Auditory-Visual Interactions Subserving Goal-Directed Saccades in a Complex Scene

B. D. CORNEIL,1 M. VAN WANROOIJ,2 D. P. MUNOZ,1 AND A. J. VAN OPSTAL 2

1 *Centre for Neuroscience Studies, Department of Physiology, Queen's University, Kingston, Ontario K7L 3N6, Canada; and* 2 *Department of Biophysics, University of Nijmegen, Geert Grooteplein 21, 6525 EZ Nijmegen, The Netherlands*

Corneil, B. D., M. Van Wanrooij, D. P. Munoz, and A. J. Van Opstal. Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J Neurophysiol* 88: 438–454, 2002; 10.1152/jn.00699.2001. This study addresses the integration of auditory and visual stimuli subserving the generation of saccades in a complex scene. Previous studies have shown that saccadic reaction times (SRTs) to combined auditory-visual stimuli are reduced when compared with SRTs to either stimulus alone. However, these results have been typically obtained with high-intensity stimuli distributed over a limited number of positions in the horizontal plane. It is less clear how auditory-visual interactions influence saccades under more complex but arguably more natural conditions, when low-intensity stimuli are embedded in complex back-grounds and distributed throughout two-dimensional (2-D) space. To study this problem, human subjects made saccades to visual-only (V-saccades), auditory-only (A-saccades), or spatially coincident auditory-visual (AV-saccades) targets. In each trial, the low-intensity target was embedded within a complex auditory-visual background, and subjects were allowed over 3 s to search for and foveate the target at 1 of 24 possible locations within the 2-D oculomotor range. We varied systematically the onset times of the targets and the intensity of the auditory target relative to back-ground [i.e., the signal-to-noise (S/N) ratio] to examine their effects on both SRT and saccadic accuracy. Subjects were often able to localize the target within one or two saccades, but in about 15% of the trials they generated scanning patterns that consisted of many saccades. The present study reports only the SRT and accuracy of the first saccade in each trial. In all subjects, A-saccades had shorter SRTs than V-saccades, but were more inaccurate than V-saccades when generated to auditory targets presented at low S/N ratios. AV-saccades were at least as accurate as V-saccades but were generated at SRTs typical of A-saccades. The properties of AV-saccades depended systematically on both stimulus timing and S/N ratio of the auditory target. Compared with unimodal A-and V-saccades, the improvements in SRT and accuracy of AV-saccades were greatest when the visual target was synchronous with or leading the auditory target, and when the S/N ratio of the auditory target was lowest. Further, the improvements in saccade accuracy were greater in elevation than in azimuth. A control experiment demonstrated that a portion of the improvements in SRT could be attributable to a warning-cue mechanism, but that the improvements in saccade accuracy depended on the spatial register of the stimuli. These results agree well with earlier electrophysiological results obtained from the midbrain superior colliculus (SC) of anesthetized preparations, and we argue that they demonstrate multisensory integration of auditory and visual signals in a complex, quasi-natural environment. A conceptual model incorporating the SC is presented to explain the observed data.