DOI: 10.61782/fa.2023.1081

CONTEXTUAL LOCALIZATION BIAS FOR LOW- AND HIGH-FREQUENCY STIMULI

Bernhard Laback[1*] **Jonas Schültke**
[1] Acoustics Research Institute, Austrian Academy of Sciences, Austria

ABSTRACT

The perceived azimuth of a sound source is biased by a preceding source (precursor), typically, towards midline (medial) by a lateral and towards the side (lateral) by a central precursor. Little is known about effects of intermediate precursor azimuths and the contribution of low and high frequency regions. We tested the hypothesis that for a certain intermediate precursor azimuth, lateral and central biases cancel each other out. Ten normal-hearing listeners localized 300-ms targets following 600-ms precursors using a head-pointing task in a virtual audio-visual environment. Both target and precursor azimuths were systematically varied across the entire azimuth range. Stimuli were white noises, filtered with listener-specific head-related transfer functions. Low-pass (0.5-2 kHz), high-pass (2.8-16 kHz), and broadband (0.5-16 kHz) stimuli were tested to study effects in frequency regions dominated by different localization cues: interaural time differences in low-pass, interaural level differences and spectral shape in high-pass, and all three cues in broadband stimuli. Precursor effects were overall strongest for target azimuths of +/-70°. Cancellation of lateral and central biases was found only for +/-70°-targets, at a precursor azimuth of 58.3°. Importantly, the data showed selective spatial contrast enhancement for targets preceded by azimuthally matched Ps. Patterns of precursor effects were relatively similar across frequency regions.

Keywords: sound *sound localization, context, bias, azimuth, frequency region*

1. INTRODUCTION

Azimuthal sound localization is pivotal in everyday life, e.g., when determining the direction of an approaching car. Nevertheless, it is known for a long time that absolute azimuthal localization of a target sound (T) is biased by the azimuth of a preceding sound (P, precursor). This is called the localization aftereffect (LA, e.g., [1]). Best known effects are a medial bias (towards midline) by a lateral P and a lateral bias (towards sides) by a central P. Since these effects represent repulsion of T from P, it has been argued that ecological pressure may favor azimuthal segregation of competing sound sources over absolute localization [2]. According to this interpretation, LA corresponds to spatial contrast enhancement of subsequent sounds ([1-3]). However, before contrast enhancement can be accepted as a “driving force” underlying the LA, its properties need to be characterized further.

First, the literature on LA is mostly confined to P stimuli being either fully lateral or central. Thus, effects of intermediate P azimuths are largely unknown. To that end, the current study systematically varied the azimuth of P and T stimuli across the full azimuthal range from left (-90°) to right (+90°). This allowed, for the first time, to rigorously address the question if LA can indeed be described as a generally repulsive mechanism that always increases spatial contrast, or if also attraction is at work in certain spatial configurations (see e.g., [4]). Further, it allowed to test the hypothesis that for a certain intermediate precursor azimuth (between central and fully lateral) the lateral and medial biases cancel each other, resulting in an absence of bias.

Second, LAs may depend on the frequency region and, therefore, on the localization cues available in each region. Medial and lateral biases have been reported for both low- and high frequency stimuli where interaural time difference (ITD) and interaural level difference (ILD) cues dominate, respectively [5-6]. Different frequency regions have not been compared, however, in the same study. To study the frequency-range-specificity of the LAs, the

---

[*] ***Corresponding author***: Bernhard.Laback@oeaw.ac.at

current study compared LAs between low-frequency, high-frequency, and broadband conditions.

2. METHODS

Ten normal-hearing participants participated in the experiment. Prior to the experiment, listener specific head-related-transfer-functions (HRTFs) were measured using the method described in [7]. Using virtual acoustics technique, the HRTFs were then used to generate virtual stimulus locations along the frontal azimuthal range. All combinations of P and T azimuths of 0, ±10, ±30, ±50, ±70 and ±90° were tested. As reference, each T azimuth was also tested without precursor (NoP). Listeners were instructed to indicate the perceived azimuth of T stimuli by head-pointing in a virtual audio-visual environment using headphones and a head-mounted display (see [6]). P and T stimuli had a duration of 600 and 300 ms, respectively, and were separated by a silent gap of 10 ms. Other aspects of the paradigm were as described in [8]. Stimuli were random white noises that were bandpass-filtered in three frequency regions, referred to as low-pass (LP, 0.5-2 kHz), high-pass (HP, 2.8-16 kHz), and broadband (BB, 0.5-16 kHz). Each combination of P/T azimuth and frequency region was tested six times. The different frequency regions were tested in separate blocks, with balanced order across participants. Before each block, listeners received localization training based on visual reinforcement as described in [6].

3. RESULTS

The reference conditions without precursor (NoP, not shown) revealed minimal overall bias (0.22° leftward) and showed mean root-mean-square localization error (RMSE) of 4.8° at midline and of 11.1° across all T azimuths. These values are very close to corresponding errors in a free-field experiment (4.6° and 11.1°, respectively [9]. The RMSE was slightly but significantly higher ($p = 0.005$) for condition LP (12.3°) than condition BB (9.9°), with condition HP lying in between (11°).

The LA was defined as the shift in perceived azimuth of T resulting from adding a P [i.e., the response azimuth in trials with P minus the response azimuth in trials without P (condition NoP)]. In the following subsections, we address the idea of cancellation of medial and lateral biases (section 2.2) and the idea of spatial contrast enhancement around the azimuth of a P.

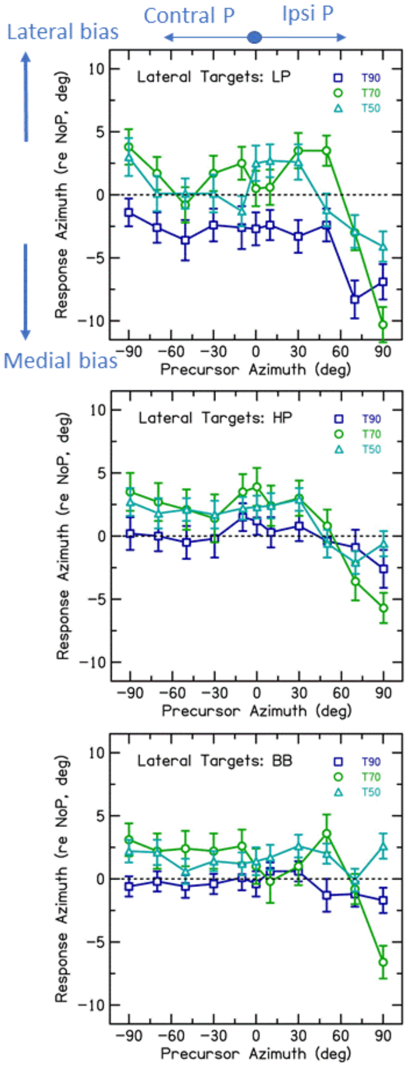

Figure 1. Localization aftereffect (LA), quantified as the difference between response azimuth in trials with P *re* in trials without P (NoP), as a function of precursor (P) azimuth for the three most lateral target (T) azimuths (±50, ±70, and ±90°). The three panels from top to bottom show data for the three frequency conditions (Low-pass [P], high-pass [HP], and bandpass [BB]). Data below the horizontal dotted line indicate medial bias and data above that line indicate lateral bias. Data on the right side (positive P azimuth) and left side (negative P azimuth) indicate ipsilateral and contralateral conditions, respectively. Error bars indicate 95% confidence intervals.

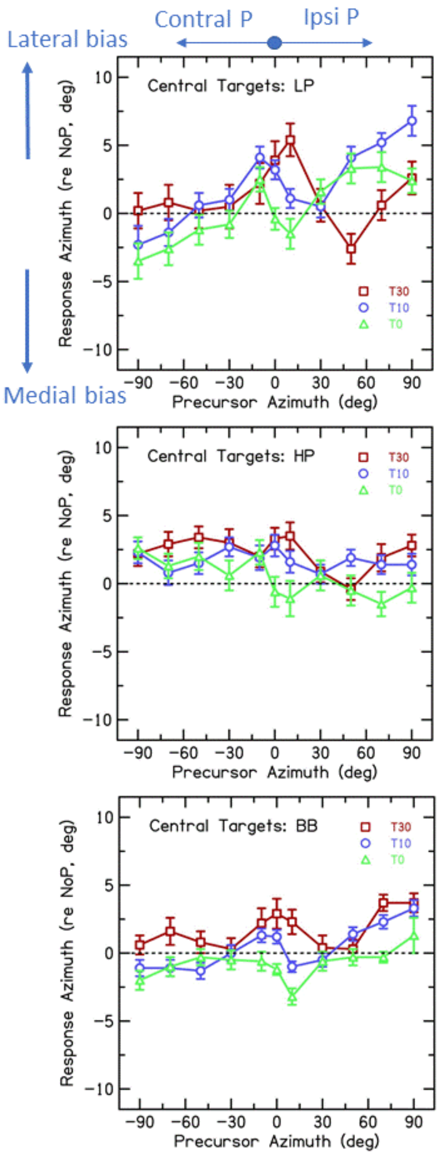

Figure 2. Same as Fig. 1, but for the three smaller T azimuths (0, ±10, and ±30°).

3.1 Is there an intermediate P azimuth for which medial and lateral LAs cancel each other?

Fig. 1 shows the LA as a function of P azimuth for the three most lateral T azimuths (±50, ±70, and ±90°). The three panels from top to bottom show data for the three frequency conditions (LP, HP and BB). In each panel, data are collapsed and averaged across left- and right-side T azimuths, such that on the right side of the graph (positive P azimuths) Ts are ipsilateral and on the left side of the graph (negative P azimuths) Ts are contralateral to the P stimuli. Data points below the horizontal (dotted) line indicate medial bias and data points above the line indicate lateral bias. According to the hypothesis, it was expected that ipsilateral Ps cause medial bias (data points in the lower right corner), that decreasing P azimuth reduces medial bias while crossing zero bias at some intermediate azimuth, and that decreasing P azimuth further induces lateral bias that peaks for an P azimuth of zero. The actual data followed these predictions only partly. Specifically, ipsilateral Ps showed a trend towards medial bias and central Ps showed a trend towards lateral bias. This appeared to depend, however, strongly on the T azimuth and, to some extent, on the frequency condition. For the lateral-most T (90°, referred to as T90), the bias almost always remained medial and hardly crossed zero. In contrast, for T50, the bias tended to be more lateral, at least for conditions HP and BB. For T70, and possibly also for T50 in the LP condition, the curves appear to best represent the predicted transition from medial bias for ipsilateral Ps to lateral bias for more central Ps. Note, however, that the bias often oscillated, e.g., sometimes resulting in zero bias at Ps around zero. A cubic polynomial function was fit to the data for T70 and the P azimuth corresponding to zero bias was determined. The zero-bias points were quite similar across the LP, HP, and BB regions (59.3, 56.6, and 58.7°, respectively).

Fig. 2 shows the corresponding data for the three smaller T azimuths (0, ±10, and ±30°). These data show no trend at all in the expected direction. However, there appears to be a pronounced oscillation in the bias patterns, particularly for P azimuths between 0 and 70°. This pattern is analyzed in the following subsection.

Overall, the data in Fig. 1 and 2 are not consistent with the idea that there is a general transition from medial to lateral bias (when shifting the P azimuth from ipsilateral to central) that is independent of the T azimuth.

3.2 Local dilation of auditory space

The oscillations in the bias effect in Figs. 1 and 2 may suggest that the effective amount of bias depends critically on the azimuthal distance between P and T stimuli. This leads to the idea (already proposed in previous studies, e.g. [2,10-11]) that a P with a certain azimuth may selectively increase the spatial contrast in auditory space in the vicinity of that P. Note that in contrast to previous studies, P azimuth was randomly changed from trial to trial, thus, any contrast enhancement must be fast. To inspect the presence of such an effect, Fig. 3 plots LAs for ranges of T azimuths centered around each P azimuth, the latter being indicated by vertical arrows (with colors and symbols according to the legend). Only "matched" conditions are shown, where T azimuths surround a given P azimuth. The important aspect of these curves is here not the absolute LA (i.e., the position along the ordinate), but their slope. Curves with a positive slope indicate enhancement of local spatial contrast, i.e. dilation of auditory space in the spatial vicinity of P. Flat curves (slope of zero) indicate no change in contrast. Curves with a negative slope indicate reduced spatial contrast, i.e., compression of local auditory space. In 16 out of 18 cases (across the three frequency conditions), at least one segment of the functions showed a positive slope (revealed by a significant interaction of factors "T azimuth" and "Presence of P" in repeated-measures ANOVA: $p \leq 0.023$). Only for P90 the slope was found to be negative in the LP condition ($p = 0.02$). As control conditions, slopes for T azimuths remote from P azimuths ("unmatched" azimuths) were analyzed (not shown). In 339 out of the 345 cases the slope was not significantly positive ($p > 0.05$). While these statistics did not correct for alpha-error inflation, the binomial probability distribution was used to judge the likelihood that these results were obtained by chance. Obtaining 16 "successes" or more out of 18 has a chance probability $p = 0.0007$ and obtaining 339 or more out of 345 has a probability $p < 0.0001$. Thus, we can conclude that the increase in spatial contrast around the azimuth of Ps is a highly significant effect. There were no pronounced differences across frequency regions.

4. DISCUSSION AND CONCLUSIONS

The localization aftereffect (LA), although a well-known phenomenon that likely affects absolute sound localization in everyday situations, is still insufficiently studied and poorly understood. The current study relied on the paradigm of [8], measuring the within-trial LA induced by Ps whose azimuth was randomly varied from trial to trial and employing a stimulus timing that roughly approximates random switching of sound sources (like speakers) in a multi-source environment such as a cafeteria. The first main goal was to study a much wider range of P azimuths than previously employed and, thus, to obtain a broader picture of the effects involved and mechanism underlying the LA. Specifically, we tested the hypothesis that for a particular intermediate P azimuth (between central and fully lateral) the lateral and medial biases would cancel each other. Some evidence for such a cancellation was found only for target azimuths of ±70° and the point of bias cancellation occurred at 58.2°, on average across the three frequency conditions tested. For all other T azimuths, the bias function either did not even reach the cancellation point (remaining medial or lateral across all P azimuths) or oscillated around the cancellation point. This outcome suggests that the LA is

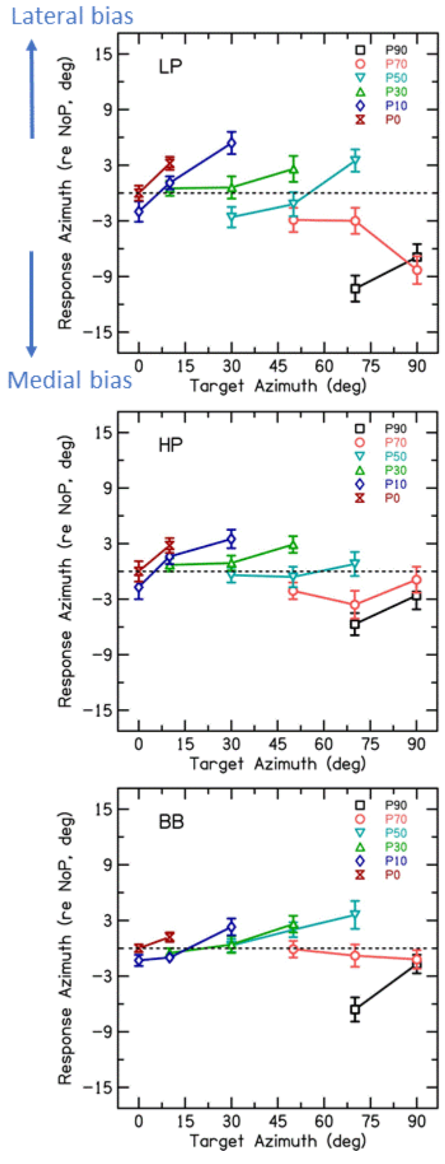

Figure 3. LAs for ranges of T azimuths that surround each P azimuth, as indicated by vertical arrows (with colors and symbols as defined in the legend). Curves with a positive slope indicate contrast enhancement, flat curves (slope of zero) indicate no change in contrast, and curves with a negative slope indicate contrast reduction. All other aspects are as in Fig. 1.

spatially restricted, i.e. occurs predominantly for Ts azimuthally matched to a given P. Note that this differs from the results of LA studies using either ILD cues [8] or ITD cues [4], where the LA was much more spread across azimuths.

This outcome inspired a follow-up analysis on the effect of a P on the local spatial contrast, i.e. on the slope of the lateralization function at the P azimuth. While previous studies have already found support for the idea of spatial contrast enhancement (e.g., [2, 10-12]; but see also [13]), those studies tested only a few P azimuths, allowing no generalization. The present results showed contrast enhancement for 16 out of 18 "matched" conditions (6 P-azimuths $*$ 3 frequency conditions) where the T and P azimuths were matched. In contrast, for the 345 unmatched conditions, only 339 showed contrast enhancement. These results suggest that at least under the conditions tested in this study, the LA is predominantly a repulsive mechanism that helps to increase azimuthal separation between subsequent sounds. Furthermore, the effect appears to be quite fast, given that the P azimuth was varied from trial to trial. For completeness, note that also attractive LA has been reported for rather short P and T stimuli which might explained by perceptual integration of target and distractor [14].

The second main goal was to study the impact of the frequency region on the LA. Overall, the patterns of LAs were quite comparable between the LP, HP and BB regions. Overall, this supports previous findings that both ITD cues (dominant at low frequencies) and ILD cues (dominant at high frequencies) are susceptible to LA. On a more detailed level, the LA might be slightly stronger for the LP region than for the HP and BB regions, which, however, does not seem to be significant (but respective analyses are only preliminary so far). If LA for conditions including the high frequency region (HP and BB) should indeed turn out to be smaller, this might be related to the availability of spectral localization cues for those conditions. However, answering this question thoroughly would require controlled manipulation of the availability of HRTF cues while keeping all other parameters (like the frequency region) constant.

Finally, it is important to note that the conclusions drawn here are preliminary and that more detailed and rigorous quantitative and statistical analyses of the data are required in follow-up work.

5. ACKNOWLEDGMENTS

We would like to thank Michael Mihocic for implementing some of the software routines used in the experiment.

6. REFERENCES

[1] S. Carlile, S. Hyams, and S. Delaney: "Systematic distortions of auditory space perception following prolonged exposure to broadband noise," *J. Acoust. Soc. Am.* 110(1), pp. 416-424, 2001.

[2] A. Lingner, M. Pecka, C. Leibold, and B. Grothe: "A novel concept for dynamic adjustment of auditory space," *Sci. Rep.* 29 (8), pp. 8335, 2018.

[3] M. Kashino and S. Nishida: "Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect," *J. Acoust. Soc. Am.* 103, pp. 597-604, 1998.

[4] D. Phillips, E. Mew, and S. Hall: "Selective adaptation in sound lateralization is not due to a repulsion effect," *J. Acoust. Soc. Am.* 136, EL424, 2014.

[5] E. A. Macpherson and J. C. Middlebrooks: "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am. 111*(5 Pt 1), pp. 2219–2236, 2002.

[6] M. Klingel and B. Laback: "Binaural-cue Weighting and Training-Induced Reweighting Across Frequencies," *Trends in Hear. 26*, 23312165221104870, 2022.

[7] P. Majdak, P. Balazs, and B Laback: "Multiple exponential sweep method for fast measurement of head-related transfer functions," *J. Audio Eng. Soc.* 55, pp. 623-637, 2007.

[8] B. Laback: "Contextual Lateralization Based on Interaural Level Differences Is Preshaped by the Auditory Periphery and Predominantly Immune Against Sequential Segregation," *Trends Hear.* 27, 23312165231171988, 2023.

[9] W. H. Slattery and J. C. Middlebrooks: "Monaural sound localization: acute versus chronic unilateral impairment," *Hear Res.* 75(1-2), pp. 38-46, 1994.

[10] J. Dahmen P. Keating, F. Nodal, A. Schulz, and A. King: "Adaptation to stimulus statistics in the perception and neural representation of auditory space," *Neuron* 24, pp. 937-948, 2010.

[11] S. Getzmann: "Spatial discrimination of sound sources in the horizontal plane following an adapter sound," *Hear Res.* 191(1-2), pp.14-20, 2004.

[12] H. Gleiss, J. Encke, A. Lingner, T. R. Jennings, S. Brosel, L. Kunz, B. Grothe, and M. Pecka: "Cooperative population coding facilitates efficient sound-source separability by adaptation to input statistics," *PLoS Biol.* 29, 17 (7), e3000150, 2019.

[13] J. K. Maier, P. Hehrmann, N. P. Harper, G. M. Klump, D. Pressnitzer, and D. McAlpine: "Adaptive coding is constrained to midline locations in a spatial listening task," *J. Neurophysiol.* 108(7), pp.1856-1868, 2012.

[14] N. Kopčo, G. Andrejková, V. Best, and B. Shinn-Cunningham: "Streaming and sound localization with a preceding distractor," *J. Acoust. Soc. Am.* 141(4), EL331, 2017.