Assessing the Influence of Sound Parameters on Crossmodal Cuing in Different Regions of Space

Jae Lee & Charles Spence

Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Oxford, UK

DATE: February, 2018

CORRESPONDENCE TO: Jae Lee, Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Oxford, OX1 3UD, UK; E-MAIL: jae.lee3@psy.ox.ac.uk.

Introduction

Research on audiovisual exogenous spatial cuing has demonstrated that the presentation of a task-irrelevant auditory cue typically leads to a short-lasting exogenous shift of visual attention to the cued region of space (see Spence & McDonald, 2004; Spence, McDonald, & Driver, 2004, for reviews). Such crossmodal spatial cuing effects are typically manifested in terms of faster reaction times (RTs) to visual targets presented from the same, rather than opposite, hemifield as the cue. Furthermore, the presentation of an auditory cue exogenously orients visual attention not only to the cued hemifield, but also narrowly to the cued region of space within the hemifield[1] (Lee & Spence, 2017; though see also Lee & Spence, 2015). Exogenous spatial cuing effects have now been documented between all possible combinations of auditory, visual, and tactile stimuli (see Spence et al., 2004, for a review).

Despite the extensive body of research on the topic of exogenous audiovisual spatial attention that has been published to date, most studies have investigated how the positioning of auditory cues modulates spatial attention crossmodally. That is, there has been little consideration as to how, and even whether, sound parameters such as duration, intensity change, and waveform structure modulate the crossmodal spread of attention. That said, there is mounting evidence to show that ecologically meaningful sounds, such as looming (i.e., rising-intensity over time) auditory cues, elicit a stronger perceptual bias than other sounds such as receding (i.e., decreasing-intensity) or constant-intensity sounds (see Bach et al., 2008; Cappe, Thut, Romei, & Murray, 2009; Ghazanfar, Neuhoff, & Logothetis, 2002; Leo, Romei, Freeman, Ladavas, & Driver, 2011; Maier, Neuhoff, Logothetis, & Ghazanfar, 2004; Morrongiello, Hewitt, & Gotowiec, 1991; Romei, Murray, Cappe, & Thut, 2009). Interestingly, however, such perceptual biases have been reported from structured tones (i.e., triangular waveforms), but not from pure tone or white noise looming sounds (e.g., Ghazanfar et al., 2002; Leo et al., 2011; Maier et al., 2004; Romei et al., 2009).

To date, the perceptual biases towards structured looming sounds in preference to structured receding sounds have been shown in terms of, for instance, the extended duration of

[1] The presentation of an auditory cue exogenously orients visual attention narrowly to the cued region of space. As a result, RTs are maximally facilitated to visual targets presented there rather than from a different position in the cued hemifield. It should, however, be noted that RTs to visual targets depend not only on spatial cuing but also the eccentricity of the visual targets in the horizontal plane. As a result, RTs to targets presented from the cued region of space, despite the maximum facilitation effect, may still be slower than those from a different target position within the cued hemifield.

behavioural orienting responses (in rhesus monkeys; Ghazanfar et al., 2002), increased sensitivity when discriminating visual orientation (in human participants; Leo et al., 2011), overestimating the change in sound intensity (with human participants; Neuhoff, 1998, 2001), and underestimating the arrival time of an approaching sound (with human participants; Neuhoff, Planisek, & Seifritz, 2009). Leo et al.'s (2011) study, in particular, demonstrated that visual orientation sensitivity to Gabor patches was larger when structured looming sounds were presented from the same side as the patches than from the opposite side of fixation. Based on these findings, Leo et al. suggested that the perceptual bias towards structured looming sounds in preference to receding sounds can occur in a spatially-specific manner, and influence the perception of visual stimuli. Such findings suggest that the magnitude of a spatial cuing effect would be larger following the presentation of a structured looming auditory cue than a structured receding cue (this will be our first hypothesis). On the other hand, the presentation of a white noise cue, regardless of whether the cue (intensity) is looming or receding, would elicit the same (i.e., not statistically different) magnitude of spatial cuing effects (this will be our second hypothesis; cf. Ghazanfar et al., 2002; Leo et al., 2011; Romei et al., 2009).

In addition to the magnitude of a spatial cuing effect, we were also interested in knowing whether the presentation of an auditory cue would exogenously orient crossmodal spatial attention narrowly to the cued region of space. Although Leo et al. (2011) argued that the perceptual bias from structured looming sounds led to the *spatially-specific* enhancement of visual orientation sensitivity, the stimuli were presented only from either left or right side, and in the frontal region of space. As a result, it is unclear whether structured looming sounds presented from the rear would have elicited a statistically smaller crossmodal effect for the frontal visual stimuli than structured looming sounds presented from the front. Lee and Spence (2015) reported that the presentation of a constant-intensity auditory cue (either pure tone or white noise, with a duration of 100ms) facilitated the perception of the frontal visual targets presented ipsilaterally as compared to those presented contralaterally, regardless of whether the cue was presented from the front or rear (hereafter, this will be referred to as the rear-to-front crossmodal spatial cuing effect; see Spence, Lee, & Van der Stoep, 2018, for a review). Therefore, our third hypothesis was that the presentation of an auditory cue either in front or rear would elicit the same magnitude of crossmodal spatial cuing effects.

The present study design involved four within-participants factors: Cue Position (front vs. rear), Cue Type (looming vs. receding),[2] Cue Structure (structured vs. white noise), and Spatial Cuing (cued if the cue and target are on the same left or right side vs. uncued if not), and one between-participants factor: Cue Duration (250 vs. 500ms).[3] The RT data with the four within-participants factors were entered into a repeated measures analysis of variance (RM-ANOVA) for each cue duration condition. We expected a significant three-way interaction between Cue Type, Cue Structure, and Spatial Cuing based on the first and second hypotheses, and no significant two-way interaction between Cue Position and Spatial Cuing based on the third hypothesis. All of the main effects and any interactions involving the factor of Spatial Cuing were reported; any other interactions were ignored.

The three hypotheses were tested using the RT data for each cue condition. In order to test the first hypothesis, a planned paired sample *t*-test was conducted between the magnitude of the cuing effect in the structured looming cue condition and that in the structured receding cue condition for each cue duration. An equivalent planned paired sample *t*-test was conducted between the magnitude of the cuing effect in the white noise looming cue condition and that in the white noise receding cue condition for each cue duration. In order to investigate the third hypothesis, a planned paired sample *t*-test was conducted between the magnitude of the cuing effect in the frontal cue condition and that in the rear cue condition for each cue duration. Following the analysis of the RT data, a RM-ANOVA for each cue duration was conducted with the error rate data in order to investigate whether there was any speed-accuracy trade-off in any of the spatial cuing effects. All of the main effects were reported, as well as any significant interactions involving the factor of Spatial Cuing was involved.

---

[2] The intensity of looming and receding auditory cues varied between 55-75dB(A) as measured from the participant's head position. Since auditory stimuli above 15dB SPL are audible (see Sabin, Macpherson, & Middlebrooks, 2005), the perceived onset timing of a receding cue should have been identical to that of a looming cue.

[3] It is often suggested that crossmodal spatial cuing effects typically last for 300ms or less from the onset of a brief cue (Spence et al., 2004; see also Fuentes & Campoy, 2008). However, there is no clear evidence as to exactly when they dissipate. In Spence and Driver's (1997) study, for instance, audiovisual crossmodal spatial cuing effects were documented at the stimulus onset asynchronies (SOAs) of 100ms and 200ms, but not at the 700ms SOA. In Lee and Spence's (2015) study, audiovisual spatial cuing effects were documented at the SOAs of 100, 200, and 700ms (when the stimuli were presented from frontal space), although reduced at the 700ms SOA as compared to those at the shorter SOAs. Since the duration of crossmodal spatial cuing effects is not so clear-cut, we did not necessarily expect to see any modulation of Spatial Cuing as a function of Cue Duration.

Methods

Participants

Forty-four (15 males and 29 females) took part in this study.[4] They were recruited via the Crossmodal Research Laboratory mailing list, the Oxford Psychology Research participant recruitment scheme, and the Oxford University Experimental Psychology Research participant recruitment scheme. Their average age was 26 years, ranging from 18 to 49 years. All of the participants reported normal (or corrected-to-normal) vision and hearing. All were right-handed by self-report. The participants were randomly assigned to one of two cue duration conditions. The experiment lasted for approximately 30-40 minutes depending on the duration of the cue. At the end of the study, the participants were either given two course credits, or else paid £7 if assigned in the 500ms cue duration condition or £5 in the 250ms cue duration condition for having taken part in the study. The study reported in this manuscript was approved by the Medical Sciences Interdivisional Research Ethics Committee and the University of Oxford (MSD-IDREC-C1-2014-019) and was conducted in line with the guidelines provided.

Apparatus and materials

The experiment was conducted in a darkened room using MATLAB R2014a with PSYCHTOOLBOX 3.0.12 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). The participants were seated facing a red LED (12v 5mm with a luminance of 8,000 millicandelas) as a fixation point with a computer keyboard on their lap. There was a loudspeaker (Ricco 2.0 Channel Wooden Speaker Home Hifi System, model number: T2018) on each side of the fixation point at eye-level (117cm above the floor). Two additional loudspeakers were placed behind the participant's head, parallel to the front loudspeakers. Each loudspeaker was equipped with a single-cone, capable of producing frequencies between 80Hz and 20kHz.

[4] Lee and Spence (2015) successfully demonstrated exogenous crossmodal spatial cuing effects with twenty-five participants (and 432 trials per participant) with an effect size, partial eta squared (η_p^2) = .514. A priori power analysis using G*Power 3 (Fault, Erdfelder, Lang, & Buchner, 2007) revealed that, given the effect size equal to .514, a sample of eleven participants would provide a statistical power of 83% to detect a main effect of Spatial Cuing. Therefore, a sample size of twenty or more for each between-participants factor (23 participants in the 250ms cue duration condition; 21 participants in the 500ms cue duration condition) and a total of 384 trials should provide enough power to detect any spatial cuing effects.

The front and rear loudspeakers were separated by 128cm, and the left and right loudspeakers by 74cm, both measured from the centres of each single-cone (see Figure 1).

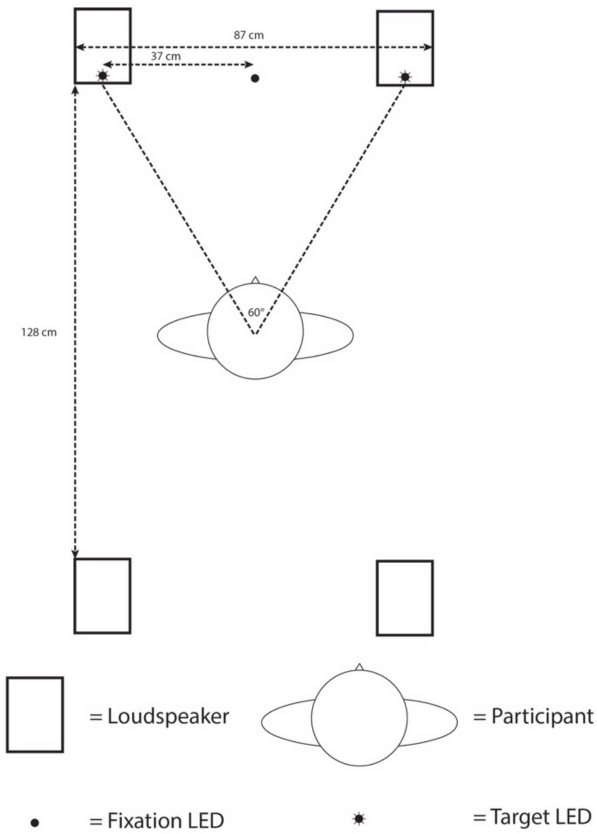

Figure 1. Bird's-eye view of the placement of the four loudspeakers, visual target location (only top two of four shown), and the fixation point in relation to the participant.

Two red target LEDs (the same as the fixation LED) were placed on each side, 9cm above and below each loudspeaker, as measured from the centre of the single-cone. Auditory stimuli, sampled at 44.1kHz, were generated using Adobe Audition CC 2015. Structured tonal cues were 400Hz triangular waveforms (similar to those used by Leo et al., 2011; Romei et al., 2009). Looming sounds had an exponential intensity change from 55 to 75dB(A), and receding sounds from 75 to 55dB(A) as measured from the participant's head position. Each auditory cue was generated separately using the same procedure described above. The auditory cues had 5ms onset and offset ramps to avoid clicks.

Design

The study design consisted of four within-participants factors: Cue Position, Cue Type, Cue Structure, and Spatial Cuing, and a between-participant factor: Cue Duration. Crossing the four within-participants factors yielded 16 possible conditions, with each condition being presented pseudo-randomly 24 times. Each participant completed three blocks of 128 trials. The participants were given a chance to take a short break between blocks.

Procedure

Each trial started with the illumination of the fixation LED. After a random delay of 400-650ms following the onset of the fixation, a spatially-nonpredictive auditory cue was presented from one of four possible locations: front-left, front-right, rear-left, or rear-right. Following the offset of the auditory cue, there was a random delay of 100-200ms before the onset of a visual target. The duration of any random delay on each trial was chosen from the discrete uniform distribution. The SOA was therefore 350-450ms and 600-700ms in the 250ms and 500ms cue duration conditions, respectively. The visual target was presented for 140ms from one of four possible corner locations: top-left, top-right, bottom-left, or bottom-right. Using the computer keyboard, the participants pressed the *up* arrow key if the target was presented from either the top-left or the top-right, or the *down* arrow key if the target was presented from either the bottom-left or the bottom-right as rapidly and accurately as possible. Once the participant’s response had been registered, or else if the participant had not responded in the 2,000ms after the onset of the visual target, the illumination of the fixation light was terminated and the next trial began. The instructions were provided on the computer screen in front of the participant, below the fixation light. Following the end of the instruction screen, the participants completed ten practice trials. The experimenter answered any questions that the participant might have and left the experimental room before the actual experimental trials began.

Results

First, participants who failed to respond to over 95% of the total trials were excluded from the data analysis. One participant failed to respond to 10% of the total trials in the 500ms cue duration condition, and was therefore excluded from the data analysis. Second, any outliers were identified based on Tukey's (1977) method[5] for each cue duration. A box-plot of the participants' mean RTs in the 250ms cue duration condition revealed a median of 365ms, with 349ms and 385ms as Q_1 and Q_3, respectively. Another box-plot of the participants' error rates in the 250ms cue duration condition revealed a median of 1.0%, with 0.1% and 2.9% as Q_1 and Q_3, respectively. Based on the Tukey's method, two participants' RTs (Ms = 476ms, 580ms) and two other participants' error rates (Ms = 6.5%, 48.5%) in the 250ms cue duration condition were above the upper limits (Ms = 441ms and 6.4%, for the mean RT and the mean error rate, respectively), and were therefore removed from further data analysis. Lastly, the following trials were further excluded from the remaining data: trials with incorrect responses, responses immediately following an incorrect response, no response trials, and trials with target RTs slower than 1,500ms or faster than 150ms. The application of the third exclusion criterion resulted in the removal of 214 trials (2.9% of the remaining data) in the 250ms cue duration condition.

A box-plot of the participants mean RTs in the 500ms cue duration revealed a median of 366ms, with 332ms and 441ms as Q_1 and Q_3, respectively. Another box-plot of the participants' error rates in the 500ms cue duration revealed a median of 1.3%, with 0.4% and 2.5% as Q_1 and Q_3, respectively. No outliers were identified in the 500ms cue duration condition based on the Tukey's (1977) method. The application of the third exclusion criterion used in the 250ms cue duration condition resulted in the removal of 246 trials (3.2% of the remaining data). The mean RTs and error rates for each condition are summarised in Table 1.

[5] Outliers were identified, first in the RT data, as those participants whose mean RTs across all experimental trials were faster than the lower limit or slower than the upper limit following the Tukey's (1977) method. According to Tukey, the lower limit was defined as 25-percent quartile (Q_1) subtracted by the interquartile range (IQR) multiplied by 1.5 and the upper limit as 75-percent quartile (Q_3) added by the IQR multiplied by 1.5 (see Tukey, 1977). Outliers in the error rate data were also identified using the same method.

Table 1.

Mean Reaction Times (RTs) in milliseconds (ms), Their Within-participants Standard Errors based on Morey's (2008) method, and Mean Error Rates in percentages.

Cue duration (ms)	Cue position	Cue type	Cue structure	Spatial cuing	RTs (ms)	Standard error (ms)	Error rates (%)
250	Front	Looming	Structured	Cued	345	4.1	1.1
				Uncued	364	3.6	1.5
			White noise	Cued	353	3.8	0.7
				Uncued	374	5.2	3.5
		Receding	Structured	Cued	350	3.7	0.9
				Uncued	361	2.8	1.5
			White noise	Cued	347	3.9	1.3
				Uncued	364	3.8	0.9
		Looming	Structured	Cued	345	3.9	1.8
				Uncued	360	3.0	1.5
			White noise	Cued	359	4.1	0.9
				Uncued	362	3.0	1.3
		Receding	Structured	Cued	356	3.5	1.3
				Uncued	361	3.7	1.8
			White noise	Cued	353	4.2	1.5
				Uncued	366	4.0	1.5
500	Rear	Looming	Structured	Cued	368	3.6	1.0
				Uncued	384	3.4	2.3
			White noise	Cued	377	4.0	1.3
				Uncued	396	4.8	1.3
		Receding	Structured	Cued	387	3.5	1.7
				Uncued	387	3.5	2.5
			White noise	Cued	379	4.7	0.4
				Uncued	398	3.9	1.0
		Looming	Structured	Cued	381	3.1	2.3
				Uncued	382	4.0	2.1
			White noise	Cued	382	3.7	0.8
				Uncued	395	5.4	1.7
		Receding	Structured	Cued	388	4.5	1.7
				Uncued	389	4.8	1.9
			White noise	Cued	392	4.2	2.1
				Uncued	398	3.9	1.7

RT data analysis – 250ms cue duration

A RM-ANOVA was conducted with the within-participants factors of Cue Position, Cue Type, Cue Structure, and Spatial Cuing. The analysis revealed a significant main effect of Spatial Cuing, $F(1, 18) = 40.701$, $MSE = 381.034$, $p < .001$, $\eta_p^2 = .693$, with RTs to targets being faster when the cue and target were presented from the same (M = 351ms) rather than from different sides (M = 364ms). A significant two-way interaction was found between Cue Position and Spatial Cuing, $F(1, 18) = 4.670$, $MSE = 313.231$, $p = .044$, $\eta_p^2 = .206$. Paired-comparisons with Bonferroni correction ($\alpha = .025$) revealed that in the frontal Cue Position condition, RTs to visual targets were significantly faster (by 17ms) in the cued (M = 349ms) than in the uncued condition (M = 366ms), $t(18) = -5.886$, $p < .001$. In the rear Cue Position condition, on the other hand, RTs to targets were significantly faster (by 9ms) in the cued (M = 354ms) than in the uncued (M = 362ms) condition, $t(18) = -3.008$, $p = .008$. Unexpectedly, there was no significant three-way interaction between Cue Type, Cue Structure, and Spatial Cuing, $F(1, 18) = 2.952$, $MSE = 220.806$, $p = .103$, $\eta_p^2 = .141$.

A planned paired sample *t*-test was conducted between the magnitude of the crossmodal spatial cuing effect reported in the structured looming cue condition and that reported in the structured receding cue condition. This analysis revealed that the magnitude of the cuing effect in the structured looming cue condition (M = 17ms) was significantly larger than that reported in the structured receding cue condition (M = 8ms), $t(18) = -2.241$, $p = .038$. Therefore, the first hypothesis was confirmed (see Figure 2). Another planned paired sample *t*-test was conducted between the magnitude of the crossmodal cuing effect reported in the white noise looming cue condition and that reported in the white noise receding cue condition. The analysis revealed that the magnitude of the cuing effect in the white noise looming cue condition (M = 12ms) was no different from that reported in the white noise receding cue condition (M = 15ms), $t(18) = .567$, $p = .578$. Therefore, the second hypothesis was confirmed. The magnitude of the cuing effect in the frontal cue condition (M = 17ms) and that in the rear cue condition (M = 9ms) were entered into a planned paired sample *t*-test. The analysis revealed that the magnitude of the cuing effect in the frontal cue condition was significantly larger than that in the rear cue condition, $t(18) = -2.122$, $p = .048$. Therefore, the third hypothesis was rejected; the presentation of a looming or receding auditory cue with a duration of 250ms elicited spatially-specific crossmodal cuing effects (see Figure 3).

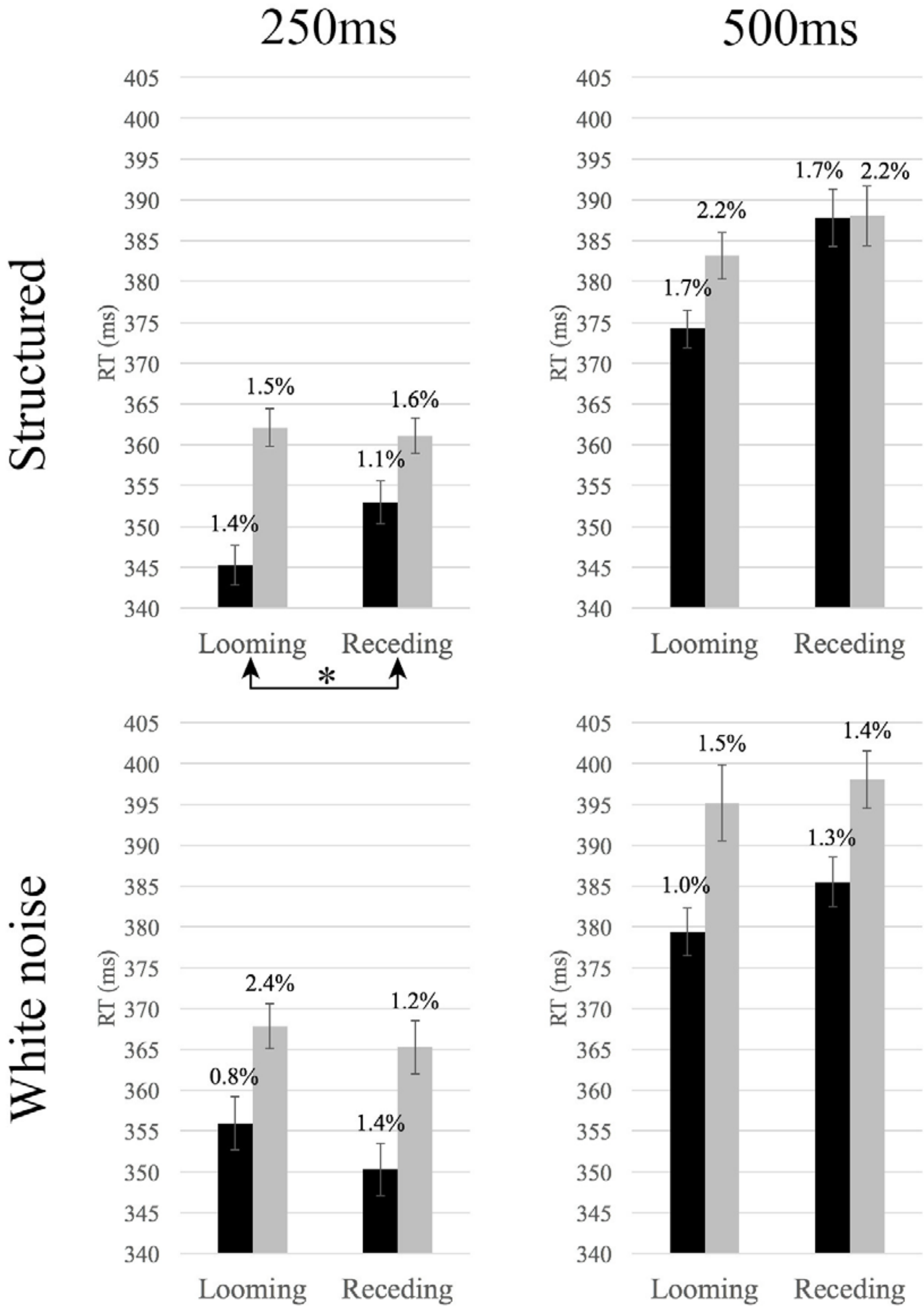

Figure 2. Mean reaction times (RTs) in milliseconds (ms) and error rates (%) as a function of the Cue Type for each cue structure condition. Black bars represent the mean RTs in the condition in which the cue and target were presented from the same side. Grey bars represent the mean RTs in the condition in which the cue and target were presented from different sides.

The left and right columns represent the mean RTs in the 250ms and 500ms cue duration conditions, respectively. The asterisk indicates a significant perceptual bias towards looming in preference to receding sounds from the RT data based on a planned pairwise comparison (*t*-test) with an alpha level of .05. Vertical lines indicate standard errors of mean RTs.

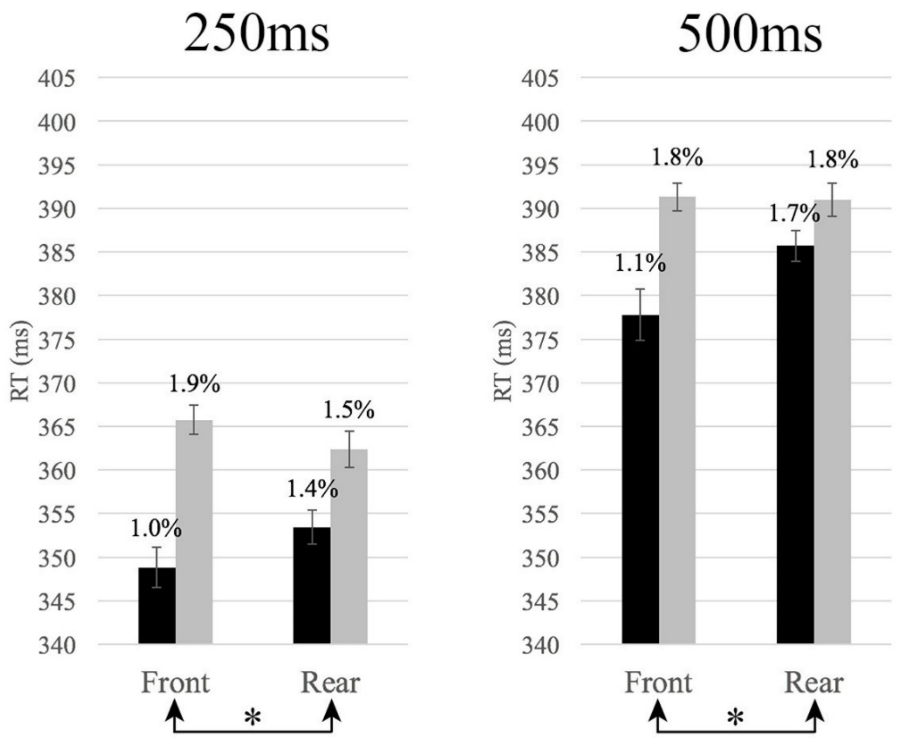

Figure 3. Mean reaction times (RTs) in milliseconds (ms) and error rates (%) as a function of the Cue Position for each cue duration. Black bars represent the mean RTs from trials in which the cue and target were presented from the same side. Grey bars represent the mean RTs from trials in which the cue and target were presented from opposite sides. The asterisks indicate the magnitude of the cuing effect in the frontal cue condition being significantly larger than that in the rear cue condition, from the RT data based on a planned pairwise comparison (*t*-test) with an alpha level of .05. Standard errors of mean RTs are indicated by the vertical lines.

Error rate data analysis – 250ms cue duration

An equivalent four-way RM-ANOVA was conducted with the error rate data in the 250ms cue duration condition. There was a three-way interaction between Cue Type, Cue Structure, and Spatial Cuing, *F*(1, 18) = 11.744, *MSE* < .001, *p* = .003, η_p^2 = .395. In order to break the interaction down, two RM-ANOVAs were conducted, one for each cue structure condition. The analysis with the structured cue condition revealed no significant terms. An equivalent RM-ANOVA with the white noise cue condition revealed a significant main effect of Spatial Cuing, *F*(1, 18) = 6.006, *MSE* < .001, *p* = .025, η_p^2 = .250, with participants making fewer errors when the cue and target were presented ipsilaterally (M = 1.1%) than when they were presented contralaterally (M = 1.8%), *t*(18) = -2.460, *p* = .024. There was also a significant two-way interaction between Cue Type and Spatial Cuing, *F*(1, 18) = 5.260, *MSE* < .001, *p* = .034, η_p^2 = .226. Paired-comparisons with Bonferroni correction (α = .025) revealed that, in the white noise looming cue condition, error rates in response to targets were significantly lower when the cue and target were presented ipsilaterally (M = 0.8%) than when they were presented contralaterally (M = 2.4%), *t*(18) = -3.521, *p* = .002. In the white noise receding cue condition, on the other hand, error rates in response to targets were not significantly different whether the cue and target were presented ipsilaterally (M = 1.4%) or contralaterally (M = 1.2%), *t*(18) = .427, *p* = .675. Taken together, the spatial cuing effects reported in the 250ms cue duration condition reflect a genuine attentional facilitation rather than a speed-accuracy trade-off (see Duncan, 1980; Spence, Pavani, & Driver, 2000).

RT data analysis – 500ms cue duration

A RM-ANOVA with the within participants factors of Cue Position, Cue Type, Cue Structure, and Spatial Cuing revealed a significant main effect of Cue Position, *F*(1, 19) = 4.677, *MSE* = 293.564, *p* = .044, η_p^2 = .198, with participants responding more rapidly to the visual targets following the presentation of an auditory cue in the front (M = 383ms) than from the rear (M = 388ms). The analysis also revealed a significant main effect of Cue Type, *F*(1, 19) = 7.502, *MSE* = 525.893, *p* = .013, η_p^2 = .283, with participants responding more rapidly to the targets following the presentation of a looming auditory cue (M = 382ms) than a receding cue (M = 389ms). More importantly, there was a significant Spatial Cuing effect, *F*(1, 19) = 11.374, *MSE* = 580.550, *p* = .003, η_p^2 = .374, with RTs to targets being faster when the cue and target were presented from the same (M = 381ms) rather than from different sides (M = 390ms).

The analysis revealed a two-way interaction between Cue Position and Spatial Cuing, $F(1, 19) = 5.107$, $MSE = 210.072$, $p = .036$, $\eta_p^2 = .212$. Paired-comparisons with Bonferroni correction ($\alpha = .025$) revealed that in the frontal cue condition, RTs to visual targets were significantly faster (by 13ms) in the cued (M = 377ms) than in the uncued condition (M = 390ms), $t(19) = -3.488$, $p = .002$. In the rear cue condition, RTs to targets were not significantly different between the cued (M = 385ms) and uncued (M = 390ms) conditions, $t(19) = -2.171$, $p = .043$. There was also a two-way interaction between Cue Structure and Spatial Cuing, $F(1, 19) = 5.916$, $MSE = 301.186$, $p = .025$, $\eta_p^2 = .237$. Paired-comparisons with Bonferroni correction ($\alpha = .025$) revealed that in the structured cue condition, RTs to visual targets were significantly faster (by 4ms) in the cued condition (M = 380ms) than in the uncued condition (M = 384ms), $t(19) = -2.449$, $p = .024$. In the white noise cue condition, RTs to targets were significantly faster (by 14ms) in the cued condition (M = 382ms) than in the uncued condition (M = 396ms), $t(19) = -3.151$, $p = .005$. There was no three-way interaction between Cue Type, Cue Structure, and Spatial Cuing, $F(1, 19) = .428$, $MSE = 271.021$, $p = .521$, $\eta_p^2 = .022$.

A planned paired sample *t*-test was conducted between the magnitude of the crossmodal spatial cuing effect in the structured looming cue condition and that in the structured receding cue condition. The analysis revealed that the magnitude of the cuing effect in the structured looming cue condition (M = 8ms) was larger, although not significantly so, than that in the structured receding cue condition (M = 0ms), $t(19) = -1.722$, $p = .092$. Therefore, the first hypothesis was rejected (see Figure 2). Another planned paired sample *t*-test was conducted between the magnitude of the cuing effect in the white noise looming cue condition and that in the white noise receding cue condition. The analysis revealed that the magnitude of the cuing effect in the white noise looming cue condition (M = 15ms) was not significantly different from that in the white noise receding cue condition (M = 12ms), $t(19) = -.554$, $p = .586$. Therefore, the second hypothesis was confirmed. The magnitude of the cuing effect in the frontal cue condition (M = 13ms) and that in the rear cue condition (M = 6ms) were entered into a planned paired sample *t*-test. The analysis revealed that the magnitude of the cuing effect in the frontal cue condition was significantly larger than that in the rear cue condition, $t(19) = -2.222$, $p = .039$. Therefore, the third hypothesis was rejected; the presentation of an auditory cue with its intensity exponentially changing elicited spatially-specific crossmodal cuing effects (see Figure 3).

Error rate data analysis – 500ms cue duration

An equivalent four-way RM-ANOVA was conducted with the error rate data in the 500ms cue duration condition. There was a main effect of Cue Structure, $F(1, 19) = 5.952$, $MSE = .001$, $p = .025$, $\eta_p^2 = .239$, with participants responding more accurately to visual targets following the presentation of a white noise cue (M = 1.3%) than a structured cue (M = 1.9%). There were no other significant terms involving the factor of Spatial Cuing. Therefore, the crossmodal spatial cuing effects reported in the RT data analysis would appear to reflect a genuine attentional facilitation rather than a speed trade-off (see Duncan, 1980; Spence et al., 2000).

Discussion

The present study was conducted in order to investigate whether, and how, sound parameters such as cue duration, cue type, and cue structure would modulate spatial cuing effects. We hypothesised that the presentation of a structured looming cue would elicit a more pronounced spatial cuing effect (i.e., larger magnitude of the cuing effect) than that of a structured receding cue. We also expected that there would be no such asymmetric perceptual bias in the white noise cue condition. Hence, it was hypothesised that the magnitude of the cuing effect in the white noise looming cue condition would be statistically no different from that in the white noise receding cue condition. The first hypothesis was confirmed in the 250ms cue duration condition. In the 500ms cue duration condition, the difference in the magnitude of the cuing effect between the structured looming cue condition and the structured receding cue condition, although trending in the right direction, failed to reach significance ($p = .092$). As hypothesised, the magnitude of the cuing effect in the white noise looming cue condition was not statistically different from that in the white noise receding cue condition in the 250ms and in the 500ms cue duration conditions. Therefore, the results confirmed the second hypothesis. Such findings suggest that when white noise cues are used, receding auditory cues can be as effective as looming cues in the exogenous orienting of crossmodal spatial attention (see Figure 2).

The interesting result to emerge from the data reported here was that the magnitude of the cuing effect was significantly larger when auditory cues were presented from the front than from the rear space in both cue duration conditions (see Figure 3). As a result, the third hypothesis was rejected. Such results contrast with previous findings of the rear-to-front crossmodal spatial cuing effect (Lee & Spence, 2015). Lee and Spence, with a similar apparatus set-up as that used in the present study, investigated whether the presentation of an auditory cue (at a constant-intensity of 100ms) in the rear space would modulate the perception of frontal visual targets. Their study revealed that the front-back location of the cue did not modulate the magnitude of crossmodal spatial cuing effect (see also Spence et al., 2018, for a review). What is more, the participants who took part in the spatial cuing task also subsequently performed a sound location discrimination task, in which the auditory cues in the spatial cuing task now became auditory targets. The task required the participants to indicate the location of an auditory target pseudo-randomly presented from either front-left, front-right, rear-left, or rear-right. The findings suggested that the participants did not confuse the front-back location of auditory targets. Therefore, Lee and Spence were able to rule out front-back confusion (Makous & Middlebrooks, 1990; Middlebrooks & Green, 1991; see also Shinn-Cunningham, Santarelli, & Kopco, 2000) as a potential explanation for the documented rear-to-front crossmodal spatial cuing effect.

Taken together, the findings reported by Lee and Spence (2015) and those reported in this present study suggest that the exogenous orienting of crossmodal attention, at least via auditory stimuli, is not spatially-specific initially. Instead, the spatial focusing of the attentional gradient narrows down to the cued region of space over time. Indeed, the data from an event-related potential (ERP) study also demonstrated that the focus of auditory spatial attention narrows to the perceived location of the sound source over time (see Teder-Sälejärvi & Hillyard, 1998). More specifically, the spatial focusing of auditory attention is rather broad in the beginning (i.e., between 80ms and 200ms from the sound onset), but narrows down to the sound location approximately 250ms after the onset of the sound.[6] It should be noted, however, that the rear-to-front spatial cuing effects reported in Lee and Spence's study at the 100ms and 200ms SOAs did not become spatially-specific at the 700ms SOA. Instead, the cuing effect failed to reach significance in the 700ms SOA condition (see Spence & Driver, 1997, for similar findings).

[6] It should be noted that Teder-Sälejärvi and Hillyard (1998) used brief (83ms) pink-noises. The time course of spatial orienting from the onset of an 83ms-long pink-noise may be different from that from the onset of a 250ms or 500ms auditory cue.

Based on the findings reported from Lee and Spence (2015), the spatially-specific cuing effects documented in the present study cannot be attributed to the SOAs that were used (i.e., 350-450ms in the 250ms cue duration condition, and 600-700ms in the 500ms cue duration condition). Furthermore, the lack of a three-way interaction between Cue Position, Cue Structure, and Spatial Cuing in the present study suggests that the cue structures, at least between triangular waveforms and white noise, do not modulate the spatial specificity of exogenous spatial cuing. Therefore, the spatially-specific cuing effects documented in the present study can be attributed to the extended presence and, potentially, the exponential intensity changes (i.e., looming or receding) of the auditory cues. It is, however, unclear from the data reported here whether the exponential increase or decrease of the cue intensity was necessary to elicit the spatially-specific cuing effects, because the study design did not include a constant-intensity cue condition. In order to further understand how the duration and intensity changes of auditory cues modulate spatial cuing effects, future studies would want to investigate whether constant-intensity cues with a duration of 250ms or 500ms would also elicit spatially-specific cuing effects.

To date, the differential effect of the looming and receding sounds from structured auditory stimuli has often been interpreted as an evolutionary adaptation that increases the chance of an organism's survival (see Ghazanfar et al., 2002; Leo et al., 2011; Neuhoff, 1998, 2001). For instance, it was argued that the structured looming sounds are ecologically more meaningful (e.g., potentially signalling an approaching predator), "*whereas white noise is associated with* […] *dispersed phenomena such as wind or rain*" (Maier & Ghazanfar, 2007, p. 4099). Clearly, the sound of, for example, an approaching lion provides more ecologically important information than the perception of wind or rain. Nonetheless, the presentation of a white noise auditory cue in the present study could elicit crossmodal spatial cuing effects, just as well as a structured looming cue. What is more, the visual inspection of Figure 2 indicates that the crossmodal spatial cuing effects may dissipate after around 300ms of the onset of a structured receding auditory cue. On the other hand, white noise cues seem to be effective in exogenously orienting crossmodal spatial attention regardless of the cue intensity and duration, at least based on the particular settings used in the present study. The results reported here therefore suggest that white noise looming and white noise receding sounds may lack the ecological meaning, but they can be as effective as (or more effective than) structured looming sounds in the crossmodal exogenous orienting of spatial attention (see McCarthy & Olsen, 2017, for similar findings).

In summary, the results in this study suggest that sound parameters modulate the exogenous orienting of crossmodal spatial attention. The findings of the present study may be limited due to the small cuing effects (i.e., less than 20ms in magnitude) compared to, for example, those reported in Spence and Driver (1997; i.e., over 40ms in magnitude). However, it should be noted that the lateral eccentricity between the left and right visual targets was 96° in Spence and Driver's study, as compared to 60° in the apparatus set-up used here. Since visual stimuli that are closer to the fixation would be perceived more rapidly than those presented more eccentrically (see Lee & Spence, 2017), the small cuing effects could perhaps simply be reflecting the well-known eccentricity effect (Carrasco, Evert, Chang, & Katz, 1995; see also Carrasco & Frieder, 1997). More importantly, even though the magnitude of the cuing effects documented here may seem trivial, it is worth remembering that the magnitude of facilitation would likely increase if the cues were to be made spatially-informative (e.g., Ho & Spence, 2005). Therefore, any crossmodal spatial cuing effects in the laboratory setting, whether the magnitudes are large or small, can be effective, for instance, in the real-life driving situations.

REFERENCES

Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Di Salle, F., Lehmann, C., Herdener, M., Scheffler, K., & Seifritz, E. (2008). Rising sound intensity: An intrinsic warning cue activating the amygdala. *Cerebral Cortex, 18*(1), 145-150.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2009). Selective integration of auditory-visual looming cues by humans. *Neuropsychologia, 47*(4), 1045-1052.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics, 57*(8), 1241-1261.

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research, 37*(1), 63-82.

Duncan, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology, 12*(1), 75-96.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Fuentes, L. J., & Campoy, G. (2008). The time course of alerting effect over orienting in the attention network test. *Experimental Brain Research,* 185(4), 667-672.

Ghazanfar, A. A., Neuhoff, J. G., & Logothetis, N. K. (2002). Auditory looming perception in rhesus monkeys. *Proceedings of the National Academy of Sciences of the USA, 99*(24), 15755-15757.

Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied, 11*(3), 157-174.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception, 36*(14), 1-89.

Lee, J., & Spence, C. (2015). Audiovisual crossmodal cuing effects in front and rear space. *Frontiers in Psychology*, *6*:1-10.

Lee, J., & Spence, C. (2017). On the spatial specificity of audiovisual crossmodal exogenous cuing effects. *Acta Psychologica*, *177*, 78-88.

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. *Experimental Brain Research*, *213*(2-3), 193-201.

Maier, J. X., & Ghazanfar, A. A. (2007). Looming biases in monkey auditory cortex. *Journal of Neuroscience*, *27*(15), 4093-4100.

Maier, J. X., Neuhoff, J. G., Logothetis, N. K., & Ghazanfar, A. A. (2004). Multisensory integration of looming signals by rhesus monkeys. *Neuron*, *43*(2), 177-181.

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *Journal of the Acoustical Society of America*, *87*(5), 2188-2200.

McCarthy, L., & Olsen, K. N. (2017). A “looming bias” in spatial hearing? Effects of acoustic intensity and spectrum on categorical sound source localization. *Attention, Perception, & Psychophysics, 79*(1), 352-362.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42*(1), 135-159.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61-64.

Morrongiello, B. A., Hewitt, K. L., & Gotowiec, A. (1991). Infants’ discrimination of relative distance in the auditory modality: Approaching versus receding sound sources. *Infant Behavior and Development*, *14*(2), 187-208.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123-124.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology, 13*(2), 87-110.

Neuhoff, J. G., Planisek, R., & Seifritz, E. (2009). Adaptive sex differences in auditory motion perception: Looming sounds are special. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 225-234.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442.

Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology, 19*(21), 1799-1805.

Sabin, A. T., Macpherson, E. A., & Middlebrooks, J. C. (2005). Human sound localization at near-threshold levels. *Hearing Research*, *199*(1-2), 124-134.

Shinn-Cunningham, B. G., Santarelli, S., & Kopco, N. (2000). Tori of confusion: Binaural localization cues for sources within reach of a listener. *Journal of the Acoustical Society of America, 107*(3), 1627-1636.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics, 59*(1), 1-22.

Spence, C., Lee, J., & Van der Stoep, N. (2018). Responding to sounds from unseen locations: Crossmodal attentional orienting in response to sounds presented from the rear. *European Journal of Neuroscience*. doi: 10.1111/ejn.13733

Spence, C., & McDonald, J. (2004). The cross-modal consequences of the exogenous spatial orienting of attention. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processing* (pp. 3-25). Cambridge, MA: MIT Press.

Spence, C., McDonald, J., & Driver, J. (2004). Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 277-320). Oxford, UK: Oxford University Press.

Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*(4), 1298-1319.

Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). The gradient of spatial auditory attention in free field: An event-related potential study. *Perception & Psychophysics, 60*(7), 1228-1242.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.