



Asymmetries in behavioral and neural responses to spectral cues demonstrate the generality of auditory looming bias

Robert Baumgartner^{a,b,1}, Darrin K. Reed^c, Brigitta Tóth^d, Virginia Best^{a,e}, Piotr Majdak^b, H. Steven Colburn^{a,c}, and Barbara Shinn-Cunningham^{a,c}

^aHearing Research Center, Boston University, Boston, MA 02215; ^bAcoustics Research Institute, Austrian Academy of Sciences, Vienna 1040, Austria; ^cDepartment of Biomedical Engineering, Boston University, Boston, MA 02215; ^dInstitute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences, Budapest 1117, Hungary; and ^eDepartment of Speech, Language & Hearing Sciences, Boston University, Boston, MA 02215

Edited by Dale Purves, Duke University, Durham, NC, and approved July 28, 2017 (received for review February 25, 2017)

Studies of auditory looming bias have shown that sources increasing in intensity are more salient than sources decreasing in intensity. Researchers have argued that listeners are more sensitive to approaching sounds compared with receding sounds, reflecting an evolutionary pressure. However, these studies only manipulated overall sound intensity; therefore, it is unclear whether looming bias is truly a perceptual bias for changes in source distance, or only in sound intensity. Here we demonstrate both behavioral and neural correlates of looming bias without manipulating overall sound intensity. In natural environments, the pinnae induce spectral cues that give rise to a sense of externalization; when spectral cues are unnatural, sounds are perceived as closer to the listener. We manipulated the contrast of individually tailored spectral cues to create sounds of similar intensity but different naturalness. We confirmed that sounds were perceived as approaching when spectral contrast decreased, and perceived as receding when spectral contrast increased. We measured behavior and electroencephalography while listeners judged motion direction. Behavioral responses showed a looming bias in that responses were more consistent for sounds perceived as approaching than for sounds perceived as receding. In a control experiment, looming bias disappeared when spectral contrast changes were discontinuous, suggesting that perceived motion in distance and not distance itself was driving the bias. Neurally, looming bias was reflected in an asymmetry of late event-related potentials associated with motion evaluation. Hence, both our behavioral and neural findings support a generalization of the auditory looming bias, representing a perceptual preference for approaching auditory objects.

auditory looming bias | electroencephalography | distance motion perception | sound externalization | head-related transfer functions

Imagine yourself alone in the wilderness. Suddenly, a threatening sound permeates the darkness. Is it approaching? This is a critical question when it comes to your survival because approaching objects usually pose a greater threat than receding objects (1). The phenomenon that approaching sounds are more salient than receding sounds is commonly termed “auditory looming bias.” Looming bias is reflected in a broad variety of psychophysical tasks related to salience and alertness: bias in loudness-change estimates (2–4) and judgments of duration (5), improved discriminability of motion speed (6), underestimated distances for egocentrically moving (4) or bypassing sounds (7, 8), and reduced reaction time for auditory (3, 9) and visual (3) targets preceded by looming sounds. In animals, looming biases result in faster learning speed during associative conditioning (10) and longer duration of attention (11). This list shows that looming bias triggers a variety of percepts across a wide range of psychoacoustic tasks. Despite its broad behavioral significance, the mechanisms underlying auditory looming bias are still poorly understood.

A universal problem of previous research on auditory looming bias is that source distance was manipulated using overall sound

intensity (sounds increasing in intensity perceived as approaching). Hence, it is unclear whether looming bias actually reflects a bias in the perception of intensity changes or distance motion. Moreover, it is hard to disentangle the contribution of simple neural nonlinearities in response to intensity gradients (12, 13) from higher-level perceptual asymmetries of motion in distance. Previous studies addressed this issue by comparing sounds of different spectral structure: complex tones elicited stronger behavioral looming bias (2, 11) and larger neural differences (14–16) than did noise with equal intensity increase, indicating that intensity changes did not directly cause looming bias. However, studies with a different spatial task (17) or species (10) found the opposite effect (i.e., stronger looming bias for noise than tonal sounds), calling into question these results.

An early fMRI study that contrasted static sounds with sounds increasing or decreasing in intensity found activity in the right planum temporale (4), which is associated with processing of auditory motion in any spatial direction (18, 19), consistent with the idea that sound intensity changes can be perceived as motion in distance. The contrast between sounds increasing vs. decreasing in intensity, however, revealed a different network, including superior temporal sulcus and amygdala, which may reflect a warning process for approaching objects (4, 9, 20). A more recent fMRI

Significance

Previous studies demonstrated “auditory looming bias” exclusively by manipulating overall sound intensity. Hence, it is not clear whether this bias truly reflects perceptual differences in sensitivity to motion direction rather than changes in intensity. We manipulated individualized spectral cues to create stimuli that were perceived as either approaching or receding, while controlling loudness. We assessed discrimination of motion direction and analyzed simultaneously recorded neural responses using electroencephalography. Our results show both behavioral and neural evidence of looming bias. Therefore, our study demonstrates that the bias is truly about perceived motion in distance, not intensity changes.

Author contributions: R.B., D.K.R., B.T., V.B., P.M., H.S.C., and B.S.-C. designed research; R.B. performed research; R.B. analyzed data; and R.B., D.K.R., B.T., V.B., P.M., H.S.C., and B.S.-C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: Code for stimulus creation (sig_baumgartner2017looming), HRTFs, and experimental results (both data_baumgartner2017looming) are integrated in the Auditory Modeling Toolbox (amtoolbox.sourceforge.net). EEG recordings and analysis scripts are provided via Zenodo (<https://dx.doi.org/10.5281/zenodo.832899>).

¹To whom correspondence should be addressed. Email: robert.baumgartner@oeaw.ac.at.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1703247114/-DCSupplemental.

study (21) and neural recordings in animals (10, 14) reported differences related to looming bias in low-level auditory cortical areas. Whether these differences are caused by lateral or top-down processes is unclear. The processing latencies reported in previous studies appear to be inconsistent: some found multisensory enhancements in visual cortical areas at early sensory levels (80-ms postsound onset) (15, 22), whereas others found no significant differences in global field strength before 600 ms (16). A more instantaneous change of distance cues (rather than gradual changes of sound intensity) might be better suited to identify the time course of auditory looming bias.

Auditory event-related potential (ERP) studies of horizontal and vertical motion perception support a stereotypical “motion onset response” that consists of two consecutive stages associated with different functions (23–26): a negative deflection occurring roughly 100 ms after the evoking event, denoted by N1, represents an early, presumably automatic, sensory detection of spatial changes, and a positive deflection occurring roughly 200 ms after the evoking event, denoted by P2, likely represents evaluation of motion direction. Looming bias arguably relies on the evaluation of motion direction. Hence, if the motion onset response can be generalized to motion in distance, then looming bias may be reflected in P2.

Here, we simultaneously measured behavior and neural responses to explore whether (i) changes in overall sound intensity and (ii) stimulus continuity are required to activate looming bias, and (iii) whether auditory looming bias is linked to early sensory or the later stages of neural processing. Based on the hypothesis that looming bias reflects a gain in perceptual detectability of approaching objects relative to receding objects, we measured looming bias in a motion direction discrimination task. Hence, we aimed to assess the origin of looming bias more directly than previous studies that measured consequential effects of the bias related to stimulus salience and alertness. We used spectral changes to create sounds perceived as approaching or receding between external, internal, and intermediate positions (colored red, blue, and green, respectively, in Fig. 1A and B) by reducing the contrast of measured high-frequency spectral cues from the natural acoustics of each individual listener. Spectral cues are particularly suitable for this investigation because (i) in contrast to gradual intensity changes, spectral cues can evoke instantaneous changes of distance percepts (27–30) and thus allow precise analysis of processing latencies; (ii) spectral cue manipulations can create transitions to internal (infinitely close) auditory percepts within peripersonal space where looming bias is most prominent (31); and (iii) spectral cues can be manipulated independently of overall stimulus intensity.

Results

Loudness Predictions. We used predictions of a loudness model (32) to assure that loudness changes cannot explain the looming bias caused by spectral contrast switches. Fig. 1B shows the effect of our contrast manipulation technique on the stimulus magnitude spectrum (Top) and on the corresponding predictions of relative loudness changes (Bottom) for all stimuli of experiment (Exp.) I (individual data provided in *Interindividual Comparison of Frequency-Specific Loudness Changes* and Fig. S1). Flattening the reference magnitude spectrum ($C = 1$, red) decreased frequency-specific loudness in the lower two to three octaves of the stimulus (1–6 kHz) and increased it only in the highest octave (8–16 kHz). On average across frequencies (Lower Right), flattening the reference magnitude spectrum decreased predicted overall loudness. Hence, changes in loudness perception, if they occurred, would be expected to oppose the typical loudness-induced distance percept.

Interaural level differences (ILDs) are also known to act as cues for distance perception (18, 29). Relative changes in frequency-specific ILDs can be easily extracted from loudness predictions (Fig. 1B, Bottom) as the difference between ipsilateral (solid lines)

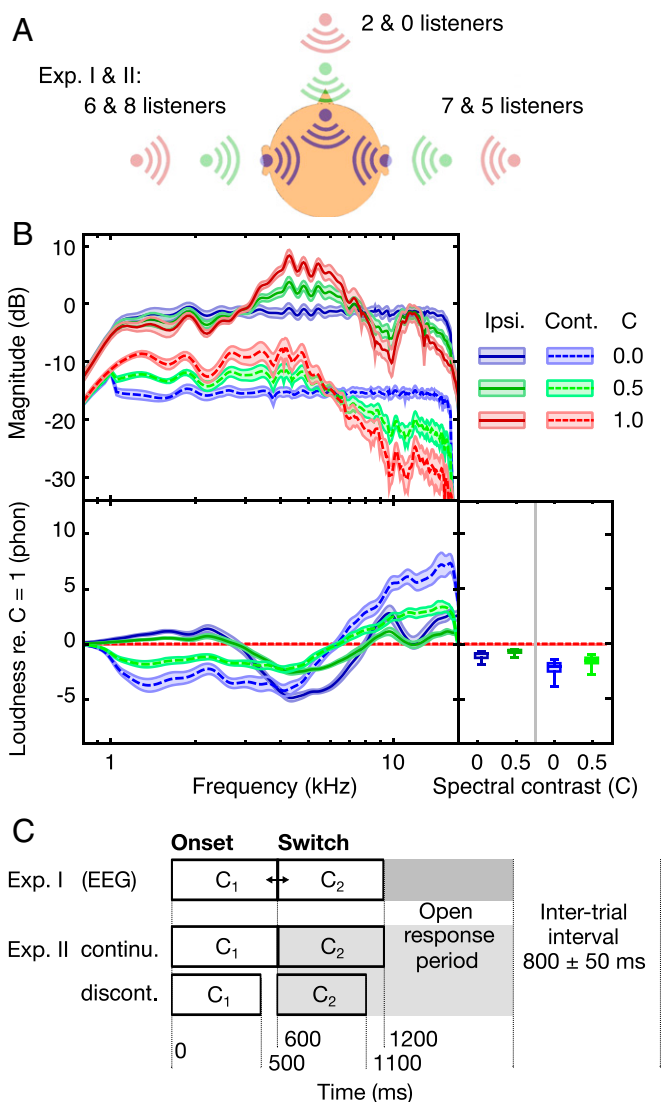


Fig. 1. Contrast of spectral cues naturally induced by the pinnae and torso was manipulated to create approaching and receding sounds. (A) Spatial configuration of hypothesized spatial percepts with number of listeners tested per source angle (left, front, and right) in Exps. I and II. (B) Effect of spectral contrast manipulation according to factor C on magnitude responses of listener-specific stimuli of Exp. I (B, Top) as well as their frequency-specific (B, Bottom Left) and overall (B, Bottom Right) loudness changes relative to $C = 1$. Shaded areas denote ± 1 SEM ($n = 15$). Note that changes in overall loudness oppose the intended effect of contrast switch. (C) Schematic representation of an experimental trial in Exp. I and Exp. II: starting with noise filtered according to C_1 , then cross-faded to C_2 , followed by an open response period, and a jittered intertrial interval. The left–right arrow in Exp. I represents a temporal jitter of ± 50 ms. Listeners were instructed to report whether the sound was perceived as approaching, receding, or static in Exp. I and only approaching or receding in Exp. II.

and contralateral (dashed lines) loudness. The predictions show that a decrease in spectral contrast was accompanied by an ILD increase at the lower two octaves and an ILD decrease at the higher two octaves.

Behavioral Results. In two experiments, listeners judged switches between spectral contrasts (C_1 and C_2) in forced choice paradigms (Fig. 1C). Experiment I allowed subjects three response alternatives [3-alternative forced-choice (AFC): “approaching,” “receding,” or “static”] and tested instantaneous spectral contrast switches

[using an interstimulus interval (ISI) of zero] while neural responses were measured using electroencephalography (EEG). Experiment II tested speeded responses for two alternatives (2-AFC: approaching or receding) and presented trials with either instantaneous but continuous (ISI of 0 ms) or discontinuous (ISI of 100 ms) changes in spectral shape across stimulus pairs. Experiment II served as a control to test whether looming bias is truly linked to the perception of sounds moving in distance rather than the mere percept of source proximity. Some listeners perceived spectral contrast manipulations as elevation changes at some source angles. Therefore, for each individual, we selected the source angle that led to the most consistent percept of a change in distance, based on a short initial run of the discrimination task (pretest) using three source angles (left, front, and right). Fig. 1A shows the source angles that were selected for the subsequent looming experiments.

Although listeners never received feedback about the hypothesized distance change, decreasing spectral contrasts ($C_1 > C_2$, solid orange lines in Fig. 2) were more likely perceived as approaching (filled orange triangles) and increasing spectral contrasts ($C_1 < C_2$, dashed green lines) as receding (open green triangles). These dominant associations were used for subsequent statistical analyses and were also observed in the pretest results (*Behavioral Results of Pretest for Individual Source Angle Selection* and Fig. S2). Trials without spectral changes ($C_1 = C_2$) in Exp. I were almost always perceived as “static,” confirming that the listeners remained vigilant and performed the task reliably.

In Exp. I (Fig. 2, *Left*), response consistency (percentage of responses for the same percept category), used here as a measure of cue salience, differed significantly between contrast pairs, i.e., the combination of spectral contrasts independent of presentation order ($0 \leftrightarrow 1$ vs. $0 \leftrightarrow 0.5$ vs. $0.5 \leftrightarrow 1$) ($F_{1,4,20} = 14$, $P < 0.01$, $\eta_p^2 = 0.51$). Listeners were most consistent for the largest contrast switches ($C: 0 \leftrightarrow 1$) and least consistent for pairs consisting of the

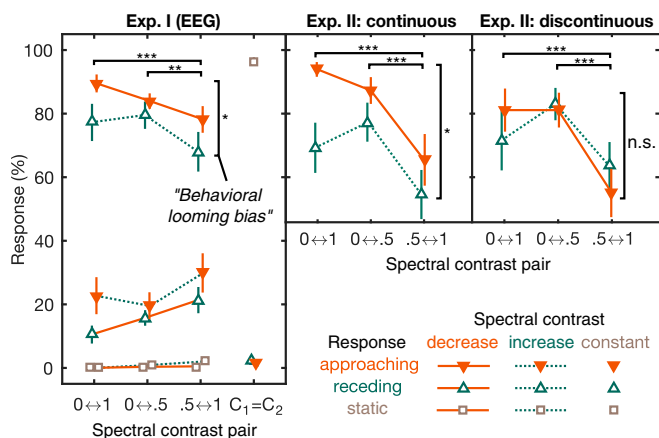


Fig. 2. Behavioral responses were more consistent for sounds perceived as approaching compared with sounds perceived as receding if instantaneous spectral changes were presented in continuous stimuli. Mean behavioral responses in the 3-AFC motion discrimination task of Exp. I (*Left*; $n = 15$) and the 2-AFC motion discrimination task of Exp. II (*Middle and Right*; $n = 13$). Results of Exp. II are separated between trials presenting instantaneous but continuous (*Middle*; as in Exp. I) and discontinuous (*Right*) spectral contrast switches. Decreasing spectral contrast switches (orange lines) were predominantly perceived as approaching (orange triangles), increasing spectral contrast switches (green lines) as receding (green triangles), and constant spectral contrasts (no lines) as static (gray squares). Statistical analyses focused on these predominant response associations. Behavioral looming bias was observed in terms of significantly higher consistency in approaching responses compared with receding responses only if spectral switches occurred within a continuous stimulus. Feedback was provided only on the detection accuracy of constant stimuli ($C_1 = C_2$) after blocks in Exp. I. Values reflect mean ± 1 SEM. Levels of significance: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

two higher contrast values ($C: 0.5 \leftrightarrow 1$; Fig. 2 shows significance levels of paired comparisons). Importantly, response consistency also differed significantly between switch directions, i.e., the presentation order within contrast pairs ($C_1 < C_2$ vs. $C_1 > C_2$) ($F_{1,14} = 6.9$, $P = 0.020$, $\eta_p^2 = 0.33$). Listeners were more consistent in labeling decreasing spectral contrasts than increasing spectral contrasts; this pattern is consistent with looming bias in that approaching sounds elicited more consistent identification responses than did receding sounds.

Response consistency in Exp. II also differed significantly between contrast pairs ($F_{1,9,22} = 40$, $P < 0.001$, $\eta_p^2 = 0.77$). As in Exp. I, judgments were least consistent for pairs consisting of the two higher contrast values ($C: 0.5 \leftrightarrow 1$). Overall response consistency was not affected by stimulus continuity (ISI = 0 ms vs. ISI = 100 ms) ($F_{1,12} = 0.55$, $P = 0.47$, $\eta_p^2 = 0.044$), but there were tendencies of a main effect of switch direction ($F_{1,12} = 4.7$, $P = 0.052$, $\eta_p^2 = 0.28$) and an interaction between stimulus continuity and switch direction ($F_{1,12} = 3.9$, $P = 0.072$, $\eta_p^2 = 0.25$). While interindividual differences in the amount of looming bias were generally large [continuity (cont.): $M = 15\%$, $SD = 19\%$; discontinuity (discont.): $M = -0.32\%$, $SD = 19\%$], only 1 of 13 listeners showed a markedly larger looming bias for discontinuous stimuli compared with continuous stimuli (data provided in *Outlier Evaluation for Experiment II* and Fig. S3). The interaction became highly significant after removing this one outlier from the analysis ($F_{1,12} = 15$, $P < 0.01$, $\eta_p^2 = 0.57$). Removal of this outlier had no significant effect on the other statistics.

Response times were measured with respect to the onset of spectral contrast switches and were evaluated only for the dominant response associations (data provided in *Response Times in Experiments I and II* and Fig. S4). In Exp. I, early response times (25% percentiles) were slightly faster for decreasing compared with increasing contrasts ($F_{1,14} = 5.2$, $P = 0.039$, $\eta_p^2 = 0.27$) and for larger contrast switches ($F_{1,5,21} = 6.0$, $P = 0.014$, $\eta_p^2 = 0.30$). Differences in response times for decreasing vs. increasing contrasts were smaller for later responses (30–90% percentiles). In Exp. II with speeded responses, variance in response times dramatically increased and neither early nor late response times differed significantly.

Event-Related Potentials. We investigated the neural underpinnings of auditory looming bias by means of ERPs. Fig. 3 shows the ERPs elicited by stimulus onset and spectral switch. The stimulus onset (Fig. 3A) was accompanied by a frontocentral negativity (onset-N1) followed by a central positivity (onset-P2). Average onset-N1 and onset-P2 amplitudes were measured for each spectral contrast (C_1) at the Cz site within the intervals of 80–140 ms and 140–280 ms, respectively, as determined by zero-crossings of the grand average. Onset-N1 amplitudes were not significantly different across spectral contrasts ($F_{1,6,23} = 1.9$, $P = 0.18$, $\eta_p^2 = 0.12$). Onset-P2 amplitudes were significantly larger for the original contrast ($C_1 = 1$) compared with the flattened contrast ($C_1 = 0$) but not to the intermediate contrast ($C_1 = 0.5$; see Fig. 3 for significance levels) ($F_{1,9,26} = 6.7$, $P < 0.01$, $\eta_p^2 = 0.32$).

The latencies of ERP components elicited by switches in spectral contrast were similar to the stimulus onset and were chosen at 90–150 ms for switch-N1 and 150–290 ms for switch-P2 (Fig. 3B). As for the onset-ERPs, these time ranges were determined on the basis of zero-crossings of the grand average; the only exception was the P2 end point, which had to be extrapolated because no final zero-crossing occurred. Switch trials ($C_1 \neq C_2$) all elicited N1–P2 complexes. Constant trials with no contrast switch ($C_1 = C_2$) elicited neither N1 nor P2 activity (dash-dotted line in Fig. 3B, *Lower Left*); these trials were therefore excluded from statistical analyses. Switch-N1 amplitudes differed significantly across contrast pairs ($F_{1,9,26} = 4.6$, $P = 0.022$, $\eta_p^2 = 0.25$) but not switch direction ($F_{1,14} = 1.8$, $P = 0.20$, $\eta_p^2 = 0.12$). After multiple comparison correction of post hoc paired comparisons, none of the switch-N1 amplitudes were significantly different across contrast

of ILDs (35), which is arguably a prominent auditory distance cue (18, 29, 36). Further experiments and/or modeling analyses will be required to clarify the perceptual weighting of these potential distance cues.

Perception of Motion in Distance but Not Changes in Overall Loudness Are Required to Induce Looming Bias. As summarized earlier, looming bias was previously represented in a broad variety of psychoacoustic measures that are either only remotely (e.g., ratings of loudness change or duration) or indirectly (e.g., absolute distance estimation) related to motion perception in distance. Instead of using more inferential measures, here, we directly measured response consistency in a forced-choice motion discrimination task because it was not necessarily clear that perceived distance, and thus induced looming biases, would be consistently affected using our spectral manipulation technique.

Our analysis shows that there was a small concomitant change in predicted overall loudness that opposed the dominant percept of motion direction. Thus, our data demonstrate that changes in overall sound intensity are not required to elicit looming bias. Moreover, in previous looming studies, overall sound intensity was usually changed slowly over time to simulate motion in distance. Here, we elicited looming bias by a near instantaneous (cross-fade of 10 ms) change of spectral cues in an ongoing stimulus. If the stimulus was interrupted by a short gap (100 ms) during which spectral cues changed, looming bias did not occur; these stimuli were likely perceived as two different auditory objects at different distances, rather than a single auditory object approaching or receding the listener. Distance perception itself was, however, negligibly affected by stimulus discontinuity as indicated by similar overall response consistencies.

Previous studies also argued that noise stimuli are less or not at all effective in inducing looming bias compared with tone complexes, because noise stimuli might be less likely perceived as a single auditory object (2, 11). The fact that we found looming bias based on noise stimuli suggests that this stimulus dependence might be specific to intensity-based distance simulations.

Taken together, our behavioral findings suggest that looming bias does not require specific tasks, distance cues, or stimuli for its activation; it rather constitutes a general increase of salience whenever sounds are perceived as moving toward vs. away from a listener.

Cortical Activity After 120 ms Reflected Salience of Spectral Cues and Perceptual Looming Bias. At stimulus onset, the similarity of onset-N1 amplitudes across conditions underscores that sound intensities were similar across spectral contrast conditions (37). Unlike onset-N1, central onset-P2 amplitudes increased with spectral contrast, i.e., salience of spectral cues. For sound localization, spectral cues are more important in the median plane compared with the horizontal plane. Consistent with our findings, a previous magnetoencephalography (MEG) study found that spatial deviants in the median plane were associated with increased neural response amplitudes within a similar latency range (200–250 ms) as our onset-P2, whereas spatial deviants in the frontal horizontal plane elicited earlier differences (100–150 ms) (38). Moreover, a previous ERP study tested spectral changes comparable to ours in a left- vs. rightward motion discrimination task and did not find any differences in onset-P2 amplitudes (25). For this task, however, spectral cue evaluation was not essential, whereas in our experiment, listeners had to evaluate changes in distance based on spectral cues. This suggests that in our study, the onset-P2 reflects early distance-evaluation processes based on spectral cues that may lead the percept of sound source distance.

At the onset of the spectral switch, larger spectral changes elicited larger switch-N1 magnitudes, independent of motion direction. Conversely, the switch-P2 magnitude depended on motion direction and reflected auditory looming bias. These results are in

line with the current functional understanding of how vertical and horizontal motion is reflected in ERPs. Specifically, switch-N1 responses are associated with the process of spectral change detection and switch-P2 responses with motion direction evaluation (23–26). Hence, our ERP results suggest a generalization of the motion onset response to the distance dimension and further support the conclusion that motion perception is required to elicit auditory looming bias.

Regarding the timing of looming perception, it has been shown that harmonic sounds increasing in intensity lead to enlarged visual cortical activity as early as 80 ms from stimulus onset (15, 22). Our results find looming bias reflected after 120 ms from the switch, primarily at central sites associated with auditory processing, which suggests that activation of looming bias requires higher-order processing of auditory motion. This apparent discrepancy may reflect the fact that processing of increments in intensity is less demanding and faster than sound motion inferences based on spectral analysis. Another possibility is that, at early processing stages, sound intensity increments are not associated with changes in distance but are simply salient events. For instance, interaural spatial cues occurring at intensity increments are weighted strongly when computing left–right location (39, 40). A strong effect of left–right congruence between increasing-intensity sounds and visual targets (41) further supports this reasoning.

Listeners showed large interindividual differences in the amount of behavioral looming bias they exhibited. Preliminary analyses showed that switch-P2 amplitudes and measures of global field power might be able to explain these interindividual differences with a potential specificity in gender (see *Correlation Between Interindividual Neural and Behavioral Looming Biases* and Fig. S5). While our sample size was sufficient to manifest the strong effect of looming bias within subjects ($\eta_p^2 \geq 0.33$), future investigations with larger and gender-balanced sample sizes are needed to explore the marked interindividual differences in the amount of behavioral looming bias.

In summary, our results demonstrate that distance cues other than overall sound intensity can elicit auditory looming bias and that its activation requires higher-level auditory processing related to motion perception. Previous studies on auditory looming bias were exclusively based on sound intensity as a distance cue and might have overinterpreted the effect of this particular stimulus feature. Future studies should review the cortical origins of auditory looming bias based on different modes of activation.

Materials and Methods

Subjects. We tested 15 paid volunteers within the age of 20–29 y ($M = 24$, $SD = 3.7$; 10 females, 5 males) in Exp. I and 12 paid volunteers plus one co-author within the age of 20–42 y ($M = 29$, $SD = 5.7$; 7 females, 6 males) in Exp. II. A subset of five listeners participated in both experiments. None of the listeners had any indication of a hearing loss greater than 20 dB relative to the threshold of the normal-hearing population, as confirmed by audiometric thresholds for frequencies from 500 Hz to 8 kHz. All subjects gave informed consent as overseen by the Boston University Institutional Review Board.

Stimuli and Procedure. Sound source locations were simulated based on individualized measurements of head-related transfer functions (HRTFs) (measurement procedure described in *Measurements of Head-Related Transfer Functions*). Stimuli consisted of two consecutively presented Gaussian white noises filtered by the band-limited HRTFs with a specific spectral contrast factor, $C \in \{0, 0.51\}$. The magnitude spectrum, $M_1(f)$, in decibels of the measured HRTFs was manipulated within a frequency range between 1 and 16 kHz (N_f frequency bins) according to C as:

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

with a frequency weighting function, w' , that approximates auditory frequency resolution by the across-frequency derivative of equivalent rectangular bandwidths (ERBs) (42). The measured phase spectrum remained unmodified. Filtered noise stimuli were band limited between 1 kHz and 16 kHz by a fourth-order Butterworth filter. Two filtered noise representations

were cross-faded by 10-ms sine ramps to form a stimulus pair. This stimulus pair was faded in and out by 50-ms squared sine ramps. Details on stimulus delivery are provided in [Stimulus Presentation and Experimental Procedure](#).

In Exp. I, 840 trials were presented per listener. The spectral contrast switched in 86% of all trials (14% constant trials). In Exp. II, 192 trials were presented and spectral contrast always switched (no constant trials). The presentation order of trials was completely randomized in both experiments. Listeners received block-based feedback only on their detection accuracy of constant trials in Exp. I (no feedback in Exp. II). Feedback was never provided on switch trials to prevent listeners from learning to use cues other than perceived distance (e.g., timbre) for discrimination. Further details on the experimental procedure are also provided in [Stimulus Presentation and Experimental Procedure](#).

EEG Recordings. The EEG of 32 scalp electrodes (Activetwo system with Activeview acquisition software, Biosemi B.V.) was recorded in Exp. I with standard 10/20 montage. In addition, one vertical and two lateral eye electrodes were recorded to better capture eye blinks and saccades, and external reference electrodes were placed at the mastoids (not used) and ear lobes. Real-time processing hardware (RP2.1, Tucker Davis Technologies, Inc.) marked timing of critical experimental events, which were recorded to an

additional data channel alongside the EEG data. The postprocessing procedure for EEG data analysis is described in [Processing of EEG Data](#).

Statistics. For statistical analyses, we used the Matlab Statistics Toolbox. We conducted repeated-measures analyses of variance with Greenhouse–Geisser correction (functions: `fitrm`, `ranova`). Tukey's honest significant difference tests were conducted for post hoc paired comparisons (`multcompare`). Effects were considered significant at $P < 0.05$.

Data Availability. Code for stimulus creation (`sig_baumgartner2017looming`), HRTFs, and experimental results (both `data_baumgartner2017looming`) are integrated in the Auditory Modeling Toolbox (43). EEG recordings and analysis scripts are provided via Zenodo (44).

ACKNOWLEDGMENTS. We thank Gerald Kidd Jr. and Norbert Kopčo for fruitful discussions and Tim Streeter and Leonard Varghese for technical support. This work was mainly funded by R.B.'s individual fellowship from the Austrian Science Fund (Project SpExCue, Grant J3803-N30). Additional funding was provided by the European Commission (Project ALT, Grant 691229) and the Boston University Hearing Research Center.

- Riskind JH, Kleiman EM, Seifritz E, Neuhoff J (2014) Influence of anxiety, depression and looming cognitive style on auditory looming perception. *J Anxiety Disord* 28: 45–50.
- Neuhoff JG (1998) Perceptual bias for rising tones. *Nature* 395:123–124.
- Bach DR, Neuhoff JG, Perrig W, Seifritz E (2009) Looming sounds as warning signals: The function of motion cues. *Int J Psychophysiol* 74:28–33.
- Seifritz E, et al. (2002) Neural processing of auditory looming in the human brain. *Curr Biol* 12:2147–2151.
- Grassi M (2010) Sex difference in subjective duration of looming and receding sounds. *Perception* 39:1424–1426.
- Neuhoff JG (2016) Looming sounds are perceived as faster than receding sounds. *Cogn Res Princ Implic* 1:15.
- Neuhoff JG, Planisek R, Seifritz E (2009) Adaptive sex differences in auditory motion perception: Looming sounds are special. *J Exp Psychol Hum Percept Perform* 35: 225–234.
- Neuhoff JG, Hamilton GR, Gittleston AL, Mejia A (2014) Babies in traffic: Infant vocalizations and listener sex modulate auditory motion perception. *J Exp Psychol Hum Percept Perform* 40:775–783.
- Bach DR, et al. (2008) Rising sound intensity: An intrinsic warning cue activating the amygdala. *Cereb Cortex* 18:145–150.
- Deneux T, Kempf A, Daret A, Ponsot E, Bathellier B (2016) Temporal asymmetries in auditory coding and perception reflect multi-layered nonlinearities. *Nat Commun* 7:12682.
- Ghazanfar AA, Neuhoff JG, Logothetis NK (2002) Auditory looming perception in rhesus monkeys. *Proc Natl Acad Sci USA* 99:15755–15757.
- Fay RR, Chronopoulos M, Patterson RD (1996) The sound of a sinusoid: Perception and neural representations in the goldfish (*Cassius auratus*). *Aud Neurosci* 2:377–392.
- Teghtsoonian R, Teghtsoonian R, Canévet G (2005) Sweep-induced acceleration in loudness change and the “bias for rising intensities”. *Percept Psychophys* 67:699–712.
- Maier JX, Ghazanfar AA (2007) Looming biases in monkey auditory cortex. *J Neurosci* 27:4093–4100.
- Romei V, Murray MM, Cappe C, Thut G (2009) Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Curr Biol* 19: 1799–1805.
- Bach DR, Furl N, Barnes G, Dolan RJ (2015) Sustained magnetic responses in temporal cortex reflect instantaneous significance of approaching and receding sounds. *PLoS One* 10:e0134060.
- McCarthy L, Olsen KN (2017) A “looming bias” in spatial hearing? Effects of acoustic intensity and spectrum on categorical sound source localization. *Atten Percept Psychophys* 79:352–362.
- Kopčo N, et al. (2012) Neuronal representations of distance in human auditory cortex. *Proc Natl Acad Sci USA* 109:11019–11024.
- Warren JD, Zielinski BA, Green GGR, Rauschecker JP, Griffiths TD (2002) Perception of sound-source motion by the human brain. *Neuron* 34:139–148.
- Hall DA, Moore DR (2003) Auditory neuroscience: The salience of looming sounds. *Curr Biol* 13:R91–R93.
- Tyll S, et al. (2013) Neural basis of multisensory looming signals. *Neuroimage* 65: 13–22.
- Cappe C, Thelen A, Romei V, Thut G, Murray MM (2012) Looming signals reveal synergistic principles of multisensory integration. *J Neurosci* 32:1171–1182.
- Getzmann S (2011) Auditory motion perception: Onset position and motion direction are encoded in discrete processing stages. *Eur J Neurosci* 33:1339–1350.
- Getzmann S, Lewald J (2012) Cortical processing of change in sound location: Smooth motion versus discontinuous displacement. *Brain Res* 1466:119–127.
- Getzmann S, Lewald J (2010) Effects of natural versus artificial spatial cues on electrophysiological correlates of auditory motion. *Hear Res* 259:44–54.
- Getzmann S, Lewald J (2010) Shared cortical systems for processing of horizontal and vertical sound motion. *J Neurophysiol* 103:1896–1904.
- Brimijoin WO, Boyd AW, Akeroyd MA (2013) The contribution of head movement to the externalization and internalization of sounds. *PLoS One* 8:e83068.
- Boyd AW, Whitmer WM, Soraghan JJ, Akeroyd MA (2012) Auditory externalization in hearing-impaired listeners: The effect of pinna cues and number of talkers. *J Acoust Soc Am* 131:EL268–EL274.
- Hartmann WM, Wittenberg A (1996) On the externalization of sound images. *J Acoust Soc Am* 99:3678–3688.
- Kulkarni A, Colburn HS (1998) Role of spectral detail in sound-source localization. *Nature* 396:747–749.
- Canzonieri E, Magosso E, Serino A (2012) Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PLoS One* 7:e44306.
- Moore BCJ, Glasberg BR, Baer T (1997) A model for the prediction of thresholds, loudness, and partial loudness. *J Audio Eng Soc* 45:224–240.
- Macpherson EA, Sabin AT (2013) Vertical-plane sound localization with distorted spectral cues. *Hear Res* 306:76–92.
- Baumgartner R, Majdak P, Laback B (2014) Modeling sound-source localization in sagittal planes for human listeners. *J Acoust Soc Am* 136:791–802.
- Brungart DS, Rabinowitz WM (1999) Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am* 106:1465–1479.
- Brungart DS (1999) Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am* 106:3589–3602.
- Rapin I, Schimmel H, Tourk LM, Krasnegor NA, Pollak C (1966) Evoked responses to clicks and tones of varying intensity in waking adults. *Electroencephalogr Clin Neurophysiol* 21:335–344.
- Fujiki N, Riederer KAJ, Jousmäki V, Mäkelä JP, Hari R (2002) Human cortical representation of virtual auditory space: Differences between sound azimuth and elevation. *Eur J Neurosci* 16:2207–2213.
- Dietz M, Marquardt T, Salminen NH, McAlpine D (2013) Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds. *Proc Natl Acad Sci USA* 110:15151–15156.
- Nelson BS, Takahashi TT (2010) Spatial hearing in echo environments: The role of the envelope in owls. *Neuron* 67:643–655.
- Leo F, Romei V, Freeman E, Ladavas E, Driver J (2011) Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Exp Brain Res* 213:193–201.
- Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. *Hear Res* 47:103–138.
- Søndergaard P, Majdak P (2013) The auditory modeling toolbox. *The Technology of Binaural Listening*, ed Blauert J (Springer, Heidelberg), pp 33–56, Available at [amtoolbox.sourceforge.net](#). Accessed August 10, 2017.
- Baumgartner R (2017) Data from “asymmetries in behavioral and neural responses to spectral cues demonstrate the generality of auditory looming bias.” Zenodo. Available at <https://dx.doi.org/10.5281/zenodo.832899>. Accessed August 10, 2017.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Delorme A, 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.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.

Supporting Information

Baumgartner et al. 10.1073/pnas.1703247114

Measurements of Head-Related Transfer Functions. To measure HRTFs, listeners were seated at the radial center of a half circle created by seven loudspeakers (Acoustic Research 215 PS; amplified by Crown 1002 XTi) in a semianechoic booth (IAC Acoustics, single-walled, 12 ft × 13 ft). The radius was 1.5 m and the loudspeakers were elevated approximately at ear level and provided an angular sampling of 30° ranging from −90° at the right of the listener to 90° at the left of the listener. Absorbing foam wedges were placed on the floor and the wall to the back of the listener to reduce early reflections. Binaural miniature microphones (AuPMC002; Ausim, Inc.) were inserted into the listener's ear canals, blocked by the microphone and foam seals. Transfer characteristics were identified by a maximum length sequence (MLS) of order 15, repeated 20 times. Recorded signals were amplified (Motu Ultralite) and analog-to-digital converted (Motu 24I/O) with a sampling rate of 44.1 kHz. Extracted impulse responses were windowed to 3 ms with 0.5-ms cosine ramps to remove reflections and equalized by reference measurements for each loudspeaker with the microphones placed at the center of the loudspeaker setup.

Stimulus Presentation and Experimental Procedure. Stimulus presentation was timed with Psychtoolbox (45) as illustrated in Fig. 1C. Stimuli were delivered via headphone preamps (HB7; Tucker Davis Technologies, Inc.) and tubeophones (ER-2; Etymotic Research, Inc.). Transfer properties were measured by an artificial ear coupler to calibrate the delivered sound pressure level for the frontal stimulus to be around 75 dB and to assure spectral magnitude deviations to be within ±5 dB as stated in the tubeophones' spec sheet. Further equalization of the headphone transfer functions was not applied because an informal listening test suggested that a perfect equalization does not seem to be critical to create externalized sounds with the headphones we used.

Before starting the EEG-monitored Exp. I, listeners completed a stimulus selection procedure (pretest), which consisted of two steps. First, listeners familiarized themselves with the sounds by listening to examples of the 1→0 contrast switch condition at all three different source angles and reported whether they associated the spectral contrast switch with a change in distance (alternative response options: no spatial change, change in elevation). The contrast switches were always associated with spatial changes, but for some listeners and angles, the manipulation only caused changes in perceived elevation. In total for the 15 listeners, one right source, five frontal sources, and zero left sources were excluded from further testing based on the listeners' reports. Second, listeners were tested on their ability to discriminate between approaching, receding, and static auditory percepts in a 3-AFC task to select the source angle (one out of three) for which each individual demonstrated best discriminability. All switch trials were tested six times, and all constant trials were tested two times. Contrast combinations were randomized and source angles were blocked in randomized order. Fig. 1A shows the selected distribution of source angles across listeners finally tested in the following EEG-monitored experiment. The whole stimulus selection procedure lasted up to 30 min.

In the EEG-monitored Exp. I, listeners performed the same discrimination task for the individually selected source angle with 120 repetitions of every switch trial and 40 repetitions each for the three constant trials. The presentation order of these 840 trials was completely randomized and grouped in 40 blocks of about 1 min. Listeners were able to take breaks between blocks and had to take breaks of at least 1 min after every 10th block. The whole recording session lasted about 1 h.

Before starting Exp. II, listeners performed only the first step of the pretest of Exp. I, which now also included discontinuous example stimuli with an ISI of 100 ms. If listeners reported that they associated example stimuli with changes in distance at multiple source directions, we randomly chose between left and right. Following this procedure, the left direction was selected for eight listeners and the right direction for five. The stimuli and motion discrimination task of Exp. II was identical to Exp. I, with the exceptions of including the condition of discontinuous stimuli with an ISI of 100 ms, allowing immediate responses, and removing the constant/static stimulus/response alternative. After two short blocks of 12 trials for training, listeners were tested on 12 blocks of 16 trials. Subjects completed the entire protocol for Exp. II in about 30 min.

Processing of EEG Data. Processing of EEG data was conducted with the Eeglab toolbox (version 13.6.5b) (46) in Matlab (Mathworks, Inc.). EEG signals were referenced to the two ear lobe electrodes. Three bad channels were identified manually throughout the recording sessions. A finite-impulse-response filter with Kaiser window design ($\beta=7.2$, $n=462$) limited the bandwidth of signals between 0.5 Hz and 20 Hz. Bandlimited signals were epoched and then resampled to 100 Hz. Epochs were baseline corrected to the 200-ms long time window preceding the epoching event. Epochs exceeding thresholds of −200 μ V and 800 μ V for brain channels and −800 μ V and 200 μ V for eye channels were removed. The extended Infomax approach was used to analyze independent signal components (option: binica). Up to three independent components per subject were manually selected to correct for eye artifacts. After this correction of eye artifacts, bad channels were interpolated and brain channel epochs that exceeded a final threshold of ±70 μ V were removed. For every listener and condition, at least two thirds of the trials (minimum 83 trials per condition) did not exceed the final threshold for rejection. The number of trials per condition was equalized within listeners by removing trials selected to be equally distributed across the recording session.

We used the Matlab Fieldtrip toolbox (FTP release 20160526) (47) to perform a cluster-based permutation test (function: `ft_timelockstatistics`) (48). The significance probability of a cluster was calculated as the Monte Carlo estimate of the P value under a distribution of 500 permutations. Dependent single-sample T statistics (two tailed) were thresholded at an alpha level of 0.05 and summed up to form cluster-level statistics. The maximum of the cluster-level statistics was then evaluated as the actual test statistic (alpha level of the permutation test: 0.025).

Interindividual Comparison of Frequency-Specific Loudness Changes. Fig. S1 shows that the frequency-specific loudness changes relative to the reference spectral contrast $C = 1$ are similar across listeners in Exp. I. Decreasing spectral contrast generally caused loudness to decrease at low to mid frequencies and to increase at high frequencies, and caused interaural level differences to increase at low frequencies and decrease at high frequencies. Only the two listeners—those tested for frontal sources—are exceptional because interaural differences are very small for positions in the median plane.

Behavioral Results of Pretest for Individual Source Angle Selection. Fig. S2 shows the results of the pretest for source angle selection. The experimental results were very similar to those from the EEG experiment (Exp. I) despite a generally smaller response consistency. Response consistency improved in the EEG experiment

possibly because of the individual selection of the source angle and accumulated task experience.

Outlier Evaluation for Experiment II. One listener (S09) of Exp. II showed a very strong looming bias in response consistency for the discontinuous stimuli, but a negative looming bias in response consistency for the continuous stimuli (Fig. S3). This behavior was very different from the other listeners of Exp. II and also from her own results from Exp. I, where she showed a small but positive consistency bias. For this reason, we excluded her results from statistical analyses of Exp. II.

Response Times in Experiments I and II. In contrast to the clear representation of looming bias in response consistency, its representation in response times was not clear. On the one hand, the fastest response times (25% percentiles) in Exp. I reflected the auditory looming bias (Fig. S4, *Left*), even though the listeners were not instructed to respond as fast as possible and they had to withhold their response until stimulus offset. On the other hand, later response times (30–90% percentiles) in Exp. I as well as speeded responses in Exp. II (Fig. S4, *Middle* and *Right*) were not significantly different between increasing and decreasing contrast conditions. Late responses are probably less distinctive because response certainty is better reflected in early responses. The large variance of response times in Exp. II suggests that at least some listeners had difficulties in rapidly discriminating motion direction and providing their response. More training and/or a different response method may be required to reveal differences in response times.

Correlation Between Interindividual Neural and Behavioral Looming Biases. According to our ERP analyses differences in (*i*) central switch-P2 amplitudes (“P2 amplitude bias”) and (*ii*) cluster-average potentials of a broad spatiotemporal cluster centered around 160 ms and central sites (“cluster amplitude bias”) were potential candidates for predictors of individual behavioral looming bias. We used multivariate linear regression models (Matlab Statistics Toolbox: fitlme) to assess these predictive relationships. To account for interindividual differences in overall ERP intensity, we standardized (*z* score) potentials within subjects and quantified the individual neural bias as the *z*-score difference between increasing and decreasing contrasts, averaged across

contrast pairs. Similarly, individual behavioral looming biases were quantified as the standardized percentage difference between consistent approaching and receding responses, also averaged across contrast pairs.

We first investigated predictability based on the P2 amplitude biases. A simple linear regression between P2 amplitude bias and behavioral bias would have suggested that the biases were not related to each other ($r = 0.22, P = 0.37$) (Fig. S5, *Left*). However, in line with a previous study based on skin conductance (3), we observed gender differences in the distribution of the P2 amplitude bias. Individual P2 amplitude biases were positive for every single female listener, whereas the biases of the males were widely distributed. Including gender in the regression model revealed a relationship between P2 amplitude bias and behavioral bias ($\beta = 33, t_{11} = 6.3, P < 0.001$) that depended on gender (male coded as 1: $\beta = 29, t_{11} = 5.6, P < 0.001$; interaction term: $\beta = -38, t_{11} = 6.1, P < 0.001$). Pearson’s correlation coefficients further suggested that correlation between P2 amplitude bias and behavioral looming bias was strong for females ($r = 0.95, P < 0.001$) but not statistically significant for males ($r = -0.28, P = 0.45$).

In contrast to this strong correlation between P2 amplitude bias and behavioral bias for female listeners, similar regression analyses based on the cluster amplitude biases revealed no significant relationships (Fig. S5, *Middle*).

A previous looming study showed that enhancements in reaction times due to audiovisual looming congruency were correlated with enhancements in global field power (22). Hence, we also evaluated differences in global field power within the time range of the cluster and analyzed its correlation with the behavioral looming bias (Fig. S5, *Right*). The power differences were not biased as much as the P2 amplitude or cluster amplitude differences, but revealed surprisingly high correlation with behavioral looming biases, even when both genders were pooled together.

It is unclear why the less-biased differences in global field power better reflect interindividual differences in behavioral looming bias than do the maximally biased differences in cluster amplitude. Future investigations with a larger sample size will be necessary to clarify these relationships also with respect to gender differences.

