


Responding to sounds from unseen locations: crossmodal attentional orienting in response to sounds presented from the rear

Charles Spence,¹  Jae Lee¹ and Nathan Van der Stoep²

¹Crossmodal Research Laboratory, Department of Experimental Psychology, Oxford University, Oxford, OX1 3UD, UK

²Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Keywords: crossmodal interactions, rear space, spatial attention

Abstract

To date, most of the research on spatial attention has focused on probing people's responses to stimuli presented in frontal space. That is, few researchers have attempted to assess what happens in the space that is currently unseen (essentially rear space). In a sense, then, 'out of sight' is, very much, 'out of mind'. In this review, we highlight what is presently known about the perception and processing of sensory stimuli (focusing on sounds) whose source is not currently visible. We briefly summarize known differences in the localizability of sounds presented from different locations in 3D space, and discuss the consequences for the crossmodal attentional and multisensory perceptual interactions taking place in various regions of space. The latest research now clearly shows that the kinds of crossmodal interactions that take place in rear space are very often different in kind from those that have been documented in frontal space. Developing a better understanding of how people respond to unseen sound sources in naturalistic environments by integrating findings emerging from multiple fields of research will likely lead to the design of better warning signals in the future. This review highlights the need for neuroscientists interested in spatial attention to spend more time researching what happens (in terms of the covert and overt crossmodal orienting of attention) in rear space.

Introduction

People tend to focus their visual attention in a fairly narrow region of frontal space. That is, they generally concentrate on what is happening in the space that they can see in front of them and that they will normally move towards (see Previc, 1998, 2000, for reviews; see also Moeller *et al.*, 2016). To date, auditory spatial attention research (no matter whether it be set in a unisensory or crossmodal context) and work on multisensory perception have largely focused on the restricted region of frontal space (see McDonald & Ward, 1999; Mondor & Zatorre, 1995; Spence & Driver, 1994; for unimodal auditory research, and Spence & Driver, 1997; for audiovisual crossmodal spatial cuing studies, respectively). Meanwhile, many other researchers have focused instead on the somewhat unnatural situation of headphone (i.e. monaural or binaural) presentation, a situation that results in sounds being perceived as

originating from inside the head (e.g. see Scharf, 1998, for a review). That said, head-related transfer functions can be used to filter sounds to reproduce signals as if originating from locations in the free-field (e.g. Wenzel, 1992; see Shilling & Shinn-Cunningham, 2000, for a review).

In contrast to research on spatial attention, the literature on auditory localization has often assessed performance in response to sounds presented from the rear in both humans and other species (e.g. Wightman & Kistler, 1989; Makous & Middlebrooks, 1990; Carlile *et al.*, 1997). A number of researchers have focused explicitly on trying to understand the challenges associated with front/back and up/down localization in humans (e.g. Hartmann & Rakerd, 1993; Hofman & Van Opstal, 1998; Macpherson & Middlebrooks, 2000; Zhang & Hartmann, 2010). Similar questions relating to the spatial tuning of auditory neurons have also been addressed in both anaesthetized cats and ferrets (e.g. Middlebrooks *et al.*, 1994, 1998; Reale *et al.*, 2000, 2003; Schnupp *et al.*, 2001; Mrsic-Flogel *et al.*, 2005) and in unanaesthetized cats (Mickey & Middlebrooks, 2003; Lee & Middlebrooks, 2011, 2013).

Much less frequently have attention researchers addressed the question of whether the presentation of auditory stimuli from the rear – specifically, from the space that falls outside of the current field of view (i.e. behind the head, and typically behind the body too) – will modulate a participant's spatial attention or multisensory perception more generally (see Fig. 1). One of the few exceptions

Correspondence: Charles Spence, as above.

E-mail: charles.spence@psy.ox.ac.uk

Received 6 June 2017, revised 27 September 2017, accepted 27 September 2017

Edited by Sophie Molholm

Reviewed by John Middlebrooks, University of California- Irvine, USA; and Claude Alain, Baycrest Centre for Geriatric Care, Rotman Research Institute, Canada

The associated peer review process communications can be found in the online version of this article.

here comes from Spence & Driver (1994, Experiment 1). These researchers presented spatially nonpredictive auditory cues from 90° to either side of central fixation, while auditory targets (requiring a front/back speeded discrimination response) were presented from one of four loudspeakers situated 45° and 135° to either side of fixation. Elsewhere, Spence *et al.* (2000b) also had their participants attend to a voice presented from a loudspeaker placed directly behind their head. However, the relevant point for present purposes is that in neither of these studies was the researchers' interest really in the representation of sounds in rear space.

A question that one might want to ask at this point, then, is why it is that rear auditory space has been neglected by those working in the field of attention research (if not auditory localization) for so long (see Spence & Driver, 2004). Who knows, perhaps it is because the space behind us tends to fall outside of our conscious awareness, given that we all, in some sense, 'see' forward (Krech & Crutchfield, 1958; Arnold *et al.*, 2016)? As we will see below, though, this neglect (from researchers interested in attention) is unfortunate given that there are some interesting questions to be addressed around how the map of rear auditory space is constructed, given vision's important role in calibrating spatial hearing in frontal space.

Few researchers interested in the topic of crossmodal spatial attention have presented sounds from behind their participants (note that the same criticism does not apply to those working on auditory localization). Perhaps the reason for this neglect reflects nothing more prosaic than practical/methodological issues. Alternatively, however, one might worry that there is an implicit assumption underpinning much of the research on crossmodal spatial attention that those findings obtained when sound sources are presented in frontal space can safely be generalized to sounds presented from the rear as well. Whatever the most appropriate explanation, the fact remains that one is certainly hard-pressed to find any mention of rear sounds in Spence & Driver's (2004) oft-cited edited volume on

Crossmodal Space and Crossmodal Attention. Empirically evaluating whether the same rules of crossmodal spatial orienting apply in both front and rear space becomes especially important for those applications where warning signals may need to be presented from behind an observer's (or should that be listener's) head.

Unique to the auditory modality is the fact that sound sources potentially provide continuous spatial information from the three-dimensional (3D) space that surrounds us (Perrott, 1993). Sounds, after all, typically alert us to those events that we fail to notice visually (such as a predator approaching stealthily from the rear; see Heffner & Heffner, 1992a,b). Similarly, nowadays, when biking or driving a car, we often hear vehicles approaching from the rear and use that information to determine on which side they will pass us. The benefit of being made aware of events in rear space (even if we cannot see them) becomes all the more apparent when thinking about the electric scooters that have become an increasingly common presence in many cities in recent years. Have not many of us had the startling experience of pizza delivery drivers sometimes suddenly appearing next to us in traffic without the scooter's engine having provided any auditory cue about their presence? No wonder, then, that many people have been worrying about the potential dangers of silent cars for other road users. Indeed, according to European Union regulations, all electric vehicles will have to make some kind of noise by 2021 (see 'Electric Cars', 2014; Petiot *et al.*, 2013).

One of the most important potential application domains, then, for spatial attention research in rear space is in the design of auditory (and/or possibly tactile/haptic) signals. Such signals are often designed to alert drivers to the presence of a vehicle situated in their blind spot, say, or elsewhere in the unseen space that lies behind them (e.g. Lee *et al.*, 2004; Ho & Spence, 2005; Ho *et al.*, 2006; McKeown & Isherwood, 2007; Chun *et al.*, 2013). Intuitively, based on the evidence highlighting the existence of robust crossmodal links in spatial attention (Baldwin *et al.*, 2012; see Spence & Driver,

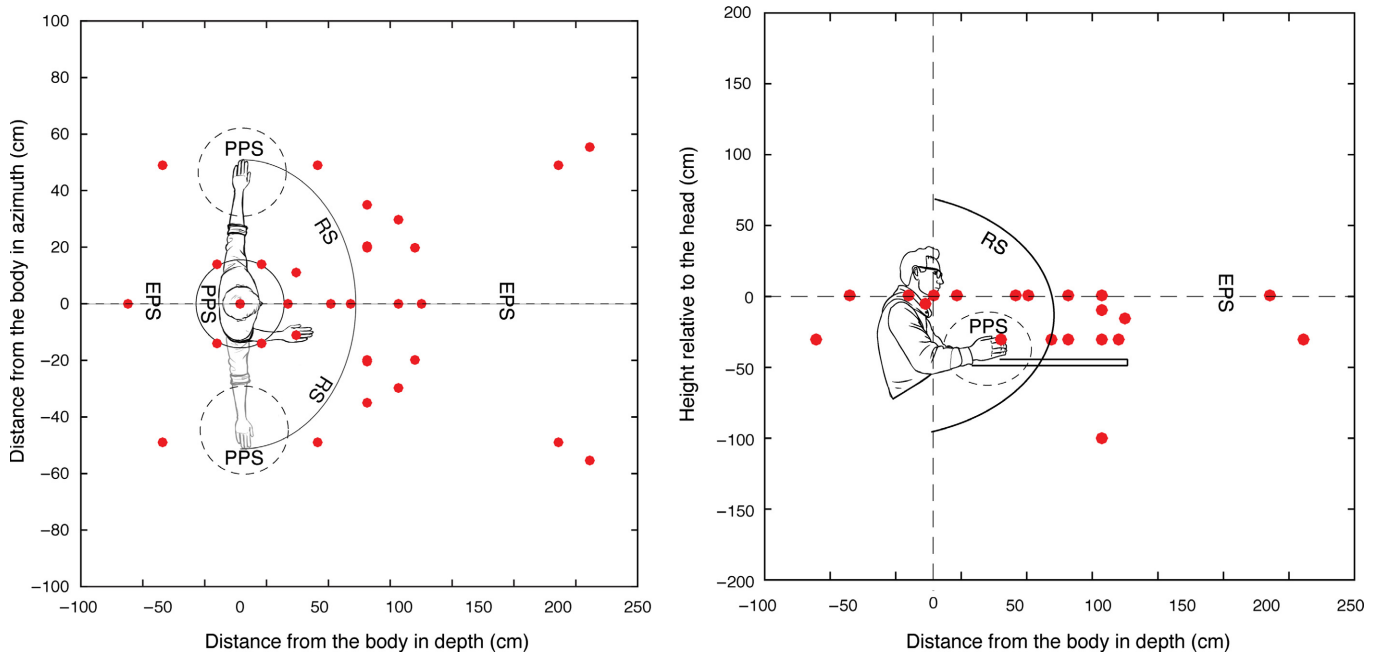


FIG. 1. Summary figure from Van der Stoep *et al.*'s (2016a) review highlighting how multisensory studies tend to involve the presentation of stimuli in frontal space. The dots indicate the locations from which the stimuli were presented in a representative selection of 12 studies. (Reprinted with permission from Van der Stoep *et al.*, 2016a).

2004, for a review), one might simply propose that auditory (or tactile) cues should be presented from the direction of the vehicle in the blind spot in order to warn the driver of an object/event that they might otherwise be unaware of, for example. In fact, a number of commentators have already suggested that rear auditory cues should be used in cars for just this purpose (e.g. Ho & Spence, 2005, Ferris & Sarter, 2008; Ho & Spence, 2008, 2009; Spence & Ho, 2015). However, as we will see below, the latest laboratory evidence suggests that things might not always be quite so simple, given the way in which we tend to respond crossmodally to those sounds whose source is situated behind us.

In the sections that follow, we discuss those studies of spatial hearing, crossmodal spatial attention and other multisensory interactions that have involved sounds presented from rear space. Often, the findings are compared to what goes on in frontal space. Our goal here is to highlight how the perception of, and more importantly the orienting response towards, auditory and multisensory stimuli from the rear is actually often quite different, both qualitatively and quantitatively, from what happens in frontal space. And, given that (as we have just seen) an awareness of what is going on behind us can play an important role in everyday activities such as safe driving, we would like to encourage more of those in the attention/perception research community to broaden their horizons (as it were) and consider the perception of, and responses to, stimuli in rear space in their research. At the same time, though, it is important to recognize the many challenges that are associated with trying to extend one's research from the laboratory to the real world (see Spence & Ho, 2015, for a review). Ultimately, we will conclude that the perception of, and responses towards, auditory stimuli in rear space differs much from what is seen in frontal space, and therefore has important consequences for the design of auditory (and multisensory) warning signals (e.g. as increasingly used in vehicles).

Localizing sounds in 3D space: mechanisms and systematic errors/biases

When thinking about the contribution that our ability to localize sounds presented from the rear makes to spatial orienting, it is important to consider the many errors/biases that have been identified in our judgement of the source of sounds presented in 3D space (i.e. 360° in azimuth and elevation; see Carlile *et al.*, 1997). For instance, differences in the localizability of sounds presented in various regions of 3D space may be expected to affect how precisely attention is focused spatially. That is, harder-to-localize sounds may potentially lead to a broader spread of exogenous spatial attention, while at the same time giving rise to less perceptual facilitation at any given location (see Spence *et al.*, 2004, for a discussion). Similarly, any difficulties in correctly localizing auditory stimuli may also affect the various multisensory behavioural enhancement effects that have been reported in recent years (see Calvert *et al.*, 2004; though see also Spence, 2013).

Localizing sounds in the horizontal plane involves the use of information from binaural cues, including interaural level and time differences (ILDs and ITDs, respectively; Blauert, 1997; Grothe *et al.*, 2010; Middlebrooks & Green, 1991; Rayleigh, 1907; Schnupp *et al.*, 2011). Binaural cues are encoded in the superior olivary complex (SOC), which then projects primarily to the inferior colliculus (IC; Yin, 2002, for a review). However, binaural cues mostly provide information about the likely lateral location of a sound source (though see Perrott & Saberi, 1990). Sounds originating from a hypothetical cone centred on the interaural axis (known as the cone of confusion; Shinn-Cunningham *et al.*, 2000) exhibit near-

identical ILDs and ITDs. Hence, based on binaural cues, it is only possible to specify the shape of the cone of confusion from which a sound may have originated.

Other than binaural cues, the spectral cues, generated by the external ear (pinna; see Butler, 1975, for a review) as well as the head and the torso (Algazi *et al.*, 2001), are regarded as providing some of the most salient information concerning sound localization (see Wightman & Kistler, 1997). Spectral cues are crucial for sound localization on the vertical plane and resolving front/back ambiguity (Ovcharenko *et al.*, 2007; Slattery & Middlebrooks, 1994; Talagala *et al.*, 2014; Wenzel *et al.*, 1993; see also Van Wanrooij & Van Opstal, 2005). The peaks and dips in the spectrum (i.e. intensity as a function of the frequency response) systematically shift as the sound source location changes (Mehrgardt & Mellert, 1977; Musicant *et al.*, 1990; see Middlebrooks & Green, 1991, for a review). The pinnae are known to modulate sound amplitude or gain in the frequency range from 3 to 4 kHz and above, and the head diffraction and torso reflections also influence the spectrum change that is observed, particularly for frequencies lower than 3 kHz (Algazi *et al.*, 2001). Spectral cues between 4 and 16 kHz are essential for sound localization with broadband noises (Langendijk & Bronkhorst, 2002; see also Butler & Humanski, 1992; Hebrank & Wright, 1974; Musicant *et al.*, 1990), although the low-frequency spectrum below 2 kHz is also particularly important as far as the resolution of front/back ambiguities is concerned (Asano *et al.*, 1990; see also Musicant & Butler, 1984). Two of the most important brain regions processing spectral cues are the inferior colliculus (IC; Davis *et al.*, 2003; see also Davis, 2005, for a review) and the auditory cortex (Zatorre & Belin, 2001; Schreiner *et al.*, 2011, for a review).

Resolution of any front/back ambiguity in the localization of a sound source can also be achieved by the movement (e.g. rotation) of the listener's head and/or the sound source (Wallach, 1940; Wightman & Kistler, 1999). Perhaps unsurprisingly, front/back confusions in the localization of sound sources occur more frequently for those stimuli that contain only a limited range of frequencies (e.g. sinusoids or narrowband noise) rather than broadband noise (Butler, 1986; see Middlebrooks & Green, 1991, for a review). Studies of absolute auditory localization abilities in the vertical plane have shown that performance is nearly as accurate as in the horizontal plane (e.g. Carlile *et al.*, 1997; Makous & Middlebrooks, 1990; Parise *et al.*, 2014; though see also Grusser, 1983; Parise *et al.*, 2012, 2013, for somewhat worse performance under a subset of conditions).

Reflections can aid determining the distance of a sound source in naturalistic enclosed environments, such as represented by the typical experimental laboratory (see Bronkhorst & Houtgast, 1999). The ratio between the direct and the reverberant energy of a sound provides a cue to its distance (i.e. direct-to-reverberant ratio). That said, it has also been reported that people sometimes find it surprisingly difficult to localize the source of sounds that have been presented within particularly reverberant environments, such as inside a car (see Fitch *et al.*, 2007). Fitch *et al.* had their participants verbally report the location from which an auditory signal had been presented. Eight equispaced sound locations (delivered using an array of four loudspeakers) were presented around the participants ($N = 32$) who were seated inside a car. Surprisingly, however, the mean accuracy of auditory discrimination responses was a mere 32% (with chance level performance coming in at 12.5%). This contrasted with 84% correct when the participants had to discriminate the location of one of eight vibrotactile stimulators situated in the base of their seat. Sound localization acuity can differ very

dramatically, then, depending on the environment in which listeners are tested.

It is perhaps also worth noting here that the surprisingly poor auditory localization performance documented in Fitch *et al.*'s (2007) study occurred despite the fact that the participants did not have any other task to perform in this study. Such poor localization performance does not necessarily mesh with the sense that we know where the voice of the back-seat passenger is coming from while we drive. However, it is important to bear in mind that the latter might be more a matter of 'knowing' where our passenger's voice originate from, rather than necessarily directly perceiving the source location. Of course, spatial ventriloquism might also have a role to play here (e.g. Jackson, 1953; Alais & Burr, 2004; Charbonneau *et al.*, 2013).

On the multisensory calibration of spatial hearing

Those of us with normal vision are provided with high spatial resolution visual input that can be used to help calibrate our spatial hearing in frontal space (Brainard & Knudsen, 1993; Röder *et al.*, 1999; Münte *et al.*, 2001; King, 2009). Knowing where a sound comes from requires the calibration of auditory input based on its physical location. Specifically, when it comes to horizontal sound localization, ILDs and ITDs need to be associated with physical spatial locations (Brainard & Knudsen, 1993; Knudsen, 2002; King, 2009). Given the high spatial resolution of the visual modality, this sense is an ideal candidate when it comes to calibrating spatial hearing, at least in frontal space (King, 2004).

To date, many studies have shown that vision naturally guides the (re-)calibration of spatial hearing. The dominance of vision in recalibrating auditory spatial representations has, for example, been repeatedly demonstrated in experiments that have investigated the spatial ventriloquist aftereffect (Canon, 1970; Recanzone, 1998; Frissen *et al.*, 2003; Zwiers *et al.*, 2003; Frissen *et al.*, 2005; Bertelson *et al.*, 2006). That is, after multiple exposures to sound–light pairs presented simultaneously from different locations (with, say, the sound always presented from 5° to the left of the light), sound localization (even in the absence of a visual stimulus) shifts in the direction of the visual stimulus. The existence of such an aftereffect has been taken to show that spatial hearing is constantly being updated in order to resolve any spatial conflict that may be detected between auditory and visual sources that belong together.

A dominant model in the literature regarding multisensory integration of spatial locations is the maximum-likelihood estimation account of optimal cue integration (e.g. Ernst & Banks, 2002; Alais & Burr, 2004). According to the account, the multisensory estimate represents the average of the cues weighted by the inverse of the reliability of the individual sensory estimates. Interestingly, though, it is the accuracy, rather than the reliability, of the individual sensory estimates that determines unsupervised recalibration (Zaidel *et al.*, 2011; see Zaidel *et al.*, 2013, for more on the notion of supervised recalibration).

A conflict between the senses is also sometimes introduced experimentally by plugging a participant's ear to simulate acute asymmetrical hearing loss. This generally results in a shift in the perceived location of a sound source towards the unplugged ear (Slattery & Middlebrooks, 1994), thus creating a conflict between the senses. After multisensory exposure to spatially and temporally aligned audiovisual stimuli, though, the accuracy of spatial hearing is typically improved under such conditions, indicating that recalibration has taken place (e.g. Hofman *et al.*, 1998; Strelnikov *et al.*, 2011; see Mendonça, 2014, for a review).

Calibrating the location of sound sources situated to the rear

At this point, the more interesting question, at least in the context of the present review, concerns what happens when vision is not available to help calibrate spatial hearing. How exactly is the spatial calibration of sound sources achieved in the space that cannot be seen? This might either be because an individual is blind or, for sighted individuals, because sounds originate from the space they normally cannot see (i.e. because the sound source is situated behind their head). Here it has been suggested that head movements may play an important role in localizing and calibrating rear auditory spatial perception (Wallach, 1940; Perrett & Noble, 1997; Hofman *et al.*, 1998). So, for example, take the following from Hofman *et al.* (1998, p. 420), when describing the results of one of their studies: 'active head movements may also have contributed to the calibration of the auditory localization system, especially for those spatial regions where vision has a poor resolution (that is, in the far retinal periphery) or is even absent (for example, for rear stimulus positions and in darkness)' (see also Brimijoin & Akeroyd, 2012).

Additionally, or alternatively, however, the calibration of spatial hearing for sounds originating from the rear might also be achieved by applying the same calibration of auditory space obtained in frontal space (i.e. relating ILDs and ITDs to spatial locations) to rear space. If this were to be the case, the accuracy and precision of sound localization would only be expected to be limited by the physical changes to sounds when they originate from the rear, rather than the front. As mentioned earlier, the spectral content of sounds changes given that the ears are directed in a frontal direction (e.g. Van Wanrooij & Van Opstal, 2005). This change in the acoustic spectrum of sounds presented from the rear may well affect the accuracy of sound localization.

In recent years, it has been shown that audio-motor feedback helps calibrate spatial hearing for unseen sound sources positioned in frontal space (Finocchietti *et al.*, 2015, 2017). Finocchietti *et al.* (2017) had a group of blindfolded sighted participants as well as a group of congenitally blind participants localize 500-Hz, 180-bpm intermittent sounds in frontal space before, and then again after, audio-motor training. In this case, training consisted of the participants moving the speaker playing the sound with their hand in order to explore the space around their body. A control group received no such training. Absolute sound localization performance was worse in the congenitally blind participants than in the blindfolded sighted participants prior to training (i.e. the localization error was larger). Intriguingly, however, sound localization performance improved dramatically in the congenitally blind participants following training. Although Finocchietti *et al.*'s results indicate that audio-motor feedback can be used to recalibrate spatial hearing in frontal space, this is presumably an unlikely candidate for the calibration of spatial hearing in either far frontal space or rear space (either near or far). This is because we mainly move our limbs around within a limited distance of the body in front of us. That is, we have only limited degrees of movement behind us (cf. Kóbor *et al.*, 2006, for differences in tactile discrimination performance between front and rear space).

More relevant here concerning the calibration of rear auditory space are findings reported by Gori *et al.* (2014). Specifically, these researchers measured auditory spatial bisection thresholds in blindfolded sighted participants before and after audio-tactile training. There was a total of two feedback training sessions in between three threshold measurements and there were three training groups in their study. One group received tactile spatially congruent feedback presented on the forearm 200 ms after the

presentation of a horizontal auditory sequence (originating from loudspeakers placed in front of them). Another group received verbal feedback regarding the speaker sequence, while a third group received no feedback at all during the auditory stimulation. The auditory sequence consisted of three sounds presented from left to right consecutively. The first and third sounds were presented from fixed positions, separated by a visual angle of 35°, whereas the second sound was presented in between. In the bisection threshold task, the participants reported whether the second sound was closer to the first or third sound.

Only the provision of tactile feedback led to improved auditory spatial localization performance, thus hinting at the occurrence of audio-tactile recalibration. In a control training phase, Gori *et al.*'s (2014) participants were rotated by 180°, thus making the tactile feedback that they received on their arm now spatially and directionally incongruent with regard to the sound sequences that were presented from behind them (e.g. auditory left to right, and tactile right to left). Perhaps unsurprisingly, no improvement in sound localization performance was documented under such conditions. However, in another group, the participants were rotated 180° and the direction of the tactile sequence reversed, such that the tactile feedback on the arm was again directionally congruent with the sound sequence during the training phase. Under such conditions, the spatial bisection threshold performance for sounds presented from the rear improved, albeit less than when the loudspeakers (and tactile feedback training) were presented from in front of the participants. These results therefore suggest that tactile spatial information can recalibrate spatial hearing, even when sounds are presented from the rear (and tactile stimulation is provided from in front of the participant). Presumably, visual information should be able to serve a similar function as touch in Gori *et al.*'s study (of calibrating spatial hearing in rear space) given the typically higher resolution of the visual system.

Potentially relevant regarding the calibration of spatial hearing in rear space (in both the horizontal and depth dimensions), several studies have shown that tactile stimuli are integrated with sounds that originate from sources placed directly behind the head, but not with sounds that are presented from further away (see Occelli *et al.*, 2011; Van der Stoep *et al.*, 2015a, for reviews). So, to the extent that somatosensation plays a role in the calibration of spatial hearing in rear space, one could imagine that this is most likely limited to those sounds originating from close to the body, or else those sound sources that approach the body. Such a distinction would potentially allow for a binary assessment of whether a sound source is close to, or far from, the back of the head (i.e. a crude calibration of auditory distance perception in the rear; see Graziano *et al.*, 1999). Sound localization in rear space (in both the horizontal and vertical dimensions) could presumably be calibrated based on somatosensory stimulation as well, as a function of where on the body tactile stimulation is experienced during the presentation of sound in rear space.

Interim summary

To summarize what we have seen thus far. We have reviewed the literature on the localization of sound sources in 3D space, and have identified some of the systematic biases/errors in sound location that have been documented across 3D space (see Carlile *et al.*, 1997, for a summary of the literature on errors in sound localization). We have also highlighted some of the ways the presumed location of sounds from the rear may be calibrated. As such, we are now in a

much better position to look at the literature concerning the spatial cuing of attention following the presentation of rear auditory cues.

Directing spatial attention to the rear

Using spatially nonpredictive auditory cues to draw attention to the rear exogenously

Over the last two decades or so, a number of studies have demonstrated that the presentation of auditory cues, even those that are spatially nonpredictive with regard to the likely location of an upcoming target, typically lead to a short-lasting shift of spatial attention to the cued region of space (e.g. see Spence, 2001, 2010, for reviews). Such auditory cues have been repeatedly shown to facilitate the processing of auditory, visual and tactile stimuli subsequently presented from the cued location for a short period after the presentation of the cue (e.g. Spence & Driver, 1994, 1997; Spence *et al.*, 1998). At longer cue-target intervals, though, these facilitatory effects often reverse to become a longer-lasting inhibitory aftereffect, known as 'inhibition of return', at least in simple detection tasks (see Klein, 2000; Spence *et al.*, 2000a). Crossmodal exogenous spatial cuing effects tend to be focused fairly narrowly around the location of the cue (Lee & Spence, 2017; see Spence *et al.*, 2004, for a review). However, given what we have seen so far, it should come as no surprise to find that the majority of the spatial cuing studies that have been published to date have involved the presentation of sounds from a fairly narrow region of visible frontal space.

At Oxford University's Crossmodal Research Laboratory, though, we have started to investigate the spatial distribution of crossmodal exogenous spatial cuing effects following the presentation of sounds originating from behind the head (see Fig. 2, for a bird's-eye view of the experimental set-up used in Lee & Spence's 2015, study). Lee and Spence used a variant of the orthogonal spatial cuing paradigm (see Spence & Driver, 1994, 1997) in which a spatially nonpredictive auditory cue was presented to either side of central fixation prior to the presentation of a visual target requiring participants to make a speeded elevation discrimination response. The visual target consisted of the brief illumination of an LED on either the left or right. Note that there were two LEDs on each side, one placed 80 cm (33.7°) directly above the other. The participant's task is orthogonal in the sense that the direction in which the cue is presented (left vs. right) is orthogonal to the direction in which participants need to discriminate (up vs. down), thus ruling out a simple response bias explanation of any cuing effect obtained. Note that such an alternative explanation has plagued the interpretation of many previous studies of crossmodal attentional cuing (see Spence & McDonald, 2004, for an overview).

The results of Lee & Spence's (2015) study revealed that responses to frontal visual targets were not only faster, but also somewhat more accurate, when an auditory cue (i.e. either a pure tone or a burst of white noise) was presented from the same, rather than the opposite hemifield (this is in line with prior studies; e.g. Spence & Driver, 1997). Note that this basic same-side spatial cuing effect fits with the literature showing that auditory cortical neurons tend to have hemifield receptive fields (RFs) that tend to fill the space contralateral to the cortical neuron. Both psychophysical data in humans (Boehnke & Phillips, 1999) and the latest neuroimaging research in awake and anaesthetized macaque monkeys (Ortiz-Rios *et al.*, 2017) suggest that the location of sound sources appears to be represented by opponent activity of two (or a small number) of

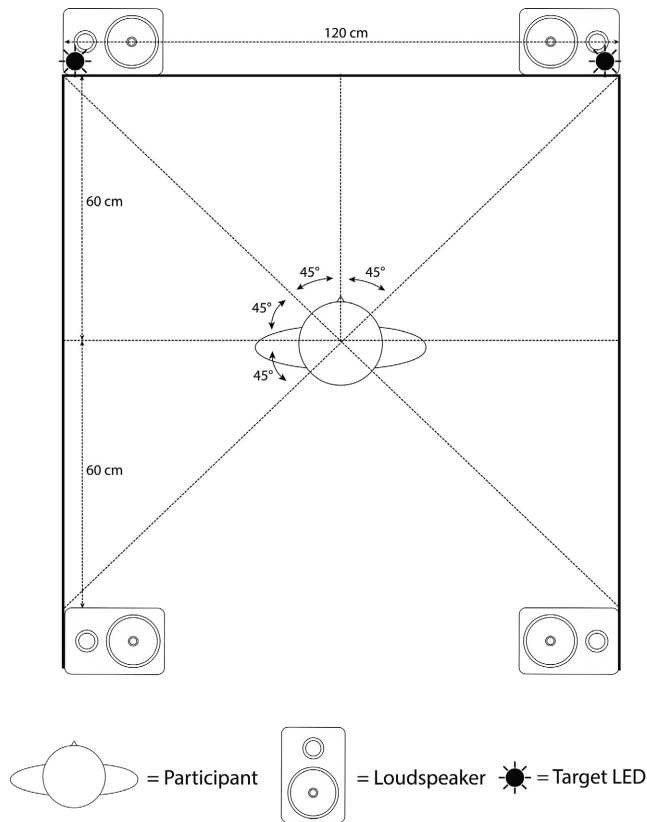


FIG. 2. Schematic bird's-eye view of the positions of the cue loudspeakers and visual targets in relation to the centrally positioned participant in Lee & Spence's (2015) study comparing the audiovisual spatial cuing effects elicited by spatially nonpredictive sounds presented in front vs. rear space. In the experiment itself, the loudspeaker cones were oriented so as to face the participant. Note also that only one of the visual targets is shown on each side. Reprinted with permission from Lee & Spence, 2015.

cortical neuron populations having broad, hemifield, spatial RFs. Nevertheless, the more surprising result to emerge from this study was that this cuing effect was observed no matter whether the spatially nonpredictive auditory cues were presented from the front or from the rear. That is, rear auditory cues facilitated frontal target discrimination performance just as much as auditory cues presented from exactly the same lateral eccentricity as the target. The latter result has been referred to as the rear-to-front crossmodal spatial cuing effect. Importantly, this pattern of results was obtained despite the fact that the rear loudspeakers were separated from the frontal visual targets by an angular separation of 90° .

A separate sound location discrimination test confirmed that the rear-to-front crossmodal spatial cuing effect could not simply be attributed to participants being confused about whether the auditory cues had come from the front or rear. In fact, they were able to correctly discriminate the front/back locations of the auditory stimuli, with an accuracy of 92 and 88% for sounds presented from front and rear space, respectively. Of course, the participant's attention would likely have been directed endogenously to the front, given that that is where all of the targets were presented. However, the evidence suggests no effect of spatial attention on the ventriloquism effect (see Bertelson *et al.*, 2000; Spence & Driver, 2004). And, by contrast, the ventriloquism effect has been shown to influence the exogenous (i.e. voluntary) allocation of spatial attention (see Spence & Driver, 2000; Vroomen *et al.*, 2001). Relevant here, though, the presence of a stimulus onset asynchrony (SOA; of 100, 200 or

700 ms) between the presentation of the cue and that of the visual target will likely also have reduced the magnitude of any spatial ventriloquism effects (see Slutsky & Recanzone, 2001; Wallace *et al.*, 2004). The key point to bear in mind, though, is that Lee & Spence (2015) were able to rule out any kind of front/back confusion as a potential explanation for the rear-to-front crossmodal spatial cuing effect.

We have now replicated the rear-to-front crossmodal spatial cuing effect in a separate study (J. Lee, C. Spence, unpublished), thus suggesting that the phenomenon is reasonably robust. In particular, in our replication study, the visual targets (i.e. white circles) were now projected onto a sound-transparent cloth screen, situated in front of the loudspeakers (see Fig. 3). Auditory cues consisted of a 50-ms white noise burst presented at 80 dBA, as measured from directly in front of the loudspeaker. Following the onset of an auditory cue in each trial, a visual target was presented on the screen for 50 ms after one of the four stimulus onset asynchronies (SOAs: 100, 200, 400, or 700 ms). Once again, the presentation of the task-irrelevant peripheral white noise auditory cue was shown to lead to faster responses to frontal visual targets on the same, rather than opposite, side of fixation, regardless of whether the cues were presented from the front or rear.

The magnitude of the crossmodal spatial cuing effect was smaller than that reported in Lee & Spence (2015; 5 ms as compared to 11 ms). It should, however, be noted that the lateral separation between the centre of a visual target on one side and the fixation point was only 9.2° in this study, as compared to 45° in Lee & Spence's (2015) previous study. Therefore, the small magnitude of the crossmodal spatial cuing effect reported in this replication study might be attributable to the close proximity of the two cue locations (at the front and at the rear; see Carrasco *et al.*, 1995; Carrasco & Frieder, 1997; Lee & Spence, 2017, on eccentricity effects).

Accounting for the rear-to-front crossmodal cuing effect

The evidence demonstrates the existence of a rear-to-front crossmodal spatial cuing effect. At first glance, this would appear to suggest that simply presenting an auditory warning signal in the direction of a vehicle (e.g. to alert the driver to the presence of an unseen vehicle in the blind spot, say) may not necessarily lead to the exogenous orienting of a driver's attention to the cued direction/region of space, as might have been expected based on the prior attentional cuing research that has been conducted in frontal space. But how should such an unexpected cuing effect be explained? An analysis of the spatial RF properties of auditory and visual neurons in the superior colliculus (SC), a midbrain structure known to be involved in overt and covert orienting of spatial attention (e.g. Kustov & Robinson, 1996; Moore *et al.*, 2003; Ignashchenkova *et al.*, 2004; Spence, 2014), might help to provide an answer here. Indeed, the role of the spatial collocation of the cue and target stimuli in the crossmodal spatial cuing effect has often been explained in terms of the response properties of the multisensory neurons in the SC (e.g. Spence & Driver, 1997; Leo *et al.*, 2008; Spence, 2013; Lee & Spence, 2017).

The multiple layers of the SC are divided into superficial and deep layers (see Wallace *et al.*, 1996), with the neurons in the superficial layers of this midbrain structure responsive exclusively to visual stimuli (Lovejoy & Krauzlis, 2010). By contrast, neurons in the deep layers respond to auditory and somatosensory stimuli as well. In the animal model (i.e. testing on anaesthetized cats), the size of the RFs of multisensory neurons in the rostral SC

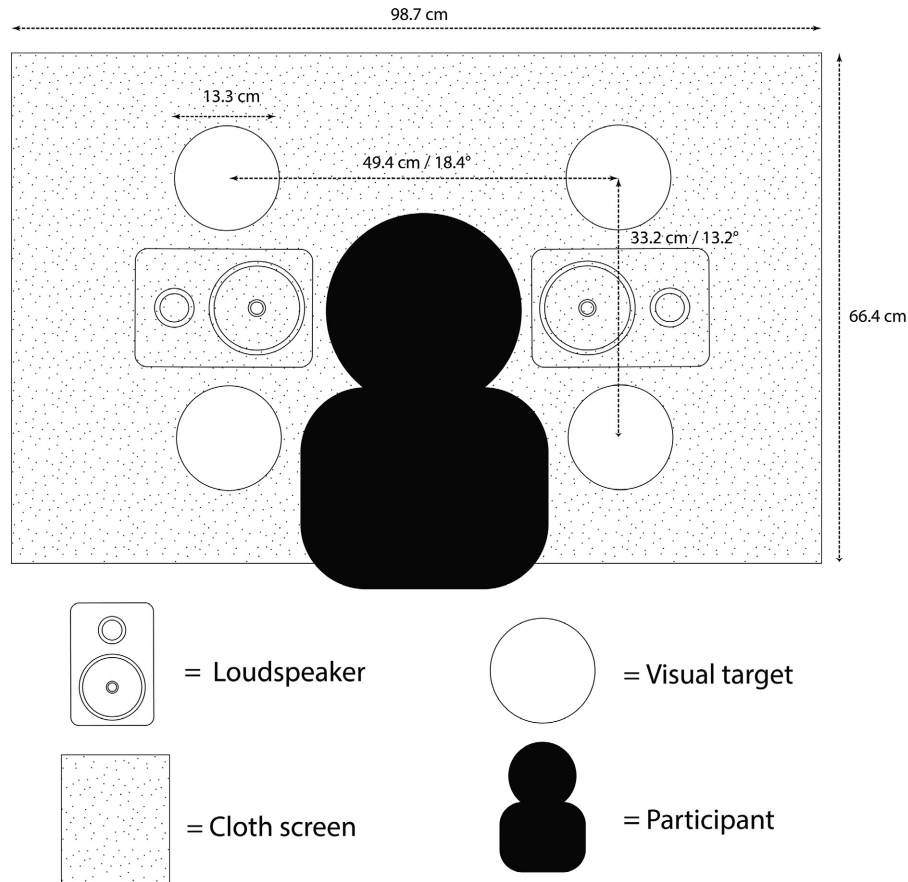


FIG. 3. Schematic view of the loudspeaker positions behind the white cloth screen and the four possible visual target locations, as seen from behind the participant's head in J. Lee and C. Spence's (unpublished) study. The two loudspeakers positioned at the rear are not shown.

(responsive to the frontal/nasal space) is considerably smaller than those found in the caudal (responsive to the peripheral/temporal space) SC (Kadunce *et al.*, 2001). In the rostral SC, visual RFs range from less than 10 to 40° of visual angle in diameter, while auditory RFs range from 20 to 60° in diameter. However, in the caudal SC, these figures jump to 40–100° and 60–135° for visual and auditory stimuli, respectively. Given the size of auditory RFs in the caudal SC, auditory RFs often extend into rear space, well beyond the limits of the visual RF (see also Middlebrooks & Knudsen, 1984; Meredith & Stein, 1996; Wallace *et al.*, 1996; Kadunce *et al.*, 1997). As such, an auditory cue presented in the rear-left position in Lee & Spence's (2015) study (e.g. 135° in azimuth from central fixation), for instance, might well still fall within the RF of an audiovisual SC neuron that has a visual RF that is responsive to visual stimuli in the front-left region of space.

At this point, though, a few notes of caution are in order. One is that the single-cell neurophysiological data based on anaesthetized cats can obviously only provide a hypothetical explanation of the crossmodal spatial cuing effects from the rear auditory cues based on awake human participants (see also Populin & Yin, 1998, 2002). However, that said, similar properties, such as RF alignment and multisensory integration, have, on occasion, also been reported in the monkey SC (see Wallace *et al.*, 1996). We therefore believe that the neurophysiological perspective concerning the multisensory interactions taking place in the SC that underlie both covert and overt attention shifts provides a viable account for the spatially nonspecific audiovisual cuing effect reported by Lee & Spence (2015).

Second, it needs to be noted that while the SC may well be a critical structure for multisensory interactions (see Stein & Meredith, 1993, for a review), it is certainly not a crucial neural structure for sound localization. Intriguingly, lesioning the SC in the cat has been shown to compromise unimodal visual and auditory orientating for a period of only 2–5 weeks. Following the recovery period, post-lesioned cats appear to exhibit relatively normal orientation behaviours towards auditory stimuli (e.g. Burnett *et al.*, 2004; see also Thompson & Masterton, 1978; and Wagner, 1993; for similar results after lesions in the external nucleus of the inferior colliculus [ICx] or the optic tectum [OT]). Such findings show that sound localization must (or at least can) be performed elsewhere than the SC. Candidate structures here include the IC and/or the auditory cortex (see Chabot *et al.*, 2013; Lomber *et al.*, 2007; for the neural network between the SC and the auditory cortex). Indeed, cortical lesions have been shown to produce clear deficits in volitional sound localization (e.g. Heffner & Masterton, 1975; Heffner & Heffner, 1990; Wagner, 1993; Malhorta *et al.*, 2004, 2008), and there is a sizeable literature on cortical representation of azimuthal sound source location (e.g. Furukawa & Middlebrooks, 2002; Middlebrooks *et al.*, 1994; see Middlebrooks, 2014, for a review).

That said, it is also worth bearing in mind that the amplitude of the binaural interaction component (BIC) of the auditory brainstem response (ABR) relates directly to the perceived location of sounds (see Laumen *et al.*, 2016, for a review). Furthermore, given that ILD and ITD are calculated in nuclei in the brainstem (MSO and LSO; see Celia, 2015), and given the fact that the IC contains a map of

auditory space, it would seem reasonable to suggest that the rapid orienting to sounds is mainly driven by brainstem activity. The subcortical structures are clearly not independent from cortical processing, but they are largely responsible for reflexive crossmodal orienting.

Exogenously directing attention to rear visual events seen in the rear-view mirror

Whatever the neural substrates of auditory localization and orienting, it is interesting to note that a rather different conclusion to that emerging from Lee & Spence (2015) studies has, though, emerged from those studies in which the sound of a car horn has been presented from directly behind (or in front of) a driver (e.g. Ho & Spence, 2005, 2008, 2009). Such warning signals have been used to alert the driver to an approaching car on the road behind them, visible via the rear-view mirror. Meanwhile, a car horn sound was presented from the front (or behind) to signal the rapid deceleration of a vehicle seen on the road ahead of the driver. (Though note that for methodological reasons, the location of the car horn sound in these studies was often made nonpredictive with regard to the location of the target car).

The results of a number of such spatial cuing studies conducted over the last decade or so have demonstrated that the presence of such semantically meaningful (or iconic; Gaver, 1986, 1993a,b) auditory cues can facilitate people's responses to visual targets presented from the cued region of rear space as compared to those presented from the opposite direction (Ho & Spence, 2005, 2008; see also Ho *et al.*, 2005, 2006; for similar results). The existence of such robust crossmodal exogenous spatial cuing effects suggests that rear sounds do indeed facilitate visual discrimination of rear visual events seen via the rear-view mirror. Therefore, Ho *et al.*'s results do, at least at first glance, seem to stand in contrast to Lee & Spence's (2015; J. Lee, C. Spence, unpublished) findings concerning the rear-to-front crossmodal cuing effect. So what exactly could be going on here to explain this discrepancy?

It is important to note that the spatially nonpredictive auditory cues used in Ho *et al.*'s studies, while presented from behind the participant/driver, actually facilitated discrimination latencies for visual targets that were seen via mirror reflection from a mirror placed directly in front of the participant. Therefore, it can be argued that at least from the perspective of the visual system, these targets were, in fact, 'presented' from frontal space. As such, one can question whether the visual attention of the participants was actually directed behind the driver in Ho *et al.*'s various simulated driving studies. That said, when the auditory cue was presented from in front of the participant/driver, it facilitated discrimination latencies for visual targets presented from the front more than those seen via the rear-view mirror (and originating from the rear). Further evidence that rear auditory events really can be linked to rear visual events, as we will see below, comes from another study (C. Spence, J. Driver, unpublished) in which the participants had to direct the auditory (and visual) attention endogenously to the rear in order to try and make out what a speaker was saying.

Endogenously attending to visual events from the rear

One of the few studies to have assessed the consequences of endogenously (i.e. voluntarily) attending to rear space for audiovisual integration was based on Driver & Spence's (1994) audiovisual shadowing studies (see also Spence & Read, 2003). Driver and Spence demonstrated that people found it significantly easier to report what a speaker was saying when the sound of their voice and

the sight of the associated lip movements came from the same, rather than different, external position(s). C. Spence and J. Driver (unpublished) took this idea to address the question of where best to position a monitor displaying the lip movements when listening to someone speaking from behind (see Fig. 4 for the experimental set-up). The results of the latter study revealed that speech intelligibility was significantly better for relevant auditory speech presented from the front when the lip movements were presented on the collocated monitor at the front than when viewing the mirror reflection (showing the monitor to the rear), despite the fact that the visual images projected onto the retina were identical in both cases. The magnitude of this frontal auditory speech advantage was, however, reduced when participants attended to the lip movements presented from behind (and seen via mirror reflection; i.e. the rear monitor image viewed indirectly via the reflection on mirror situated in front). These results therefore demonstrate that audiovisual attention can, in some meaningful sense, be directed endogenously to rear space. Here, it is worth noting that researchers have also used a similar mirror set-up to investigate the role of common spatial origin in visuotactile integration too (see Maravita *et al.*, 2002; Helbig & Ernst, 2007; Sambo & Forster, 2011).

Using spatially predictive auditory cues to direct attention to the rear endogenously

There is good evidence to show that informative auditory cues can be used to facilitate goal-directed saccades to visual targets in frontal space (e.g. Perrott *et al.*, 1991; Corneil *et al.*, 2002). Furthermore, many published human factors studies have demonstrated that such auditory cues enhance the performance of those interface operators wanting to detect, localize and/or identify visual targets (e.g. see Perrott *et al.*, 1996; Bolia *et al.*, 1999; Rudmann & Strybel, 1999; Vu *et al.*, 2006; McIntire *et al.*, 2010; Hancock *et al.*, 2013). The

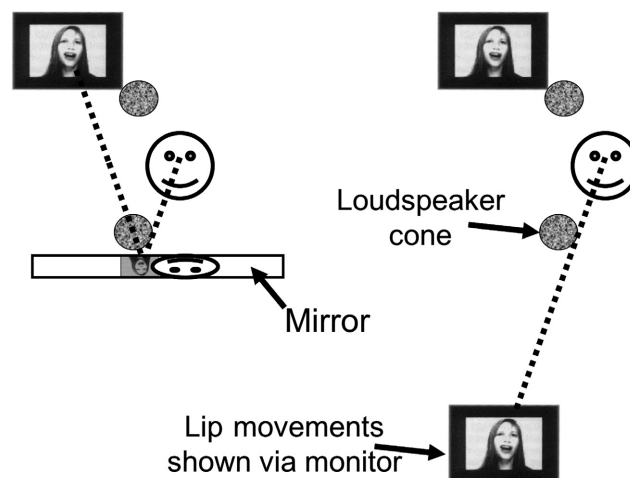


FIG. 4. Set-up used by C. Spence and J. Driver (unpublished; see also Driver & Spence, 1994; Spence & Read, 2003, for similar work) to study cross-modal links in spatial attention in rear space. A relevant speech stream was presented from one loudspeaker cone, while an irrelevant distractor speech stream was presented from the other (note that the two loudspeakers were equidistant from the participant, one placed in front and the other behind). The visual lip movements corresponding to the relevant speech stream were presented from a monitor that was either placed in front of, or behind, the participant. White noise was presented continuously from under the participant's seat in order to make the shadowing task more difficult, and hence increase the likelihood that the participants would have to rely on lip-reading information.

presentation of spatially informative auditory cues facilitates reaction times (RTs) in visual tasks (such as search, detection or discrimination) by as much as 40% or more without necessarily increasing error rates, compared to performance in the condition from which no cues were provided (Hancock *et al.*, 2013).

Here, though, it is important to note that the presentation of an auditory cue might be informative in one of two ways. Its occurrence might simply alert an interface operator to the presence of a visual event (i.e. without providing any information as to where that visual event is located in space; e.g. Diederich & Colonius, 2008; Los & Schut, 2008; Los & Van der Burg, 2013; Spence & Driver, 1997; Van der Stoep *et al.*, 2015b). Additionally, however, the location of the auditory cue may itself indicate where the relevant visual event can be found. Both types of cue have been shown to facilitate human performance (e.g. Diederich & Colonius, 2008). Furthermore, it should also be noted that the magnitude of spatial cuing effects, no matter whether the cues are informative or not, will likely depend on the eccentricity of the visual targets, with greater cuing effects seen in the periphery (see Perrott *et al.*, 1990; Lee & Spence, 2017).

Importantly, such spatially correlated (e.g. informative) auditory cues turn out to be more effective in facilitating visual performance when the targets (and hence also the auditory cues) are presented in rear space rather than in the frontal space. For instance, the participants in a study by Perrott *et al.* (1990) had to discriminate visual letter targets (i.e. 'L' or 'R') that were presented randomly in the free-field 0–130° azimuth to either side of central fixation. Note that the participants in this study were allowed to move their heads freely following stimulus onset allowing them to (eventually) see the visual targets that were presented at 130°. The visual targets were presented at the same time as an auditory cue. In the spatially correlated condition, the auditory and visual stimuli were always presented from the same location on each trial. By contrast, in the spatially uncorrelated condition, the auditory cue was always presented from directly in front of the participants while a visual target was presented from a random location on each trial. Perhaps unsurprisingly, the results of this early study revealed that the participants responded significantly faster to visual targets in the spatially correlated condition than in the uncorrelated condition. However, the slightly more intriguing result to emerge from this study was that the magnitude of this crossmodal facilitation effect increased as a function of increasing target eccentricity in the former condition.

Interim summary

Taken together, then, the limited evidence concerning the consequences for crossmodal links in audiovisual attention of presenting sounds from the rear portrays a somewhat complex story. On the one hand, the presentation of spatially nonpredictive auditory cues from the blind spot leads to an exogenous shift of spatial attention to frontal locations on the same side (Lee & Spence, 2015; and J. Lee, C. Spence, unpublished data). On the other, a number of studies have demonstrated that spatially nonpredictive car horn sounds, when presented from directly behind a participant in a driving simulator, facilitate people's responses to visual events occurring at the rear (rather than the front). Furthermore, the results of a shadowing study have also demonstrated that people find it easier to listen to someone speaking from behind them (rather than in front) if simultaneously paying attention to the sight of someone speaking behind them (J. Driver, C. Spence, unpublished). On top of that, those studies that have presented spatially collocated (and hence 100% predictive) auditory cues highlight the sometimes dramatic visual search

benefits that may ensue when people have to try and localize unseen visual targets (e.g. Perrott *et al.*, 1990).

One possible explanation here as to why sounds from the rear do not always direct a person's attention towards the location of their source may be that it depends on the context of the behaviour underlying the effect. That is, if targets can appear at any location in 360°, then the location of the source of the auditory cue becomes spatially pertinent with regard to the front/rear location of the visual target (e.g. Perrott *et al.*, 1990). However, if visual targets only ever appear in the frontal hemifield, then the position of the source of the auditory cue is no longer spatially predictive with regard to the actual location of the visual target (e.g. as in Lee & Spence, 2015; J. Lee, C. Spence, unpublished studies). Put simply, perhaps we only pay attention to rear space when what is going on there is potentially relevant to our current task demands. This explanation entails a flexible (or context-dependent) system, just as has been proposed previously when trying to explain the conflicting findings regarding the importance of the principle of spatial alignment in multisensory interactions (see Spence, 2013, for a review). One might also be reminded here of the principle of spatial relevance in auditory exogenous spatial cuing (see McDonald & Ward, 1999), which itself bears some similarity to earlier notions around contingent capture of attention (see Folk *et al.*, 1992).

Nevertheless, given this continued uncertainty concerning the conditions under which sounds from the rear do/do not direct a person's attention towards their source, further research is clearly needed. Addressing this issue will likely also require clarification concerning what exactly it means to say that someone is 'attending visually to the rear'. Obviously, stimuli that are presented out of the current field of view can only be seen via mirror reflection (e.g. Ho & Spence, 2005), video feedback (cf. Tipper *et al.*, 1998, 2001) or else their presence may be inferred by the way in which they illuminate the environment, or the shadows they cast; cf. Pavani & Castiello, 2004; Pavani & Galfano, 2007). To date, two kinds of solutions to the study of human information processing for stimuli presented in rear space have been used in studies of crossmodal spatial attention. Some studies have had their participants overtly orient to find visual targets that have deliberately been presented out of the current field of view (i.e. turn their head, see Ho & Spence, 2009; Perrott *et al.*, 1990). Meanwhile, other studies have presented visual stimuli whose origin is located out of the current field of view, but which can be seen via mirror reflection (e.g. Ho & Spence, 2005; Ho *et al.*, 2005, 2006; C. Spence, J. Driver, unpublished).

The situation is, though, somewhat simpler when it comes to the case of the audio-tactile interactions taking place in front vs. rear space, and it is on that research that we will focus briefly in the final section of our review.

Audio-tactile multisensory interactions in rear space

In recent years, there has been a sudden growth of interest in the study of audio-tactile interactions in rear space. It turns out that the nature of the interactions between these two spatial senses differs qualitatively, and not just quantitatively, as a function of whether sounds are presented close to, or further away from, the back of the head. The region that is known as near-rear peripersonal space is thought to extend for about 20–60 cm out from the back of the head (see Occelli *et al.*, 2011, for a review). Several studies have demonstrated that multisensory spatial interactions tend to be much more pronounced in this region than when the auditory stimuli (especially when they consist of broadband stimuli like white noise) are presented from further away behind the head or from in front of an

observer. While support for the distinction between different regions of rear space initially came from animal neurophysiology (Graziano *et al.*, 1999, 2004), subsequent support has come from a number of sources, including both neuropsychological patient studies (Farnè & Làdavas, 2002) and human psychophysics (e.g. Kitagawa *et al.*, 2005; Occelli *et al.*, 2010; though see also Zampini *et al.*, 2007).

In our own behavioural work here at the Crossmodal Research Laboratory in Oxford, we have demonstrated robust differences in the magnitude of the spatial modulation of unsped audio-tactile temporal order judgements (Kitagawa *et al.*, 2005; Experiment 1), spatial speeded response compatibility effects (Kitagawa *et al.*, 2005; Experiment 2) and the Colavita effect (Occelli *et al.*, 2010). The latter refers to the phenomenon whereby people fail to respond to one of two simultaneously presented stimuli perceived via different senses (see Spence *et al.*, 2011, for a review). So, for example, Occelli *et al.* reported that the presentation of an auditory target sometimes suppressed participants' responses to a simultaneously presented vibrotactile target. However, this form of sensory dominance was only observed when the sounds were presented from close to the back of the participant's head on the same side as the tactile stimulus that they extinguished. No such crossmodal effects were observed when the auditory targets were presented from frontal space instead.

Meanwhile, Ho & Spence (2009) took Graziano *et al.*' suggestion concerning the existence of a near-rear peripersonal defensive brain circuit and investigated people's speeded head-turning responses in a simulated driving task following the presentation of a near-rear auditory warning signal. The presentation of a burst of white noise from just behind a driver's head gave rise to significantly faster head-turning responses than when the same cue was presented from far frontal locations, or when visual/tactile cues were used to convey the same information (i.e. warning signal) instead. While Ho and Spence's study confounded distance (near/far) with front/back location, other research in this area has clearly demonstrated that it is specifically near-rear auditory cues that are needed to trigger these head-turning effects (see Occelli *et al.*, 2011, for a review). Given their results, Ho and Spence argued that auditory warning signals should be presented in different regions of space depending on the particular behavioural response that was desired from the participant/driver (e.g. head-turning vs. braking to avoid a front-to-rear-end collision, say; see also Spence, 2012; on this theme). Once again, though, the key point to take away from the research that has been reviewed in this section is that people respond in a qualitatively different manner to those sounds that are presented in near-rear peripersonal space than to those sounds that originate from further away (see Van der Stoep *et al.*, 2014, 2015a, 2016a,b, for the effects of stimulus-observer distance and spatial alignment in depth on crossmodal cuing and multisensory integration in front and rear space).

Conclusions

In this review, we have discussed the literature concerning how attention is oriented to sounds (and other stimuli) that happen to be presented from locations falling outside of the current field of view. It should be clear, hopefully, that there is now growing evidence to show that qualitatively different patterns of crossmodal interaction can, and often are, observed in the various different regions of space (see Occelli *et al.*, 2011; Van der Stoep *et al.*, 2015a, 2016a, for reviews). One therefore needs to be careful about assuming that the responses to auditory stimuli that have been documented in frontal space will necessarily also be seen with sound sources positioned in rear space instead. While such differences between front and rear space may, in part, be related to differences in the localizability of

sound sources (as a function of their positions relative to a listener's head), this does not appear to be the whole story. Rather, based on the emerging neurophysiological literature, it would appear that different brain circuits are responsible for detecting, and responding to, sounds presented in different regions of space (see Previc, 1998, Graziano *et al.*, 1999, Previc, 2000; Graziano *et al.*, 2004; Ho & Spence, 2009; Occelli *et al.*, 2011). These neural representations appear to have somewhat different response properties, as documented in studies of attentional orienting, overt responding (i.e. head-turning) and multisensory integration. That said, it can be difficult to discriminate between the effects of crossmodal spatial attention and multisensory integration (see McDonald *et al.*, 2001; and Van der Stoep *et al.*, 2015b, on this theme). Intriguingly though, for those who are interested in this distinction, it has recently been suggested that attentional effects may show an asymmetry not present for multisensory integration (see Chen & Spence, 2017, for a review). In the future, therefore, it will be interesting to try and determine the extent to which lateral asymmetries in crossmodal attention vary in rear space in a similar manner as they have been hypothesized to do in frontal space. As such, in the future, applied researchers will need to exercise caution in those real-world situations in which an interface operator's spatial attention needs to be directed to regions of space that they cannot currently see.

Another important point to take from this review is the need for close communication, or at least interconnectedness, between different fields of research. It can be argued that combining insights from audiology, neuroscience and experimental psychology will be key to gaining a better understanding of the intricate relation between and the relative contributions of the senses to perception in 3D space. With this review, we hope to have provided the reader with a better understanding of and new insights concerning the perceptual processes and neural mechanisms at play here. As such, moving forward, our hope is that 'out of sight' will no longer imply 'out of mind'.

Acknowledgements

C.S. wishes to thank the AHRC Rethinking the Senses grant (AH/L007053/1) for supporting his research.

Conflict of interest

The authors confirm that they have no potential conflicts of interest with respect to this review article.

Author contributions

All authors contributed to the writing of this manuscript. Furthermore, all authors reviewed the final version of the manuscript.

Abbreviations

IC, Inferior Colliculus; ICx, External Nucleus of the Inferior Colliculus; OT, Optic Tectum; RF, Receptive Field; SC, Superior Colliculus.

References

- Alais, D. & Burr, D. (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.*, **14**, 257–262.
- Algazi, V.R., Avendano, C. & Duda, R.O. (2001) Elevation localization and head-related transfer function analysis at low frequencies. *J. Acoust. Soc. Am.*, **109**, 1110–1122.
- Arnold, G., Spence, C. & Auvray, M. (2016) Taking someone else's spatial perspective: natural stance or effortful decentring? *Cognition*, **148**, 27–33.

- Asano, F., Suzuki, Y. & Sone, T. (1990) Role of spectral cues in median plane localization. *J. Acoust. Soc. Am.*, **88**, 159–168.
- Baldwin, C.L., Spence, C., Bliss, J.P., Brill, J.C., Wogalter, M.S., Mayhorn, C.B. & Ferris, T.K. (2012) Multimodal cueing: the relative benefits of the auditory, visual, and tactile channels in complex environments. *Hum. Fac. Erg. Soc. P.*, **56**, 1431–1435.
- Bertelson, P., Vroomen, J., De Gelder, B. & Driver, J. (2000) The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept. Psychophys.*, **62**, 321–332.
- Bertelson, P., Frissen, I., Vroomen, J. & De Gelder, B. (2006) The aftereffects of ventriloquism: patterns of spatial generalization. *Atten. Percept. Psycho.*, **68**, 428–436.
- Blauert, J. (1997) *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA.
- Boehnke, S.E. & Phillips, D.P. (1999) Azimuthal tuning of human perceptual channels for sound location. *J. Acoust. Soc. Am.*, **106**, 1948–1955.
- Bolia, R.S., D'Angelo, W.R. & McKinley, R.L. (1999) Aurally aided visual search in three-dimensional space. *Hum. Factors*, **41**, 664–669.
- Brainard, M.S. & Knudsen, E.I. (1993) Experience-dependent plasticity in the inferior colliculus: a site for visual calibration of the neural representation of auditory space in the barn owl. *J. Neurosci.*, **13**, 4589–4608.
- Brimijoin, W.O. & Akeroyd, M.A. (2012) The role of head movements and signal spectrum in an auditory front/back illusion. *I-Perception*, **3**, 179–182.
- Bronkhorst, A.W. & Houtgast, T. (1999) Auditory distance perception in rooms. *Nature*, **397**, 517–520.
- Burnett, L.R., Stein, B.E., Chapanis, D. & Wallace, M.T. (2004) Superior colliculus lesions preferentially disrupt multisensory orientation. *Neuroscience*, **124**, 535–547.
- Butler, R.A. (1975) The influence of the external and middle ear on auditory discriminations. In Keidel, W.D. & Neff, W.D. (Eds), *Handbook of Sensory Physiology*. Springer-Verlag, Berlin, pp. 247–260.
- Butler, R.A. (1986) The bandwidth effect on monaural and binaural localization. *Hearing Res.*, **21**, 67–73.
- Butler, R.A. & Humanski, R.A. (1992) Localization of sound in the vertical plane with and without high-frequency spectral cues. *Percept. Psychophys.*, **52**, 182–186.
- Calvert, G., Spence, C. & Stein, B.E. (Eds) (2004) *The Handbook of Multisensory Processing*. CMIT Press, Cambridge, MA.
- Canon, L.K. (1970) Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J. Exp. Psychol.*, **84**, 141–147.
- Carlile, S., Leong, P. & Hyams, S. (1997) The nature and distribution of errors in sound localization by human listeners. *Hearing Res.*, **114**, 179–196.
- Carrasco, M., Evert, D.L., Chang, I. & Katz, S.M. (1995) The eccentricity effect: target eccentricity affects performance on conjunction searches. *Percept. Psychophys.*, **57**, 1241–1261.
- Carrasco, M. & Frieder, K.S. (1997) Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Res.*, **37**, 63–82.
- Celesia, G.G. (2015) Hearing disorders in brainstem lesions. *Handb. Clin. Neurol.*, **129**, 509–536.
- Chabot, N., Mellott, J.G., Hall, A.J., Tichenoff, E.L. & Lomber, S.G. (2013) Cerebral origins of the auditory projection to the superior colliculus of the cat. *Hearing Res.*, **300**, 33–45.
- Charbonneau, G., Véronneau, M., Boudrias-Fournier, C., Lepore, F. & Collignon, O. (2013) The ventriloquist in periphery: impact of eccentricity-related reliability on audio-visual localization. *J. Vision*, **13**, 20, 1–14.
- Chen, Y.-C. & Spence, C. (2017) Hemispheric asymmetry: looking for a novel signature of the modulation of spatial attention in multisensory processing. *Psychon. B. Rev.*, **24**, 690–707.
- Chun, J., Lee, I., Park, G., Seo, J., Choi, S. & Han, S.H. (2013) Efficacy of haptic blind spot warnings applied through a steering wheel or a seatbelt. *Transport. Res. F Traf.*, **21**, 231–241.
- Corneil, B.D., van Wanrooij, M., Munoz, D.P. & van Opstal, A.J. (2002) Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J. Neurophysiol.*, **88**, 438–454.
- Davis, K.A. (2005) Spectral processing in the inferior colliculus. *Int. Rev. Neurobiol.*, **70**, 169–205.
- Davis, K.A., Ramachandran, R. & May, B.J. (2003) Auditory processing of spectral cues for sound localization in the inferior colliculus. *J. Assoc. Res. Oto.*, **4**, 148–163.
- Diederich, A. & Colonius, H. (2008) Crossmodal interaction in saccadic reaction time: separating multisensory from warning effects in the time window of integration model. *Exp. Brain Res.*, **186**, 1–22.
- Driver, J. & Spence, C.J. (1994) Spatial synergies between auditory and visual attention. In Umiltà, C. & Moscovitch, M. (Eds), *Attention and Performance XV: Conscious and Nonconscious Information Processing*. MIT Press, Cambridge, MA, pp. 311–331.
- Electric Cars (2014) Electric cars must make noise under new EU rules. Available <http://www.bbc.co.uk/news/world-europe-26857743>. Accessed May 26, 2017.
- Ernst, M.O. & Banks, M.S. (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, **415**, 429–433.
- Farnè, A. & Ladavas, E. (2002) Auditory peripersonal space in humans. *J. Cognitive Neurosci.*, **14**, 1030–1043.
- Ferris, T.K. & Sarter, N.B. (2008) Cross-modal links among vision, audition, and touch in complex environments. *Hum. Factors*, **50**, 17–26.
- Finocchietti, S., Cappagli, G., Baud-Bovy, G., Magnusson, C., Caltenco, H., Wilson, G., Brewster, S. & Rogge, A.K. et al. (2015) ABBi, a new technology for sensory-motor rehabilitation of visual impaired people. In *International Conference on Enabling Access for Persons with Visual Impairment*, pp. 80–84.
- Finocchietti, S., Cappagli, G. & Gori, M. (2017) Auditory spatial recalibration in congenital blind individuals. *Front. Neurosci.*, **11**, 76.
- Fitch, G.M., Kiefer, R.J., Hankey, J.M. & Kleiner, B.M. (2007) Toward developing an approach for alerting drivers to the direction of a crash threat. *Hum. Factors*, **49**, 710–720.
- Folk, C.L., Remington, R.W. & Johnston, J.C. (1992) Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psychol. Human*, **18**, 1030–1044.
- Frissen, I., Vroomen, J., de Gelder, B. & Bertelson, P. (2003) The aftereffects of ventriloquism: are they sound-frequency specific? *Acta Physiol.*, **113**, 315–327.
- Frissen, I., Vroomen, J., de Gelder, B. & Bertelson, P. (2005) The aftereffects of ventriloquism: generalization across sound-frequencies. *Acta Physiol.*, **118**, 93–100.
- Furukawa, S. & Middlebrooks, J.C. (2002) Cortical representation of auditory space: information-bearing features of spike patterns. *J. Neurophysiol.*, **87**, 1749–1762.
- Gaver, W.W. (1986) Auditory icons: using sound in computer interfaces. *Hum. Comput. Int.*, **2**, 167–177.
- Gaver, W.W. (1993a) What in the world we hear? An ecological approach to auditory event perception. *Ecol. Psychol.*, **5**, 1–29.
- Gaver, W.W. (1993b) How do we hear in the world? Explorations in ecological acoustics. *Ecol. Psychol.*, **5**, 285–313.
- Gori, M., Vercillo, T., Sandini, G. & Burr, D. (2014) Tactile feedback improves auditory spatial localization. *Front. Psychol.*, **5**, 1121.
- Graziano, M.S.A., Reiss, L.A.J. & Gross, C.G. (1999) A neuronal representation of the location of nearby sounds. *Nature*, **297**, 428–430.
- Graziano, M.S.A., Gross, C.G., Taylor, C.S.R. & Moore, T. (2004) A system of multimodal areas in the primate brain. In Spence, C. & Driver, J. (Eds), *Crossmodal Space and Crossmodal Attention*. Oxford University Press, Oxford, UK, pp. 51–67.
- Grothe, B., Pecka, M. & McAlpine, D. (2010) Mechanisms of sound localization in mammals. *Physiol. Rev.*, **90**, 983–1012.
- Grusser, O.-J. (1983) Multimodal structure of the extrapersonal space. In Hein, A. & Jeannerod, M. (Eds), *Spatially Oriented Behavior*. Springer-Verlag, New York, NY, pp. 327–352.
- Hancock, P.A., Mercado, J.E., Merlo, J. & Van Erp, J.B.F. (2013) Improving target detection in visual search through the augmenting multi-sensory cues. *Ergonomics*, **56**, 729–738.
- Hartmann, W.M. & Rakerd, B. (1993) Auditory spectral discrimination and the localization of clicks in the sagittal plane. *J. Acoust. Soc. Am.*, **94**, 2083–2092.
- Hebrank, J. & Wright, D. (1974) Spectral cues used in the localization of sound sources on the median plane. *J. Acoust. Soc. Am.*, **56**, 1829–1834.
- Heffner, H.E. & Heffner, R.S. (1990) Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *J. Neurophysiol.*, **64**, 915–931.
- Heffner, R.S. & Heffner, H.E. (1992a) Evolution of sound localization in mammals. In Webster, D.B., Fay, R.R. & Popper, A.N. (Eds), *The Evolutionary Biology of Hearing*. Springer-Verlag, New York, NY, pp. 691–715.
- Heffner, R.S. & Heffner, H.E. (1992b) Visual factors in sound localization in mammals. *J. Comp. Neurol.*, **317**, 219–232.
- Heffner, H.E. & Masterton, R.B. (1975) The contributions of auditory cortex to sound localization in the monkey. *J. Neurophysiol.*, **38**, 1340–1358.
- Helbig, H.B. & Ernst, M.O. (2007) Knowledge about a common source can promote visual-haptic integration. *Perception*, **36**, 1523–1533.

- Ho, C. & Spence, C. (2005) Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *J. Exp. Psychol. Appl.*, **11**, 157–174.
- Ho, C. & Spence, C. (2008) *The Multisensory Driver: Implications for Ergonomic car Interface Design*. Ashgate, Aldershot, HA.
- Ho, C. & Spence, C. (2009) Using peripersonal warning signals to orient a driver's gaze. *Hum. Factors*, **51**, 539–556.
- Ho, C., Spence, C. & Tan, H.Z. (2005) Warning signals go multisensory. *Proceedings of HCI International 2005*, 9, Paper No. 2284, 1–10.
- Ho, C., Tan, H.Z. & Spence, C. (2006) The differential effect of vibrotactile and auditory cues on visual spatial attention. *Ergonomics*, **49**, 724–738.
- Hofman, P.M. & Van Opstal, A.J. (1998) Spectro-temporal factors in two-dimensional human sound localization. *J. Acoust. Soc. Am.*, **103**, 2634–2648.
- Hofman, P.M., Van Riswick, J.G. & Van Opstal, A.J. (1998) Relearning sound localization with new ears. *Nat. Neurosci.*, **1**, 417–421.
- Ignashchenkova, A., Dicke, P.W., Haarmeier, T. & Their, P. (2004) Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nat. Neurosci.*, **7**, 56–64.
- Jackson, C.V. (1953) Visual factors in auditory localization. *Q. J. Exp. Psychol.*, **5**, 52–65.
- Kadunce, D.C., Vaughan, W.J., Wallace, M.T., Benedek, G. & Stein, B.E. (1997) Mechanisms of within- and cross-modality suppression in the superior colliculus. *J. Neurophysiol.*, **78**, 2834–2847.
- Kadunce, D.C., Vaughan, W.J., Wallace, M.T. & Stein, B.E. (2001) The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Exp. Brain Res.*, **139**, 303–310.
- King, A. (2004) Development of multisensory spatial integration. In Spence, C. & Driver, J. (Eds), *Crossmodal Space and Crossmodal Attention*. Oxford University Press, Oxford, UK, pp. 1–24.
- King, A.J. (2009) Visual influences on auditory spatial learning. *Philos. T. Roy. Soc. B*, **364**, 331–339.
- Kitagawa, N., Zampini, M. & Spence, C. (2005) Audiotactile interactions in near and far space. *Exp. Brain Res.*, **166**, 528–537.
- Klein, R. (2000) Inhibition of return. *Trends Cogn. Sci.*, **4**, 138–147.
- Knudsen, E.I. (2002) Instructed learning in the auditory localization pathway of the barn owl. *Nature*, **417**, 322–328.
- Kóbor, I., Füredi, L., Kovács, G., Spence, C. & Vidnyánszky, Z. (2006) Back-to-front: improved tactile discrimination performance in the space you can't see. *Neurosci. Lett.*, **400**, 163–167.
- Krech, D. & Crutchfield, R.S. (1958) *Elements of Psychology*. A. A. Knopf, New York, NY.
- Kustov, A.A. & Robinson, D.L. (1996) Shared neural control of attentional shifts and eye movements. *Nature*, **384**, 74–77.
- Langendijk, E.H.A. & Bronkhorst, A.W. (2002) Contribution of spectral cues to human sound localization. *J. Acoust. Soc. Am.*, **112**, 1583–1596.
- Laumen, G., Ferber, A.T., Klump, G.M. & Tollin, D.J. (2016) The physiological basis and clinical use of the binaural interaction component of the auditory brainstem response. *Ear Hearing*, **37**, e276–e290.
- Lee, C.C. & Middlebrooks, J.C. (2011) Auditory cortex spatial sensitivity sharpens during task performance. *Nat. Neurosci.*, **14**, 108–114.
- Lee, C.C. & Middlebrooks, J.C. (2013) Specialization for sound localization in fields A1, DZ, and PAF of cat auditory cortex. *J. Assoc. Res. Oto.*, **14**, 61–82.
- Lee, J. & Spence, C. (2015) Audiovisual crossmodal cuing effects in front and rear space. *Front. Psychol.*, **6**, 1–10.
- Lee, J. & Spence, C. (2017) On the spatial specificity of audiovisual cross-modal exogenous cuing effects. *Acta Psychol.*, **177**, 78–88.
- Lee, S., Olsen, E. & Wierwille, W. (2004) A comprehensive examination of naturalistic lane-changes. U.S. Department of Transportation, DOT HS 809 702.
- Leo, F., Bertini, C., di Pellegrino, G. & Ladavas, E. (2008) Multisensory integration for orienting responses in humans requires the activation of the superior colliculus. *Exp. Brain Res.*, **186**, 67–77.
- Lomber, S.G., Malhorta, S. & Hall, A.J. (2007) Functional specialization in non-primary auditory cortex of the cat: areal and laminar contributions to sound localization. *Hearing Res.*, **222**, 31–45.
- Los, S.A. & Schut, M.L. (2008) The effective time course of preparation. *Cognitive Psychol.*, **57**, 20–55.
- Los, S.A. & Van der Burg, E. (2013) Sound speeds vision through preparation, not integration. *J. Exp. Psychol. Human*, **39**, 1612–1624.
- Lovejoy, L.P. & Krauzlis, R.J. (2010) Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments. *Nat. Neurosci.*, **13**, 261–267.
- Macpherson, E.A. & Middlebrooks, J.C. (2000) Localization of brief sounds: effects of level and background noise. *J. Acoust. Soc. Am.*, **108**, 1834–1849.
- Makous, J.C. & Middlebrooks, J.C. (1990) Two-dimensional sound localization by human listeners. *J. Acoust. Soc. Am.*, **87**, 2188–2200.
- Malhorta, S., Hall, A.J. & Lomber, S.G. (2004) Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. *J. Neurophysiol.*, **92**, 1625–1643.
- Malhorta, S., Stecker, G.C., Middlebrooks, J.C. & Lomber, S.G. (2008) Sound localization deficits during reversible deactivation of primary auditory cortex and/or the dorsal zone. *J. Neurophysiol.*, **99**, 1628–1642.
- Maravita, A., Spence, C., Sergent, C. & Driver, J. (2002) Seeing your own touched hands in a mirror modulates cross-modal interactions. *Psychol. Sci.*, **13**, 350–356.
- McDonald, J.J. & Ward, L.M. (1999) Spatial relevance determines facilitatory and inhibitory effects of auditory covert spatial orienting. *J. Exp. Psychol. Human*, **25**, 1234–1252.
- McDonald, J.J., Teder-Sälejärvi, W.A. & Ward, L.M. (2001) Multisensory integration and crossmodal attention effects in the human brain. *Science*, **292**, 1791.
- McIntire, J.P., Havig, P.R., Watamaniuk, S.N. & Gilkey, R.H. (2010) Visual search performance with 3-D auditory cues: effects of motion, target location, and practice. *Hum. Factors*, **52**, 41–53.
- McKeown, D. & Isherwood, S. (2007) Mapping candidate within-vehicle auditory displays to their referents. *Hum. Factors*, **49**, 417–428.
- Mehrgardt, S. & Mellert, V. (1977) Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.*, **61**, 1567–1576.
- Mendonça, C. (2014) A review on auditory space adaptations to altered head-related cues. *Front. Neurosci.*, **8**, 219.
- Meredith, M.A. & Stein, B.E. (1996) Spatial determinants of multisensory integration in cat superior colliculus neurons. *J. Neurophysiol.*, **75**, 1843–1857.
- Mickey, B.J. & Middlebrooks, J.C. (2003) Representation of auditory space by cortical neurons in awake cats. *J. Neurosci.*, **23**, 8649–8663.
- Middlebrooks, J.C. (2014) Distributed cortical representation of sound locations. In Popper, A.N. & Fay, R.R. (Eds), *Perspectives on Auditory Research*. Springer Science+Business Media, New York, NY, pp. 361–378.
- Middlebrooks, J.C. & Green, D.M. (1991) Sound localization by human listeners. *Annu. Rev. Psychol.*, **42**, 135–159.
- Middlebrooks, J.C. & Knudsen, E.I. (1984) A neural code for auditory space in the cat's superior colliculus. *J. Neurosci.*, **4**, 2621–2634.
- Middlebrooks, J.C., Clock, A.E., Xu, L. & Green, D.M. (1994) A panoramic code for sound location by cortical neurons. *Science*, **264**, 842–844.
- Middlebrooks, J.C., Xu, L., Eddins, A.C. & Green, D.M. (1998) Codes for sound-source location in nontopographic auditory cortex. *J. Neurophysiol.*, **80**, 863–881.
- Moeller, B., Zoppke, H. & Frings, C. (2016) What a car does to your perception: distance evaluations differ from within and outside of a car. *Psychon. B. Rev.*, **23**, 781–788.
- Mondor, T.A. & Zatorre, R.J. (1995) Shifting and focusing auditory spatial attention. *J. Exp. Psychol. Human*, **21**, 387–409.
- Moore, T., Armstrong, K.M. & Fallah, M. (2003) Visuomotor origins of covert spatial attention. *Neuron*, **40**, 671–683.
- Mrsic-Flogel, T.D., King, A.J. & Schnupp, J.W.H. (2005) Encoding of virtual acoustic space stimuli by neurons in ferret primary auditory cortex. *J. Neurophysiol.*, **93**, 3489–3503.
- Münste, T.F., Kohlmetz, C., Nager, W. & Altenmüller, E. (2001) Superior auditory spatial tuning in conductors. *Nature*, **409**, 580.
- Musicant, A.D. & Butler, R.A. (1984) The influence of pinnae-based spectral cues on sound localization. *J. Acoust. Soc. Am.*, **75**, 1195–1200.
- Musicant, A.D., Chan, J.C. & Hind, J.E. (1990) Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *J. Acoust. Soc. Am.*, **87**, 757–781.
- Ocelli, V., O'Brien, J.H., Spence, C. & Zampini, M. (2010) Assessing the audiotactile Colavita effect in near and rear space. *Exp. Brain Res.*, **203**, 517–532.
- Ocelli, V., Spence, C. & Zampini, M. (2011) Audiotactile interactions in front and rear space. *Neurosci. Biobehav. R.*, **35**, 589–598.
- Ortiz-Rios, M., Azevedo, F.A.C., Kuśmierk, P., Balla, D.Z., Munk, M.H., Keliris, G.A., Logothetis, N.K. & Rauschecker, J.P. (2017) Widespread and opponent fMRI signals represent sound location in macaque auditory cortex. *Neuron*, **93**, 971–983.
- Ovcharenko, A., Cho, S.J. & Chong, U.P. (2007) Front-back confusion resolution in three-dimensional sound localization using databases built with a dummy head. *J. Acoust. Soc. Am.*, **122**, 489–495.

- Parise, C.V., Spence, C. & Ernst, M. (2012) When correlation implies causation in multisensory integration. *Curr. Biol.*, **22**, 46–49.
- Parise, C.V., Harrar, V., Ernst, M.O. & Spence, C. (2013) Cross-correlation between auditory and visual signals promotes multisensory integration. *Multisens. Res.*, **26**, 307–316.
- Parise, C.V., Knorre, K. & Ernst, M.O. (2014) Natural auditory scene statistics shapes human spatial hearing. *Proc. Natl. Acad. Sci. USA*, **111**, 6104–6108.
- Pavani, F. & Castiello, U. (2004) Binding personal and extrapersonal space through body shadows. *Nat. Neurosci.*, **7**, 13–14.
- Pavani, F. & Galfano, G. (2007) Self-attributed body-shadows modulate tactile attention. *Cognition*, **104**, 73–88.
- Perrett, S. & Noble, W. (1997) The contribution of head motion cues to localization of low-pass noise. *Atten. Percept. Psycho.*, **59**, 1018–1026.
- Perrott, D.R. (1993) Auditory and visual localization: Two modalities and one world. *Proceedings of the Audio Engineering Society 12th International Conference*. Sneekkersten, Copenhagen, pp. 221–231.
- Perrott, D.R. & Saberi, K. (1990) Minimum audible angle thresholds varying in both elevation and azimuth. *J. Acoust. Soc. Am.*, **87**, 1728–1731.
- Perrott, D.R., Saberi, K., Brown, K. & Strybel, T.Z. (1990) Auditory psychomotor coordination and visual search performance. *Percept. Psychophys.*, **48**, 214–226.
- Perrott, D.R., Sadralodabai, T., Saberi, K. & Strybel, T.Z. (1991) Aurally aided visual search in the central visual field: effects of visual load and visual enhancement of the target. *Hum. Factors*, **33**, 389–400.
- Perrott, D.R., Cisneros, J., McKinley, R.L. & D'Angelo, W.R. (1996) Aurally aided visual search under virtual and free-field listening conditions. *Hum. Factors*, **38**, 702–715.
- Petiot, J.-F., Kristensen, B.G. & Maier, A.M. (2013) How should an electric vehicle sound? User and expert perception. *Proceedings of the ASME 2013 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2013*. August 4–7, 2013, Portland, Oregon, USA.
- Populin, L.C. & Yin, T.C.T. (1998) Sensitivity of auditory cells in the superior colliculus to eye position in the behaving cat. In Palmer, A.R., Rees, A., Summerfield, A.Q. & Meddis, R. (Eds), *Psychophysical and Physiological Advances in Hearing*. Whurr, London, UK, pp. 441–448.
- Populin, L.C. & Yin, T.C.T. (2002) Bimodal interactions in the superior colliculus of the behaving cat. *J. Neurosci.*, **22**, 2826–2834.
- Previc, F.H. (1998) The neuropsychology of 3-D space. *Psychol. Bull.*, **124**, 123–164.
- Previc, F.H. (2000) Neuropsychological guidelines for aircraft control stations. *IEEE Eng. Med. Biol.*, **19**, 81–88.
- Rayleigh, L. (1907) XII. On our perception of sound direction. *London Edinburgh Dublin Philos. Mag. J. Sci.*, **13**, 214–232.
- Reale, R.A., Brugge, J.F. & Hind, J.E. (2000) Directional sensitivity of neurons in the primary auditory (AI) cortex of the cat to successive sounds ordered in time and space. *J. Neurophysiol.*, **84**, 435–450.
- Reale, R.A., Jenson, R.L. & Brugge, J.F. (2003) Directional sensitivity of neurons in the primary auditory (AI) cortex: effects of sound-source intensity level. *J. Neurophysiol.*, **89**, 1024–1038.
- Recanzone, G.H. (1998) Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc. Natl. Acad. Sci. USA*, **95**, 869–875.
- Röder, B., Teder-Saljärvi, W., Sterr, A., Rösler, F., Hillyard, S.A. & Neville, H.J. (1999) Improved auditory spatial tuning in blind humans. *Nature*, **400**, 162–166.
- Rudmann, D.S. & Strybel, T.Z. (1999) Auditory spatial facilitation of visual search performance: effect of cue precision and distractor density. *Hum. Factors*, **41**, 146–160.
- Sambo, C.F. & Forster, B. (2011) When far is near: ERP correlates of cross-modal spatial interactions between tactile and mirror-reflected visual stimuli. *Neurosci. Lett.*, **500**, 10–15.
- Scharf, B. (1998) Auditory attention: the psychoacoustical approach. In Pashler, H. (Ed.), *Attention*. Psychology Press, London, UK, pp. 75–117.
- Schnupp, J.W.H., Masic-Flogel, T.D. & King, A.J. (2001) Linear processing of spatial cues in primary auditory cortex. *Nature*, **414**, 200–204.
- Schnupp, J., Nelken, I. & King, A. (2011) *Auditory Neuroscience: Making Sense of Sound*. MIT Press, Cambridge, MA.
- Schreiner, C.E., Froemke, R.C. & Atencio, C.A. (2011) Spectral processing in auditory cortex. In Winer, J.A. & Schreiner, C.E. (Eds), *The Auditory Cortex*. Springer, New York, NY, pp. 275–308.
- Shilling, R.D. & Shinn-Cunningham, B. (2000) Virtual auditory displays. In Stanney, K. (Ed.), *Handbook of Virtual Environments: Design, Implementation, and Applications*. Lawrence Erlbaum, Mahwah, NJ, pp. 65–92.
- Shinn-Cunningham, B.G., Santarelli, S. & Kopco, N. (2000) Tori of confusion: binaural localization cues for sources within reach of a listener. *J. Acoust. Soc. Am.*, **107**, 1627–1636.
- Slattery, W.H. & Middlebrooks, J.C. (1994) Monaural sound localization: acute versus chronic unilateral impairment. *Hearing Res.*, **75**, 38–46.
- Slutsky, D.A. & Recanzone, G.H. (2001) Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, **12**, 7–10.
- Spence, C. (2001) Crossmodal attentional capture: a controversy resolved? In Folk, C. & Gibson, B. (Eds), *Attention, Distraction and Action: Multiple Perspectives on Attentional Capture*. Elsevier Science BV, Amsterdam, pp. 231–262.
- Spence, C. (2010) Crossmodal spatial attention. *Ann. N. Y. Acad. Sci.*, **1191**, 182–200.
- Spence, C. (2012) Drive safely with neuroergonomics. *Psychologist*, **25**, 664–667.
- Spence, C. (2013) Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Ann. N. Y. Acad. Sci.*, **1296**, 31–49.
- Spence, C. (2014) Orienting attention: a crossmodal perspective. In Nobre, A.C. & Kastner, S. (Eds), *The Oxford Handbook of Attention*. Oxford University Press, Oxford, UK, pp. 446–471.
- Spence, C.J. & Driver, J. (1994) Covert spatial orienting in audition: exogenous and endogenous mechanisms. *J. Exp. Psychol. Human*, **20**, 555–574.
- Spence, C. & Driver, J. (1997) Audiovisual links in exogenous covert spatial orienting. *Percept. Psychophys.*, **59**, 1–22.
- Spence, C. & Driver, J. (2000) Attracting attention to the illusory location of a sound: reflexive crossmodal orienting and ventriloquism. *Neuroreport*, **11**, 2057–2061.
- Spence, C. & Driver, J. (Eds) (2004) *Crossmodal Space and Crossmodal Attention*. Oxford University Press, Oxford, UK.
- Spence, C. & Ho, C. (2015) Crossmodal attention: from the laboratory to the real world (and back again). In Fawcett, J.M., Risko, E.F. & Kingstone, A. (Eds), *The Handbook of Attention*. MIT Press, Cambridge, MA, pp. 119–138.
- Spence, C. & McDonald, J. (2004) The cross-modal consequences of the exogenous spatial orienting of attention. In Calvert, G.A., Spence, C. & Stein, B.E. (Eds), *The Handbook of Multisensory Processes*. MIT Press, Cambridge, MA, pp. 3–26.
- Spence, C. & Read, L. (2003) Speech shadowing while driving on the difficulty of splitting attention between eye and ear. *Psychol. Sci.*, **14**, 251–256.
- Spence, C., Nicholls, M.E.R., Gillespie, N. & Driver, J. (1998) Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Percept. Psychophys.*, **60**, 544–557.
- Spence, C., Lloyd, D., McGlone, F., Nicholls, M.E.R. & Driver, J. (2000a) Inhibition of return is supramodal: a demonstration between all possible pairings of vision, touch and audition. *Exp. Brain Res.*, **134**, 42–48.
- Spence, C., Ranson, J. & Driver, J. (2000b) Crossmodal selective attention: ignoring auditory stimuli presented at the focus of visual attention. *Percept. Psychophys.*, **62**, 410–424.
- Spence, C., McDonald, J. & Driver, J. (2004) Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In Spence, C. & Driver, J. (Eds), *Crossmodal Space and Crossmodal Attention*. Oxford University Press, Oxford, pp. 277–320.
- Spence, C., Parise, C. & Chen, Y.-C. (2011) The Colavita visual dominance effect. In Murray, M.M. & Wallace, M. (Eds), *Frontiers in the Neural Bases of Multisensory Processes*. CRC Press, Boca Raton, FL, pp. 523–550.
- Stein, B.E. & Meredith, M.A. (1993) *The Merging of the Senses*. MIT Press, Cambridge, MA.
- Strelnikov, K., Rosito, M. & Barone, P. (2011) Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One*, **6**, e18344.
- Talagala, D.S., Zhang, W., Abhayapala, T.D. & Kamineni, A. (2014) Binaural sound source localization using the frequency diversity of the head-related transfer function. *J. Acoust. Soc. Am.*, **135**, 1207–1217.
- Thompson, G.C. & Masterton, R.B. (1978) Brain stem auditory pathways involved in reflexive head orientation to sound. *J. Neurophysiol.*, **41**, 1183–1202.
- Tipper, S.P., Lloyd, D., Shorland, B., Dancer, C., Howard, L.A. & McGlone, F. (1998) Vision influences tactile perception without proprioceptive orienting. *Neuroreport*, **9**, 1741–1744.
- Tipper, S.P., Phillips, N., Dancer, C., Lloyd, D., Howard, L.A. & McGlone, F. (2001) Vision influences tactile perception at body sites that cannot be viewed directly. *Exp. Brain Res.*, **139**, 160–167.

- Van der Stoep, N., Nijboer, T.C. & Van der Stigchel, S. (2014) Exogenous orienting of crossmodal attention in 3-D space: support for a depth-aware crossmodal attentional system. *Psychon. B. Rev.*, **21**, 708–714.
- Van der Stoep, N., Nijboer, T.C.W., Van der Stigchel, S. & Spence, C. (2015a) Multisensory interactions in the depth plane in front and rear space: a review. *Neuropsychologia*, **70**, 335–349.
- Van der Stoep, N., Spence, C., Nijboer, T.C.W. & Van der Stigchel, S. (2015b) On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. *Acta Physiol.*, **162**, 20–28.
- Van der Stoep, N., Serino, A., Farnè, A., Di Luca, M. & Spence, C. (2016a) Depth: the forgotten dimension in multisensory research. *Multisens. Res.*, **29**, 493–524.
- Van der Stoep, N., Van der Stigchel, S., Nijboer, T.C.W. & Van der Smagt, M.J. (2016b) Audiovisual integration in near and far space: effects of changes in distance and stimulus effectiveness. *Exp. Brain Res.*, **234**, 1175–1188.
- Van Wanrooij, M.M. & Van Opstal, A.J. (2005) Relearning sound localization with a new ear. *J. Neurosci.*, **25**, 5413–5424.
- Vroomen, J., Bertelson, P. & de Gelder, B. (2001) Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Physiol.*, **108**, 21–33.
- Vu, K.P.L., Strybel, T.Z. & Proctor, R.W. (2006) Effects of displacement magnitude and direction of auditory cues on auditory spatial facilitation of visual search. *Hum. Factors*, **48**, 587–599.
- Wagner, H. (1993) Sound-localization deficits induced by lesions in the barn owl's auditory space map. *J. Neurosci.*, **13**, 371–386.
- Wallace, M.T., Wilkinson, L.K. & Stein, B.E. (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. *J. Neurophysiol.*, **76**, 1246–1266.
- Wallace, M.T., Roberson, G.E., Hairston, W.D., Stein, B.E., Vaughan, J.W. & Schirillo, J.A. (2004) Unifying multisensory signals across time and space. *Exp. Brain Res.*, **158**, 252–258.
- Wallach, H. (1940) The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psychol.*, **27**, 339–368.
- Wenzel, E.M. (1992) Localization in virtual acoustic displays. *Presence Teleop. Virt.*, **1**, 80–107.
- Wenzel, E.M., Arruda, M., Kistler, D.J. & Wightman, F.L. (1993) Localization using nonindividualized head-related transfer functions. *J. Acoust. Soc. Am.*, **94**, 111–123.
- Wightman, F.L. & Kistler, D.J. (1989) Headphone stimulation of free-fields listening. II: psychophysical validation. *J. Acoust. Soc. Am.*, **85**, 868–878.
- Wightman, F.L. & Kistler, D.J. (1997) Monaural sound localization revisited. *J. Acoust. Soc. Am.*, **101**, 1050–1063.
- Wightman, F.L. & Kistler, D.J. (1999) Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J. Acoust. Soc. Am.*, **105**, 2841–2853.
- Yin, T.C. (2002) Neural mechanisms of encoding binaural localization cues in the auditory brainstem. In Oertel, D., Fay, R.R. & Popper, A.N. (Eds), *Integrative Functions in the Mammalian Auditory Pathway*. Springer, New York, NY, pp. 99–159.
- Zaidel, A., Turner, A.H. & Angelaki, D.E. (2011) Multisensory calibration is independent of cue reliability. *J. Neurosci.*, **31**, 13949–13962.
- Zaidel, A., Ma, W.J. & Angelaki, D.E. (2013) Supervised calibration relies on the multisensory percept. *Neuron*, **80**, 1544–1557.
- Zampini, M., Torresan, D., Spence, C. & Murray, M.M. (2007) Audiotactile multisensory interactions in front and rear space. *Neuropsychologia*, **45**, 1869–1877.
- Zatorre, R.J. & Belin, P. (2001) Spectral and temporal processing in human auditory cortex. *Cereb. Cortex*, **11**, 946–953.
- Zhang, P.X. & Hartmann, W.M. (2010) On the ability of human listeners to distinguish between front and back. *Hearing Res.*, **260**, 30–46.
- Zwiers, M.P., Van Opstal, A.J. & Paige, G.D. (2003) Plasticity in human sound localization induced by compressed spatial vision. *Nat. Neurosci.*, **6**, 175–181.