



OPEN ACCESS

EDITED BY

Achim Klug,
University of Colorado Anschutz Medical
Campus, United States

REVIEWED BY

Benjamin Stodt,
Leibniz Research Centre for Working
Environment and Human Factors (IfADo),
Germany
Katarina Poole,
Imperial College London, United Kingdom

*CORRESPONDENCE

Tobias Greif
✉ tobias.greif@oeaw.ac.at
Robert Baumgartner
✉ robert.baumgartner@oeaw.ac.at

RECEIVED 12 June 2025

ACCEPTED 01 August 2025

PUBLISHED 25 August 2025

CITATION

Greif T, Barumerli R, Ignatiadis K, Tóth B and
Baumgartner R (2025) The role of spatial
perception in auditory looming bias:
neurobehavioral evidence from impossible
ears. *Front. Neurosci.* 19:1645936.
doi: 10.3389/fnins.2025.1645936

COPYRIGHT

© 2025 Greif, Barumerli, Ignatiadis, Tóth and
Baumgartner. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

The role of spatial perception in auditory looming bias: neurobehavioral evidence from impossible ears

Tobias Greif^{1*}, Roberto Barumerli^{1,2}, Karolina Ignatiadis¹,
Brigitta Tóth³ and Robert Baumgartner^{1*}

¹Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria, ²Department of
Neurosciences, Biomedicine and Movement, University of Verona, Verona, Italy, ³Institute of Cognitive
Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary

Introduction: Spatial hearing enables both voluntary localization of sound sources and automatic monitoring of the surroundings. The auditory looming bias (ALB), characterized by the prioritized processing of approaching (looming) sounds over receding ones, is thought to serve as an early hazard detection mechanism. The bias could theoretically reflect an adaptation to the low-level acoustic properties of approaching sounds, or alternatively necessitate the sound to be localizable in space.

Methods: To investigate whether ALB reflects spatial perceptual decisions or mere acoustic changes, we simulated ears that disrupted spectrospatial associations on the perceptual level while maintaining the original spectrospatial entropy on the acoustic level. We then assessed sound localization, ALB and distance ratings.

Results: Compared to native ears, these novel ears impaired sound localization in both the direction and ego-centric distance dimensions. ALB manifestation also differed significantly between native and novel ears, as evidenced by behavioral discrimination performance and early cortical activity (N1 latency). Notably, the N1 electroencephalographic response closely resembled distance ratings, suggesting a strong link between spatial perception and ALB-related neural processing. Integrating this neural marker into a hierarchical perceptual decision-making model improved explanatory power, underscoring its behavioral relevance.

Discussion: These findings suggest a strong link between the localizability of sounds and their ability to elicit ALB.

KEYWORDS

EEG, spatial hearing, auditory distance, externalization, auditory looming bias, sound localization, perceptual decision-making

1 Introduction

Spatial hearing is a natural and integral part of the human sensory experience. Being able to make inferences about the location of a sound source is an important skill to navigate throughout the world (Middlebrooks, 2015; Majdak et al., 2020), as is the automatic and continuous monitoring of our surrounding auditory space in order to detect potential incoming dangers. One key example for this threat detection functionality is the auditory looming bias (ALB), which describes a perceptual preference for looming, i.e., approaching, sounds, compared to receding ones, presumably caused by evolutionary adaptation (Neuhoff, 2001; Ignatiadis et al., 2024a; Neuhoff, 2018). The auditory system, in a sense, “warns” the organism that something is approaching and enables it to react

quickly and appropriately by, for instance, moving out of the way. This adaptive nature is also highlighted by traditional ALB experiments showing that looming, compared to receding stimuli, exhibit significantly faster reaction times as well as better discrimination accuracy (Bach et al., 2009; Baumgartner et al., 2017; Bidelman and Myers, 2020; Ignatiadis et al., 2024a; McCarthy and Olsen, 2017; Neuhoﬀ, 2001, 1998; Neuhoﬀ et al., 2009; Seifritz et al., 2002). It has also been suggested that ALB elicitation is contingent on the perception of a continuous auditory object moving in distance (Baumgartner et al., 2017), but the exact details behind it are yet poorly understood. In particular, current evidence is inconclusive as to whether low-level acoustic features, typically emitted by approaching sound sources, trigger the bias or whether higher-order perceptual processes drive it instead.

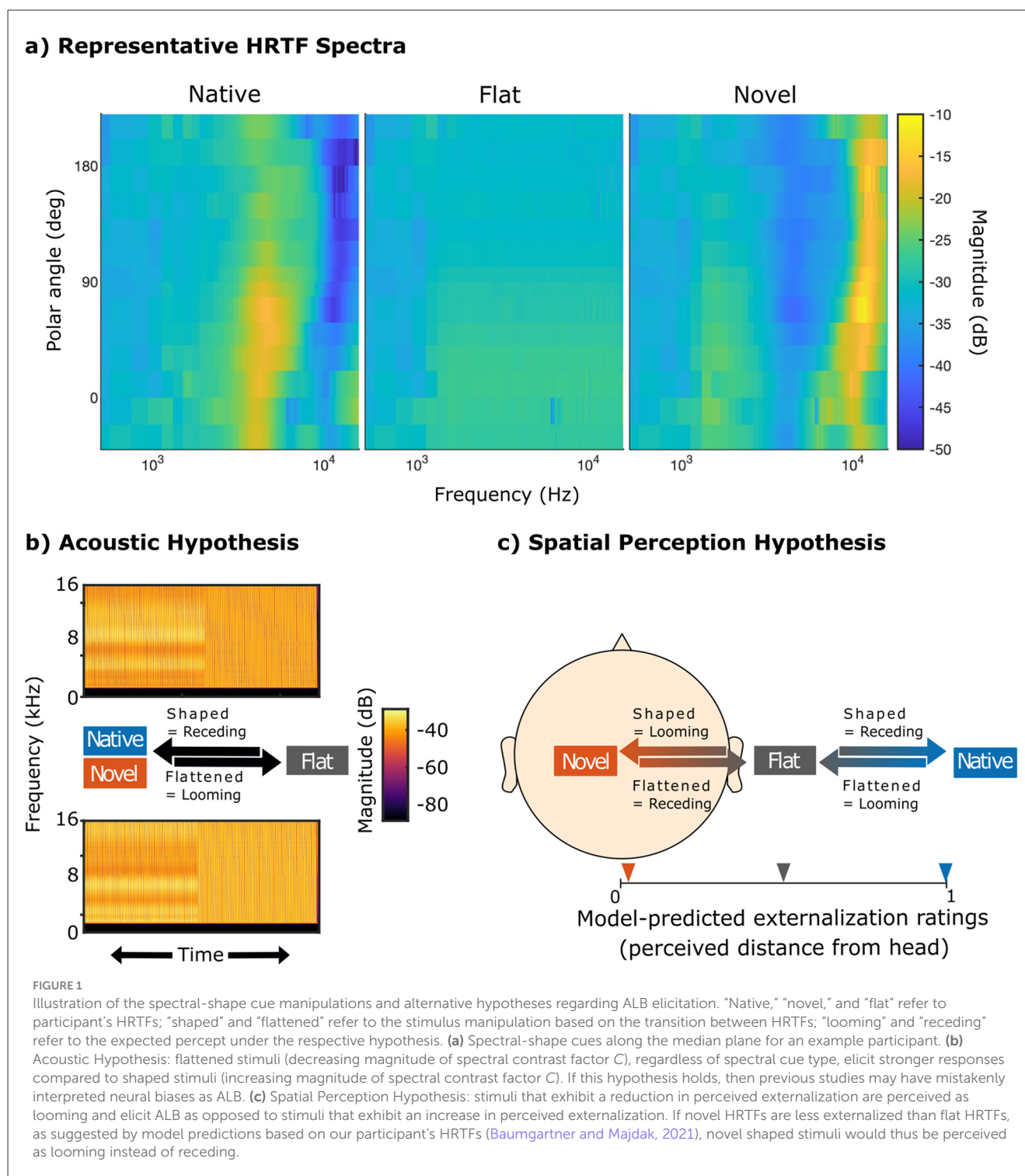
Most ALB studies (Bach et al., 2009; Bidelman and Myers, 2020; McCarthy and Olsen, 2017; Neuhoﬀ, 2001, 1998; Neuhoﬀ et al., 2009; Seifritz et al., 2002) approximated the sensation of looming and receding sound sources solely by manipulating overall sound intensity, i.e., increasing intensity to signal looming; decreasing intensity to signal receding. A neurobehavioral study expanded on this by demonstrating that the ALB could also be elicited via spectral changes while keeping broadband intensity constant (Baumgartner et al., 2017). By transitioning between two sounds, one filtered with the listener-specific (native) head-related-transfer-function (HRTF) and the other with a spectrally flat HRTF, they were able to create looming and receding sensations and elicit the ALB. HRTFs capture spectral cues at high frequencies that strongly depend on sound-source location and are mainly induced by the complex morphology of the listener's ears (Blauert, 1996). A loss of spectral features, i.e., a flat HRTF, impoverishes cues that are central for perceptual externalization, the sensation that a sound is originating from a certain ego-centric distance in "the outside world," as opposed to from "within-the-head"; an essential skill for accurate environmental awareness and immersive auditory perception, particularly in complex listening environments. A degradation of externalization is often observed when listening through headphones (Hartmann and Wittenberg, 1995; Jeffress and Taylor, 1961) or hearing aids (Best et al., 2020; Best and Roverud, 2024). Transitioning from a native, successfully externalized sound, into a flat, moreso internalized sound, could thus elicit the perception of an approaching, looming, object- or vice versa a receding object.

Subsequent work has since replicated ALB elicitation via spectral shaping and demonstrated that neural markers of ALB could be elicited even in a passive listening mode with diverted attention (Ignatiadis et al., 2024a). Interestingly, the observed differences between looming and receding stimuli are represented neurally rather early, already 80 ms and extending up 250 ms after motion onset. This is mostly in line with latency differences found between the processing of natural and artificial spatial cues in auditory motion perception (Getzmann and Lewald, 2010) but rather early compared to other aspects of spatial hearing relying on spectral cues. Changes in elevation, for instance, are only neurally distinct 200–400 ms after elevation-change onset (Fujiki et al., 2002; Bialas et al., 2023). In human newborns ALB markers were only found using intensity cues, not spectral cues (Ignatiadis et al., 2024a), indicating that ALB elicitation based on spectral cues seemingly depends on spatial associative

learning over lifetime, suggesting a developmental reliance on more salient, low-level acoustic features prior to the maturation of spectro-spatial integration. Over the course of development, the auditory system receives input from diverse spatial locations and gradually forms associations between spectral characteristics and spatial positions (Hofman et al., 1998; Wanrooij and Opstal, 2005), likely forming an internal spectro-spatial map that guides spatial perception. However, neonates exhibit a range of unique EEG characteristics that make it substantially different from an adult EEG, such as diffuse slowing, discontinuity, asynchrony and minimal reactivity (Louis et al., 2016), which complicates the interpretation of ERP markers like N1 or MMN that are typically used to assess spatial auditory processing.

Previous work has also argued that not only spatial perceptual associations but also the identifiability of the sound contribute to ALB elicitation. When presenting intensity ramps diotically over headphones, natural tonal/harmonic sounds have been reported to elicit ALB more effectively than noise, with some authors arguing that this might be due to the more realistic nature of tonal stimuli, perhaps resembling the sounds of a predator more closely and, generally, being more easily identifiable (Neuhoﬀ, 1998, 2001; Ponsot et al., 2015). An alternative interpretation of these results, however, would be that the diotic presentation negatively impacted bias elicitation with noise. Broadband sounds, such as noise, are expected to provide spectral localization cues to listeners. When these cues are absent due to diotic presentation, the spatial percept is expected to degrade (Baumgartner and Majdak, 2021), hindering ALB elicitation since the sound is less externalized to begin with and little to any motion would be perceived. Tonal sounds, lacking sufficient bandwidth, cannot convey spectral localization cues in the first place and therefore do not lead to such conflicts that might interfere with bias elicitation. Therefore, the reported stimulus dependencies favoring tonal/harmonic diotically presented sounds may instead reflect the necessity of spatial perceptual associations to elicit ALB. In line with this, other studies have since reported stimulus dependencies in the opposite direction, favoring noise over tonal/harmonic stimuli, when investigating ALB elicitation using different tasks (McCarthy and Olsen, 2017) or species (Deneux et al., 2016), suggesting that perhaps the additional spectral localization cues aided in ALB elicitation. However, these inconsistencies in the literature make it difficult to decisively conclude whether the (spectral) ALB is truly a spatial phenomenon and depends on learned associations between spatial positions and spectral cues, driven by a change in externalization, or whether spectral properties/changes themselves are sufficient to facilitate the bias. The previously reported early processing latencies suggests that the ALB might constitute a pre-attentive mechanism that possibly reflects rapid sensory prioritization. However, the discrepancies in terms of stimulus dependencies, as well as the absence of a spectral ALB in newborns, call this into question.

The present study aimed to investigate how spectral cues of novel, physically impossible ears, intended to disrupt established spectrospatial associations while providing a comparable amount of spectral information (Figure 1a), affect directional localization, externalization, and ALB elicitation. We expected novel HRTFs to acutely degrade localization performance (Middlebrooks, 1999) and externalization perception (Hartmann and Wittenberg,



1995). We hypothesized ALB elicitation to be in-line with changes in externalization rather than the amount of spectral shaping (Baumgartner et al., 2017). If the ALB is driven by the amount of spectral shaping rather than spatial associations, any transition into a flat HRTF should elicit the ALB. That is, both novel into flat as well as native into flat stimuli should elicit the ALB (Acoustic Hypothesis, Figure 1b). If, on the other hand, the bias relies on spatial associations, it should be consistent with the change

in externalization perception. As native spectral cues are thought to form the reference and increasing deviations from this reference are expected to degrade externalization, the externalization model proposed by Baumgartner and Majdak (2021) predicts that flat HRTFs are less externalized than native HRTFs, but also that novel HRTFs are even less externalized than flat HRTFs (Spatial Perception Hypothesis, Figure 1c). If this prediction holds, stimulus transitions from novel to flat would instead be perceived as

receding, and thus elicit a bias in the opposite direction than the native case. Third, we hypothesized early neurophysiological markers of the ALB, such as ERP components associated with motion onset, to be behaviorally relevant, suggesting a shared underlying mechanism. The relationship between behavioral and neural ALB measures was assessed via joint hierarchical modeling (Turner et al., 2013; Brown and Heathcote, 2008).

2 Methods

Normal-hearing young adults performed three spatial hearing tasks. First, behavioral localization experiments were conducted in virtual reality (Majdak et al., 2010), then ALB was tested using a neurobehavioral (EEG) paradigm (Baumgartner et al., 2017), and finally perceived sound externalization was rated on a continuous scale (Boyd et al., 2012; Leclère et al., 2019). We contrasted neurobehavioral responses to native spectral cues against responses to novel spectral cues that provided a similar amount of spectral change but were designed to disrupt established spectrospatial associations.

2.1 Participants

We recruited 15 young adults (10 male; mean age: 25.6 ± 2.19) from the institutes' internal participant database. The sample size was chosen identical to a previous study (Baumgartner et al., 2017). One participant was excluded from the ALB and localization analyses due to a trigger box malfunction during the EEG recording and missing data. Externalization data were only collected from the last 11 participants. All participants had normal or corrected-to-normal vision as well as normal hearing, meaning that audiometric hearing thresholds did not differ more than 20 dB from their age-specific norm (Rodríguez Valiente et al., 2014) in the high-frequency frequency range of 1 to 16 kHz. All participants were monetarily compensated for their participation. The study applied the standard methodology of the Acoustics Research Institute, which the Ethics Board of the Austrian Academy of Sciences has approved.

2.2 Paradigms

The order of tasks was fixed across participants. Everyone began with directional localization followed by the looming bias task and, finally, the externalization paradigm. This order was chosen deliberately because naive participants experience high difficulties when performing a spatial task in anechoic conditions. Since the localization paradigm was the most interactive, it served as a good starting point for the testing session. Externalization was only evaluated after the looming bias paradigm as we aimed to avoid priming effects that would have come about by participants actively evaluating distance perception with novel spectral cues prior to the looming bias paradigm. All participants, but not the experimenter, were blinded regarding which experimental conditions were tested at what point within each paradigm.

2.2.1 Directional localization

To evaluate localization abilities for our different stimulus manipulations, we adapted an established VR-localization paradigm (Majdak et al., 2010). The experiment itself was conducted in a three-dimensional virtual environment (Oculus Rift CV1, Oculus VR). Inside the environment, a gray empty space was shown, with a red sphere indicating a central reference point in front of the participant. Around the median and coronal plane, 29 smaller blue spheres were displayed in order to allow for easier navigation in the otherwise empty space. Participants were allowed to complete the experiment either seated, or standing up. However, their movement was limited such that they could only rotate but not translate through the real physical space.

Participants initiated the start of every trial via a button press on a hand-held controller. Their head had to be oriented toward the front and kept still during stimulus presentation, as to not introduce conflicting proprioception. Once a stimulus had been presented from a randomly drawn position in space, participants used the hand-pointing device (visible to them) to indicate the direction the sound came from. Participants first completed six training blocks with native cues only (50 trials each), to become familiar with the VR setup and task procedures. Training blocks also incorporated visual and auditory feedback (Majdak et al., 2010), as well as sensorimotor feedback. Here, after giving their initial response, a red cube appeared at the actual sound-source position (visual feedback). Participants were then required to acknowledge this “true” position by selecting the cube with the hand pointer, before they were asked to fixate back on the initial mid-front position. The sound was then repeated (auditory feedback) and subsequently played continuously, allowing participants to localize it while turning their head in any direction until they confirmed its position by selecting the cube once again (sensorimotor feedback). After familiarization was completed, participants continued with the testing procedure. Here, stimuli were presented in blocks of 100 trials each, with a fixed order of native, flat and then novel. This was repeated thrice, resulting in 300 trials per condition and 900 trials in total. No feedback was provided during testing blocks and participants were free to take breaks between blocks at any point but were forced to do so after three experimental blocks to limit fatigue. Depending on the speed and break timing of individual participants, the experiment took around 3 h to complete.

2.2.2 Looming bias

Participants performed a two-alternative forced-choice motion discrimination task while EEG was recorded. They were seated in a dimly lit room in front of a black computer screen with a central white fixation dot and were presented the above-mentioned spectrally shaped or flattened sounds. After each stimulus played in its entirety, participants had to judge via a keyboard button press (right index finger), whether they perceived the sound as looming or receding. Each participant completed eight experimental blocks with 100 trials each. Block order was fixed and participants alternated between native-only and novel-only blocks, starting with

native-only because participants expressed difficulties with judging motion direction when starting with novel-only blocks and also when native and novel trials were interleaved during the pilot phase of the experiment. This was probably caused by the transitions between spectral cues being quite subtle and thus making the task rather difficult, especially when interleaving trials. Within each block, 50 shaped and 50 flattened stimuli were presented in a randomized fashion. In total, the experiment took around 90 min and participants received no feedback on their performance. We opted for a forced-choice task without a non-response option in order to facilitate the analysis based on linear ballistic accumulator modeling (Section 2.4.4). Although this left the possibility of trials without distance motion percepts uncontrolled within the looming bias task, we subsequently collected perceptual externalization ratings for the start and end phase of each stimulus.

2.2.3 Externalization

To assess the perceived externalization of sounds filtered with different spectral contrasts, participants were seated in front of the same experimental screen. They judged every sound's perceived degree of externalization on a continuous visual sliding scale. The scale was anchored by the image of a loudspeaker on the left side and the representation of a head on the right side. The range between the ear and midpoint of the head was represented by 15% of the scale. Each combination of position and spectral contrast was judged three times, resulting in 36 ratings given. Participants were presented with 12 pages on the experimental screen containing three trials each. On each page, the three spectral contrasts were presented together, from one of the four possible positions. Participants could listen to all three sounds on one page as many times as they wanted and could adjust each scale as needed. Once they were satisfied with their judgments, they could move on to the next page, where the next position was presented. The order of positions and the order of spectral contrasts on each page was randomized and participants were blind to the experimental conditions. Overall, the externalization task took between 20 and 30 min to complete.

2.3 Stimuli

We first recorded the acoustic transmission properties for each participant in terms of listener-specific HRTFs that take into account the complex listener-specific pinna, head and torso geometry [apparatus and procedure described in detail in Ignatiadis et al. (2024a)]. Thus, we captured the native spectrospatial properties of sounds arriving at the ear canal, which allowed us to subsequently manipulate and create sounds with distinct spectral shapes to our needs. The measured HRTF magnitude spectrum (spectral contrast factor $C = 1$; at 1.2 m distance from the listener) was spectrally smoothed (gammatone filter with bandwidth factor of 1) (Hassager et al., 2016) to obtain native spectral cues and then either mathematically inverted ($C = -1$) to obtain novel spectral cues or deprived of spectral spatial information ($C = 0$) to

obtain flat spectra by applying a previously proposed manipulation method (Baumgartner et al., 2017):

$$M_c(f) = CM_1(f) + (1 - C) \frac{1}{N_f} \sum_{k \in f} w'(k) M_1(k), \quad (1)$$

where $M_1(f)$ denotes the magnitude spectrum of the native HRTFs, N_f denotes the frequency bins within the manipulated frequency range of 1–16 kHz, and w' denotes a frequency weighting function that approximates auditory frequency resolution by the across-frequency derivative of equivalent rectangular bandwidths (Glasberg and Moore, 1990). Representative HRTF spectra are shown in Figure 1a.

For the directional localization experiment, 500 ms Gaussian white noise bursts with 10 ms on- and offset ramps were manipulated using the spectral contrast procedure described above. This resulted in sets of stimuli that contain either the native spectral information, novel spectral information, or are devoid of any spectral information (i.e., flat).

For the externalization ratings, harmonic tone complexes (12 different fundamental frequencies evenly spaced between 100 and 140 Hz, bandwidth: 1–16 kHz, phase curvature: 0.5) (Schroeder, 1970) were corrected for constant loudness by iterative model-based adjustments (Moore et al., 2016) and subsequently filtered with either native, novel or flat HRTFs from four possible incidence angles, i.e. positions: Front-Up (FU) = 45°, Front-Down (FD) = 135°, Back-Up (BU) = 225° and Back-Down (BD) = 270° (Figure 2). We varied the fundamental frequency and the position of our stimuli, in order to avoid making the task too monotonous as well as enforcing listeners to base their judgements on spatial, rather than specific timbre, cues.

Looming bias was assessed using the same tonal stimuli as for the externalization ratings but combined in pairs. By transitioning between two stimuli with different spectral manipulations, we created flattened and shaped stimuli. The transition was initiated 600 ms after the onset of the first sound, with a jitter of ± 50 ms, and was implemented as a short 10 ms crossfade with raised cosine faders to ensure constant intensity for the highly correlated tonal stimuli. The transition was kept short to obtain a temporally clearly defined motion event for the ERP analysis (Baumgartner et al., 2017; Ignatiadis et al., 2024a).

All stimuli across all paradigms were presented binaurally using tubephones (ER-2, Etymotic Research) at an approximate sound pressure level of 70 dB at the ipsilateral ear.

2.4 Data analysis

2.4.1 Directional localization

To assess localization performance of our participants, we separately analyzed the following error measures: lateral error, local polar error and quadrant error rate (global polar error) (Middlebrooks, 1999). Lateral and local polar errors were defined as the difference in angular coordinates between target and response, along both the lateral dimension ranging from left (−90°) to right (90°) and the polar dimension ranging from below (−90°) via front (0°) and above (90°) to back (180°). The local polar

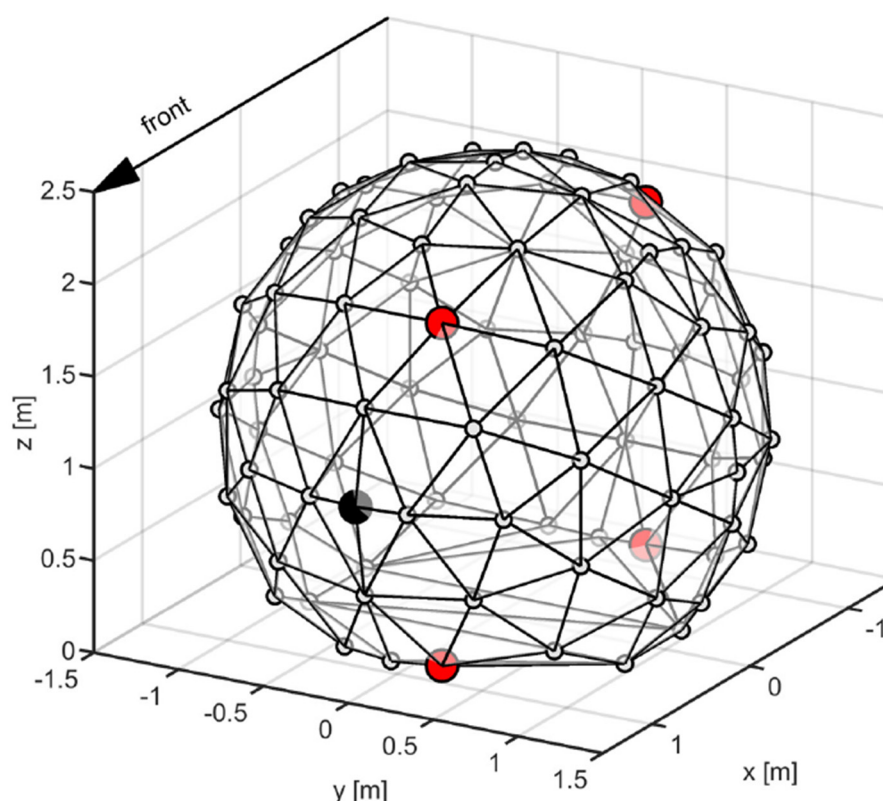


FIGURE 2

Illustration of the loudspeaker array system with 91 source positions. Participants are facing the speaker highlighted in black during the localization paradigm. Positions used for generating the ALB and externalization stimuli are highlighted in red.

error excluded all polar errors exceeding 90° in absolute terms. The quadrant error rate was defined as the percentage of trials where the polar error exceeded 90° .

In order to reveal the effect of our stimulus manipulations on participants' error rates, we conducted repeated measure ANOVAs for each error measure separately. Due to the long testing sessions we additionally included an "order" factor in these analyses to check whether performance differed between the first and second half of responses for each cue type. Partial eta squared was calculated as a measure of effect size. If a significant difference between at least two groups was found, we conducted Tukey-HSD *post-hoc* tests to reveal which groups differed from each other. Sphericity assumptions were tested using Mauchly's test for sphericity, implemented in MATLAB. In case of violations, a Greenhouse–Geisser correction was applied.

2.4.2 Externalization

Externalization rating data were collected on a continuous scale ranging from 0 to 1, representing a percept localized at the center of the listener's head to a fully externalized sound image at the distance of the loudspeakers used for the HRTF measurement, respectively. The data were averaged over fundamental frequencies to achieve normality for beta regression modeling. Using the R (Version 4.4.1) packages `data.table` (Barrett et al., 2024), `emmeans` (Lenth, 2024), and `ggplot` (Wickham, 2016),

we tested five different model alternatives taking into account the influence of cue-type (native, novel, flat) and position (FU, FD, BU, BD) on externalization ratings: (1) Intercept-only; (2) main effect of cue-only; (3) main effect of position-only; (4) main effects of both cue and position, and (5) main effects of cue and position and their interaction. All model alternatives were fitted using the maximum likelihood approach and a logit-link implemented in the R package `betareg` (Grün et al., 2012). Model comparison was based on Bayesian information criterion (BIC). To determine the significance of main effects, a Bayes Factor (BF) approximation based on the difference in BIC between an intercept-only model and a main effect model, and the associated posterior probabilities, was computed to express evidence in favor of, or evidence against, a model including a main effect (Wagenmakers, 2007). *Post-hoc* pairwise comparisons were conducted using estimated marginal means and the *p*-values were adjusted using Tukey's multiplicity adjustment.

2.4.3 EEG

EEG data were recorded using a high-density 128-channel EEG cap (actiCAP with actiCHamp; Brain Products GmbH, Gilching, Germany) and the Brain-Vision Recorder software. Pre-processing was conducted in EEGLAB (Delorme and Makeig, 2004). As a first step, bad channels (i.e., channels with high impedance ($\geq 50\text{k}\Omega$) and visible noise during the recording)

that were marked as such during data collection were removed. Subsequently, the data was bandpass-filtered (0.5–30 Hz) using an FIR filter with Kaiser window ($\beta = 7.2$, $n = 462$). Since our event of interest is the change/transition between the sounds with different spectral contrasts, the data was epoched around this changepoint (600 ± 50 ms after stimulus onset) into segments of $[-100, 1,500]$ and subsequently resampled to 100 Hz. We then performed an Independent Component Analysis (ICA) and removed components indicating eye movement-related artifacts or power-line-noise using EEGLAB's ICLabel plugin. Components that were categorized as eye movements or power-line-noise with a high probability ($\geq 80\%$) were manually checked and removed. Manual visual inspection also took place on a trial-by-trial basis, removing trials that contained apparent noise artifacts such as malfunctioning channels that were not noted previously. No automated trial and/or channel rejection took place. A maximum of ten malfunctioning EEG channels and 5% excluded trials were accepted per participant. On average, 2.14 (SD: 1.46) channels and 7.43 (SD: 6.55) trials were rejected per participant. After data cleaning was completed, the missing channels were interpolated using EEGLAB's default spherical interpolation function. The data were then re-referenced to the average. Next, we equalized the number of trials to retain the same amount across all conditions within participants. This was done by finding the minimum number of trials present within all conditions, and then rejecting trials from the other conditions, equally spaced across the session. This resulted in, on average, 0.7 (SD: 1.12) trials being removed per condition per participant and ensured an unbiased, blind trial equalization procedure.

For statistical inference, we used Fieldtrip (Oostenveld et al., 2010). Baseline correction was performed on each individual trial, using the 100 ms before the event of interest. The data was organized into group-level structures and average ERPs for our four conditions were computed with `ft_timelockanalysis`. To address our research question, we first compared the evoked potentials for flattened and shaped stimuli for the native and novel conditions separately. Next, we compared the evoked flattened-shaped contrast for native cues with the one for novel cues, by subtracting the shaped condition from the flattened condition, separately for native and novel, and then comparing those differences. To address whether there were any significant spatiotemporal differences between our conditions, we used non-parametric cluster-based permutation testing (Maris and Oostenveld, 2007). Neighbors were defined using Fieldtrip's "distance" method. The test-statistic of choice was a two-tailed dependent-samples *t*-test, using the Monte Carlo re-sampling method (10,000 permutations; α : 0.05). Our analysis time-window was restricted to 0–300 ms. To account for the two-tailed test, the resulting probabilities were corrected using Fieldtrip's `correcttail` function set to `prop`. Aggregated cluster scores were calculated by averaging over the whole cluster epoch and all channels that contributed to the cluster at any point. Finally, we conducted a two-way repeated measures ANOVA to investigate the effects of spectral cue and position on the evoked native-novel difference within the obtained cluster. Sphericity assumptions were again tested using Mauchly's test for sphericity. In case of violations, a Greenhouse–Geisser correction would have been applied.

2.4.4 Neurobehavioral model

We evaluated the ALB magnitude in the behavioral data of the looming bias paradigm by combining response choices and reaction times within a linear ballistic accumulator (LBA) model (Brown and Heathcote, 2008), implemented in R (Version 4.4.1) using the packages EMC2 (Stevenson et al., 2024), `data.table` (Barrett et al., 2024), `msm` (Jackson, 2011), `coda` (Plummer et al., 2006), and `ggplot2` (Wickham, 2016). The LBA model is designed for multi-alternative forced choice tasks, where each possible response option competes with the others by accumulating evidence at a specific rate, known as the drift rate v_i . Each accumulator starts from a random starting point drawn from a uniform distribution within the range $[0, A]$. A response is generated when the first accumulator crosses a decision threshold B . To account for residual variance in total response times, such as motor or sensory latencies, the model also includes a non-decision time parameter t_0 . Performance in forced choice tasks has been demonstrated to reflect the combined effects of task difficulty and response caution; hence, the LBA model jointly accounts for both speed and accuracy (or the likelihood of a certain response) and can thus control for the speed-accuracy trade-off present in forced choice tasks (Brown and Heathcote, 2008). In contrast to process-oriented models such as the LBA, analyzing accuracy and reaction times separately might lead to misleading results when interpreting changes in distributional parameters (e.g., mean location) as variations of the underlying cognitive processes (Matzke and Wagenmakers, 2009).

For parameter estimation, we employed Bayesian statistics with a Markov Chain Monte Carlo (MCMC) approach, enabling hierarchical parameter estimation. This method provided robust estimates and facilitated the analysis of group-level parameter variations (Gunawan et al., 2020). In particular, we relied on the library EMC2 (Stevenson et al., 2024), which implements the particle Metropolis within Gibbs (PMwG) approach for MCMC. This approach achieves faster convergence, requiring fewer samples compared to earlier methods (Gunawan et al., 2020). Convergence was ensured through multiple chains (independent sequences of samples), allowing for automatic convergence assessment based on the Gelman-Rubin diagnostic (Brooks and Gelman, 1998). We used default EMC2 prior values and, after parameter estimation, inspected the contraction statistics (i.e., one minus the ratio of posterior sample variance to prior sample variance) to verify minimal prior influence.

We employed linear modeling features to examine how experimental factors influenced model parameters across participants. Our experimental design incorporated spectral cue type (native or novel), motion direction (flattened or shaped), stimulus-response match (flattened \rightarrow looming or shaped \rightarrow looming), and position. For position, we partitioned the levels into two factors: vertical (up or down) and horizontal (front or back). This allowed us to test multiple configurations of experimental factors and their relationships with model parameters. Quality of fit was compared across models using the Bayesian Predictive Information Criterion (BPIC) metric (Ando, 2011), which aims to balance model fit and parsimony. By penalizing models with excessive free parameters, BPIC helps mitigate overfitting.

In addition to using only behavioral data, we employed a joint modeling approach to test whether integrating EEG data in the

model's parameter estimation improves its quality of fit (Turner et al., 2013). For this configuration, we extracted the mean amplitudes of the EEG signal within a significant spatiotemporal cluster of difference (see Section 3.3) on a trial-by-trial basis and included them as an additional regressor to predict the drift rates. The method follows a simple directed approach that includes the cluster amplitude as a linear predictor of cognitive parameters. If the inclusion of this EEG-based regressor improves the model's predictive power, it supports the assumption that the extracted EEG activity is relevant to the behavioral outcome. Thereby, the joint approach enables a more comprehensive understanding of how perceptual decisions originate from neural activity (Turner et al., 2013).

To assess differences in drift rates between factors, we sampled drift rates from the group-level posterior distributions and computed the mean difference between conditions, denoted Δ . This quantity served as a summary statistic to capture the direction and magnitude of drift rate changes. We then estimated its 95% credible interval and the probability of Δ being positive or negative using a one-tailed credible interval approach (McElreath, 2018).

3 Results

3.1 Directional localization

Figure 3 illustrates group-level lateral and polar localization response patterns for each cue-type in order to convey a first qualitative impression of the results. For the lateral angle, the scatter plots indicate high correspondence between target and response lateral angle (responses close to the main diagonal) for all three conditions. Larger differences across conditions are visible for the polar angle. In this dimension, high correspondence between target and response polar angle are only present for the native condition, whereas the flat and novel conditions led to a substantial degradation in relationship between target and response angles. For both flat and novel conditions, listeners exhibited a strong response bias toward the horizontal plane, primarily at their rear (around 180°).

For statistical evaluation, localization performance was quantified as lateral error, quadrant error rate, and local polar error (Figure 4). Spectral cue conditions significantly influenced the lateral error [$F_{(2,26)} = 12.89$, $p < 0.01$, partial $\eta^2 = 0.50$]. The effect of order (first half presentation vs. second half presentation) failed to reach significance [$F_{(1,13)} = 1.10$, $p = 0.31$], as did the interaction between order and cue type [$F_{(2,26)} = 0.17$, $p = 0.83$]. *Post-hoc* comparisons revealed that the mean lateral error measure was significantly lower in the native condition compared to both the flat [$d = -2.07^\circ$, 95% CI $(-3.26^\circ, -0.88^\circ)$, $p < 0.01$] and novel [$d = -1.84^\circ$, 95% CI $(-2.87^\circ, -0.81^\circ)$, $p < 0.001$] conditions. No significant difference was observed between flat and novel [$d = 0.23^\circ$, 95% CI $(-1.08^\circ, 1.53^\circ)$, $p = 0.89$] spectral cues.

Considering the quadrant error rate, we again observe a large effect of cue manipulation [$F_{(1.04,13.56)} = 20.07$, $p < 0.001$, partial $\eta^2 = 0.61$]. The effect of order again failed to reach significance [$F_{(0.52,6.78)} = 0.02$, $p = 0.90$], as did the interaction between order and cue type [$F_{(1.05,13.56)} = 3.16$, $p = 0.08$]. *Post-hoc* comparisons showed that performance in the native condition was

significantly better than in both the flat condition [$d = -14.25\%$, $p < 0.01$, 95% CI $(-24.13\% - 4.38\%)$] and the novel condition [$d = -18.99\%$, $p < 0.001$, 95% CI $(-27.79\% - 10.19\%)$]. There was no significant difference in performance between flat and novel cue manipulations [$d = -4.73^\circ$, 95% CI $(-10.10^\circ, 0.63^\circ)$, $p = 0.08$].

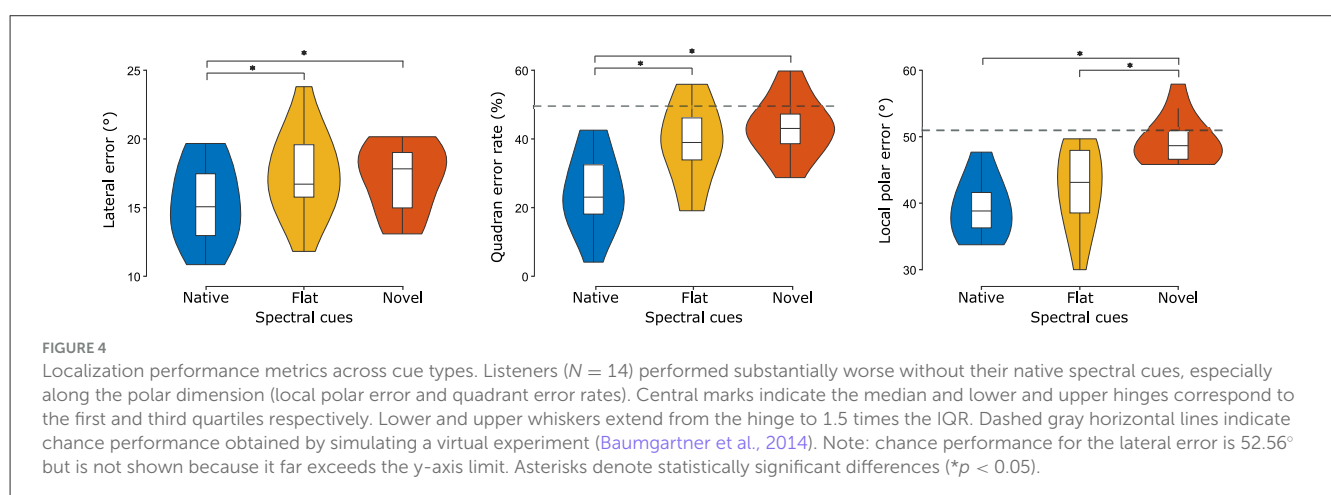
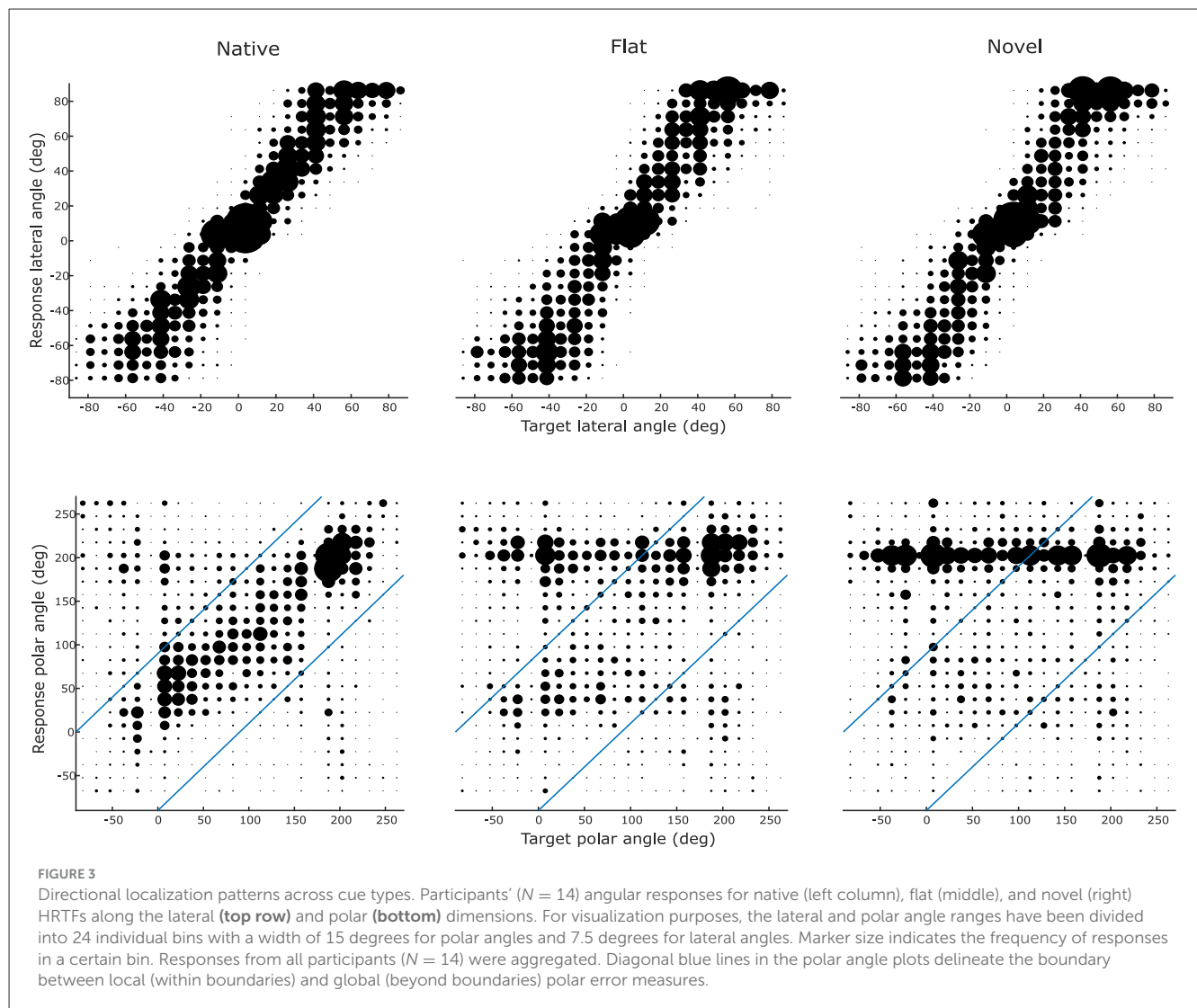
Regarding the local polar error, cue manipulation again had a large effect [$F_{(2,26)} = 27.15$, $p < 0.001$, partial $\eta^2 = 0.68$]. There was no significant effect of order [$F_{(1,13)} = 0.44$, $p = 0.52$], nor the interaction between order and cue type [$F_{(2,26)} = 0.70$, $p = 0.50$]. *Post-hoc* comparisons found that the local polar error was significantly lower for native, compared to novel [$d = -10.05^\circ$, 95% CI $(-12.87^\circ, -7.23^\circ)$, $p < 0.001$] and also lower for flat, compared to novel [$d = -6.96^\circ$, 95% CI $(-10.56^\circ, -3.35^\circ)$, $p < 0.001$]. There was no significant difference between native and flat [$d = -3.09^\circ$, 95% CI $(-7.55^\circ, 1.37^\circ)$, $p = 0.20$] spectral cues.

We also investigated whether, and to what extent, performance was better (or worse) than chance-level. To that end, we simulated a virtual localization experiment with fully randomized responses (Baumgartner et al., 2014) and used this as a reference for normalization. The division of real and simulated error measures thus shows how similar the performance is to chance level, with 1 indicating chance performance and 0 indicating lack of any errors. When simulating the virtual localization experiment, an equal response distribution across all possible response locations inside the target range was assumed, while considering the actual target locations of the experiment. For the lateral error, all cue-types performed better than chance-level (native: 0.29; flat: 0.33; novel: 0.33). Regarding the local polar error, performance with native and flat spectral cues was, albeit a bit worse overall, again better than chance (native: 0.76; flat: 0.82). However, performance with novel cues was close to chance level (novel: 0.95). The quadrant error rate exhibits a similar pattern: evident better than chance performance for native and flat spectral cues (native: 0.49; flat: 0.77) and close to chance level with novel cues (novel: 0.87). Overall, polar-angle localization performance was poor and around chance level for novel spectral cues.

3.2 Externalization

We further analyzed the localizability of sounds under our different spectral cue manipulations and source positions in terms of the perceived degree of externalization. Figure 5a visualizes the distribution of the obtained externalization ratings across spectral cue types, indicating a gradual decay of externalization ratings with increasing deviation from the native reference. We analyzed those data by selecting the best mixed-effects model as outlined in Section 2.4.2. The model predicting externalization scores based on the two main effects of cue and position yielded the best fit (model #1 in Table 1). Compared to the intercept-only model, the BF approximation approach (Wagenmakers, 2007) revealed very strong evidence in favor of the cue-only model [$BF = 27,487.83$; $Pr(H_0|D) > 0.99$] and strong evidence in favor of the position-only model [$BF = 69.21$; $Pr(H_0|D) > 0.98$], confirming a significant contribution of our main effects to explaining the data.

Regarding the main effect of spectral cue type, the estimated marginal means suggest a gradual decay of externalization



perception from native, over flat to novel spectral cues. That is, native stimuli were rated as the most externalized [native = 0.67, 95-CI(0.61, 0.74)], followed by flat stimuli [flat = 0.48, 95-CI(0.41, 0.55)] and finally novel stimuli [novel

= 0.28, 95-CI(0.22, 0.34)]. *Post-hoc* pairwise comparisons revealed that the differences between all three levels of cue types were significant (native—flat: $z = 3.84$, $p < 0.01$; native—novel: $z = 8.26$, $p < 0.01$; flat—novel: $z = 4.10$,

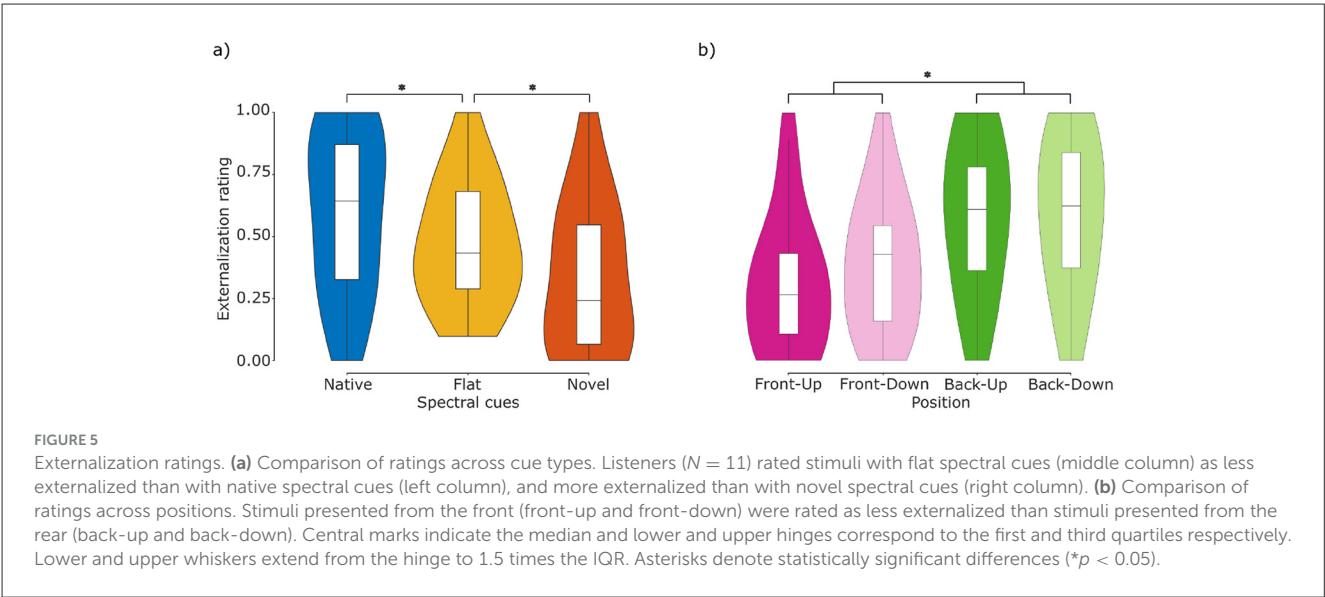


TABLE 1 Externalization model summary.

Models			dBIC		
1. $Ext \sim (cue + pos) sbj$			0		
2. $Ext \sim (cue * pos) sbj$			14.1		
3. $Ext \sim (cue) sbj$			15.1		
4. $Ext \sim (pos) sbj$			27.1		
5. $Ext \sim 1 sbj$			35.1		
Fixed effects coefficients (model 1)	Est	SE	z	$Pr(> z)$	
(Intercept) reference cue = nat, pos=FU	0.06	0.23	0.26	0.80	
cue = flat	-0.88	0.24	-3.76	<0.01	***
cue = nov	-1.83	0.25	-7.43	<0.01	***
pos = FD	0.28	0.27	1.05	0.30	
pos = BU	1.34	0.27	4.88	<0.01	***
pos = BD	1.33	0.27	4.85	<0.01	***
Predictions (model 1) for Ext	Pos:	FU	FD	BU	BD
cue = nat		0.51	0.58	0.80	0.80
cue = flat		0.30	0.37	0.63	0.62
cue = nov		0.15	0.18	0.39	0.39

Several models were estimated and compared relative to the winning model #1 in terms of BIC difference. Smaller BIC values indicate better model fit. For the selected model #1, fixed effects are reported with model estimates (Est), standard errors (SE), z -ratio, corresponding p -values and model predictions. Fixed effects are relative to the respective reference levels (native for cue, FU for position). Ext = externalization rating (0–1); cue = spectral cue type (native, flat, novel); pos = position (FU, FD, BU, BD); sbj = participant ID (1–11). Asterisks denote statistically significant differences ($***p < 0.001$).

$p < 0.01$). These differences in externalization perception between different spectral cue types are in-line with the model-predicted ratings (Figure 1c) and thus support the

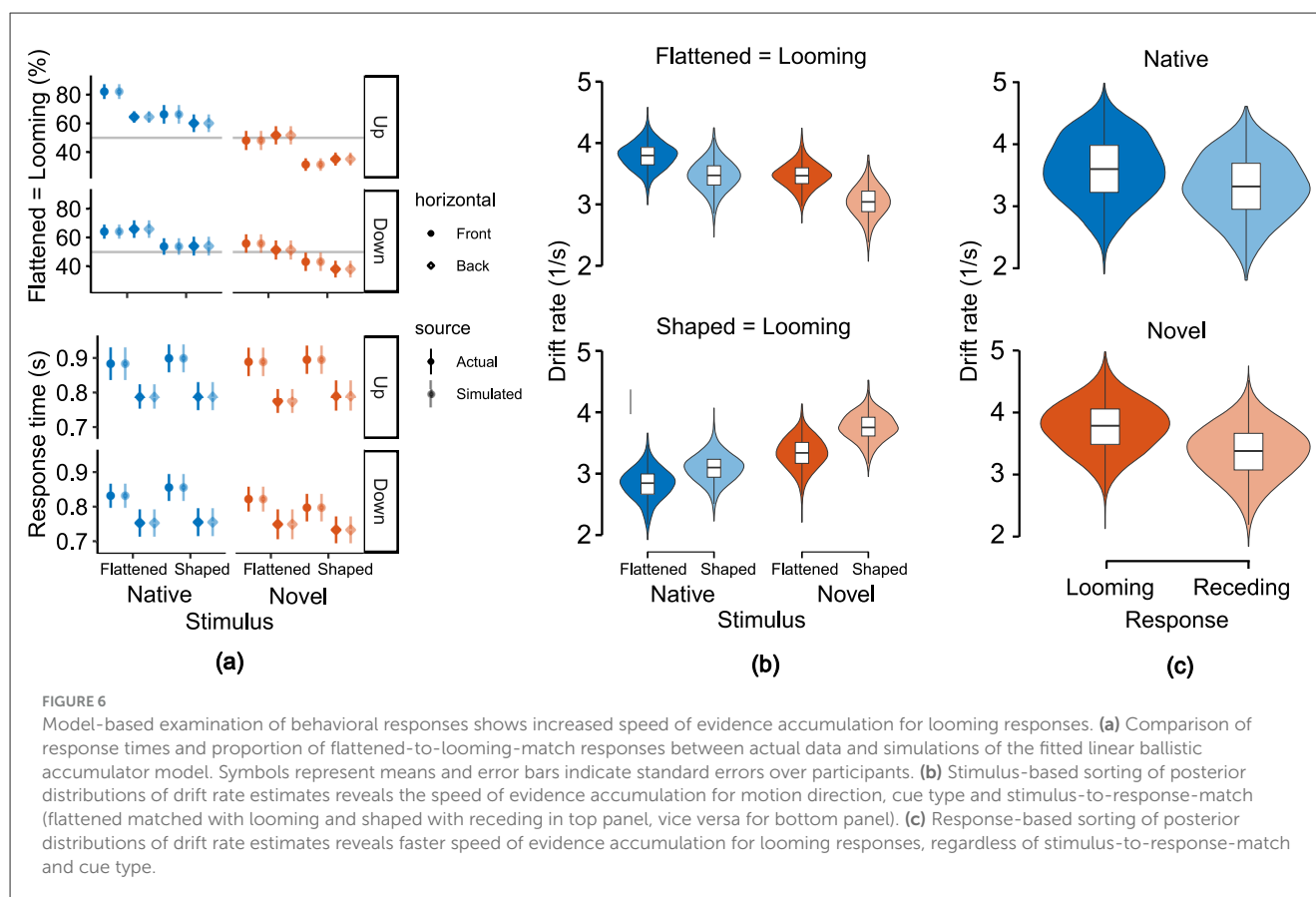
expected motion perception in the looming bias paradigm. This enables to test the Spatial Perception Hypothesis against the Acoustic Hypothesis.

Results regarding the main effect of position are shown in Figure 5b. The estimated marginal means indicated that stimuli presented from the frontal hemifield were rated as less externalized [FU = 0.32, 95-CI(0.25, 0.40); FD = 0.38, 95-CI(0.30, 0.46)] than stimuli presented from the rear hemifield [BU = 0.61, 95-CI(0.53, 0.69); BD = 0.61, 95-CI(0.53, 0.68)]. *Post-hoc* pairwise comparisons showed no significant differences between the positions within frontal and rear hemifields (FU–FD: $z = -1.05$, $p = 0.72$; BU–BD: $z = 0.04$, $p > 0.99$), but significant differences between them (FU–BU: $z = -5.13$, $p < 0.01$; FU–BD: $z = -5.09$, $p < 0.01$; FD–BU: $z = -4.02$, $p < 0.01$; FD–BD: $z = -3.98$, $p < 0.01$).

3.3 Looming bias

Figure 6a shows the mean response times and the proportion of flattened-to-looming-match responses. In the native flattened condition, participants indicated “looming” in most cases (60%–80%) and in the native shaped condition, they primarily responded “receding” (50%–60%). In the novel flattened condition, however, participants responded mostly random (50% “looming” and 50% “receding”), whereas in the novel shaped condition they mostly responded with “looming” (60%–70%), indicating a perceptual switch between native and novel spectral cue conditions.

Before constructing the neurobehavioral model, results of the ALB paradigm were first investigated based on EEG data only. Cluster-based permutation tests showed no significant differences between flattened and shaped stimuli within cue-type ($p = 0.072$ for a negative cluster in the native condition ranging from 100 to 180 ms after correction for two-tailed testing). Comparing the differences between flattened and shaped stimuli between cue types, however, revealed a significant amplitude difference between the native and novel flattened-shaped contrasts in the N1-related range



of 80–170 ms (clusterstat = -294.0344, $p = 0.04$) after motion onset. The native condition was characterized by a more negative N1-related-cluster amplitude for flattened, compared to shaped stimuli. In contrast, the novel condition was instead characterized by a more negative amplitude for shaped, compared to flattened stimuli (Figure 7a), in line with the externalization-based ALB elicitation hypothesis (Figure 1c) and participant's externalization ratings for the different spectral cue manipulations (Figure 5). Topographically, this difference across cue types emerged over left occipital electrodes and subsequently propagated toward central electrodes. At around 140 ms, the cluster became spatially bimodal and fronto-temporally lateralized (Figure 7b, right-most topographies). We further confirmed that both the native and novel evoked differences between flattened and shaped stimuli within the N1-related cluster of difference were significantly different from zero (native: $t = -3.46, p < 0.01$; novel: $t = 3.16, p < 0.02$, see Figure 7b, right panel). A two-way repeated measures ANOVA further revealed a significant main effect of spectral cue on the cluster amplitude of the evoked native-novel difference [$F_{(1,13)} = 16.4, p < 0.001$, partial $\eta^2 = 0.56$]. There was no significant effect of position [$F_{(3,39)} = 1.52, p = 0.22$], nor was there a significant interaction between spectral cue and position [$F_{(3,39)} = 0.46, p = 0.71$].

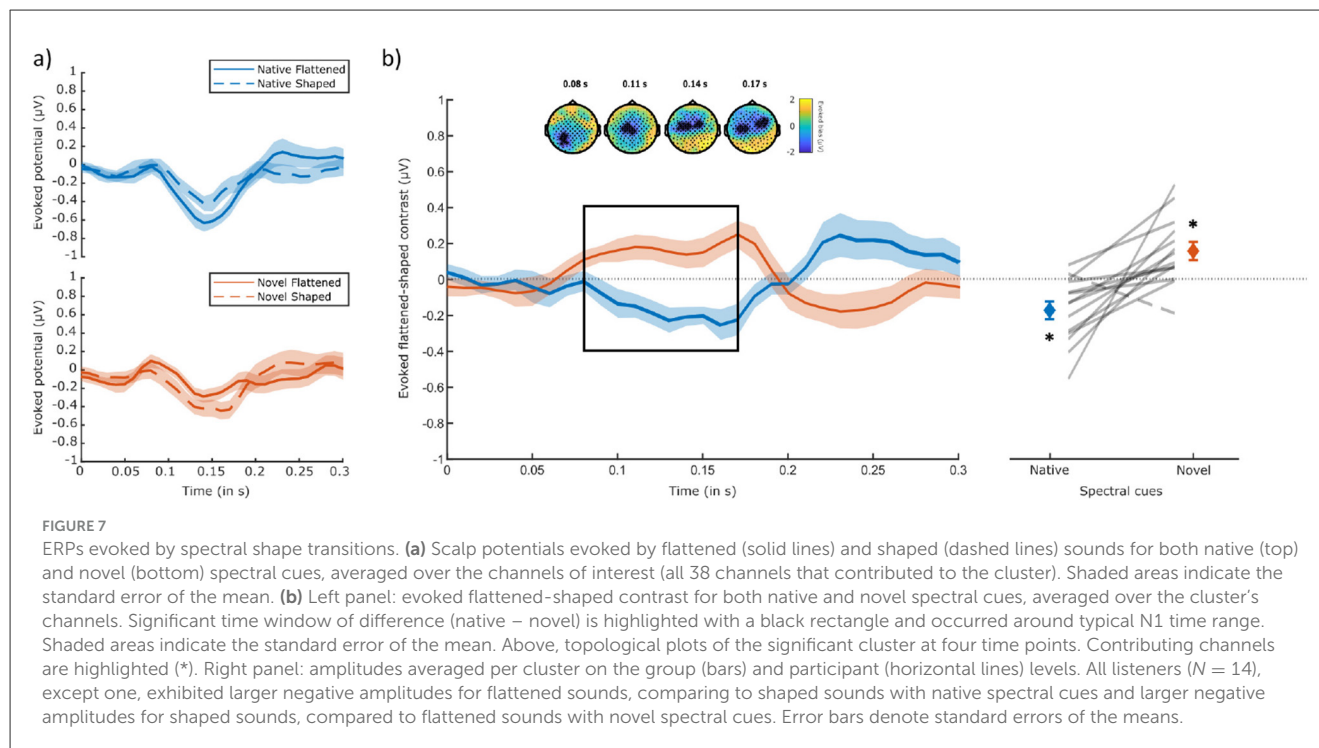
We then investigated the behavioral results of the motion discrimination task by comparing different LBA models used to describe the speed of evidence accumulation in terms of drift rates. We started with identifying the best purely behavioral LBA

model and then tested the benefit of adding the trial-by-trial cluster amplitudes (see Section 3.3) as an additional neural regressor:

$$v \sim match * direction * cue + vertical * horizontal + eeg \quad (2)$$

Our comparison identified this full model where the drift rate was modulated by all five acoustic and behavioral factors (i.e., flattened-stimulus-to-looming-response match, cue type, motion direction, vertical and horizontal positions) and the additional neural marker as the best fit. The quality of fit was lower without the neural marker (dBPIC = 2) and the preceding selection of the purely behavioral reference model showed that it clearly outperformed simpler alternatives (intercept only: dBPIC = 2,999; match only: dBPIC = 2,701; match and direction: dBPIC = 2,318; match, direction and cue: dBPIC = 1,129) as well as an alternative regressand (response threshold instead of drift rate: dBPIC = 99). The benefit of adding the neural marker was also verified against surrogate data matching the mean and SD of the cluster scores (dBPIC = 12). In Figure 6a, we report a direct comparison between the real and simulated behavioral data obtained from the winning model.

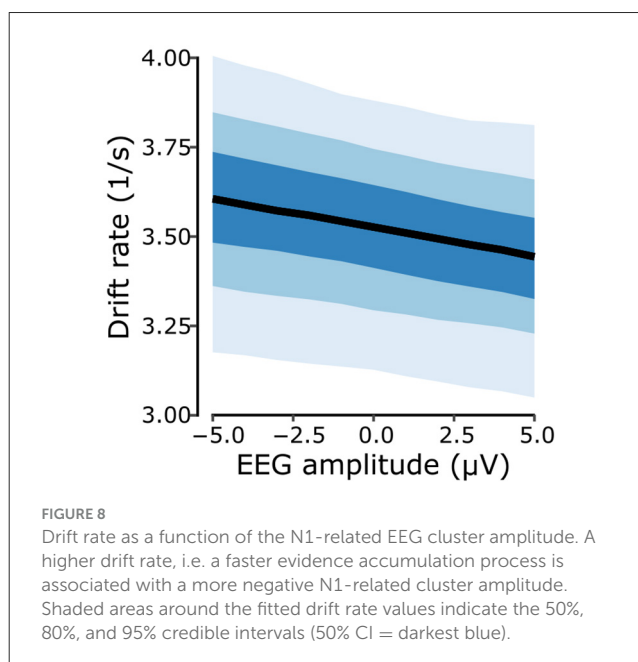
We then used the model predictions to assess the strength of the ALB in terms of drift rates. The results demonstrate a dependence on perceived motion direction (Figures 6b, c). For correct responses, drift rates resulted higher for flattened compared to shaped stimuli [$P_{\Delta > 0} = 0.995$, with $\Delta = 0.362s^{-1}$, 95%-CI (0.093, 0.658)]. For incorrect responses, drift rates resulted smaller



for flattened compared to shaped stimuli [$P_{\Delta < 0} = 0.970$, with $\Delta = -0.339s^{-1}$, 95%-CI $(-0.648, 0.017)$], indicating that looming responses consistently exhibited higher drift rates compared to receding responses. Furthermore, the analysis revealed that cue type modulates drift rates differently based on response match. For correct responses, the native condition yielded a greater drift rate compared to the novel condition [$P_{\Delta > 0} = 0.951$, with $\Delta = 0.356s^{-1}$, 95%-CI $(-0.079, 0.793)$]. Conversely, for incorrect responses, drift rates for the native condition reported smaller values than those from the novel condition [$P_{\Delta < 0} = 0.991$, with $\Delta = -0.590s^{-1}$, 95%-CI $(-1.088, -0.106)$]. Finally, we inspected the magnitude of the fitted linear coefficient associated with the magnitude of the N1-related EEG cluster for the prediction of the model's drift rate which corresponded to $-0.017\ 1/\mu Vs$ with a 95%-CI $(-0.039, 0.005)$, and with a 0.942 probability of it being a negative value. Figure 8 illustrates the linear relationship between the EEG predictor (N1-related EEG cluster amplitude) and the fitted drift rates, revealing that faster evidence accumulation is associated with a more negative EEG amplitude. This result, alongside the model comparison, further supports the notion that there is a coherent relationship between the neurophysiological marker and the behavioral outcome.

4 Discussion

The present study set out to clarify whether the biases observed in previous looming studies are truly a spatial phenomenon. Do sounds need to be localizable in order to elicit the (spectral) ALB or are mere stimulus characteristics sufficient? To answer this question, we first proposed an HRTF manipulation method to obtain novel spectral cues that provided the same amount of spectral shaping as native cues while significantly deteriorating



sound localizability in terms of directional localization and externalization perception. Then, we demonstrated that biases in early perceptual processing (N1-related cluster amplitudes) evoked by novel spectral cues are consistent with the direction of externalization change rather than stimulus shaping. Finally, we confirmed that the behavioral speed of evidence accumulation is generally higher for looming responses and showed that this behavioral effect can better be explained by integrating the EEG marker.

4.1 Novel spectral cues degrade directional localization performance and perceptual externalization

We demonstrated that localization performance with novel spectral cues was significantly impaired compared to performance with native cues across all error measures (lateral error; local polar error; quadrant error rate). This indicates that established spatial associations between stimulus characteristics and locations were successfully broken apart. Furthermore, we also observed diminished local polar localization performance for novel cues compared to flat cues, which is also reflected in the fact that performance with novel cues was close to chance performance, whereas performance with flat cues was still better than chance.

This raises the question of why participants were still able to localize sounds in the polar dimension reliably at least to some degree with flat spectral cues. A possible explanation is that our HRTF manipulation only targeted frequencies above 1 kHz (see Section 2.3), thus low frequency information provided by torso reflections was still available to participants and possibly provided some usable elevation (Algazi et al., 2001), externalization (Zotter and Frank, 2018) and front-back discriminatory cues (Asano et al., 1990). The additional presence of contradictory, i.e., inverted, high frequency spectral cues in the novel cue condition could have potentially interfered with the information provided below 1 kHz, thus, leading to worse performance and a stronger degradation of externalization than in the flat condition, where low frequency cues were the only available ones. We also observed a response bias toward the horizontal plane with flat and novel spectral cues, primarily at participant's rear (around 180°). A response bias such as this is, also referred to as low elevation gain (slope of the stimulus-response relation) (Zwiers et al., 2003) under perceptual uncertainty, i.e. with impoverished spectral cues, is in line with previous work reporting similar patterns (Wanrooij and Opstal, 2005; Zonooz et al., 2018, 2019). A second possible explanation could arise from an implementation issue in the experimental design. Due to a miscommunication, participants were tested using a fixed order of spectral cue conditions—native, flat, and novel—which coincides with the observed performance ranking. However, since this fixed order was repeated three times, it also included native after novel blocks and thus is rather unlikely to account for the pronounced differences between conditions.

Moreover, lateral localization performance and quadrant error rate was also significantly reduced with flat spectral cues, compared to native spectral cues. Even though the impairment was rather small, reduced lateral localization abilities with flat spectral cues is in-line with a more recent, updated, view on the contribution of spectral cues for the lateral dimension. Traditionally, lateral sound localization was thought to almost exclusively rely on binaural cues, such as interaural time- and level differences, with monaural spectral cues thought to only contribute very little (Macpherson and Middlebrooks, 2002). However, recent work (Stevenson-Hoare et al., 2022) has challenged this idea by demonstrating an apparent front/back asymmetry in sound localization precision measures, which is especially strong for oblique azimuths. That is, the pinna introduces a high-frequency spectral cue that enhances lateral sound localization in the frontal hemifield only. Hence, contemporary computational models on human sound localization

include both monaural and binaural cues as joint contributors to lateral and polar localization performance predictions (Barumerli et al., 2023). By flattening individualized HRTFs, these additional cues were no longer available to our participants, potentially explaining their diminished lateral localization performance with flat HRTFs compared to native HRTFs.

Furthermore, we also demonstrated that the spectral manipulation of individualized HRTFs degraded the perceived externalization of sound sources. An externalization breakdown was observed for both flat and novel HRTFs. Stimuli based on novel HRTFs were rated as even less externalized than stimuli based on flat ones, indicating a shift toward an internalized percept with novel, spectrally inverted HRTFs and a gradual decay of externalization. This decay is consistent with a template-based model of externalization under static listening conditions (Baumgartner and Majdak, 2021). Under this framework, perceptual externalization depends on comparing observed auditory cues and listener-specific templates in memory. When incoming sounds deviate from the established reference template, externalization perception is predicted to gradually decrease with increased deviation magnitude. In the present study this assumption would thus also predict a significant, but smaller, difference between native (e.g., template-matching) and flat (e.g., template-deviation) stimuli, compared to the difference between native and novel stimuli consistent with our observation.

4.2 ALB elicitation conforms with spatial percept

Based on the present results we conclude that ALB elicitation was consistent with the listeners' perceived changes in externalization rather than triggered by inherent acoustical features independent of auditory space. In line with the template-based model predictions (Baumgartner and Majdak, 2021), participants reported less externalization for novel stimuli compared to flat and native stimuli. Furthermore, they tended to exhibit a low rate of shaped-stimulus-to-receding-response match, indicating that their perception of novel stimuli seems to have been reversed compared to native stimuli. Stimuli that exhibited a reduction of perceptual externalization during the short stimulus transition were perceived as looming (e.g. native flattened and novel shaped), and vice versa, stimuli that contained an increase of perceptual externalization were perceived as receding (e.g. native shaped and novel flattened). The speed of evidence accumulation, i.e. the drift rate, was also consistently higher for looming responses, regardless of stimulus-to-response-match. Taken together, our data support the hypothesis that ALB elicitation relies on spatial perceptual associations, rather than purely acoustic properties (see Figures 1b, c).

Although the current study demonstrates that ALB elicitation conformed with participant's spatial perception, due to its correlational nature we cannot clearly pinpoint the causal dependencies between bias elicitation, sound localizability and externalization perception in particular. A follow-up learning study investigating the generalizability from sound localization to ALB elicitation with novel spectral shape cues (Greif et al., 2024) will address this.

4.3 Early perceptual processing biases affect speed of evidence accumulation

Neural markers of ALB were identified between 80 ms and 170 ms after motion onset via cluster-based permutation testing and confirmed to be behaviorally relevant via LBA modeling. The latency and topography of that marker closely relates to the N1 ERP component and is consistent with previously reported results on ALB and motion perception (Ignatiadis et al., 2021, 2024a). The N1 component is an established marker of early perceptual processing, implicated to play a central role in attention allocation via sensory gain control (Getzmann, 2011; Getzmann and Lewald, 2010; Hillyard et al., 1998). That is, attended stimuli tend to exhibit an increased N1 amplitude compared to unattended stimuli. In the context of our study, sounds exhibiting a decay of externalization were predominantly perceived as looming, captured attention, and thus facilitated ALB elicitation. Interestingly, previous work has demonstrated this neural bias toward looming sounds also while attention was diverted, in a passive listening mode (Ignatiadis et al., 2024a), highlighting the high salience and attentional capture potential of approaching sound sources.

Incorporating the trial-by-trial N1-related EEG cluster amplitudes in the LBA model improved its explanatory power. Yet, the overall benefit was rather small ($\text{dBIC} = 2$), suggesting that this single-trial marker was quite noisy and supposedly incomplete. Alternative neurophysiological markers such as ERP latency, oscillatory power, or a joint metric, should thus be considered. Previous corticocortical connectivity analyses have identified bottom-up projections between temporal, parietal and the frontal regions as a possible neural network involved in the elicitation of the ALB (Ignatiadis et al., 2021). Specifically, connectivity among regions along the ventral stream, traditionally coined the “what” stream in visual and auditory processing (Saur et al., 2008), seems to play an essential role in ALB elicitation (Ignatiadis et al., 2024b). This makes it a suitable target for future work incorporating more sophisticated neural ALB markers (e.g. phase locking value). Implementing more flexible joint modeling methods with non-linear relationship assumptions (Palestro et al., 2018; Nunez et al., 2024) could also further improve the present model. Translating the present work into more immersive environments, e.g. via VR-based applications using more realistic sounds, could also lead to a stronger perceptual impression of spectral looms and recedes, potentially increasing the observed effects (Deng et al., 2019; Getzmann and Lewald, 2010) and strengthening the relationship between behavioral measurement and neural marker.

4.4 Limitations

A key limitation of the present study is the relatively small sample size, which may reduce the statistical power to detect small or moderate effects. This limitation implies that non-significant findings, such as the lack of EEG differences between flattened and shaped stimuli within cue-type, should be interpreted with caution, as they may reflect insufficient power rather than a true absence of effect. Nevertheless, the study provides initial evidence that the emergence of auditory looming bias requires spatial perceptual associations to be established.

Another limitation concerns the low ecological validity of stimuli. As in previous studies (Baumgartner et al., 2017; Ignatiadis et al., 2024a), we here isolated spectral from intensity cues to hone in on the contribution of spatial perception to ALB, but it is worth noting that this isolation of cues does not happen outside of the laboratory and is expected to limit the extent of ALB (Bach et al., 2009) and attention control (Deng et al., 2019). Secondly, the spectral cue manipulations to elicit looming and receding sensations happened on a very rapid time scale. Such a rather instantaneous change in perceived distance has the advantage of providing a clearly defined event of interest for the analysis of EEG data, but it may limit the generalization of our results to the real world where changes in distance usually occur more gradually. Finally, our experimental setting did not allow participants to move during stimulus presentation because this would render the stimuli even more implausible. Listeners naturally engage in self-motion to resolve perceptual ambiguities (Wallach, 1940), consistent with principles of active inference (Friston et al., 2016). Future work is needed to develop more naturalistic experimental settings with high stimulus control for neuroscientific investigations.

4.5 Outlook

The here presented results establish a strong link between spectrospatial associations and looming/receding perception, implying that any disruption of natural spectral cues would also disrupt motion perception in general, leading to potentially dangerous situations for anyone wearing hearing aid or hearing protection devices. Furthermore, the preservation of natural cues, or at least providing a sufficiently good approximation, seems integral for developing immersive virtual realities that provide realistic auditory experiences. Actively manipulating spectral cues provided to listeners could also provide unique and interesting opportunities for the development of gamified environments and/or audio design.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://osf.io/f5wx2/>.

Ethics statement

The studies involving humans were approved by Commission for Science Ethics of the Austrian Academy of Sciences. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

TG: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. RBar: Formal analysis, Methodology, Software, Writing – original draft, Writing –

review & editing. KI: Methodology, Software, Writing – review & editing. BT: Conceptualization, Funding acquisition, Writing – review & editing. RBau: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This project has received funding from the Austrian Science Fund (FWF, Grant-DOIs 10.55776/I4294 and 10.55776/ZK66) and the European Union's Horizon Europe research and innovation programme under grant agreement No. 101129903.

Acknowledgments

We thank Regina Pfennigschmidt for collecting the data.

References

- Algazi, V. R., Avendano, C., and Duda, R. O. (2001). Elevation localization and head-related transfer function analysis at low frequencies. *J. Acoust. Soc. Am.* 109, 1110–1122. doi: 10.1121/1.1349185
- Ando, T. (2011). Predictive Bayesian model selection. *Am. J. Math. Manag. Sci.* 31, 13–38. doi: 10.1080/01966324.2011.10737798
- Asano, F., Suzuki, Y., and Sone, T. (1990). Role of spectral cues in median plane localization. *J. Acoust. Soc. Am.* 88, 159–168. doi: 10.1121/1.399963
- Bach, D. R., Neuhoﬀ, J. G., Perrig, W., and Seifritz, E. (2009). Looming sounds as warning signals: the function of motion cues. *Int. J. Psychophysiol.* 74, 28–33. doi: 10.1016/j.ijpsycho.2009.06.004
- Barrett, T., Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Hocking, T., et al. (2024). *data.table: Extension of 'data.frame'.*
- Barumerli, R., Majdak, P., Geronazzo, M., Meijer, D., Avanzini, F., Baumgartner, R., et al. (2023). A Bayesian model for human directional localization of broadband static sound sources. *Acta Acust.* 7:12. doi: 10.1051/aacus/2023006
- Baumgartner, R., and Majdak, P. (2021). Decision making in auditory externalization perception: model predictions for static conditions. *Acta Acust.* 5:59. doi: 10.1051/aacus/2021053
- Baumgartner, R., Majdak, P., and Laback, B. (2014). Modeling sound-source localization in sagittal planes for human listeners. *J. Acoust. Soc. Am.* 136, 791–802. doi: 10.1121/1.4887447
- Baumgartner, R., Reed, D. K., Tóth, B., Best, V., Majdak, P., Colburn, H. S., et al. (2017). Asymmetries in behavioral and neural responses to spectral cues demonstrate the generality of auditory looming bias. *Proc. Nat. Acad. Sci.* 114, 9743–9748. doi: 10.1073/pnas.1703247114
- Best, V., Baumgartner, R., Lavandier, M., Majdak, P., and Kopčo, N. (2020). Sound externalization: a review of recent research. *Trends Hear.* 24:2331216520948390. doi: 10.1177/2331216520948390
- Best, V., and Roverud, E. (2024). Externalization of speech when listening with hearing aids. *Trends Hear.* 28:23312165241229572. doi: 10.1177/23312165241229572
- Bialas, O., Maess, B., and Schönwiesner, M. (2023). Evoked responses to localized sounds suggest linear representation of elevation in human auditory cortex. *bioRxiv.* doi: 10.1101/2023.05.03.539222
- Bidelman, G. M., and Myers, M. H. (2020). Frontal cortex selectively overrides auditory processing to bias perception for looming sonic motion. *Brain Res.* 1726:146507. doi: 10.1016/j.brainres.2019.146507
- Blauert, J. (1996). *Spatial Hearing: The Psychophysics of Human Sound Localization.* Cambridge, MA: The MIT Press. doi: 10.7551/mitpress/6391.001.0001
- Boyd, A. W., Whitmer, W. M., Soraghan, J. J., and Akeroyd, M. A. (2012). Auditory externalization in hearing-impaired listeners: the effect of pinna cues and number of talkers. *J. Acoust. Soc. Am.* 131, EL268–EL274. doi: 10.1121/1.3687015
- Brooks, S. P., and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* 7, 434–455. doi: 10.1080/10618600.1998.10474787
- Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178. doi: 10.1016/j.cogpsych.2007.12.002
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Deneux, T., Kempf, A., Daret, A., Ponsot, E., and Bathellier, B. (2016). Temporal asymmetries in auditory coding and perception reflect multi-layered nonlinearities. *Nat. Commun.* 7:12682. doi: 10.1038/ncomms12682
- Deng, Y., Choi, I., Shinn-Cunningham, B., and Baumgartner, R. (2019). Impoverished auditory cues limit engagement of brain networks controlling spatial selective attention. *Neuroimage* 202:116151. doi: 10.1016/j.neuroimage.2019.116151
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., Pezzulo, G. (2016). Active inference and learning. *Neurosci. Biobehav. Rev.* 68, 862–879. doi: 10.1016/j.neubiorev.2016.06.022
- Fujiki, N., Riederer, K. A., Jousmäki, V., Mäkelä, J. P., and Hari, R. (2002). Human cortical representation of virtual auditory space: differences between sound azimuth and elevation. *Eur. J. Neurosci.* 16, 2207–2213. doi: 10.1046/j.1460-9568.2002.02276.x
- Getzmann, S. (2011). Auditory motion perception: onset position and motion direction are encoded in discrete processing stages. *Eur. J. Neurosci.* 33, 1339–1350. doi: 10.1111/j.1460-9568.2011.07617.x
- Getzmann, S., and Lewald, J. (2010). Effects of natural versus artificial spatial cues on electrophysiological correlates of auditory motion. *Hear. Res.* 259, 44–54. doi: 10.1016/j.heares.2009.09.021
- Glasberg, B. R., and Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hear. Res.* 47, 103–138. doi: 10.1016/0378-5955(90)90170-T
- Greif, T., Ignatiadis, K., Baier, D., Toth, B., and Baumgartner, R. (2024). *From sound Localization to Hazard Detection: Adaptation to Novel Spectral Cues.* OSF.
- Grün, B., Kosmidis, I., and Zeileis, A. (2012). Extended beta regression in r: shaken, stirred, mixed, and partitioned. *J. Stat. Softw.* 48, 1–25. doi: 10.18637/jss.v048.i11
- Gunawan, D., Hawkins, G. E., Tran, M.-N., Kohn, R., and Brown, S. D. (2020). New estimation approaches for the hierarchical linear ballistic accumulator model. *J. Math. Psychol.* 96:102368. doi: 10.1016/j.jmp.2020.102368

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Hartmann, W. M., and Wittenberg, A. (1995). On the externalization of sound images. *J. Acoust. Soc. Am.* 99, 3678–3688. doi: 10.1121/1.414965
- Hassager, H. G., Gran, F., and Dau, T. (2016). The role of spectral detail in the binaural transfer function on perceived externalization in a reverberant environment. *J. Acoust. Soc. Am.* 139, 2992–3000. doi: 10.1121/1.4950847
- Hillyard, S. A., Vogel, E. K., and Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. B Biol. Sci.* 353, 1257–1270. doi: 10.1098/rstb.1998.0281
- Hofman, P. M., Van Riswick, J. G. A., and Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nat. Neurosci.* 1, 417–421. doi: 10.1038/1633
- Ignatiadis, K., Baier, D., Barumerli, R., Sziller, I., Tóth, B., Baumgartner, R., et al. (2024a). Cortical signatures of auditory looming bias show cue-specific adaptation between newborns and young adults. *Commun. Psychol.* 2, 1–15. doi: 10.1038/s44271-024-00105-5
- Ignatiadis, K., Baier, D., Tóth, B., and Baumgartner, R. (2021). Neural mechanisms underlying the auditory looming bias. *Audit. Percept. Cogn.* 4, 60–73. doi: 10.1080/25742442.2021.1977582
- Ignatiadis, K., Barumerli, R., Deco, G., Tóth, B., and Baumgartner, R. (2024b). Threat-related corticocortical connectivity elicited by rapid auditory looms. *bioRxiv*. doi: 10.1101/2024.06.21.600106
- Jackson, C. H. (2011). Multi-state models for panel data: the msm package for r. *J. Stat. Softw.* 38, 1–29. doi: 10.18637/jss.v038.i08
- Jeffress, L. A., and Taylor, R. W. (1961). Lateralization vs localization. *J. Acoust. Soc. Am.* 33, 482–483. doi: 10.1121/1.1908697
- Leclère, T., Lavandier, M., and Perrin, F. (2019). On the externalization of sound sources with headphones without reference to a real source. *J. Acoust. Soc. Am.* 146, 2309–2320. doi: 10.1121/1.5128325
- Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. Available online at: <https://rvinlenth.github.io/emmeans/>
- Louis, E. K. S., Frey, L. C., Britton, J. W., Frey, L. C., Hopp, J. L., Korb, P., et al. (2016). “The developmental EEG: premature, neonatal, infant, and children,” in *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants* (Chicago, IL: American Epilepsy Society).
- Macpherson, E. A., and Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J. Acoust. Soc. Am.* 111, 2219–2236. doi: 10.1121/1.1471898
- Majdak, P., Baumgartner, R., and Jenny, C. (2020). “Formation of three-dimensional auditory space,” in *The Technology of Binaural Understanding, Modern Acoustics and Signal Processing*, eds. J. Blauert, and J. Braasch (Cham: Springer International Publishing), 115–149. doi: 10.1007/978-3-030-00386-9_5
- Majdak, P., Goupell, M. J., and Laback, B. (2010). 3-d localization of virtual sound sources: effects of visual environment, pointing method, and training. *Attent. Percept. Psychophys.* 72, 454–469. doi: 10.3758/APP.72.2.454
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- Matzke, D., and Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-gaussian and shifted wald parameters: a diffusion model analysis. *Psychon. Bull. Rev.* 16, 798–817. doi: 10.3758/PBR.16.5.798
- McCarthy, L., and Olsen, K. N. (2017). A “looming bias” in spatial hearing? Effects of acoustic intensity and spectrum on categorical sound source localization. *Attent. Percept. Psychophys.* 79, 352–362. doi: 10.3758/s13414-016-1201-9
- McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: Chapman and Hall/CRC. doi: 10.1201/9781315372495
- Middlebrooks, J. C. (1999). Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *J. Acoust. Soc. Am.* 106, 1493–1510. doi: 10.1121/1.427147
- Middlebrooks, J. C. (2015). “Chapter 6 - Sound localization,” in *Handbook of Clinical Neurology, Volume 129 of The Human Auditory System*, eds. M. J. Aminoff, F. Boller, and D. F. Swaab (Amsterdam: Elsevier), 99–116. doi: 10.1016/B978-0-444-62630-1.00006-8
- Moore, B. C. J., Glasberg, B. R., Varathanathan, A., and Schlittenlacher, J. (2016). A loudness model for time-varying sounds incorporating binaural inhibition. *Trends Hear.* 20:2331216516682698. doi: 10.1177/2331216516682698
- Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature* 395, 123–124. doi: 10.1038/25862
- Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecol. Psychol.* 13, 87–110. doi: 10.1207/S15326969ECO1302_2
- Neuhoff, J. G. (2018). “Adaptive biases in visual and auditory looming perception,” in *Spatial Biases in Perception and Cognition*, ed. T. L. Hubbard (Cambridge: Cambridge University Press), 180–190. doi: 10.1017/9781316651247.013
- Neuhoff, J. G., Planisek, R., and Seifritz, E. (2009). Adaptive sex differences in auditory motion perception: looming sounds are special. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 225–234. doi: 10.1037/a0013159
- Nunez, M. D., Fernandez, K., Srinivasan, R., and Vandekerckhove, J. (2024). A tutorial on fitting joint models of m/EEG and behavior to understand cognition. *Behav. Res. Methods* 56, 6020–6050. doi: 10.3758/s13428-023-02331-x
- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2010). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:e156869. doi: 10.1155/2011/156869
- Palestro, J. J., Bahg, G., Sederberg, P. B., Lu, Z.-L., Steyvers, M., Turner, B. M., et al. (2018). A tutorial on joint models of neural and behavioral measures of cognition. *J. Math. Psychol.* 84, 20–48. doi: 10.1016/j.jmp.2018.03.003
- Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News* 6, 7–11.
- Ponsot, E., Meunier, S., Kacem, A., Chatron, J., and Susini, P. (2015). Are rising sounds always louder? Influences of spectral structure and intensity-region on loudness sensitivity to intensity-change direction. *Acta Acust. United Acustica* 101, 1083–1093. doi: 10.3813/AAA.918902
- Rodríguez Valiente, A., Trinidad, A., García Berrocal, J. R., Górriz, C., and Ramírez Camacho, R. (2014). Extended high-frequency (9–20 kHz) audiometry reference thresholds in 645 healthy subjects. *Int. J. Audiol.* 53, 531–545. doi: 10.3109/14992027.2014.893375
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., et al. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci.* 105, 18035–18040. doi: 10.1073/pnas.0805234105
- Schroeder, M. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (corresp.). *IEEE Trans. Inf. Theory* 16, 85–89. doi: 10.1109/TTT.1970.1054411
- Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., et al. (2002). Neural processing of auditory looming in the human brain. *Curr. Biol.* 12, 2147–2151. doi: 10.1016/S0960-9822(02)01356-8
- Stevenson, N., Donzallaz, M. C., Innes, R., Forstmann, B., Matzke, D., Heathcote, A., et al. (2024). EMC2: an R package for cognitive models of choice. *PsyArXiv Preprints*. doi: 10.31234/osf.io/2e4dq
- Stevenson-Hoare, J. O., Freeman, T. C. A., and Culling, J. F. (2022). The pinna enhances angular discrimination in the frontal hemifield. *J. Acoust. Soc. Am.* 152, 2140–2149. doi: 10.1121/10.0014599
- Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Sederberg, P. B., Steyvers, M., et al. (2013). A Bayesian framework for simultaneously modeling neural and behavioral data. *Neuroimage* 72, 193–206. doi: 10.1016/j.neuroimage.2013.01.048
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804. doi: 10.3758/BF03194105
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psychol.* 27, 339–368. doi: 10.1037/h0054629
- Wanrooij, M. M. V., and Opstal, A. J. V. (2005). Relearning sound localization with a new ear. *J. Neurosci.* 25, 5413–5424. doi: 10.1523/JNEUROSCI.0850-05.2005
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag. doi: 10.1007/978-3-319-24277-4_9
- Zonooz, B., Arani, E., Kording, K. P., Aalbers, P. A. T. R., Celikel, T., Van Opstal, A. J., et al. (2019). Spectral weighting underlies perceived sound elevation. *Sci. Rep.* 9:1642. doi: 10.1038/s41598-018-37537-z
- Zonooz, B., Arani, E., and Opstal, A. J. V. (2018). Learning to localise weakly-informative sound spectra with and without feedback. *Sci. Rep.* 8:17933. doi: 10.1038/s41598-018-36422-z
- Zotter, F., and Frank, M. (2018). *Low-Frequency Trick to Improve Externalization With Non-Individual HRIRs*. Berlin: DEGA, 940–943.
- Zwiers, M. P., Van Opstal, A. J., and Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nat. Neurosci.* 6, 175–181. doi: 10.1038/nn999