Localization of visual and auditory stimuli during smooth pursuit eye movements

Kerstin Königs
Department of Neurophysics, Philipps-University Marburg, Karl-von-Frisch-Straße 8a, Marburg, Germany

Frank Bremmer
Department of Neurophysics, Philipps-University Marburg, Karl-von-Frisch-Straße 8a, Marburg, Germany

Humans move their eyes more often than their heart beats. Although these eye movements induce large retinal image shifts, we perceive our world as stable. Yet, this perceptual stability is not complete. A number of studies have shown that visual targets which are briefly presented during such eye movements are mislocalized in a characteristic manner. It is largely unknown, however, if auditory stimuli are also mislocalized, i.e. whether or not perception generalizes across senses and space is represented supramodally. In our current study subjects were asked to localize brief visual and auditory stimuli that were presented during smooth pursuit in the dark. In addition, we measured auditory and visual detection thresholds. Confirming previous studies, perceived visual positions were shifted in direction of the pursuit. This shift was stronger for the hemifield the eye was heading towards (foveopetal). Perceptual auditory space was compressed towards the pursuit target (ventriloquism effect). This perceptual error was slightly reduced during pursuit as compared to fixation and differed clearly from the mislocalization of visual targets. While we found an influence of pursuit on localization, we found no such effect on the detection of visual and auditory stimuli. Taken together, our results do not provide evidence for the hypothesis of a supramodal representation of space during active oculomotor behavior.

Keywords: pursuit, visual, auditory, psychophysics, sensorimotor integration


Introduction

Primates perform thousands of eye movements every day. With each of these movements the image of the static environment on the retina gets displaced but nonetheless we perceive the outer world as being stable. Yet, different from introspection, spatial perception during eye movements is not veridical. A number of studies have shown in recent years that voluntary eye movements (saccades and smooth pursuit) modulate the perceived localization of briefly flashed targets. During visually guided saccades, the error pattern depends on the exact experimental settings. In darkness, stimuli flashed briefly before saccade onset and during the first half of the saccade are mislocalized in the direction of the eye movement (Cai, Pouget, Schlag-Rey, & Schlag, 1997; Honda, 1989). Perceived locations of stimuli presented during the second half or shortly after the end of the saccade are shifted in the direction opposite to the saccade leading to an overall bimodal error pattern. In light, perisaccadic stimuli are localized near the endpoint of the eye movement leading to a perceptual compression of space (Awater, Burr, Lappe, Morrone, & Goldberg, 2005; Kaiser & Lappe, 2004; Lappe, Awater, & Krekelberg, 2000; Ross, Morrone, & Burr, 1997).

Not only saccades but also smooth eye movements modulate the spatial perception of briefly flashed visual targets. It has been shown that targets are perceptually shifted in the direction of smooth pursuit. This mislocalization is stronger in the hemifield the eye is heading towards (foveopetal) (Kerzel, Aivar, Ziegler, & Brenner, 2006; Mitrani & Dimitrov, 1982; Rotman, Brenner, & Smee, 2004; van Beers, Wolpert, & Haggard, 2001). Recent studies showed that such mislocalization is not only found during visually guided smooth pursuit eye movements but also during the smooth phase of a reflexive eye movement, namely optokinetic nystagmus (OKN) (Kaminiarz, Krekelberg, & Bremmer, 2007; Tozzi, Morrone, & Burr, 2007). The hemifield-asymmetry of localization as described during pursuit, however, was not found during OKN. Also the error pattern during optokinetic afternystagmus (OKAN) differed from the one found during smooth pursuit (Kaminiarz, Krekelberg, & Bremmer, 2008).

Natural targets in the outside world usually stimulate multiple senses, e.g. vision and hearing. It is well known that there is a difference in the spatial resolution of the visual and auditory systems. However, this does not argue against the processing of visual and auditory stimuli using similar mechanisms and integrating them into a common, supramodal spatial representation. Recent neurophysiological studies have provided evidence for a common supramodal spatial encoding of visual and tactile or visual and auditory stimuli in the macaque posterior parietal...
cortex (Avillac, Deneve, Olivier, Pouget, & Duhamel, 2005; Bremmer, Schlack, Duhamel, Graf, & Fink, 2001a; Duhamel, Colby, & Goldberg, 1998; Schlack, Sterbing-D’Angelo, Hartung, Hoffmann, & Bremmer, 2005). In these studies single neurons in the ventral intraparietal area (VIP) of the monkey were shown to respond not only to visual, but also to tactile or auditory stimuli. Interestingly, the receptive fields for visual and tactile or visual and auditory stimuli as mapped during fixation were spatially congruent. Schlack et al. (2005) furthermore showed that visual and auditory signals tended to be encoded in the same spatial frame of reference: visual and auditory receptive fields were either anchored with respect to the retina (eye centered encoding), with respect to the head (head-centered encoding) or with respect to an intermediate frame of reference. These data suggest a supramodal encoding of spatial information in macaque posterior parietal cortex. fMRI studies suggest that a functional equivalent of macaque area VIP exists also in humans (Bremmer, Schlack, Shah, Zafiris, Kubischik, Hoffmann et al., 2001b; Konen, Kleiser, Seitz, & Bremmer, 2005).

The above mentioned results almost automatically lead to the question if, like in macaques, such a supramodal spatial representation might also be found in humans. If such a modality independent representation existed and if localization was based on the neuronal readout of such a ‘supramodal map’, the perception of auditory targets should also be influenced by eye position and eye movements. Results from a few studies already suggested that this indeed might be the case. Lewald and colleagues showed that the position of the eyes in the orbit influences the perceived location of an auditory stimulus (Lewald, 1998; Lewald & Getzmann, 2006). In these experiments a noise burst was presented while gaze was directed either straight ahead or to an eccentric (horizontal or vertical) position. The authors found a systematic mislocalization of the auditory stimulus away from the current eye position: gaze directed to the right induced a leftward mislocalization and vice versa. Other studies, however, found the opposite, i.e. a co-alignment of effects: gaze shifts in one given direction induced a shift of perceived auditory location in the same direction (Lewald & Ehrenstein, 1996, 1998).

In one of our own studies, we showed that the perceived location of a noise burst presented during OKN is shifted in the direction of the slow phase eye movement (Königs, Knöll, & Bremmer, 2007). This shift is temporally modulated around the OKN fast phase. Like for visual targets presented during visually guided saccades, perceived locations are first shifted in the direction of the fast phase and then into the opposite direction. Localization of auditory targets during visually guided saccades is also not veridical. The observed error pattern, however, differs from the one found for visual localization (Klingenhoef er & Bremmer, 2009).

Taken together, results from previous studies suggest that spatial perception of auditory targets might be modulated during voluntarily controlled smooth pursuit eye movements. Accordingly, in our present study we asked subjects to report the perceived location of an auditory stimulus presented during smooth pursuit of a small moving visual target. We compared these results with those obtained for the localization of briefly flashed visual targets.

It has been proposed that during pursuit, the focus of attention is shifted ahead of the pursuit target (van Donkelaar & Drew, 2002). This attentional shift might lower detection thresholds for stimuli presented ahead of the pursuit target which, in turn, could influence localization. Accordingly, we were interested to see whether the detection of briefly presented stimuli is modulated during smooth eye movements.

**Methods**

**Subjects**

5 subjects (2 male and 3 female, 19–29 years, mean age 23.8) participated in the auditory and in the first visual localization experiment. We measured visual localization performance from another set of 5 subjects (5 female, 24–33 years, mean age 27.6) in a second visual localization experiment. We obtained detection thresholds from 5 subjects (2 male and 3 female, 27–33 years, mean age 29.6). All subjects had normal hearing and normal or corrected to normal vision. All were naïve with respect to the goal of the experiments except for one of the authors (KK). All subjects gave informed written consent prior to the experiments. All procedures used in this study conformed to the Declaration of Helsinki.

**Set-up**

All experiments were performed in a dark, sound attenuated chamber (2.85 × 1.85 × 2.45 m³). Visual stimuli were either presented on a 22" (40° horizontal, 30° vertical) Iiyama Vision Master Pro 513 monitor or via an Electrohome Marquee 8000 CRT projector on a 79 inch (70° horizontal, 52.5° vertical) sound transparent projection screen. Both projection systems ran at 100 Hz with a spatial resolution of 1152 × 864 pixels. A loudspeaker (Visaton SC 8, 8Ω) was mounted on a rail system. When using the monitor, the rail system was placed in front and slightly below the monitor and was hidden from view by a sound transparent cloth. When using the projection screen, the rail system was placed behind the center of the screen and visual stimuli were (front-) projected onto the screen, i.e. the video-projector was placed behind the subject (outside the experimental chamber) and in front of the projection screen. A computer-operated motorized feedback servo system controlled the speaker position. The
movement of the rail system prior to each trial was acoustically masked by playing white noise from a second speaker placed beneath the rail system.

Subjects were seated 57 cm (monitor) or 114 cm (projection screen) away from the screen with their eyes leveled at the center of the screen and their heads stabilized by a chin rest. Eye movements were recorded at 500 Hz with an infrared eye tracking system (SR Research: Eye Link II) and analyzed off-line. Trials were excluded from further analysis if the subjects either had made a saccade towards the visual or auditory stimulus which had to be localized or terminated pursuit or fixation before the offset of the target. Trials containing catch-up saccades within 100 ms before or after presentation of the stimulus were also excluded. Subjects responded to their specific tasks by using a standard computer keyboard.

Paradigms

We tested localization and detection thresholds for auditory and visual stimuli during three oculomotor tasks: pursuit to the right, pursuit to the left and fixation. For localization tasks a horizontal ruler was presented after the offset of the pursuit or fixation target. The ruler covered the entire width of the screen (70°) with a resolution of 1°. The numbers on the ruler ranged from 10 to 99 and were randomly re-arranged for each trial to prevent learning of any possible relationships between number and stimulus position. Subjects were instructed to indicate the perceived position of the auditory stimulus by entering via a keyboard the number on the ruler being closest to the perceived stimulus location. In the following, positions with positive values indicate locations to the right of the center of the screen (0°).

Auditory localization

The pursuit or fixation target was a white dot (diameter = 0.2°, luminance = 43 cd/m²) on a black background (luminance < 0.01 cd/m²). A 5 ms white noise burst (83 dB SPL, sound pressure level measured directly at the speaker) served as the auditory target stimulus. It was presented at one of 7 (−12°, −8°, −4°, 0°, 4°, 8° or 12°) positions. Stimuli were presented when the pursuit or fixation target was either at the position of the auditory stimulus or 8° to its right or left.

Prior to each trial the stimulus speaker was moved to its desired position. This positioning was followed by the onset of the visual (fixation or pursuit) target. For pursuit trials the visual target remained stationary for 800 ms and then started moving either to the left or to the right at 10°/s for 3000 ms while it remained stationary for fixation trials (see Figure 1). The pursuit or fixation target could appear at one of ten (starting) positions, all of which were

Figure 1. Procedure of the auditory and visual localization experiments. The pursuit target appeared at a given location for 800 ms. It then started moving rightwards or leftwards at 10°/s for a total of 3000 ms. During pursuit either an auditory (5 ms white noise) or a visual (10 ms flash) stimulus was presented. There were 1100, 1500 or 1900 ms between motion onset and onset of the auditory stimulus. The visual target always appeared 1500 ms after motion onset. After offset of the pursuit target a ruler appeared on the screen and subjects had to indicate the perceived stimulus position.
presented on the horizontal meridian in 8° steps from 31° to -1° and from 1° to -31°. For each of these starting positions localization was tested for three different speaker positions. These were chosen so that the three speaker positions were either all in one hemifield or evenly distributed around the midline (see Figure 2). The start and end positions of the pursuit target were chosen so that for each speaker position total pursuit duration and pursuit duration prior to sound presentation were the same for all three spatial offsets ‘dx’ between the fovea and the speaker which had to be localized (dx = -8°, 0°, and +8°). Accordingly, there was a total of 3 (pursuit paths) * 3 (speaker positions) * 3 (distance visual target/speaker) * 3 (eye movement tasks) = 81 stimulus conditions. Each condition was presented 10 times. All conditions were presented in pseudo-randomized order.

### Visual localization

In a first set of experiments we determined the perceived position of briefly flashed visual stimuli during smooth pursuit. The pursuit target was the same as in the auditory localization experiment. When the target moved through the center of the screen a second, larger dot (diameter = 0.5°, luminance = 43 cd/m²) was flashed for 10 ms at one of 25 positions, evenly spaced around the screen center on a 5 × 5 grid (horizontal and vertical spacing = 4°, see Figure 1). In total there were 75 conditions with 10 repetitions per condition. The vertical position of the ruler, which was shown after the target motion, was presented at the target’s elevation in all cases in order to prevent misjudgment due to differences in elevation between target stimulus and ruler.

In a second set of experiments we repeated the auditory localization experiment using briefly flashed visual stimuli instead of white noise bursts. All other parameters were kept identical.

### Detection thresholds

For the detection experiment we used the same stimuli as for the localization experiments (auditory stimulus: 5 ms white noise; visual stimulus: white dot presented for 10 ms). A white dot (diameter = 0.2°, luminance = 43 cd/m²) on a black background (luminance < 0.1 cd/m²) served as the pursuit or fixation target. The stimuli, which had to be detected, were presented at retinal eccentricities of dx = -8°, 0°, or 8°. Individual detection thresholds were obtained with an adaptive one-up-one-down staircase (Levitt, 1971). The initial value for the Michelson contrast of the visual stimulus was 95%, the initial value of the auditory stimulus was 52 dB SPL (measured at the speaker). The stimulus could be presented during one of two intervals. After each trial the subject had to indicate by button-press in which interval they had perceived a stimulus. In 10% of the trials no stimulus was presented (catch-trials). The step size of the staircase was decreased by 20% for each reversal point. The staircases were...
terminated when a pre-defined minimum step size (0.1 dBA for auditory and 1% Weber contrast for visual stimuli) was reached. The threshold was calculated as the arithmetic mean of the values of the last six reversal points.

**Results**

**Auditory localization**

Subjects had to localize a brief auditory stimulus presented during fixation or during pursuit to the right or to the left. The eye-movement target and the auditory stimulus either had the same horizontal position (spatially congruent) or were separated by ±8°.

We ran a 7 (speaker positions) × 3 (spatial offsets) ANOVA for each of the three eye movement tasks (fixation, pursuit to the left and pursuit to the right). Here and in the following ANOVAs were run on mean error, unless stated otherwise. In all three cases we found a significant main effect of speaker position (pursuit leftwards: F(6,24) = 55.242, p < 0.001; rightwards: F(6,24) = 62.220, p < 0.001; fixation: F(6,24) = 80.890, p < 0.001).

Perceived locations were shifted towards larger eccentricities: target locations left from straight ahead were perceived further to the left and targets right from straight ahead were perceived further to the right. This can be considered an expansion of perceptual space. This expansion saturated at more eccentric positions and could be described by a sigmoid function. Localization was also dependent on the distance between speaker and visual target. There was a significant main effect of spatial offset for pursuit to the left (F(2, 8) = 43.136, p < 0.001), right (F(2, 8) = 28.328, p < 0.001) and fixation (F(2, 8) = 68.865, p < 0.001). If the visual target was to the right (left) of the speaker, the perceived stimulus location was also more to the right (left) as compared to a situation where stimulus and eye position were co-aligned.

In order to further analyze the perceptual errors, we will firstly consider the spatially congruent conditions (Figures 3A and 3C). We fitted sigmoids with four parameters to our fixation and pursuit data to quantify the mislocalization.

\[ y = y_0 + \frac{a}{1 + e^{-\left(\frac{x - x_0}{b}\right)}}. \]

In this equation, \( y_0 \) indicates the inflection point of the sigmoid while \( y_0 + a \) gives the value of the second plateau. Furthermore, \( a \) denotes the amplitude and \( b \) the slope of the sigmoid function.

We defined half the amplitude \( a \) as the size of the mislocalization effect (mean-err). In the fixation condition (Figure 3A) mean-err was 2.88°. Pursuit induced errors were almost twice as large as errors during fixation and were about equally strong for both eye movement directions. There was a mean-err of 4.85° for pursuit to the left (Figure 3C, blue line) and of 4.73° for pursuit to the right (Figure 3C, red line). A 3 eye movement conditions ANOVA on mean-err revealed a significant main effect of eye movement (F(2, 8) = 5.911, p = 0.027). Post-hoc pairwise comparisons (Holm-Sidak, \( \alpha = 0.05 \)) revealed significant differences between fixation and pursuit in either direction, but not between the two pursuit directions. Mean-err was significantly different from zero for pursuit to the right (Paired t-test, \( t(4) = 9.093, p < 0.001 \)) or left (Paired t-test, \( t(4) = -9.055, p < 0.001 \)) and for fixation (Paired t-test, \( t(4) = -8.901, p < 0.001 \)). In addition, this perceptual expansion was modulated by pursuit direction. The midpoint values of the sigmoids (center = \( y_0 + a/2 \)) can be considered the perceived center-positions in the various conditions. We fitted both single subject data and data averaged across all observers: center\(_{fix} = -0.39° ± 0.37° (-0.42°), center\_left = -1.57° ± 0.61° (-1.86°), center\_right = 0.25° ± 0.41° (0.17°). The values given are mean ± one standard error of the midpoint values derived from the sigmoids fitted to the single subject data. In brackets, we give the midpoint values from the sigmoids fitted to the data averaged across all observers. Compared to fixation, the midpoint values for pursuit were shifted in the direction of pursuit. During pursuit to the right all perceived locations were shifted to the right by 2.04° ± 0.76° as compared to perceived locations during pursuit to the left (Paired t-test, \( t(4) = -3.213, p = 0.032 \)). Here and in the following, all values given are mean ± one standard error.

For two of the speaker positions (\( x = -4° \) and \( x = 4° \)) localization was tested after smooth pursuit of different durations. In one case the sound was presented 1100 ms after onset of target motion, in the other case 1900 ms thereafter. This approach allowed us to determine whether or not pursuit duration might have an influence on localization. Localization error between the two pursuit durations differed on average by 0.55° ± 0.44 and 0.56° ± 0.36 for pursuit to the right and left, respectively, and was statistically not different for either pursuit direction (pursuit to the right: Paired t-test, \( t(4) = -1.253, p = 0.279 \); pursuit to the left: Paired t-test, \( t(4) = -1.518, p = 0.204 \)).

In order to quantify the influence of the position of the visual target on localization of the sound we compared the congruent condition (same horizontal position for visual target and speaker) to each of the two corresponding incongruent conditions (visual target and speaker separated by 8°) for each speaker position and each of the eye movement conditions. For example, for fixation and speaker position 0° we determined the difference in perceived sound position between target position 0° and target position –8° or target position 8°, respectively. Data was then collapsed across speaker positions and, in
In the case of pursuit, across the directions of movement. We ran a 2 (spatial offsets) × 3 (presentation times) ANOVA for both fixation and pursuit. In both cases there was a significant main effect of spatial offset (fixation: $F(1,4) = 95.077, p < 0.001$; pursuit: $F(1,4) = 76.340, p < 0.001$) but no effect of timing (fixation: $F(2,8) = 4.191, p = 0.057$; pursuit: $F(2,8) = 4.056, p = 0.061$). Compared to the congruent condition, perceived sound position was shifted towards the visual target for both fixation (Figure 3B) and pursuit (Figure 3D). This shift was stronger for fixation ($dx = 2.46° ± 0.25$) as compared to pursuit ($dx = 1.68° ± 0.19$) (Paired t-test, $t(4) = 2.973, p = 0.041$).

### Visual localization

We presented visual stimuli that our subjects had to localize while they performed smooth pursuit or kept steady fixation. Data from a single subject are given in Figures 4A–4C while data averaged across all observers (the same whose data are shown in Figure 3) are given in Figures 4D–4E. Like for auditory stimuli, and confirming results from previous studies, localization during steady fixation was not veridical (Figures 4A and 4D).

We ran a 5 (horizontal) × 5 (vertical stimulus positions) ANOVA. There was a significant main effect of horizontal
stimulus position ($F(4,16) = 6.384, p = 0.003$) but no effect of vertical stimulus position ($F(4,16) = 1.321, p = 0.304$). We observed a shift of perceived stimulus location towards the fovea (foveopetal) that increased with increasing retinal eccentricity. The increase of the absolute size of the horizontal error with increasing eccentricity could be modeled by a linear regression function $y = -0.19x - 0.36$ ($R^2$-value: 0.91).

Errors during pursuit differed markedly from those made during steady fixation (Figures 4B and 4E). Confirming previous results reported in the literature (e.g. van Beers et al., 2001) we observed an overall shift in the direction of the eye movement: all perceived positions were shifted in the direction of pursuit. Averaged across all stimulus positions the error was $2.3^\circ - 0.7^\circ$. We ran a 5 (horizontal) × 5 (vertical stimulus positions) ANOVA. There were significant main effects of horizontal ($F(4,16) = 6.384, p = 0.003$) and vertical stimulus position ($F(4,16) = 6.060, p = 0.004$). Post-hoc pairwise comparisons of vertical stimulus position (Holm-Sidak, $\alpha = 0.05$) revealed significant differences for 2 of the 10 comparisons ($-4^\circ$ and $0^\circ$ vs. $8^\circ$). Errors were larger for

Figure 4. Mean perceived position of briefly flashed visual targets during pursuit and fixation. A–C: single subject, D–F: average across all observers. Stimuli were always flashed when the eye was at the center of the screen, so retinal and screen coordinates are the same. Horizontal and vertical positions are given on the abscissa and the ordinate, respectively. The dots indicate the real positions of the flash. The tip of the arrow denotes the perceived position (averaged across all subjects for D–F). Data from the two pursuit directions was combined. Arrows pointing to the right indicate shifts of perceived position in pursuit direction. A, D: Data from fixation trials (black). B, E: Data from pursuit trials (red). C, F: Baseline-corrected data (green). During fixation perceived positions were shifted towards the center of the screen while during pursuit perceived positions were shifted in the direction of pursuit. Localization error during pursuit was larger for positions in the hemifield the pursuit target was moving into than for positions in the hemifield it was leaving.
stimuli presented close to the vertical meridian. Mis-localization was stronger for flash-positions in the foveo-petal hemifield, i.e. for stimuli presented ahead of the pursuit target (Paired t-test, \(t(4) = -2.864, p = 0.046\)).

Subtracting the errors made during fixation (baseline) from those made during pursuit (Figures 4C and 4F) strengthened the difference between stimuli presented ahead and behind the pursuit target. A 5 (horizontal) \(\times\) 5 (vertical stimulus positions) ANOVA showed significant main effects of horizontal (\(F(4,16) = 36.097, p < 0.001\)) and vertical stimulus position (\(F(4,16) = 8.813, p < 0.001\)). Post-hoc pairwise comparisons of vertical stimulus position (Holm-Sidak, \(\alpha = 0.05\)) revealed significant differences for 3 of the 10 comparisons (\(-4^\circ\) and \(0^\circ\) vs. \(8^\circ\) and \(-4^\circ\) vs. \(4^\circ\)). Errors were larger for stimuli presented close to the vertical meridian. Localization errors were significantly larger for stimuli presented ahead of the pursuit target (Paired t-test, \(t(4) = -7.370, p = 0.002\)).

We also ran a visual localization experiment using the same parameters as used in the auditory localization experiment. Visual stimuli were presented 1100, 1500 or 1900 ms after fixation of pursuit onset at head-centered positions within 12° right or left from straight ahead and at three different retinal eccentricities (\(-8^\circ, 0^\circ\) or \(8^\circ\)).

During fixation we observed no head-centered effects (Figure 5A). Errors were equally large for all head-
centered positions. We ran a one-way ANOVA (9 stimulus positions). There was a significant main effect of stimulus position ($F(2,8) = 3.199, p = 0.002$). However, post-hoc pairwise comparisons (Holm-Sidak, $\alpha = 0.05$) revealed no significant differences between any of the stimulus positions.

When collapsing the data across all positions and plotting it in retinal coordinates we observed a compression towards the fovea equal to the one we had found in our first visual localization experiment (Figure 5B). A 3 (retinal eccentricities) × 3 (presentation times) ANOVA showed a significant main effect of retinal eccentricity ($F(2,8) = 31.338, p < 0.001$) but no effect of timing ($F(2,8) = 3.885, p = 0.066$).

During pursuit we observed a shift in the direction of pursuit (Figure 5C). Since stimulus position was confounded with the timing of the stimulus we ran a one-way ANOVA (3 stimulus positions) for each of the three presentation times. No significant effect of stimulus position was found for any of the three presentation times (1100 ms: $F(2,8) = 1.111, p = 0.375$; 1500 ms $F(2,8) = 2.42, p = 0.051$; 1900 ms: $F(2,8) = 0.486, p = 0.632$). Analyzing our data in a retinocentric reference frame revealed an asymmetric shift in the direction of pursuit (Figure 5D). A 3 (presentation times) × 3 (retinal eccentricities) ANOVA revealed significant main effects of timing ($F(2,8) = 132.6, p < 0.001$) and retinal position ($F(2,8) = 5.237, p = 0.035$). Errors were larger the earlier the stimuli were presented after pursuit onset. The shift was also stronger for stimuli presented in the foveopetal hemifield.

**Comparison of visual and auditory localization**

Since we ran a localization experiment with both visual and auditory stimuli using identical parameters we were able to directly compare localization performance for the two modalities. The results are summarized in Table 1.

A qualitative comparison of the observed errors shows differences for errors in retinocentric as well as craniocentric frames of reference. For auditory stimuli we observed a craniocentric expansion of perceptual space during both fixation and pursuit, respectively. On the other hand, there were no craniocentric effects for visual localization in both cases. When analyzing data in retinocentric coordinates we observed a compression of perceptual space towards the fovea for auditory data during both fixation and pursuit. For visual data we also found a compression towards the fovea for fixation, but an asymmetric shift in the direction of pursuit. Errors were larger for stimulus positions in the foveopetal as compared to the foveofugal hemifield.

### Detection thresholds

The differences in localization during smooth pursuit as compared to steady fixation might have been based on different sensory thresholds during eye movements and

---

<table>
<thead>
<tr>
<th>modality</th>
<th>reference frame</th>
<th>auditory</th>
<th>visual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>craniocentric</td>
<td>retinocentric</td>
<td>craniocentric</td>
</tr>
<tr>
<td>fixation</td>
<td>expansion</td>
<td>compression</td>
<td>no effect</td>
</tr>
<tr>
<td>pursuit</td>
<td>expansion</td>
<td>compression</td>
<td>no effect</td>
</tr>
<tr>
<td>Effects of pursuit direction</td>
<td>shift in the direction of pursuit</td>
<td></td>
<td>shift in the direction of pursuit</td>
</tr>
</tbody>
</table>

Table 1. Comparison of localization errors for auditory and visual stimuli in craniocentric and retinocentric reference frames.
fixation. If for example the threshold for detecting a stimulus was higher during pursuit as compared to fixation, the stimulus would have to be integrated over a larger time window resulting in a larger localization error. Alternatively, a shift of attention ahead of the pursuit target, as proposed by van Donkelaar and Drew (2002), could lead to lower detection thresholds for stimuli presented ahead of the pursuit target. Accordingly, we wanted to determine whether or not the detection of brief visual and auditory targets was dependent on oculomotor tasks. We used the same stimuli for the localization and the detection experiment.

Subjects were presented with auditory targets of 5 ms duration during pursuit and fixation. The visual target could be at the same horizontal position as the speaker or 8° to the right or left. We ran a 3 (stimulus positions) × 2 (types of eye movement) ANOVA on threshold. There was no significant main effect of eye movement ($F(1,4) = 1.411, p = 0.301$) or stimulus position ($F(2,8) = 0.993, p = 0.412$) (Figure 6).

We also measured visual contrast thresholds during pursuit and fixation. The stimulus was presented at three different retinal eccentricities ($-8°, 0°$ or $8°$). A 3 (stimulus positions) × 2 (types of eye movement) ANOVA on threshold revealed no significant main effect of eye movement ($F(1,4) = 1.459, p = 0.288$) or stimulus position ($F(2,8) = 0.659, p = 0.462$) (Figure 7).

**Discussion**

Neurophysiological as well as psychophysical data suggest that primates might use a supramodal representation of space in order to localize objects in the outer world. We tested for such a supramodal spatial encoding by asking human subjects to localize briefly presented auditory or visual targets, either during fixation or during smooth pursuit eye movements. Auditory localization was not veridical, neither during fixation nor during pursuit. In all conditions we observed an expansion of perceptual space. This mislocalization pattern clearly differed from the one found for visual localization, where flashed targets were mislocalized in the direction of the pursuit. This latter error was found predominantly in the hemifield the eyes were moving towards to. Detection of visual and auditory stimulation was not dependent on whether or not subjects performed eye movements.

Our localization data argue against a single, supramodal representation of sensory stimuli during eye movements. In the following, we will discuss possible mechanisms of the observed error patterns.

**Auditory localization**

The perceived position of stimuli presented in the periphery was shifted outward with respect to the actual position of the stimuli. The size of the error in our study was comparable to that found in previous studies (see e.g. Lewald, 1998). This expansion of perceptual space was modulated by eye movements and eye position. Shifts were larger during pursuit than during fixation. Localization differed between the two pursuit directions for all spatial test positions. Perceived positions during pursuit to the right were more rightward than during pursuit to the left.

Our current results are in line with the ones we had previously found for localization during optokinetic nystagmus (Königs et al., 2007) (Figure 8). In both cases we used the same stimulus velocity ($10°/s$) but very different visual stimuli to elicit eye movements. While we had used a random dot pattern (black dots on a gray background) filling the entire screen for eliciting optokinetic nystagmus (OKN), we used a single dot (white, on a black background) to drive smooth pursuit. Despite these differences the observed localization errors are remarkably similar. Thus, it seems that auditory localization during slow eye movements is independent of the visual stimulus used. This has already been demonstrated for visual localization during pursuit (Kerzel et al., 2006).

The position of the visual target also had an influence on the perceived position of the sound. When the visual target was right (left) of the auditory stimulus, the perceived stimulus position was also more to the right (left).
Such an attraction of an auditory event towards a spatially and temporally congruent visual event is known as the ventriloquist effect (Howard & Templeton, 1966). It is assumed that the ventriloquist effect is a perceptual phenomenon (Bertelson & Aschersleben, 1998). It has been proposed that bimodal stimuli are integrated in a Bayesian fashion to produce the ventriloquist effect (Battaglia, Jacobs, & Aslin, 2003; Binda, Bruno, Burr, & Morrone, 2007; Sato, Toyoizumi, & Aihara, 2007; Witten & Knudsen, 2005). Spatial information from the two modalities is combined in a way to optimize localization. Usually localization of the visual stimulus is better, thus giving it greater reliability (i.e. ‘weight’ in the Bayesian vocabulary). If the visual stimulus gets degraded, perceived sound positions will shift towards the actual sound position (Alais & Burr, 2004). In our experiments the visual stimuli were single small spots of light. Accordingly, position information had low variance and therefore a high weight resulting in a ventriloquist effect.

We found a shift of perceived sound positions towards the visual target that was reduced during pursuit as compared to fixation. This is well in line with the aforementioned experiments (Alais & Burr, 2004; Binda et al., 2007): the visual target moved and hence its spatial position changed over time. This higher positional variance probably results in a smaller weight in the Bayesian framework.

Visual localization

Fixation

Flashed stimuli were mislocalized towards the fovea during fixation. This shift was not uniform but rather increased linearly with increasing retinal eccentricities. This result is in contrast to our findings that auditory stimuli were localized further in the periphery indicating a perceptual expansion of space.

Despite extensive research the reason for this centripetal shift of visual localization remains unclear. van der Heijden, van der Geest, de Leeuw, Krikke, and Musseler (1999) proposed that the system providing commands for controlling saccades also provides localization information. It has often been observed that saccades fall short of the target (e.g. Becker, 1972; Deubel, 1991; Henson, 1978). Accordingly, this so called “saccadic undershoot” could be functionally related to errors in localization, i.e. peripheral targets would appear closer to the fovea than they actually are.

Another possible explanation is a shift of attention towards the fovea during fixation. It has been demonstrated that receptive fields in macaque area LIP shift towards the fixation point as compared to free viewing (Ben Hamed, Duhamel, Bremmer, & Graf, 2001). Such a shift of receptive field position could cause mislocalization. It is important to note, however, that such a shift at the neural level would cause a centrifugal shift of perceived positions if neural processing was based on a labeled line coding. We therefore suggest that these attentional processes do not account for perceptual mislocalization during fixation.

Pursuit

During pursuit we observed a perceptual shift in direction of the movement that was modulated by stimulus eccentricity. The shift revealed a hemifield asymmetry, i.e. the error was larger for flash positions the eye was moving towards to. This asymmetry occurred for both pursuit directions (left and right) indicating that the effect depends on the position of the stimulus relative to the pursuit target and not on its absolute position in space.

Many studies have demonstrated that the perceived position of a visual target flashed during pursuit is shifted in the direction of the movement. This mislocalization has been explained as a mismatch between (fast) retinal and (rather slow) extraretinal signals (Brenner, Smeets, & van den Berg, 2001; Hazeltown & Wiersma, 1924; Schlag & Schlag-Rey, 2002). Other studies have also observed that the shift in perceived position is more pronounced in the
foveopetal hemifield than in the foveofugal one (Kerzel et al., 2006; Mitrani & Dimitrov, 1982; Rotman et al., 2004; van Beers et al., 2001). So far there is no conclusive evidence explaining this hemifield effect. Van Beers et al. (2001) assumed that the retinal and extraretinal signals are combined differently for the two hemifields. On the other hand there might be a tendency to overestimate eccentricities during pursuit which, together with an overall shift in pursuit direction, produces the observed differences between hemifields (Kerzel et al., 2006; Rotman et al., 2004). The neural basis of such a centrifugal bias, however, remains unclear.

Comparison of visual and auditory localization

When comparing the results from the auditory and visual localization experiments, we found no evidence for a common supramodal representation of space. Subjects reliably overestimated the craniocentric eccentricity of auditory stimuli. In addition, there was a compression towards the visual target (ventriloquist effect). For visual stimuli we found no such craniocentric effects but a retinocentric compression of visual space during fixation and a shift during pursuit in direction of the eye movement. While pursuit induced a shift of the perceived position for both visual and auditory stimuli the characteristics of the errors differed considerably between the two modalities. Accordingly, data from this current study as well as data from our previous study on visual and auditory localization during OKN (Kaminiarz et al., 2007; Königs et al., 2007) argue against a supramodal spatial representation of stimuli during smooth eye movements in general.

Thresholds

Visual thresholds

We wanted to test whether a possible shift in attention ahead of the pursuit target, as hypothesized by van Donkelaar and Drew (2002), will not only manifest itself in shorter reaction times (van Donkelaar, 1999; van Donkelaar & Drew, 2002) but also in enhanced detection thresholds for stimuli presented ahead of the pursuit target. Several previous studies had found no effect of pursuit on contrast thresholds for foveally presented targets (e.g. Flipse, van der Wildt, Rodenburg, Keemink, & Knol, 1988; Murphy, 1978). Schütz, Delipetkos, Braun, Kerzel, and Gegenfurtner (2007) found slightly reduced contrast sensitivity during smooth pursuit for peripheral, moving stimuli presented above or below the pursuit or fixation target and even enhanced chromatic sensitivity. More recent studies, however, suggest that attention centers on the pursuit target and not ahead of it. Lovejoy, Fowler, and Krauzlis (2009) could demonstrate that subjects were only able to successfully perform a letter discrimination task when the stimuli were presented to the fovea. Kerzel and Ziegler (2005) found that memory capacity was reduced for peripheral targets when subjects had to perform spatial tasks. However, this was not true for color judgments. In addition, memory capacity was not impaired for spatial tasks when the targets were presented to the fovea.

In our study we presented stimuli at the fovea or 8° to the left or right while the subjects fixated or performed horizontal pursuit. Thresholds were not influenced by the eccentricity of the stimuli. In addition, there was no difference between thresholds during fixation and pursuit for any condition. This implies that the observed mis-localizations are not due to differences in visibility of the stimuli during pursuit.

Auditory thresholds

To test whether smooth pursuit eye movements influence auditory stimulus detection we measured thresholds for white noise bursts during pursuit and fixation. The auditory stimuli were presented at the same horizontal position as the pursuit or fixation target or 8° eccentric. There was no effect, however, of eccentricity or type of eye movement on auditory thresholds. Similar results had previously been obtained for fast eye movements. Harris and Lieberman (1996) demonstrated that saccades have no effect on the detection of auditory stimuli. This suggests that, since thresholds were unaffected by pursuit eye movements or eccentric stimulus presentation, the observed localization errors were not caused by differential detectability of the auditory stimuli.

Conclusion

Taken together, our data confirm previous data from experiments on visual localization (Rotman et al., 2004; van Beers et al., 2001). More important, however, they allow us to directly compare at a single subject and group level the localization of visual and auditory stimuli: it becomes obvious that visual and auditory localization are quite different. During steady fixation we observed a perceptual expansion of auditory space while we clearly observed a perceptual compression of visual space. Localization of visual targets during smooth pursuit was rather asymmetric while it was symmetric for auditory targets. We conclude that sensory spatial representation during smooth pursuit eye movements is not supramodal. Instead, humans rely on individual representations of visual and auditory target space during fixation and smooth eye movements.
Acknowledgments

Supported by Deutsche Forschungsgemeinschaft (GRK-885-NeuroAct, FOR-560) and EU-MEMORY.

We thank the two anonymous reviewers for their helpful comments.

Commercial relationships: none.
Corresponding author: Kerstin Königs.
Email: kerstin.koenigs@physik.uni-marburg.de.
Address: Department of Neurophysics, Philipps-University Marburg, Karl-von-Frisch-Straße 8a, D-35032 Marburg, Germany.

References


