Vision Research 49 (2009) 2241–2253

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Object recognition during foveating eye movements

Alexander C. Schütz *, Doris I. Braun, Karl R. Gegenfurtner

Abteilung Allgemeine Psychologie, Justus-Liebig-Universität, Otto-Behaghel-Str. 10F, 35394 Giessen, Germany

ARTICLE INFO

Article history:
Received 12 August 2008
Received in revised form 18 May 2009

Keywords:
Smooth pursuit eye movements
Saccadic eye movements
Object recognition
Visual perception
Perception and action

ABSTRACT

We studied how saccadic and smooth pursuit eye movements affect the recognition of briefly presented letters appearing within the eye movement target. First we compared the recognition performance during steady-state pursuit and during fixation. Single letters were presented for seven different durations ranging from 10 to 400 ms and four contrast levels ranging from 5% to 40%. For both types of eye movements the recognition rates increased with duration and contrast, but they were on average 11% lower during pursuit. In daily life humans use a combination of saccadic and smooth pursuit eye movements to foveate a peripheral moving object. To investigate this more natural situation, we presented a peripheral target that was either stationary or moving horizontally, above or below the fixation spot. Participants were asked to saccade to the target and to keep it foveated. The letters were presented at different times relative to the first target directed saccade. As would be expected from retinal masking and motion blur during saccades, the discrimination performance increased with increasing post-saccadic delay. If the target moved and the saccade was followed by pursuit, letter recognition performance was on average 16% lower than if the target was stationary and the saccade was followed by fixation.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Primates with a foveal region of high-spatial acuity on the retina need to use saccadic and smooth pursuit eye movements to bring and to keep an object of interest in that region. This increases the spatial resolution and reduces the retinal image motion in the case of moving objects. A major question is what happens to object recognition during the execution of eye movements. Here we study the effect of smooth pursuit and foveating saccades on the recognition of letters presented at different contrasts for different durations.

Despite the large variability of the retinal image due to illumination and viewing perspective, the human visual system recognizes objects impressively reliably and rapidly. Indeed, it has been shown that humans are able to process complex images and to recognize familiar objects very rapidly (Guyonneau, Kirchner, & Thorpe, 2006; Kirchner & Thorpe, 2006; Thorpe, Fize, & Marlot, 1996), so that 50–80 ms of cortical processing time are sufficient. Even though the categorization of large objects is still possible even in the far periphery of the visual field (Thorpe, Gegenfurtner, Fabre-Thorpe, & Bulthoff, 2001), overall performance deteriorates rapidly due to the low spatial resolution in the retinal periphery.

Therefore, in natural situations saccadic eye movements are used to bring objects of interest to the fovea and smooth pursuit eye movements are used to stabilize moving objects of interest on the fovea. Thereby spatial resolution is maximized and retinal smear minimized. Both effects obviously are beneficial for object recognition. Beside these advantages, there might be some disadvantages caused by these eye movements, which become only obvious when target images are presented at low contrasts or short durations. At present it is still unclear how efficiently and rapidly objects can be recognized during or after foveating eye movements.

High visual acuity is an important prerequisite for object recognition. Several studies have compared visual acuity during fixation and smooth pursuit. Ludvigh and Miller (1958) found in humans that dynamic visual acuity measured with Landolt rings at angular velocities ranging from 10 to 170 deg/s almost matched static visual acuity at low velocities, but declined rapidly at higher velocities. Later Methling and Wernicke (1968) concluded that retinal image movements caused by the inaccuracy of eye movement control resulted in the decrease of dynamic acuity. Brown (1972a, 1972b, 1972c) studied the effect of stimulus contrast, size and position on human dynamic acuity. His studies confirmed that dynamic visual acuity during pursuit depends solely on the retinal stabilization and is only limited by the accuracy of the eye movement which is improved by increasing the stimulus contrast (Brown, 1972a, 1972b, 1972c; Ludvigh & Miller, 1958; Methling & Wernicke, 1968).

Besides visual acuity, object recognition depends on the spatial contrast sensitivity (Chung, Legge, & Tjan, 2002), which has been shown to be influenced by eye movements. During saccades the luminance contrast sensitivity for low-spatial frequency stimuli

---

* Corresponding author. Fax: +49 641 99 26 119.
E-mail address: alexander.c.schuetz@psychol.uni-giessen.de (A.C. Schütz).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.05.022

is strongly suppressed, while the sensitivity for high-spatial frequencies and color does not seem to be affected (Burr, Morrone, & Ross, 1994). For smooth pursuit it is known that temporal contrast sensitivity depends in principle on the retinal stimulus motion and not on the physical stimulus motion (Flipse, van der Wildt, Rodenburg, Keemink, & Knol, 1988; Liu & Jiang, 1984; Murphy, 1978; Schütz, Delipetkos, Braun, Kerzel, & Gegenfurtner, 2007b). Moreover, we showed that the contrast sensitivity for high-spatial frequencies and for color is improved, probably due to an enhanced sensitivity of the parvocellular pathway (Schütz, Braun, Kerzel, & Gegenfurtner, 2008). Since color and high-spatial frequencies are important sources of information for identification of objects, this enhanced contrast sensitivity might facilitate object recognition. Interestingly this enhancement of chromatic sensitivity occurs also during optokinetic nystagmus, but not during visually-enhanced vestibulo-ocular reflex (Schütz, Braun, & Gegenfurtner, 2009).

Another important aspect influencing object recognition during voluntary eye movements is the coupling of spatial attention and eye movements. A large number of studies investigated perceptual performance right before a saccade (Castet, Jeanjean, Montagnini, Laugier, & Masson, 2006; Deubel, 2008; Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler, Anderson, Dosher, & Blaser, 1995; McPeek, Maljkovic, & Nakayama, 1999). In principle, all of these studies showed that performance at the future target location of the saccade is best. This indicates that a shift of spatial attention to the saccade goal precedes the saccadic eye movement. However, it is still unclear what happens to performance after the execution of saccades. So far performance during complex visual tasks and the behavior of spatial attention has not been measured for the time period after saccades. Also during pursuit, similar to saccades, spatial attention seems to be concentrated on the moving target, which leads to a performance decrement at peripheral locations (Kerzel & Ziegler, 2005; Khurana & Kowler, 1987; Schütz et al., 2007b). Since little is currently known about the recognition of briefly presented complex objects during and after the execution of saccadic and smooth pursuit eye movements we investigated how high-level object recognition is influenced by smooth pursuit and foveating saccades.

2. Methods

2.1. Design

We used a 20-AFC paradigm to measure the recognition performance for letters at different contrasts and presentation durations. 20 different letters appeared within a masking noise patch, which either moved horizontally (pursuit) or was stationary (fixation). We tested performance during steady-state pursuit (Experiments 1 and 2), during pursuit initiation (Experiment 3) and after a foveating saccade to the target noise patch (Experiment 4).

2.2. Participants

Six participants participated in this study: one of them was the author ACS, while the other five participants were female undergraduate students from the Justus-Liebig-University, who were naïve to the purpose of the experiments and were paid for their participation. All experiments were completed by the author and four naïve participants.

2.3. Equipment

Participants were seated in a dimly lit room facing a 21-inch CRT monitor (ELO Touchsystems, Fremont, CA, USA) driven by an Nvida Quadro NVS 285 graphics board with a refresh rate of 100 Hz non-interlaced. At a viewing distance of 47 cm, the active screen area subtended 45° in the horizontal direction, and 36° vertical on the participant's retina. With a spatial resolution of 1280 × 1024 pixels this results in 28 pixels/deg. The participant's head was kept fixed in place using a chin rest.

2.4. Eye movement recording

Eye position signals were recorded with a head-mounted, video-based eye tracker (EyeLink II; SR Research Ltd., Osgoode, Ontario, Canada) and were sampled at 250 Hz. Participants viewed the display binocularly. Stimulus display and data collection were controlled by a PC.

2.5. Visual stimuli

On a gray background a circular noise patch was used as the eye movement target. This patch had a diameter of 2° and was defined by a one octave wide, band pass filtered noise with a central frequency of three cycles per letter (Fig. 1). A previous study on letter identification showed that such a frequency masks letters optimally (Solomon & Pelli, 1994). The noise contrast amounted to 50% centered on the gray background. To measure the recognition performance for letters, we used 20 different characters of the font style Bookman Old Style. All uppercase letters but B, I, O, Q, S, Z (Sperling, Budiansky, Spivak, & Johnson, 1971) were used. The presented letter was centered within the patch, and was defined by a luminance increase. To collect the participant's decision, we presented a 4 × 5 array of all possible letters. Participants were asked to select the letter which was presented in the trial by gaze position. Then the gaze selected letter was highlighted by the computer and the participants were asked to confirm their selection by pressing an assigned key. Participants received an acoustic feedback if their response was incorrect. All participants performed one training session with high-contrast letters, to make sure that they were able to use the gaze input as response with high reliability.

2.6. Experiment 1: steady-state pursuit and fixation

Fig. 2 shows the time course of a pursuit and a fixation trial. At the beginning of each trial a black bull's-eye with an outer radius of 0.3° and an inner radius of 0.075° appeared at the screen center. The participants had to fixate the bull's-eye and press an assigned button to start the trial. With pressing the button, the EyeLink II System performed a drift correction to correct errors of headband slippage or other factors. If the drift correction succeeded, the bull's-eye was replaced by the circular noise patch. 500 ms after the drift correction the noise patch started moving with a velocity of 10.57 deg/s for 1500 ms in pursuit trials; in fixation trials the noise patch remained stationary for 2000 ms. 1250 ms after trial onset one of the randomly selected letters appeared within the center of the noise patch. Participants selected the recognized letter at the end of the trial.

We tested the recognition of single letters at four different contrast levels (5%, 10%, 20% and 40%) and four different presentation durations (10, 20, 30 and 50 ms). All contrast levels and presentation durations were combined, resulting in a 4 × 4 matrix of 16 conditions. All conditions were presented interleaved. Each participant completed at least six sessions of 160 trials.

2.7. Experiment 2: steady-state pursuit with additional masking and fixation

In order to measure even longer presentation durations, we performed an additional version of Experiment 1, in which we applied

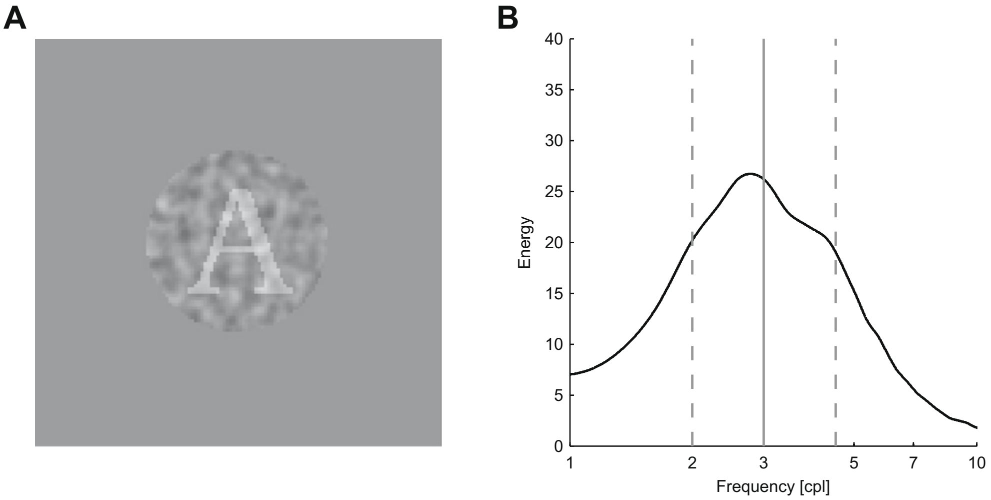

Fig. 1. Exemplary noise patch with target letter A (A) and the frequency distribution of the noise (B). The dashed-gray lines indicate the frequency pass band and the solid-gray line indicates the noise frequency which optimally masks letter recognition (Solomon & Pelli, 1994).

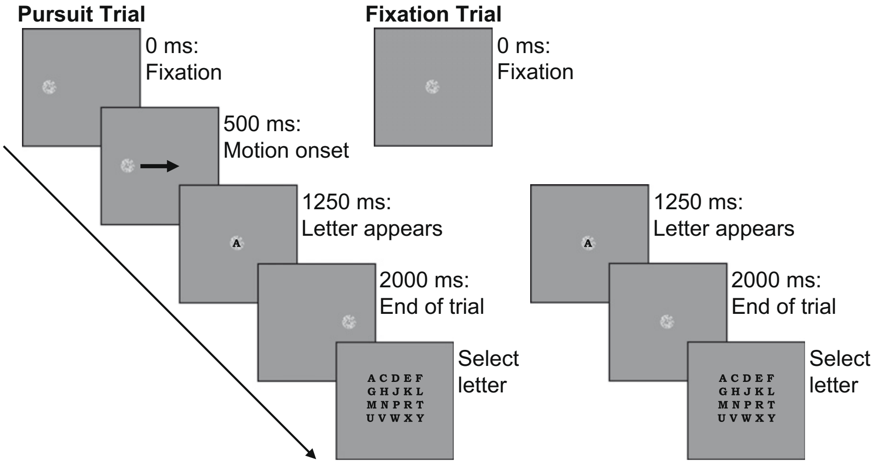

Fig. 2. Schematic diagram of a pursuit and a fixation trial in Experiment 1.

a letter mask after the target letter presentation. The mask consisted of two high-contrast letters, which were randomly selected from all available letters and randomly rotated by 45° steps. We used the same contrast levels as in Experiment 1, but the presentation durations were now 50, 100, 200 and 400 ms.

2.8. Experiment 3: pursuit initiation and fixation after saccades

In this experiment participants were asked to execute a saccadic eye movement to the noise patch which appeared 7.5° above or below the initial fixation target (Fig. 3). In pursuit trials the patch appeared above or below but displaced to the left or right of the horizontal center and started moving towards the screen center 250 ms after the drift correction. At a randomized delay between 400 and 600 ms, the central fixation spot disappeared and the participants were instructed to saccade to the moving noise patch and to pursue it. After the detection of the saccade onset, a single letter was flashed after one of eight different onsets delays (0, 50, 100, 150, 200, 250, 300 and 350 ms). In fixation trials the noise patch appeared horizontally centered and remained stationary throughout the trial. After the offset of the fixation spot participants were asked to saccade to the stationary noise patch, in which a letter appeared (see above). As we were more interested in the time course of the recognition rate during pursuit inition and fixation after a target directed saccade, all letters were presented for only 10 ms at a contrast level of 30%. Each participant completed at least eight sessions of 160 trials.

2.9. Experiment 4: saccades and fixation

In this experiment the noise patch appeared 7.5° above or below the screen center. In fixation trials the initial fixation target appeared at the same place as the following noise patch. In saccade trials, the initial fixation target appeared at the screen center. At a randomized delay between 400 and 600 ms, the fixation target disappeared and the participants were instructed to saccade to the noise patch above or below. After the detected saccade onset, a single letter appeared after one of eight different onsets delays (0, 50, 100, 150, 200, 250, 300 and 350 ms). Like in Experiment 3, the letter was presented for 10 ms at a contrast of 30%. Fig. 4 shows the time course of a saccade and a fixation trial.

2.10. Eye movement analysis

Saccades were detected online using a velocity criterion of 20 deg/s. For the offline analysis, eye position signals were filtered

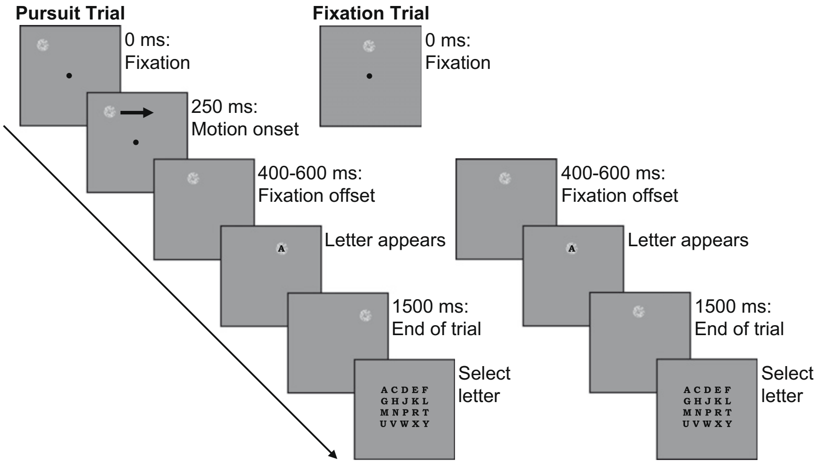

Fig. 3. Schematic diagram of a pursuit and a fixation trial in Experiment 3.

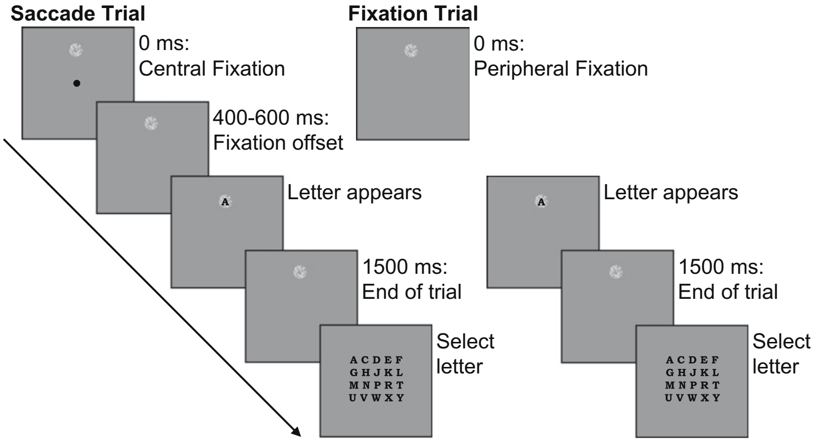

Fig. 4. Schematic diagram of a pursuit and a fixation trial in Experiment 4.

by a second-order Butterworth filter with a cut-off frequency of 30 Hz. The eye velocity signals were obtained by digital differentiation of eye position signals over time and afterwards filtered by a Butterworth filter with a cut-off frequency of 20 Hz. Saccades were detected offline by a cut-off (75.000 deg/s^3) on the third derivative of eye position (Wyatt, 1998). Retinal errors were analysed in a 60 ms (Experiments 1, 3 and 4) or a 200 ms (Experiment 2) interval starting at the onset of the letter. The retinal position error was calculated as the average Euclidian distance between eye and target position. The retinal velocity error was calculated as the average Euclidian difference between eye and target velocity during the presentation time of the letter. Position and velocity jitter were calculated as the standard deviation of the Euclidean differences. The pursuit gain was analysed during the presentation time of the letter. In Experiments 1 and 2, trials were discarded if the pursuit gain was below 0.7 or if a saccade occurred in a critical interval of 100 ms before and 100 ms after the letter presentation (4%). In Experiments 3 and 4, trials were discarded if the saccade was not detected correctly online. For onset delays larger than 100 ms, we discarded trials if a saccade occurred in a 200 ms interval centered at the time of letter presentation.

2.11. Psychophysical data analysis

To estimate the object recognition performance, we calculated the recognition rate for each combination of letter contrast and presentation duration. Theses values were arcsine transformed ($\arcsin(\sqrt{\text{proportion correct}})$), before statistical analysis was performed. Based on the recognition rates, we fitted psychometric functions (cumulative Gaussians) for each constant value of contrast or presentation duration. We used the psignifit matlab toolbox for the calculation of the psychometric functions (Wichmann & Hill, 2001a). Thresholds were defined as the value of the independent variable at which proportion correct was 50%. We used a bootstrap procedure to determine the 95% confidence interval of the thresholds (Wichmann & Hill, 2001b).

3. Results

3.1. Experiment 1: steady-state pursuit

In the first experiment, we measured object recognition performance during ongoing pursuit and during fixation at four different contrast levels and durations. As expected for pursuit as well as fixation, the recognition rate improved with increasing contrast (Fig. 5) and with increasing presentation duration (Fig. 6). The selected contrasts and presentation durations covered the whole range of performance from chance or 0% to 100%. Depending on the contrast and eye movement condition the recognition rate increased from around 0–30% at a contrast level of 5% to around 90–100% at a contrast level of 40%. For the measured range of presentation durations from 10 to 50 ms recognition rates increased with

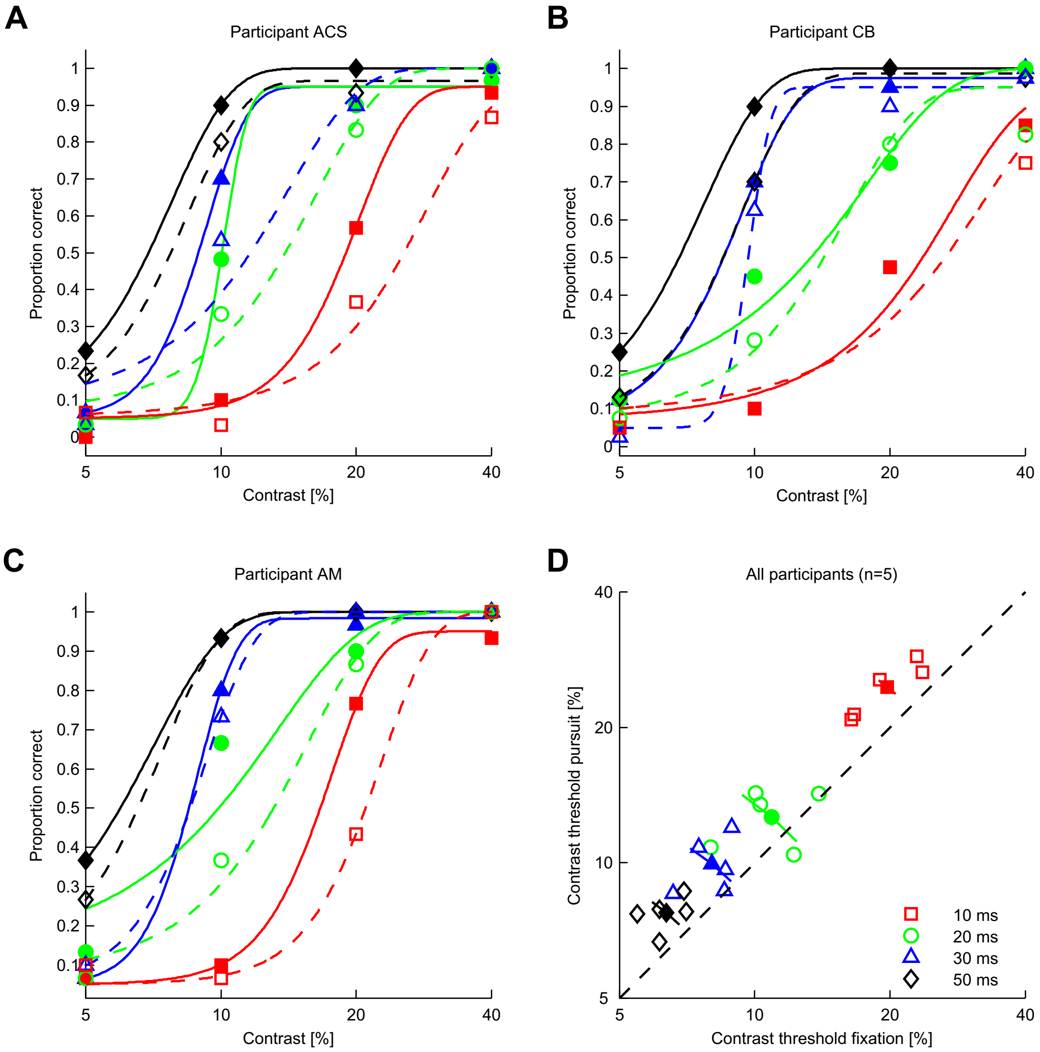

Fig. 5. Experiment 1: recognition rate over contrast. (A–C) Data of three exemplary participants. Solid lines and filled symbols indicate data for fixation and dashed lines and open symbols data for pursuit. The different colours and symbols denote different presentation durations. (D) Contrast thresholds during pursuit and fixation for all participants. The filled symbols indicate the average thresholds across participants. The diagonal error bars denote the 95% within-participants confidence interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

longer presentation durations. Astonishingly, for some contrast and eye movement conditions, recognition rate almost reached 100% in the 10 ms condition. This indicates that only a very brief presentation of the letter may be sufficient to recognize its shape properly. On the other hand recognition performance did not reach threshold performance in some conditions of the 50 ms presentation duration, suggesting a trade-off between contrast and duration. Beside the global increase of recognition performance with an increase of contrast or presentation duration, there seemed to be a difference between the eye movement conditions: in most conditions letter recognition rates were higher during fixation than during pursuit. However this difference was not very pronounced. For example for a fixed contrast of 10% for instance, the presentation duration during pursuit had to be increased by 16 ms to reach the same performance as for fixation; or for a fixed presentation duration of 20 ms, the contrast had to be increased by 8% during pursuit. Across conditions and participants the average difference in detection rate amounted to 11%.

Next we sought to test statistically the increase of recognition rate with contrast or presentation duration and the difference between the eye movement conditions. To do so, we calculated a three-way repeated-measures ANOVA with the factors: eye movement conditions, contrast and duration. We found a significant main effect for contrast ($F(3,12) = 667.428$, $MSE = 0.103$, $P < 0.001$), as well as for duration ($F(3,12) = 135.346$, $MSE = 0.106$, $P < 0.001$). This confirms the known fact that higher contrasts as well as longer presentation durations significantly facilitated object recognition. However, the main effect for eye movement condition was also significant ($F(1,4) = 25.216$, $MSE = 0.016$, $P = 0.007$). Hence the subtle difference between fixation and pursuit seem to be reliable. We also found a significant two-way-interaction for contrast and duration ($F(9,36) = 16.060$, $MSE = 0.153$, $P = 0.001$). The other two-way interactions of contrast and duration with eye movement condition and the three-way interaction were not significant, which indicates that pursuit caused a general performance deficit and not a specific influence on contrast sensitivity or temporal integration.

Another way to test for differences between fixation and pursuit is to fit psychometric functions to the data and to compare the estimated thresholds. We could obtain valid thresholds for contrast levels of 10% and 20% and for all duration levels. We found a significant difference for a contrast level of 10% ($t(4) = 8.07$; $P = 0.001$) but not for a contrast level of 20% ($t(4) = 1.97$; $P = 0.120$). The analysis of the durations showed significant differences for a duration

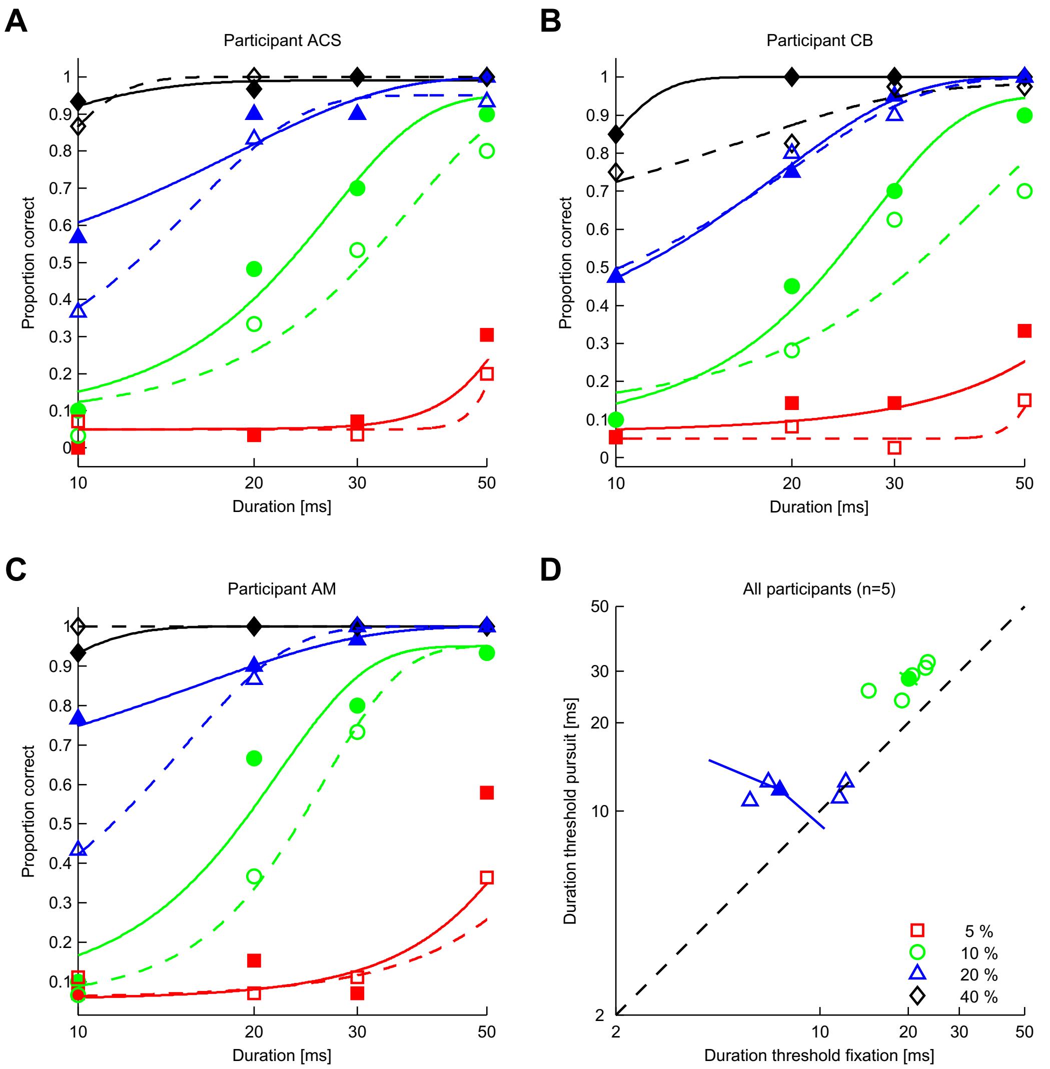

Fig. 6. Experiment 1: recognition rate over presentation duration. (A–C) Data of three exemplary participants. Solid lines and filled symbols indicate data for fixation and dashed lines and open symbols data for pursuit. The different colours and symbols denote different contrasts. (D) Duration thresholds during pursuit and fixation for all participants. The filled symbols indicate the average thresholds across participants. The diagonal error bars denote the 95% within-participants confidence interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of 10 ms ($t(4) = 7.71$; $P = 0.002$), 30 ms ($t(4) = 3.16$; $P = 0.034$) and 50 ms ($t(4) = 4.33$; $P = 0.012$) but not for a duration of 20 ms ($t(4) = 1.61$; $P = 0.183$). In sum, we found significant differences across participants for four out of six conditions, which could be fitted with a psychometric function.

Imperfect tracking might be one reason for the performance decrement during pursuit. To test this hypothesis, we compared the position and velocity error during fixation and pursuit. First we calculated cumulative probability curves for fixation and pursuit separately. As expected, the retinal errors were higher during pursuit than during fixation. The average retinal position error was 0.38° (SD 0.26) for fixation and 0.60° (SD 0.32) for pursuit. The average retinal velocity error amounted to 1.23 deg/s (SD 0.73) for fixation and 1.98 deg/s (SD 0.94) for pursuit. The position jitter was 0.02° (SD 0.01) for fixation and 0.14° (SD 0.06) for pursuit; while the velocity jitter reached 0.52 deg/s (SD 0.36) for fixation and 0.80 deg/s (SD 0.44) for pursuit.

To test if the distribution of retinal errors during fixation and pursuit was statistically significant different, we performed an ROC analysis (Fig. 7) and calculated the area under the ROC curve and its standard error (Hanley & McNeil, 1982). The position error ($A_{ROC} = 0.727$; $A_{ROC} > 0.5$: $P < 0.001$), the velocity error ($A_{ROC} = 0.779$; $A_{ROC} > 0.5$: $P < 0.001$), the position jitter ($A_{ROC} = 0.992$; $A_{ROC} > 0.5$: $P < 0.001$) as well as the velocity jitter ($A_{ROC} = 0.729$; $A_{ROC} > 0.5$: $P < 0.001$) were significantly higher during pursuit than during fixation. However, this does not necessarily mean that the differences in retinal errors caused the differences of letter recognition rates between fixation and pursuit. To determine the influence of retinal errors on perceptual performance, we calculated the retinal error distributions separately for hits and misses. Then we performed the same ROC analysis like before, but tested for differences in the distributions for hit trials and miss trials. For fixation as well as for pursuit there was no significant difference between position error, velocity error, position jitter and velocity jitter in hit trials and miss trials. This indicates that retinal errors did not have a direct influence on the letter recognition performance and therefore were probably not the source of the performance differences between fixation and pursuit.

3.2. Experiment 2: steady-state pursuit with additional masking

The applied letter mask increased effectively the contrast level (Fig. 8) and presentation duration (Fig. 9) necessary for proper recognition: for instance, the average recognition rate for a 10% con-

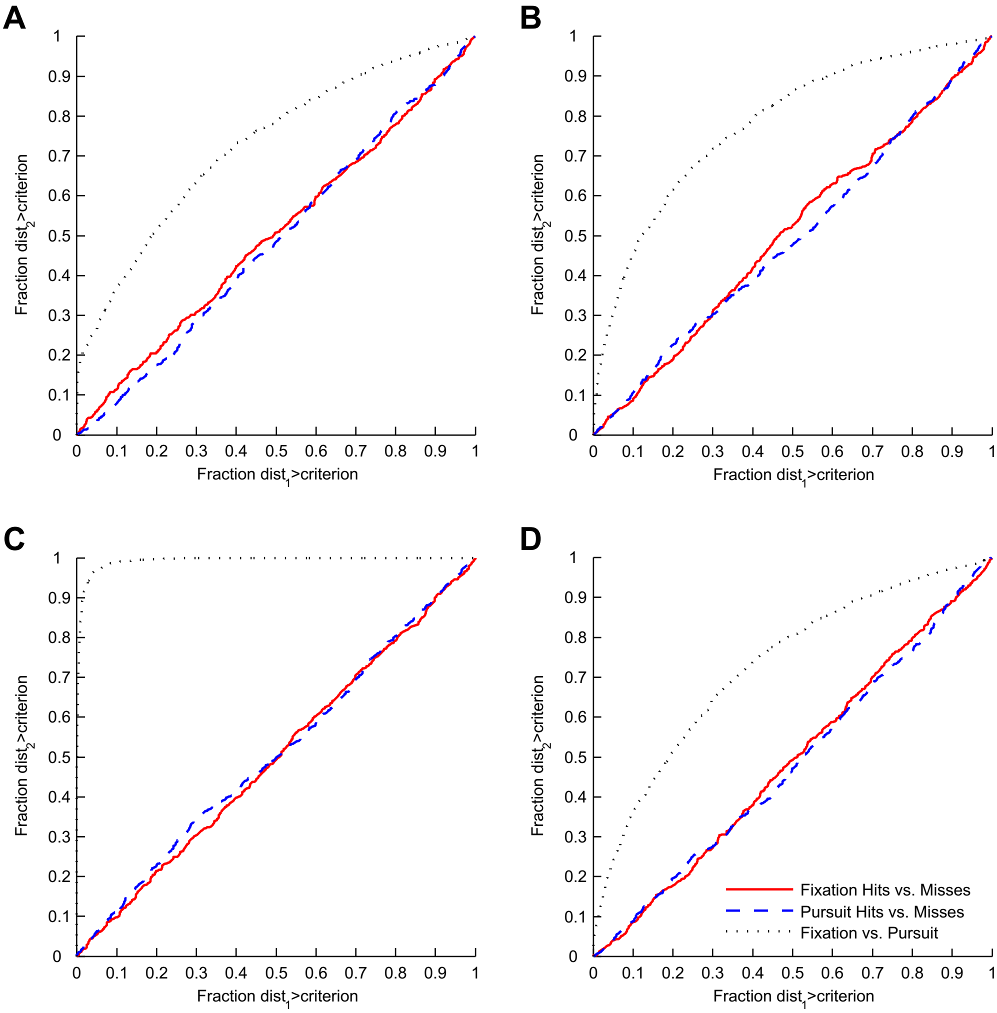

Fig. 7. Experiment 1: ROC curves for position error (A), velocity error (B), position jitter (C) and velocity jitter (D). The black dotted line indicates the difference between fixation and pursuit; the red solid line indicates the difference between hit and miss trials for fixation; the blue dashed line indicates the difference between hit and miss trials for pursuit. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

trast letter, presented for 50 ms, felt from 86% in Experiment 1 to chance level in Experiment 2. For a 20% contrast letter, presented also for 50 ms, the average recognition rate felt from 99% in Experiment 1 to 32% in Experiment 2.

Although the necessary contrasts and presentation durations were much larger in Experiment 2, we found similar differences in recognition performance between fixation and pursuit. For instance for a fixed contrast of 10% the presentation duration had to be increased by 40 ms during pursuit, to reach the same recognition rate. For a fixed duration of 100 ms, the contrast had to be increased by 11%. Across conditions and participants the average difference in detection rate amounted to 10%.

To test these effects statistically, we calculated a three-way repeated-measures ANOVA with the factors: eye movement conditions, contrast and duration. We found a significant main effect for contrast ($F(3,12) = 904.738$, MSE = 0.093, $P < 0.001$), as well as for duration ($F(3,12) = 286.343$, MSE = 0.085, $P < 0.001$). The main effect for eye movement condition was also significant ($F(1,4) = 16.410$, MSE = 0.016, $P = 0.015$). Also a significant two-way interaction between contrast and duration ($F(9,36) = 24.931$, MSE = 0.253, $P < 0.001$) as well as between contrast and eye movement condition ($F(3,12) = 6.572$, MSE = 0.019, $P < 0.036$) and duration and eye movement condition ($F(3,12) = 9.001$, MSE = 0.009, $P < 0.012$) was present. The three-way interaction was not significant.

Like in Experiment 1, we fitted psychometric functions to the data and to obtain contrast and duration thresholds. Valid psychometric functions could be fitted for the contrast level of 10% and for presentation durations of 50, 100, 200 and 400 ms. A significant difference was present for a contrast level of 10% ($t(4) = 10.06$; $P = 0.001$). The analysis of the durations showed significant differences for a duration of 100 ms ($t(4) = 10.68$; $P < 0.001$), 200 ms ($t(4) = 3.70$; $P = 0.021$) and 400 ms ($t(4) = 5.06$; $P = 0.007$) but not for a duration of 50 ms ($t(4) = 0.98$; $P = 0.381$). Hence we obtained significant differences for four out of five conditions. This shows in principle that the differences between fixation and pursuit are also present for longer presentation durations and for masking conditions.

Similar to Experiment 1, retinal errors were higher during pursuit than during fixation: the average retinal position error was 0.38° (SD 0.25) for fixation and 1.22° (SD 0.54) for pursuit. The average retinal velocity error amounted to 1.36 deg/s (SD 0.68) for fixation and 1.87 deg/s (SD 0.76) for pursuit. The position jitter was 0.03° (SD 0.02) for fixation and 0.59° (SD 0.29) for pursuit; while the velocity jitter reached 0.74 deg/s (SD 0.43) for fixation and 0.96 deg/s (SD 0.47) for pursuit. The ROC analysis obtained significant differences for the position error ($A_{ROC} = 0.965$; $A_{ROC} > 0.5$: $P < 0.001$), the velocity error ($A_{ROC} = 0.748$; $A_{ROC} > 0.5$: $P < 0.001$), the position jitter ($A_{ROC} = 0.999$; $A_{ROC} > 0.5$: $P < 0.001$) as well as the velocity jitter ($A_{ROC} = 0.688$; $A_{ROC} > 0.5$: $P < 0.001$). There were

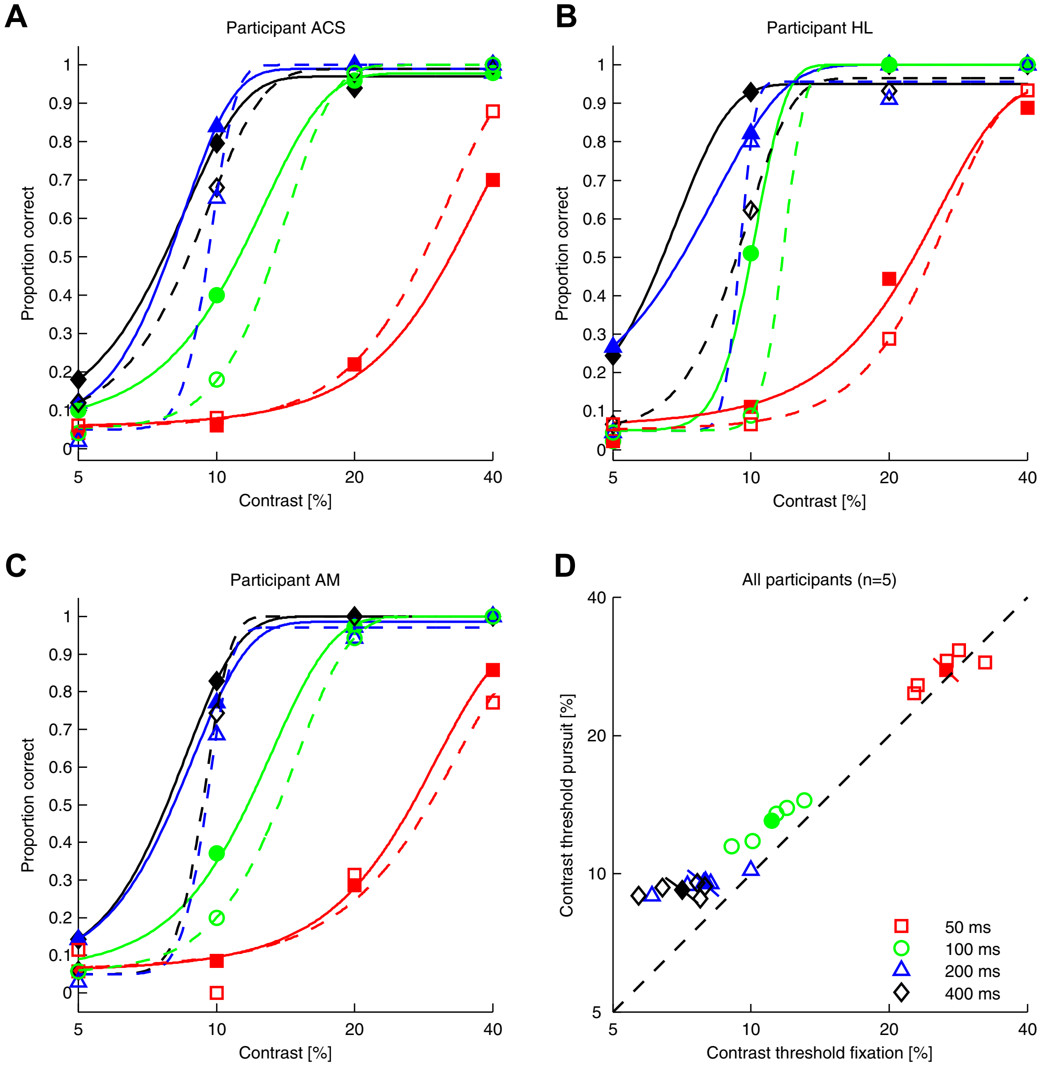

Fig. 8. Experiment 2: recognition rate over contrast. Conventions are the same as in Fig. 5. Please note that the presentation durations are different from Experiment 1.

no significant differences of retinal position error, velocity error, position jitter or velocity jitter between hit and miss trials during fixation. During pursuit, retinal position error ($A_{ROC} = 0.579$; $A_{ROC} > 0.5$: $P < 0.001$) and position jitter ($A_{ROC} = 0.585$; $A_{ROC} > 0.5$: $P < 0.001$) were significantly smaller in miss trials than in hit trials. However, the difference was on average rather small in magnitude for position error (Hits: 1.29, Misses: 1.14) as well as for position jitter (Hits: 0.63, Misses: 0.54) compared to the differences between fixation and pursuit. Also the differences were in the opposite direction as the differences between fixation and pursuit.

3.3. Experiment 3: pursuit initiation

The aim of the third experiment was to test the more natural condition that an object in the periphery appears and a saccade to the object is executed in order to recognize the object foveally. Therefore, we tested letter recognition performance right after a foveating saccade. The saccade could be made either to a stationary noise patch (saccade followed by fixation) or to a moving noise patch (saccade followed by pursuit) appearing in the periphery. In the noise patch we presented a single letter with a fixed contrast of 30% for 10 ms at different points in time after detection of the first target directed saccade. The online detection of the saccade lagged the true onset of the saccade and occurred around the time of the peak velocity of the saccade. Fig. 10 shows the recognition rate as a function of the onset asynchrony (OA) between saccade and letter. Most obvious, the recognition rate increased with increasing delay to the saccade. At the time of peak velocity of the saccade, recognition performance was at chance and increased up to an asymptotic performance level of about 80% for larger onset delays. The asymptotic performance level was reached not earlier than 150–200 ms after the saccade. The fastest increase of recognition rate occurred during the deceleration of the saccade and a more gradual increase happened after the end of the saccade. Like in the first experiment, performance seemed to be inferior during pursuit compared to fixation. On average the difference in recognition performance between pursuit and fixation amounted to 16%. We calculated a two-way repeated-measures ANOVA with the factors OA and eye movement condition on the recognition rates. The main effect of OA ($F(7,28) = 231.368$, $\text{MSE} = 0.052$, $P < 0.001$) as well as the main effect of eye movement condition were significant ($F(1,4) = 62.996$, $\text{MSE} = 0.004$, $P = 0.001$). Thus, the rise of recognition rate with increasing delay to the saccade and the performance difference between fixation and pursuit were significant. The interaction between OA and eye movement condition was also significant ($F(7,28) = 7.417$, $\text{MSE} = 0.003$, $P = 0.005$).

Like in Experiment 1, we wanted to analyze the influence of retinal errors on the recognition performance. As the initial saccades

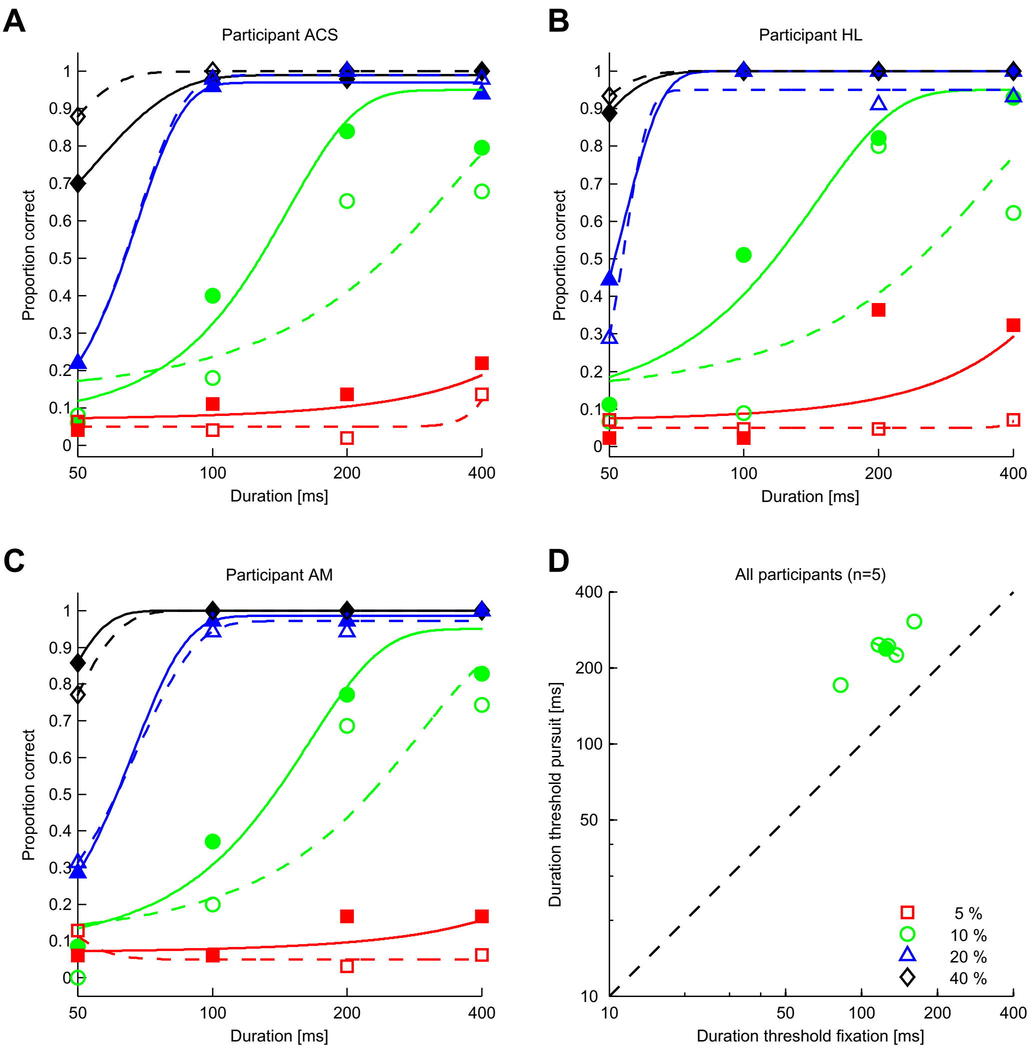

Fig. 9. Experiment 2: recognition rate over duration. Conventions are the same as in Fig. 6. Please note that the presentation durations are different from Experiment 1.

introduced large retinal errors, which we did not want to analyze, we only included trials with a OA larger than 100 ms in the following analysis. In general, the average retinal errors were comparable to the retinal errors measured in Experiments 1 and 2. In Experiment 3 the average retinal position error was 0.68° (SD 0.48) for fixation and 0.72 (SD 0.40) for pursuit; the average retinal velocity error reached 1.51 deg/s (SD 0.91) for fixation and 2.21 deg/s (SD 1.15) for pursuit. The position jitter was 0.02° (SD 0.02) for fixation and 0.12° (SD 0.06) for pursuit, while the velocity jitter was 0.63 deg/s (SD 0.43) for fixation and 0.89 deg/s (SD 0.52) for pursuit. Next we tested whether the retinal errors were different in pursuit and fixation conditions. Like in Experiment 1, the position error ($A_{ROC} = 0.545$; $A_{ROC} > 0.5$: $P < 0.001$), the velocity error ($A_{ROC} = 0.716$; $A_{ROC} > 0.5$: $P < 0.001$), the position jitter ($A_{ROC} = 0.974$; $A_{ROC} > 0.5$: $P < 0.001$) as well as the velocity jitter ($A_{ROC} = 0.683$; $A_{ROC} > 0.5$: $P < 0.001$) were significantly higher during pursuit than during fixation. Again we did not find significant differences of retinal errors in the hit trials and miss trials. Thus retinal errors were probably not the reason for the detrimental performance during pursuit.

To summarize, we found that the recognition of objects appearing in the periphery increased significantly with increasing OA after the foveating saccades and that the recognition rate was reduced when the saccade was followed by pursuit. Therefore, we replicated in Experiment 3 the finding of Experiments 1 and 2, that object recognition performance was significantly lower during pursuit than during fixation.

3.4. Experiment 4: saccades

The fourth experiment investigated the time course of object recognition performance after a foveating saccade compared to recognition performance during continuous fixation. Fig. 11 shows the recognition rate as a function of onset asynchrony between saccade and letter. In the fixation condition, the recognition rate was rather constant over time, at around 80%. In the saccade condition, recognition rate increased with increasing OA to the saccade. At the time of the peak velocity of the saccade, letter recognition performance was at chance level, however 150–200 ms later, recognition performance reached the level of the fixation condition. Like in Experiment 3, the steepest rise of recognition rate occurred during the deceleration of the target directed saccade, followed by a gradual increase after the end of the saccade. We calculated a two-way repeated-measures ANOVA with the factors OA and eye movement condition on the recognition rates. The main effect of OA ($F(7,28) = 16.562$, $\mathrm{MSE} = 0.051$, $P < 0.001$) as well as the main effect of eye movement condition were significant ($F(1,4) = 365.498$, $\mathrm{MSE} = 0.002$, $P < 0.001$). Thus, the increase of recognition rate with increasing delay to the saccade and the performance difference between fixation and saccade were significant. The interaction be-

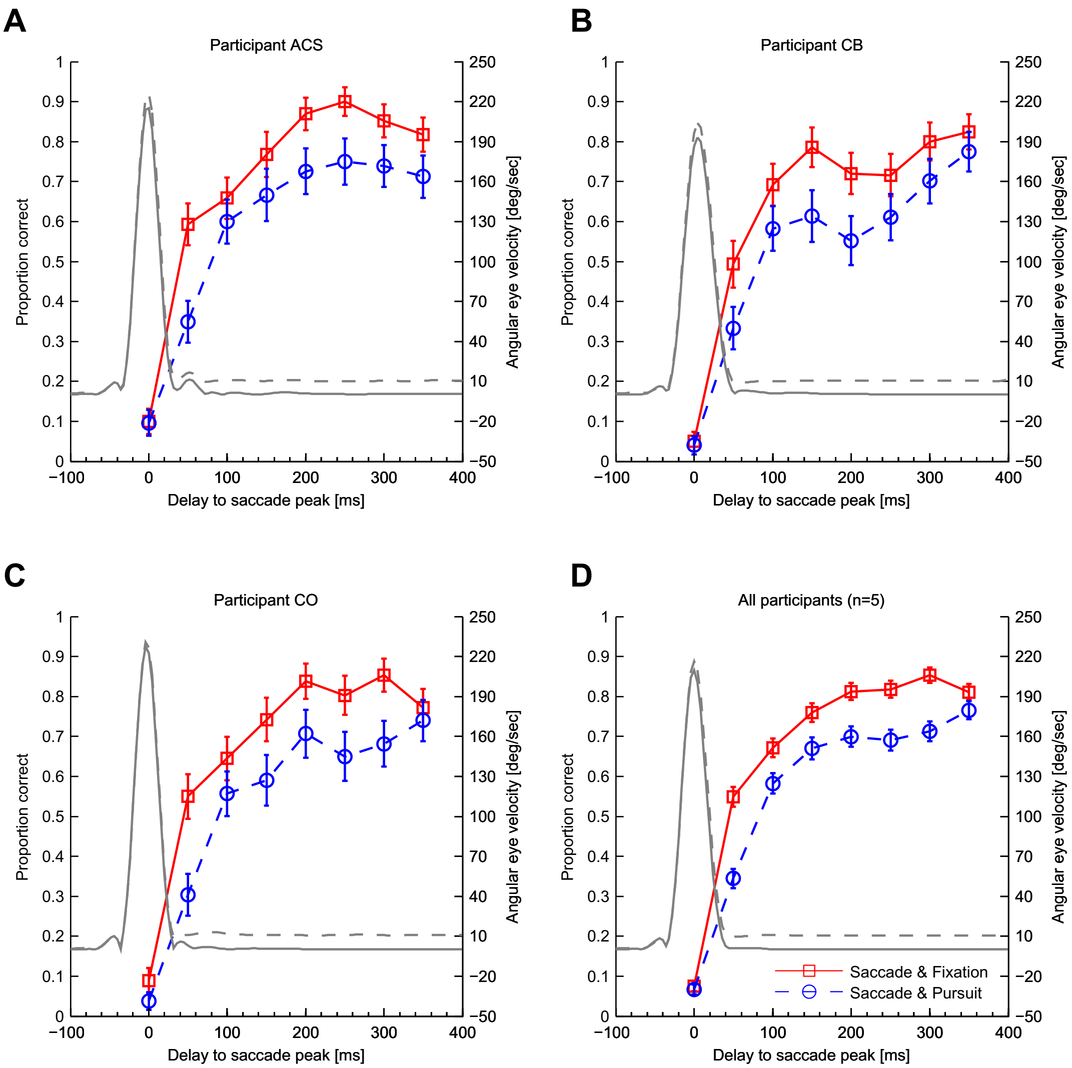

Fig. 10. Experiment 3: recognition rate during pursuit and during fixation after the first saccade to the noise patch. (A–C) Data for three exemplary participants. (D) Mean across all participants. Red solid lines and squares indicate data for fixation and blue dashed lines and circles data for pursuit. The corresponding eye velocity is plotted in grey. Error bars denote the standard error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tween OA and eye movement condition was also significant ($F(7,28) = 29.128$, MSE = 0.010, $P < 0.001$).

To determine the influence of retinal errors, we performed the same analysis like for Experiment 3. The average retinal position error was 0.35° (SD 0.25) for fixation and 0.59° (SD 0.41) post-saccadic; while the average retinal velocity error reached 1.28 deg/s (SD 0.91) for fixation and 1.40 deg/s (SD 0.85) post-saccadic. The position jitter was 0.02° (SD 0.01) for fixation and 0.02° (SD 0.01) post-saccadic and the velocity jitter was 0.54 deg/s (SD 0.40) for fixation and 0.58 deg/s (SD 0.43) post-saccadic. Next we tested whether the retinal errors were different in saccade and fixation conditions. All retinal errors were post-saccadic significantly different than during fixation. The position error ($A_{ROC} = 0.716$; $A_{ROC} > 0.5$: $P < 0.001$), the velocity error ($A_{ROC} = 0.564$; $A_{ROC} > 0.5$: $P < 0.001$), the position jitter ($A_{ROC} = 0.550$; $A_{ROC} > 0.5$: $P < 0.001$) as well as the velocity jitter ($A_{ROC} = 0.538$; $A_{ROC} > 0.5$: $P < 0.001$) were significantly different in the eye movement conditions. Again we did not find significant differences of retinal errors in the hit trials and miss trials.

4. Discussion

The aim of our study was to investigate how the recognition of complex shapes is influenced by voluntary eye movements. The first and second experiment compared letter recognition performance for different contrasts and presentation durations during steady-state pursuit and during fixation. We found that object recognition performance was significantly reduced during steady-state smooth pursuit eye movements. Across observers and experimental conditions, the recognition performance was on average 11% lower during smooth pursuit compared to fixation. Given that a secondary task, the pursuit eye movement had to be carried out and that the foveal stabilization was not perfect during pursuit, this reduction seemed to be quite small. The third experiment was designed to investigate object recognition performance during the initiation of smooth pursuit eye movements and during saccades. Recognition performance increased for about 150–200 ms after the target directed saccade. Besides this effect of the saccade, there was again a significant reduction of recognition rate during pursuit, which replicates the findings of Experiments 1 and 2. This attenuation was present for all presentation times, therefore the initiation of smooth pursuit seems not to add additional interference for object recognition compared to steady-state pursuit. The fourth experiment measured recognition performance after foveating saccades in comparison to continuous fixation. Again, like in Experiment 3 we found a long rise of recognition performance after the foveating saccade.

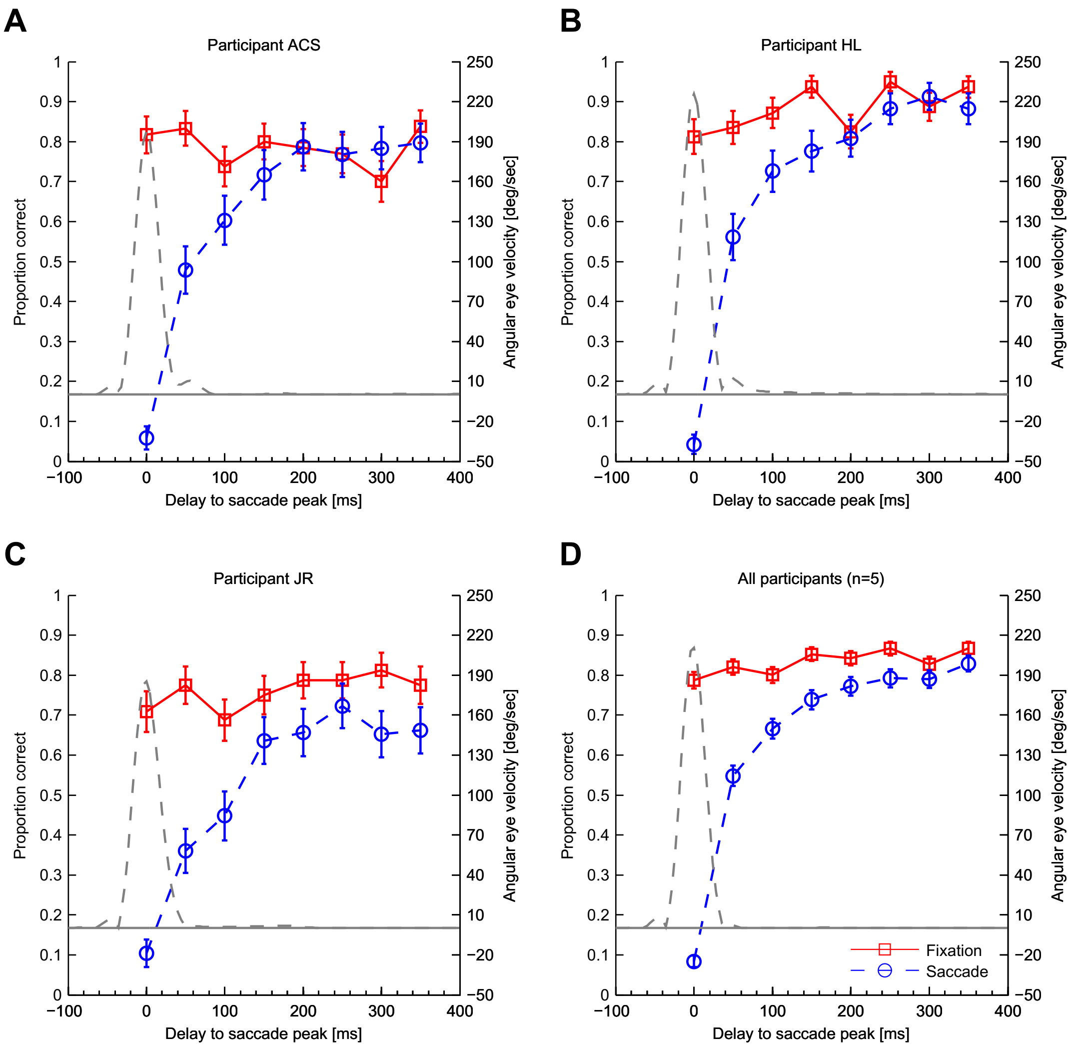

Fig. 11. Experiment 4: recognition rate during continuous fixation and after a foveating saccade to the noise patch. (A–C) Data for three exemplary participants. (D) Mean across all participants. Red solid lines and squares indicate data for fixation and blue dashed lines and circles data for saccades. The corresponding eye velocity is plotted in grey. Error bars denote the standard error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1. Steady-state pursuit

The present study shows that object recognition during steady-state pursuit is significantly impaired relative to fixation. However, with respect to the larger retinal position and velocity errors during pursuit, the impairment seems to be rather small. More specifically, amazingly short presentation durations of only 10 ms were sufficient for letter recognition if the object's contrast was high. Therefore, object recognition is possible during very short periods of pursuit or fixation. Thus longer pursuit epochs probably aim at another goal, for instance the prediction of the motion trajectory of the tracked object or the velocity estimation of moving targets (Land, 2006; Spering, Schütz, & Gegenfurtner, 2008).

In the following sections we will discuss several possible reasons for the detrimental performance during pursuit. A very obvious reason might be imperfect tracking, which leads to retinal position and velocity errors. Position errors can result in projections of the target object outside the fovea, which means that the maximum spatial resolution for optimal recognition performance is not reached. Velocity errors can impair recognition performance by introducing retinal jitter and image blur. Furthermore, temporal contrast sensitivity depends mainly on retinal speed (Flipse et al., 1988; Kelly, 1979; Liu & Jiang, 1984; Murphy, 1978; Schütz et al., 2007b), so that velocity errors directly influence contrast sensitivity. Our analysis showed that the retinal position and velocity errors were indeed higher during pursuit than during fixation. However we did not observe any clear dependency of recognition performance on retinal position or velocity error, neither for fixation, nor for pursuit. Therefore, imperfect tracking is probably not the reason for the performance difference. Furthermore the performance difference between fixation and pursuit was qualitatively similar for the short presentation durations in Experiment 1 and the long presentation durations in Experiment 2. However, briefly flashed targets are more or less stabilized on the retina and should not be affected by any retinal motion. The performance decline found also for these flash conditions argues additionally against a pure retinal effect as explanation. In fact it is quite astonishing that performance during pursuit was only reduced by 11% compared to fixation, although retinal position and velocity errors were much larger. In this respect, object recognition performance seems to be quite robust against tracking errors.

Another possible factor could be a differential sensitivity for contrasts during pursuit and fixation. Previous studies on contrast sensitivity during pursuit found no general suppression of sensitivity, but an enhancement of the sensitivity of the parvocellular retino-geniculate pathway (Schütz et al., 2008). In the present study letters were accompanied by a masking background consisting of high-spatial frequencies. A possible parvocellular enhancement may increase the energy of both, the noise and the letters in a similar way, leaving the signal-to-noise ratio unchanged.

Also visuo-spatial attention can affect object recognition performance, especially because there is ample of evidence that the allocation of attention and the execution of eye movements are somewhat correlated (Deubel & Schneider, 1996; Kowler et al., 1995). For steady-state pursuit, however it is commonly assumed that spatial attention is bound to the eye movement target (Kerzel, Souto, & Ziegler, 2008; Kerzel & Ziegler, 2005; Khurana & Kowler, 1987; Madelain, Krauzlis, & Wallman, 2005; Schütz et al., 2007b). Since the letter appeared in the center of the pursuit target, we expect no difference for the location of spatial attention during fixation and pursuit.

The impairment during pursuit may also be explained by the dual task requirements in pursuit conditions. It may be that the tracking of the moving target interferes on a higher cognitive level with the recognition of objects. In contrast to more reflexive eye movements like OKN or ocular following, smooth pursuit has to be initialized voluntarily. For pursuit, the designated eye movement target has to be segmented from the rest of the visual scene. This might hamper performance for other tasks like object recognition.

4.2. Pursuit initiation

In Experiment 3 we did not observe an additional deficit due to the initiation of pursuit. This is further evidence for the claim that only brief pursuit epochs might be sufficient for object recognition. It is long debated if the initiation of smooth pursuit requires the allocation of attention to the pursuit target. Whereas the allocation of attention to the future saccade target seems to be obligatory (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler et al., 1995), the relationship for pursuit initiation seems to be more complicated. In a previous study, we found only a slight reduction of contrast sensitivity during the initiation of pursuit (Schütz, Braun, & Gegenfurtner, 2007a). This reduction occurs at the time of the target onset and thus might rather reflect a distraction by the abrupt motion onset (Yantis & Jonides, 1984) than a preparation of the pursuit. A recent study showed that the distraction of attention only affects the latency of the closed-loop response, but not the latency of the open-loop response of pursuit (Souto & Kerzel, 2008).

4.3. Saccades

In Experiments 3 and 4, we found a long rise of letter recognition rate, up to 150–200 ms after the saccade. During normal viewing, humans execute saccades every 200–400 ms (Rayner, 1998). Given that the saccade programming is finished around 50–100 ms before the actual onset of the saccade (Ludwig, Gilchrist, McSorley, & Baddeley, 2005; Ludwig, Mildinhall, & Gilchrist, 2007), only a time period of 100–350 ms is available to gather information about the foveated object and about the future target for the following saccade. This would mean that the choice for the next saccade is made before or right after the maximum perceptual performance is reached. However, one has to consider that the relatively short fixation durations are measured during inspection of high-contrast text or pictures. Under those circumstances, recognition performance probably reaches its maximum earlier than in our conditions near the absolute threshold. Indeed it is known that fixation duration increases with decreasing visibility of the content (Hooge & Erkelens, 1996). We think that the found increase of recognition performance after saccades is compatible with an optimal information collection during fixations and an optimal timing of subsequent saccades.

We think that at least two components contribute to the recognition increase. First, during the saccade, performance is probably impaired by the fast retinal image motion induced by the rapid eye movement. This retinal motion smears out fine details such as the shape of the letters and therefore impairs their recognition. However, this effect is restricted to the duration of the saccade. After the saccade a second mechanism has to be responsible for the further recognition improvements. We want to discuss several possible explanations: One candidate is of course the strong suppression of contrast sensitivity during saccades (Ross, Morrone, Goldberg, & Burr, 2001). However, we think that saccadic suppression is not a probable explanation, since saccadic suppression was found to affect only contrast sensitivity for luminance contrast with low-spatial frequencies (Burr et al., 1994). The critical spatial frequency for letter recognition in our paradigm was well above the range of spatial frequencies for which saccadic suppression has been reported. Furthermore, saccadic suppression only lasts for about 50 ms after saccade onset (Diamond, Ross, & Morrone, 2000), which is much shorter than the long rise, observed in our study. Due to the high-spatial frequency content of our stimuli and the longer time course, we exclude saccadic suppression. Another possibility might be masking by the retinal motion of the noise. Retinal motion of a structured background has been shown to produce a similar suppression of contrast sensitivity like saccades (Diamond et al., 2000). However the effects of such a masking by retinal motion are longer present than the suppression by saccades. Hence the long rise of recognition rates after foveating saccades could be caused by the retinal motion of the noise patch in our experiments.

5. Summary

We investigated the ability to recognize briefly presented complex objects like letters during smooth pursuit and saccadic eye movements. We found a detrimental effect of steady-state pursuit eye movements compared to fixation. Depending on the experimental condition, the recognition rate was lowered by 11–16%. At the same time, retinal position and velocity errors were increased during pursuit. In the light of these retinal differences, object recognition performance seems to be quite robust during pursuit. Interestingly, recognition performance was not only interrupted during saccades but did not recover to fixation performance until 150 ms after the peak velocity of the saccade. When foveating saccades were followed by pursuit, no additional impairment of performance was found for pursuit initiation.

Acknowledgments

This work was supported by the DFG Forschergruppe FOR 560 “Perception and action” and the DFG Graduiertenkolleg GRK 885 “NeuroAct”. We thank Elisabeth Baumgartner for help with data collection and our participants.

References

Brown, B. (1972a). Dynamic visual acuity, eye movements and peripheral acuity for moving targets. *Vision Research, 12*(2), 305–321.

Brown, B. (1972b). The effect of target contrast variation on dynamic visual acuity and eye movements. *Vision Research, 12*(7), 1213–1224.

Brown, B. (1972c). Resolution thresholds for moving targets at the fovea and in the peripheral retina. *Vision Research, 12*(2), 293–304.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature, 371*(6497), 511–513.

Castet, E., Jeanjean, S., Montagnini, A., Laugier, D., & Masson, G. S. (2006). Dynamics of attentional deployment during saccadic programming. *Journal of Vision, 6*(3), 196–212.

Chung, S. T., Legge, G. E., & Tjan, B. S. (2002). Spatial–frequency characteristics of letter identification in central and peripheral vision. *Vision Research, 42*(18), 2137–2152.

Deubel, H. (2008). The time course of presaccadic attention shifts. *Psychological Research, 72*(6), 630–640.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *Journal of Neuroscience, 20*(9), 3449–3455.

Flipse, J. P., van der Wildt, G. J., Rodenburg, M., Keemink, C. J., & Knol, P. G. (1988). Contrast sensitivity for oscillating sine wave gratings during ocular fixation and pursuit. *Vision Research, 28*(7), 819–826.

Guyonneau, R., Kirchner, H., & Thorpe, S. J. (2006). Animals roll around the clock: the rotation invariance of ultrarapid visual processing. *Journal of Vision, 6*(10), 1008–1017.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*(6), 787–795.

Hooge, I. T., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics, 58*(7), 969–976.

Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69*(10), 1340–1349.

Kerzel, D., Souto, D., & Ziegler, N. E. (2008). Effects of attention shifts to stationary objects during steady-state smooth pursuit eye movements. *Vision Research, 48*(7), 958–969.

Kerzel, D., & Ziegler, N. E. (2005). Visual short-term memory during smooth pursuit eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 354–372.

Khurana, B., & Kowler, E. (1987). Shared attentional control of smooth eye movement and perception. *Vision Research, 27*(9), 1603–1618.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46*(11), 1762–1776.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25*(3), 296–324.

Liu, Y. M., & Jiang, B. C. (1984). Contrast sensitivity measured during smooth pursuit movement. *Scientia Sinica Series B, 27*(7), 710–721.

Ludvigh, E., & Miller, J. W. (1958). Study of visual acuity during the ocular pursuit of moving test objects. I. Introduction. *Journal of the Optical Society of America, 48*(11), 799–802.

Ludwig, C. J., Gilchrist, I. D., McSorley, E., & Baddeley, R. J. (2005). The temporal impulse response underlying saccadic decisions. *Journal of Neuroscience, 25*(43), 9907–9912.

Ludwig, C. J., Mildinhall, J. W., & Gilchrist, I. D. (2007). A population coding account for systematic variation in saccadic dead time. *Journal of Neurophysiology, 97*(1), 795–805.

Madelain, L., Krauzlis, R. J., & Wallman, J. (2005). Spatial deployment of attention influences both saccadic and pursuit tracking. *Vision Research, 45*(20), 2685–2703.

McPeek, R. M., Maljkovic, V., & Nakayama, K. (1999). Saccades require focal attention and are facilitated by a short-term memory system. *Vision Research, 39*(8), 1555–1566.

Methling, D., & Wernicke, J. (1968). Visual acuity in horizontal tracking movements of the eye. *Vision Research, 8*(5), 554–565.

Murphy, B. J. (1978). Pattern thresholds for moving and stationary gratings during smooth eye movement. *Vision Research, 18*(5), 521–530.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences, 24*(2), 113–121.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2007a). Contrast sensitivity during the initiation of smooth pursuit eye movements. *Vision Research, 47*(21), 2767–2777.

Schütz, A. C., Braun, D. I., Kerzel, D., & Gegenfurtner, K. R. (2008). Improved visual sensitivity during smooth pursuit eye movements. *Nature Neuroscience, 11*(10), 1211–1216.

Schütz, A. C., Delipetkos, E., Braun, D. I., Kerzel, D., & Gegenfurtner, K. R. (2007b). Temporal contrast sensitivity during smooth pursuit eye movements. *Journal of Vision, 7*(13), 1–15.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2009). Chromatic contrast sensitivity during optokinetic nystagmus, visually-enhanced vestibulo-ocular reflex and smooth pursuit eye movements. *Journal of Neurophysiology, 101*, 2317–2327.

Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature, 369*(6479), 395–397.

Souto, D., & Kerzel, D. (2008). Dynamics of attention during the initiation of smooth pursuit eye movements. *Journal of Vision, 8*(14), 1–16.

Spering, M., Schütz, A. C., & Gegenfurtner, K. R. (2008). Smooth pursuit eye movements and the prediction of visual motion. *Journal of Vision, 8*(6), 664a.

Sperling, G., Budiansky, J., Spivak, J. G., & Johnson, M. C. (1971). Extremely rapid visual search: The maximum rate of scanning letters for the presence of a numeral. *Science, 174*(6), 307–311.

Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520–522.

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bulthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience, 14*(5), 869–876.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293–1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63*(8), 1314–1329.

Wyatt, H. J. (1998). Detecting saccades with jerk. *Vision Research, 38*(14), 2147–2153.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 601–621.