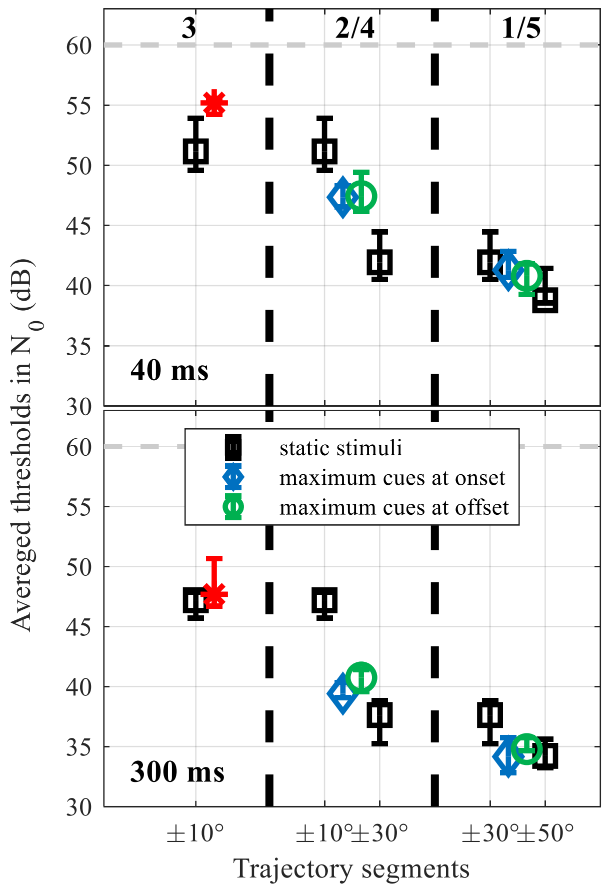

Figure 2:. Measured detection threshold of a 500 Hz sine tone in noise from the front (N_0) averaged across symmetric segments. The upper panel shows measured data for a stimulus duration of 40 ms, the lower panel for 300 ms. Static positions are plotted as black squares, blue diamonds correspond to detection thresholds of a moving stimulus with the maximum possible binaural cues at the signal onset, while green circles depict the offset. Again, in red, detection thresholds are plotted for the ±10° condition.

Conclusions

Results of this preliminary study show that if the stimulus duration is long enough, measured thresholds for a moving stimulus tend to be similar to those at the most lateral position, where the maximum binaural unmasking is obtained. With a short stimulus duration of 40 ms, the detection threshold of the moving stimulus lies between the thresholds determined at the boundary positions of the trajectory segment. These results suggest an integration time constant which is known as binaural sluggishness. It is in accordance with previous literature on binaural sluggishness that a fast moving stimulus has a higher detection threshold than a static stimulus of the same duration at the most lateral position of the trajectory. However, this effect can only be observed for a short stimulus duration in the current experiment. With a stimulus duration of 300 ms, the detection thresholds determined for the moving and the static stimulus at the most lateral position are approximately the same. This is in accordance with the time constant reported in the literature of approximately 200-300 ms for binaural sluggishness [6].

Further, the results of this pilot experiment show that the detection thresholds of a moving stimulus with maximum interaural cues at onset are approximately the same as for a stimulus with maximum cues at offset. This effect seems to be independent of the stimulus duration, since it can be observed for short (40 ms) as well as for longer (300 ms) moving stimuli. Therefore, no onset dominance for detection thresholds can be observed, as shown for lateralization or ITD detection thresholds [7, 8].

Acknowledgements

This study was funded by TUM and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 352015383 – SFB 1330 C5. The rtSOFE system was funded by BMBF 01 GQ 1004B.

Literature

[1] Blauert, J.: Spatial hearing: the psychology of human sound localization, Cambridge, Mass, MIT Press, 1999.

[2] Wightman, F.L. and Kistler, D.J.: Resolution of front–back ambiguity in spatial hearing by listener and source movement, J. Acoust. Soc. Am. 105 (1999), 2841-2853.

[3] Breebart, J., van de Par, S., and Kohlrausch, A.: The contribution of static and dynamic varying ITDs and IIDs to binaural detection, J. Acoust. Soc. Am. 106 (1999), 979-992.

[4] Kolotzek, N. and Seeber, B.U.: Spatial unmasking of circular moving sound sources in the free field, Proc. of 23rd International Congress on Acoustics, integrating 4th EAA Euroregio 2019, Aachen, Germany, German Acoustical Society e.V. (DEGA) (2019), 1609-1614.

[5] Grantham, D.W. and Wightman, F.L.: Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation, J. Acoust. Soc. Am., 65 (1979), 1509-1517.

[6] Holube, I., Kinkel, M., and Kollmeier, B.: Binaural and monaural auditory filter bandwidths and time constrants in probe tone detection experiments. J. Acoust. Soc. Am. 104 (1998), 2412-2425.

[7] Houtgast, T. and Aoki, S.: Stimulus-onset dominance in the perception of binaural information. Hear. Res. 72 (1996), 26-36.

[8] Stecker, G.C. and Hafter, E.: Temporal weighting in sound localization, J. Acoust. Soc. Am. 112 (2002), 1046-1057.

[9] F. Zotter F. and Frank, M.: Ambisonics - A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality, Heidelberg, Germany, Springer, 2019.

[10] Daniel, J.: Représentation de champs acoustiques, application à la transmission et à le reproduction de scènes sonores complexes dans un context multimèdia, Université de Paris, Paris, France, 2001

[11] Seeber, B.U., Kerber, S. and Hafter, E.R.: A system to simulate and reproduce audio-visual environments for spatial hearing research. Hear. Res. 260 (2010), 1-10.

[12] Fastl, H. and Zwicker, E.: Psychoacoustics Facts and Models, Berlin Heidelberg, Springer, 2007.

[13] Levitt, H.: Transformed Up-Down Methods in Psychoacoustics. J. Acoust. Soc. Am. 49 (1971), 467-477.