Perception, 2007, volume 36, pages 000 – 000

doi:10.1068/p5731

Last but not least

Abstract. Perceptual integration of audio – visual stimuli is fundamental to our everyday conscious experience. Eye-movement analysis may be a suitable tool for studying such integration, since eye movements respond to auditory as well as visual input. Previous studies have shown that additional auditory cues in visual-search tasks can guide eye movements more efficiently and reduce their latency. However, these auditory cues were task-relevant since they indicated the target position and onset time. Therefore, the observed effects may have been due to subjects using the cues as additional information to maximize their performance, without perceptually integrating them with the visual displays. Here, we combine a visual-tracking task with a continuous, task-irrelevant sound from a stationary source to demonstrate that audio – visual perceptual integration affects low-level oculomotor mechanisms. Auditory stimuli of constant, increasing, or decreasing pitch were presented. All sound categories induced more smooth-pursuit eye movement than silence, with the greatest effect occurring with stimuli of increasing pitch. A possible explanation is that integration of the visual scene with continuous sound creates the perception of continuous visual motion. Increasing pitch may amplify this effect through its common association with accelerating motion.

Using eye-tracking to study audio – visual perceptual integration

In our everyday perception, we integrate information across modalities. One clear example of audio – visual integration was reported by Sekuler et al (1997). Their subjects were presented with displays showing two identical objects moving in opposite directions, which could be perceived as either passing each other or colliding and reversing directions. The study showed that a brief clicking sound, played near the moment of coincidence of the two objects, biased the subjects toward perceiving a collision. While this is a convincing example for audio – visual perceptual integration, as reported by the subjects, the immediate behavioral consequences of such integration have not been thoroughly explored. Such consequences could, for example, be reflected in the patterns of an observer's eye movements. Moving sound sources are known to elicit both saccadic eye movements (Fiebig et al 1981) and smooth-pursuit eye movements (Hashiba et al 1996) even without concurrent visual stimulation. Moreover, auditory cues presented at the onset of a visual stimulus can reduce subjects' saccadic latency in visual-search tasks (Engelken and Stevens 1989). However, in these studies, the sound signal is task-relevant and can be used as an additional cue to facilitate task performance. No crossmodal perceptual integration is necessary to benefit from this cue and create the observed oculomotor effects. To study such integration, we employed a visual-tracking task with sound stimuli that were task-irrelevant, ie did not contain any information that could have helped the subjects to perform better at the tracking task.

For this experiment, we created displays showing a small blue disk (with a diameter of 0.5 deg, luminance 17.5 cd m^{-2}, CIE $x = 0.15$, $y = 0.06$) on a gray background (31.7 cd m^{-2}, $x = 0.32$, $y = 0.33$) moving horizontally (either left-to-right or right-to-left) at a constant speed of 4 deg s^{-1} for a duration of 6 s. On its trajectory, the disk was only visible during five intervals lasting 130 ms each, separated by four intervals of invisibility lasting 1170 ms each. In order to make subjects track the disk, they were instructed to manually report any color changes of the disk from blue to red (29.4 cd m^{-2}, $x = 0.64$, $y = 0.33$), which happened during half of the trials (target trials) at a random time. As indicated by preliminary studies (see figure 1), this task induced a mixture of saccades and smooth pursuit in the subjects (cf Becker and Fuchs 1985).

Our hypothesis was that a task-irrelevant sound from a stationary source would alter the subjects' perception of the displays in such a way that the proportion of their

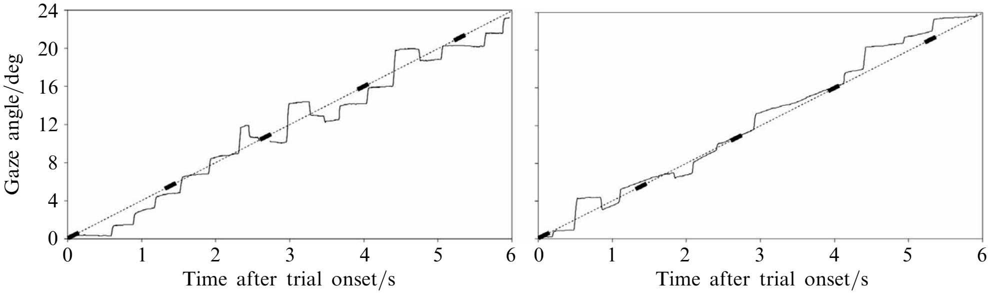

Figure 1. Example time courses of horizontal gaze position (measured at 500 Hz by an SR Research EyeLink-II system) during the visual-tracking task, subtending 24 deg, or 1300 pixels, on a 21 inch CRT monitor. In each panel, the diagonal, dotted line, indicates the interpolated trajectory of the moving disc, and its thick, solid segments represent the intervals when the disc was visible. The upper trajectory is dominated by saccades (near-vertical segments) and fixations (near-horizontal segments), whereas the lower one shows substantial smooth pursuit (slanted segments).

reports of smooth-pursuit to saccadic eye movement was systematically biased. To avoid low-level orienting effects on visual attention by abrupt sound onsets, we played one continuous sound throughout a given trial. Even playing a constant tone might bias subjects towards perceiving the display as one integrated event, that is a single disk moving across the screen as opposed to several disks appearing sequentially at different positions. This percept, in turn, could induce a greater amount of smooth pursuit. Sounds of changing pitch may lead to even stronger biases, as increasing pitch is typically associated with acceleration, and decreasing pitch is often experienced through the Doppler effect when fast objects pass by. Therefore, besides a no-sound condition, we studied three different sound conditions: linearly increasing pitch (262 Hz to 674 Hz sine wave), constant pitch (468 Hz sine wave), and linearly decreasing pitch (674 Hz to 262 Hz sine wave). The sounds were played at identical volume by two speakers positioned to the left and to the right of the stimulus monitor, so that the subjects received a sound pressure level of approximately 75 dB.

Since there is no standard method for measuring the amount of smooth pursuit in a gaze trajectory, we applied three different methods to ensure that our results were independent of any particular measurement technique. Following Spering et al (2005), we computed the average horizontal gaze velocity, that is, the average slope of the horizontal gaze position as a function of time, excluding saccades (as detected by a combined 30 deg s^{-1} velocity and 8000 deg s^{-2} acceleration threshold). This slope was measured by linear regression on groups of gaze-position samples measured between any two consecutive saccades (global velocity) or on small groups of only four successive gaze-position samples (local velocity). To derive a third measure that is independent of specific saccade detection algorithms, we also computed the number of horizontal one-pixel gaze position intervals hit by at least one of the gaze samples recorded during a given trial (pixel coverage). Increased smooth pursuit leads to a more homogeneous distribution of horizontal gaze positions, which in turn increases pixel coverage.

Each of our twenty subjects performed four practice trials, followed by five experimental trials for each of the sixteen conditions (4 sound conditions $\times$ 2 directions of motion $\times$ target presence/absence). Only target-absent trials without manual response were analyzed. Two-way ANOVAs showed that the direction of motion had no effect on any of the three smooth-pursuit measures (all $F_{1,19}$s < 1), and did not interact with the factor sound (all $F_{3,57}$s < 1). The sound conditions, however, exerted a significant main effect on all three measures (all $F_{3,57}$s > 5.93, ps < 0.005—see figure 2). The pattern of results was almost identical across these measures. There was more smooth

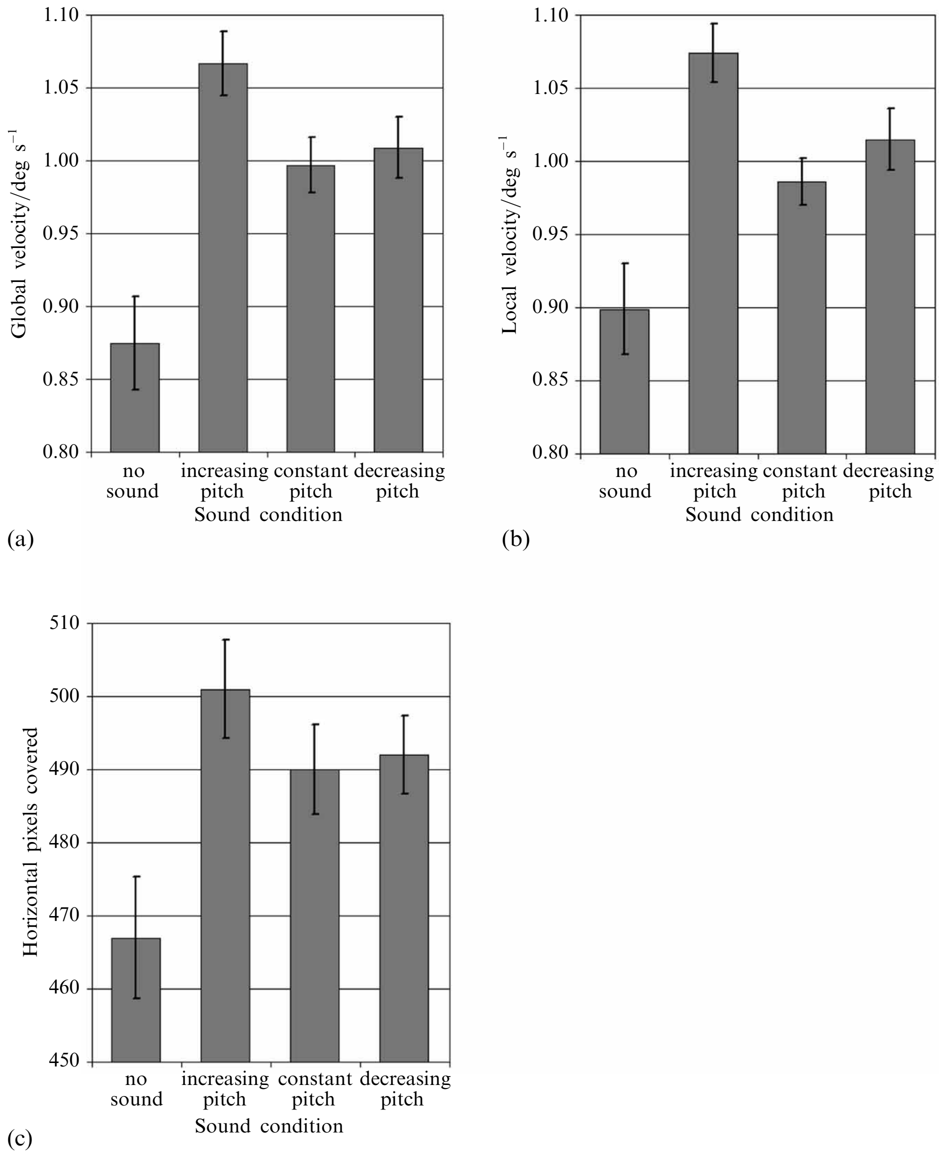

Figure 2. Amount of smooth pursuit under the four sound conditions, as measured by three different variables: global gaze velocity (a), local gaze velocity (b), and pixel coverage (c). Owing to the substantial individual differences in the absolute amount of smooth pursuit, the error bars indicate the standard error of the normalized data (zero mean across conditions for each subject).

pursuit in each of the sound conditions than in the no-sound condition (all t_{19}s > 2.22, ps < 0.05), and there was more smooth pursuit in the increasing-pitch condition than in the other two sound conditions. However, while this latter effect was significant for the local velocity measure (both t_{19}s > 2.09, ps < 0.05), global velocity revealed a significant difference only between increasing and constant pitch ($t_{19} = 2.26$, $p < 0.05$), and a tendency toward a difference between increasing and decreasing pitch ($t_{19} = 1.96$, $p = 0.065$) and pixel coverage showed only weak tendencies in both cases ($t_{19} = 1.44$, $p = 0.168$, and $t_{19} = 1.68$, $p = 0.11$, respectively).

These results demonstrate that task-irrelevant, stationary, continuous sound can bias oculomotor control toward smooth pursuit, suggesting that the crossmodal perceptual integration of continuous sound with the visual display facilitated the perception of continuous object motion. Moreover, the results indicate that this effect is more pronounced for increasing pitch than for constant or decreasing pitch. The reason for this difference might be our association of increasing pitch with accelerating motion. However, further research is necessary to study the perceptual relationship between the velocity and direction of motion with the rate of pitch increase or decrease. While the present research provided first evidence of audio – visual perceptual integration influencing low-level oculomotor control, future studies need to analyze this effect from a broader perspective.

Mei Xiao, May Wong, Michelle Umali§, Marc Pomplun¶
Department of Computer Science, University of Massachusetts at Boston, 100 Morrissey Boulevard, Boston, MA 02125-3393, USA; §Center for Neurobiology & Behavior, Columbia University, 1051 Riverside Drive, New York, NY 10032, USA; e-mail: marc@cs.umb.edu

Acknowledgments. We would like to thank Christopher D Jones for generating the sound stimuli.

References

Becker W, Fuchs A F, 1985 "Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target" *Experimental Brain Research* **57** 526 – 575
Engelken E J, Stevens K W, 1989 "Saccadic eye movements in response to visual, auditory, and bisensory stimuli" *Aviation, Space and Environmental Medicine* **60** 762 – 768
Fiebig E, Schaefer K P, Süβ K J, 1981 "Form and accuracy of voluntary ocular tracking movements in response to sinusoidally moving acoustic targets" *Journal of Neurology* **226** 77 – 84
Hashiba M, Matsuoka T, Baba S, Watanabe S, 1996 "Non-visually induced smooth pursuit eye movements using sinusoidal target motion" *Acta Otolaryngologica Supplement* **525** 158 – 162
Sekuler R, Sekuler A B, Lau R, 1997 "Sound alters visual motion perception" *Nature* **385** 308
Spering M, Kerzel D, Braun D I, Hawken M J, Gegenfurtner K R, 2005 "Effects of contrast on smooth pursuit eye movements" *Journal of Vision* **5** 455 – 465

Author: Please make sure that you return the Notice of PDF Requirements with the corrected proofs. If you have not received one, it can be downloaded from http://www.perceptionweb.com/reprints.pdf

¶ Author to whom all correspondence should be addressed.