Crossmodal Links Between Auditory and Visual Attention

Charles Spence* and Jon Driver†

*Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, England OX1 3UD and †Institute of Cognitive Neuroscience, University College London, Gower Street, London, England WC1E 6BT

Abstract: Recent research using a variety of behavioural paradigms has highlighted extensive spatial links between covert auditory and visual attention in the control of both endogenous (voluntary) and exogenous (automatic) orienting. We report a number of recent studies using our orthogonal-cuing paradigm, which demonstrate symmetrical links between audition and vision in the control of endogenous orienting, but asymmetrical crossmodal links for the exogenous case. Moreover, it appears that these crossmodal links are maintained in appropriate spatial register even when gaze is deviated with respect to the head. These findings converge with results gained from other methodologies, including ERPs, and neurophysiological single-cell recordings.

We have developed (1, 2) the ‘orthogonal-cuing paradigm’ as a method for measuring the spatial distribution of covert attention in both hearing and vision. Participants make speeded discriminations concerning the elevation (up vs. down) for each of a series of targets, presented in a random order such that target modality and target side are unpredictable (see Figure 1). Endogenous covert spatial attention can be directed by a spatial expectancy concerning the likely target side. When a target is strongly expected on a particular side in just one modality (e.g., audition), up/down discriminations improve on that side in the expected modality, but also in other modalities (e.g., vision), suggesting a tendency for common shifts in endogenous spatial attention across the modalities (1; see 3 for convergent findings from a recent audiovisual ERP study). Further studies show that people can ‘split’ their auditory and visual attention to some extent when targets in the two modalities are expected on constant but opposite sides, but covert attention effects are larger if targets can be expected at the same location in both modalities. The cost of directing auditory and visual attention to different locations at the same time can also be found in continuous dual-task situations (4).

Spence and Driver recently (2) adapted the orthogonal-cuing paradigm to investigate crossmodal links in exogenous spatial orienting, when there is no spatial expectancy. Targets were presented equally often from each of the four possible target locations in either modality, with every target now being preceded by a spatially-uninformative cue presented at an intermediate elevation on either the same (50% of trials) or opposite side as the subsequent target (intermingled 50% of trials). Presenting a peripheral auditory cue resulted in a rapid exogenous shift of both auditory and visual attention toward it. By contrast, visual cues produced only intramodal shifts of visual attention, having no effect on the spatial distribution of auditory attention, at least when eye movements were prevented (see (2) for a discussion of possible reasons for this asymmetry). We have recently extended these findings to show that spatially-nonpredictive auditory cues result in covert shifts of visual attention to a specific location within a hemifield (see Figure 2A), rather than to the entire hemifield in which the cue was presented (5). Thus, spatially nonpredictive peripheral visual cues have no effect on auditory

FIGURE 1: Schematic view of the position of possible target loudspeakers (shown by ellipses; target sounds were pulsed white noise), and target lights (black circles), plus central fixation, in Spence and Driver’s (1, 2) studies of audiovisual links in covert spatial attention. The participant's head is cartooned, and the direction of steady fixation is indicated with dashed lines.
B ~GH 2: Schematic view of Driver and Spence’s (5) gaze deviation study in which participants made speeded elevation discriminations for target lights, regardless of where the immediately preceding sound had been. In A, where participants fixated directly ahead, visual discriminations were best at the same eccentricity and side as the immediately preceding sound. In B, where participants fixated eccentrically (note that all visual events have been laterally translated along with gaze), visual discriminations were again best for lights at the same external location as the immediately preceding sound, but these now occupied different retinal locations as compared with A. This result demonstrates remapping between auditory locations in the control of exogenous crossmodal attention, to keep vision and audition in register as regards external space despite the gaze deviation.

Attention when saccades to them are prevented. If, however, participants are required to make a saccade toward the peripheral visual cue, then covert shifts of auditory attention in the direction of the cue are now seen immediately prior to saccade execution (6). These results demonstrate close crossmodal links between mechanisms of covert and overt orienting, and also highlight the importance of controlling for eye position.

To date, virtually all studies of crossmodal links in attention have been performed with the eyes and head in alignment (i.e., eyes looking straight-ahead with respect to the head). However, gaze is frequently deviated in daily life, which realigns visual receptors relative to auditory receptors. This raises the important issue of whether audiovisual links in attention are controlled by a fixed ear-retina mapping, or whether instead the relationship between the modalities gets spatially remapped whenever gaze is deviated. Recent studies using our orthogonal-cuing paradigm suggest that spatial alignment is maintained when the eyes are deviated with respect to the head via remapping (see Figure 2B). These psychological findings on audiovisual attention effects in people may be related to recent neurophysiological findings on single-cell activity in the parietal lobe of monkeys, indicating that cells there change their coding of auditory location as a function of how the eyes are deviated in the head (7).

To summarize, our results suggest that extensive crossmodal links govern spatial attention in audition and vision. However, the precise nature of these links depends upon the attentional mechanism (i.e., endogenous or exogenous) under consideration. Moreover, crossmodal links appear to be maintained under receptor misalignment, so that our attention may be focused on the same external location across the modalities irrespective of any receptor misalignment between them. The study of crossmodal links in attention appears to offer an exciting opportunity to bring together potentially converging evidence from behavioural and ERP studies in humans, plus neurophysiological data from animals on multimodal interactions in the neural substrates thought to underlie the behavioural effects.

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council (UK) and Wellcome Trust (UK).

REFERENCES