Types and tokens in transsaccadic object identification: Effects of spatial position and left–right orientation

JOHN M. HENDERSON and AMY B. C. SIEFERT
Michigan State University, East Lansing, Michigan

What types of representations support our ability to integrate information acquired during one eye fixation with information acquired during the next fixation? In Experiment 1, transsaccadic integration was explored by manipulating whether or not the relative position of a picture of an object was maintained across a saccade. In Experiment 2, the degree to which visual details of a picture are coded in a position-specific representational system was explored by manipulating whether or not both the relative position and the left–right orientation of the picture were maintained across a saccade. Position-specific and nonspecific preview benefits were observed in both experiments. Only the position-specific benefits were influenced by the number of task-relevant pictures presented in the preview display (Experiment 1) and the left–right orientation of the picture presented in the preview display (Experiment 2). The results support a model of transsaccadic integration based on two independent representational systems. One system codes abstract, prestored object types, and the other codes episodic tokens consisting of stimulus properties linked to scene- or configuration-based position markers.

Object perception is a dynamic process: The entire visual scene disappears and then reappears three to four times each second during saccadic eye movements. How does the human visual system integrate this continually changing input? The purpose of the present study was to investigate the nature of the representational systems that support the retention and integration of information about real-world objects across saccades during dynamic visual perception.

One intuitively appealing hypothesis concerning transsaccadic integration is that detailed sensory representations are created and fused from one fixation to the next within a spatiotopically organized reference frame (e.g., Breitmeyer, 1984; Davidson, Fox, & Dick, 1973; Feldman, 1985; Jonides, Irwin, & Yantis, 1982; McConkie & Rayner, 1975). In this type of sensory fusion system, the sensory information gathered during two consecutive fixations could be aligned by tracking the extent of the saccade and/or by comparing the similarity of the projected sensory patterns. The fused pattern could then be used as input to the object identification system. However, despite many attempts to find empirical support for spatiotopically based sensory fusion, most investigators now agree that it does not occur (Irwin, 1991, 1992; O'Regan, 1992; Pollatsek & Rayner,

1992; although see Deubel, Schneider, & Bridgeman, 1996, for an alternative perspective).

Although there is little evidence supporting spatiotopic sensory fusion across saccades, many studies have shown that the information acquired from a preview of a semantically meaningful visual stimulus (e.g., word or object) during one fixation can speed identification of that stimulus during the next fixation (Henderson, 1992a, 1997; Henderson, Pollatsek, & Rayner, 1987, 1989; McConkie & Rayner, 1975; Pollatsek, Rayner, & Collins, 1984; Pollatsek, Rayner, & Henderson, 1990; Rayner, 1978; Rayner, McConkie, & Zola, 1980). Further investigation showed that this preview benefit is not based on a point-by-point alignment of sensory information: Facilitation can be observed despite a 10% change in the size of the object (Henderson et al., 1987; Pollatsek et al., 1984; Pollatsek et al., 1990), despite a change in the spatial position of the object (Henderson, 1997; Pollatsek et al., 1990), and despite changes to the visual details of the object across the saccade (Henderson, 1997; Pollatsek et al., 1984).

If transsaccadic integration is not supported by the fusion of sensory information, on what is it based? Recent work suggests that transsaccadic integration may be based on two independent representational systems (Henderson, 1994; Henderson & Anes, 1994; see also Irwin & Andrews, 1996), one coding general object *types* and the other coding specific object *tokens*. In this type–token framework, when a perceptual object is attended in the visual field, object types are activated, and an object token is constructed. The purpose of the type system is to support visual identification, whereas the purpose of the token system is to pre-

serve information about the specific objects that are present in the visual environment (Driver & Kanwisher, 1992; Kahneman, Treisman, & Gibbs, 1992). Identification of an object will be facilitated across a saccade by prior type activation to the extent that the new view of the object leads to continued activation of the same subset of type representations as were activated prior to the saccade. At the same time, identification will be facilitated by the prior construction of an object token to the extent that the new view causes a retrieval of that token. Retrieval of a token, in turn, is based on the degree to which spatiotemporal continuity is maintained. Spatiotemporal continuity is taken to be based on the representation of an object's position within a spatial reference frame, where the position can be updated over time, given perceptual evidence for such a change (e.g., evidence for motion).

Two main findings support the hypothesis that two independent representational systems contribute to transsaccadic integration. First, the benefit produced by a preview of a perceptual object prior to a saccade to that object consists of two components, an *object-specific preview benefit* that is dependent on the spatiotemporal continuity of the object across the saccade and a *nonspecific preview benefit* that survives disruptions to spatiotemporal continuity across the saccade (Henderson, 1994; Henderson & Anes, 1994; see also Kahneman et al., 1992; Pollatsek et al., 1990). Second, the object-specific and nonspecific preview benefits can be dissociated by a manipulation of integration load. Specifically, the object-specific benefit is reduced (in some experiments, eliminated) when the preview display contains two task-relevant perceptual objects rather than only one, whereas the nonspecific benefit is relatively unaffected by the presence of an additional task-relevant object (Henderson, 1994; Henderson & Anes, 1994). Both of these results are consistent with the view that transsaccadic information integration is mediated by two representational systems, one that codes episodic tokens that are tied to a particular set of spatial coordinates at a given time and that are susceptible to interference from other tokens, and another that codes prestored object properties that are independent of spatiotemporal continuity and relatively unaffected by the number of individual tokens visible (at least up to two).

The purpose of the present study was to test further the hypothesis that transsaccadic integration is supported both by stored object types and by episodic object tokens.

EXPERIMENT 1

Pollatsek et al. (1990) presented evidence that transsaccadic preview benefits were not significantly greater when a picture (line drawing) of a real-world object remained in the same spatial location across a saccade than when the drawing exchanged its location with a location landmark. This result is consistent with the view that transsaccadic integration is based only on the activation of prestored type representations, and not on spatially coded episodic tokens. In contrast, in a subsequent set of experiments using letters as stimuli, we demonstrated that preview benefits were larger when a letter remained in the same location across a saccade than when the letter changed location (Henderson, 1994; Henderson & Anes, 1994). This latter result suggests that both prestored type representations and spatially coded episodic tokens can play a role in the integration process. The purpose of Experiment 1 was to investigate further whether position-specific and nonspecific preview benefits can be observed during the transsaccadic identification of pictures of real-world objects. This issue is important because, if position-specific benefits were limited to transsaccadic integration of letters, it would be difficult to argue that the phenomenon generalizes to natural perception.

To determine whether position-specific benefits can be observed with pictures of objects, we used a technique in which the display could be changed on the basis of the execution of a saccade (e.g., Henderson, 1997; Henderson et al., 1987; Pollatsek et al., 1984; Rayner, 1978; Rayner et al., 1980). Each trial consisted of three display events. First, a fixation field was presented, consisting of three fixation markers. The participant began each trial by fixating the leftmost marker. Second, a preview field was shown, consisting of two pictures of meaningful real-world objects, a picture of a meaningful object and a meaningless object-like pattern, or two copies of the meaningless pattern. The participant was instructed to execute a saccade to a location between the two items as quickly as possible once they were presented. Third, when the computer detected the execution of a saccade, a target field was displayed, containing the target picture and the meaningless pattern. The participant named the target picture as quickly as possible following the saccade.

Six preview conditions were included in the experiment, generated from a 3×2 factorial combination of the position of the target picture relative to the preview and the type of flanker shown with the preview. The flanker manipulation was included to provide an opportunity to compare the results of the present experiment with those of our earlier studies in which we have used letters as stimuli. In this experiment, the three levels of the *position continuity*[1] factor were *same position*, in which the target remained in the same spatial position from preview to target display, *switch position*, in which the target switched to the other possible spatial position from preview to target display, and *control*, in which no preview of the target was shown in the preview display. The two levels of the *flanker* condition were *meaningful*, in which another potential target picture appeared in the preview display along with the target, and *meaningless*, in which a meaningless pattern that was not associated with any real-world object appeared in the preview display with the target.

Method

Participants. Eight members of the Michigan State University participant pool participated in the experiment for course credit. All participants had normal vision or wore contact lenses. The partici-

pants had not participated in previous eye movement experiments and were naive with respect to the hypotheses under investigation.

Stimuli. The stimuli were 12 pictures (line drawings) of common real-world objects taken from Snodgrass and Vanderwart (1980) and a meaningless pattern. Each preview and target display contained two meaningful pictures, a meaningful picture and the meaningless pattern, or two copies of the meaningless pattern, centered within rectangular frames oriented vertically. The meaningless pattern was created by blending the contours of several Snodgrass and Vanderwart pictures that were not included in the experiment and then deleting a majority of the interior contour.

The pictures were about 3.6° in height and 5.1° in width, on average. The meaningless pattern was the same size as this average. There were 10.5° between the centers of the pictures and 6.9° between their nearest contours, on average. The frames subtended 7.7° vertically and 7.2° horizontally, with 2.8° separating the nearest contours of the frames. Table 1 shows the 12 preview conditions for the target picture "fox."

Twenty-four target displays were created, consisting of each of the 12 pictures in each of the two (top and bottom) frames. The meaningless pattern occupied the frame that did not contain the target picture in each target display.

Forty-nine preview displays were used. Twenty-four of these preview displays were identical to the target displays. These displays served as the preview displays for the same-frame meaningless-flanker and switch-frame meaningless-flanker conditions. An additional 24 displays contained two pictures each, 12 pairs of pictures × 2 configurations (e.g., fox above bed and bed above fox). These displays were used in the same-position meaningful-flanker condition, the switch-position meaningful-flanker condition, and the meaningful-flanker control condition. For example, the display containing the fox in the top frame and the bed in the bottom frame served as the preview for the same-frame meaningful-flanker condition when the fox appeared as the target in the top frame and when the bed appeared as the target in the bottom frame. This same display served as the preview for the switch-frame meaningful-flanker condition when the bed appeared as the target in the top frame and when the fox appeared as the target in the bottom frame. This display also served as the meaningful-flanker control condition when the football appeared as the target in the top frame. (The meaningful-flanker control for the football when it appeared as the target in the bottom frame was the bed above the fox.) The final preview display contained two meaningless pictures and was used in the meaningless-flanker control condition.

Apparatus. The stimuli were displayed at a resolution of 640 × 480 pixels on a Zenith flat-screen VGA monitor operating at 60 Hz. The contours of the stimuli appeared black (pixels off) against a gray (pixels on) background. Eye movements were monitored via an ISCAN RK-416 high-speed eyetracker. Signals were generated by the eyetracker at a frequency of 120 Hz, and the computer changed the display contingent on detecting an eye movement of greater than 0°. Display changes required a maximum of 25 msec. Because a 20° horizontal saccade has a duration of well over 50 msec (Becker, 1989; Collewijn, Erkelens, & Steinman, 1988), the display change was accomplished during the saccade when vision was suppressed. Vocal responses were collected with a voice key connected to a dedicated input–output (I/O) board; activation of the voice key stopped a millisecond clock on the I/O board and generated a system interrupt that was serviced by software. The eyetracker, display monitor, and voice key were interfaced with a 486-based microcomputer that controlled the experiment. The computer maintained a complete record of saccade and naming latencies.

Procedure. Upon arriving for a session, each participant was seated comfortably with his or her head placed in a chin and forehead rest to minimize head movements. The participants were first shown each of the target pictures and asked to name them. All the participants used the names given by Snodgrass and Vanderwart (1980). The calibration of the eyetracker then took place. After calibration, the participants took part in one short block of 16 practice trials, and two blocks of 96 test trials each. A trial consisted of the following events. First, a fixation display appeared containing three test fixation markers and a small cross that indicated the computer's estimate of the current fixation position. The participant fixated each test marker, and if the calibration was satisfactory (±⅓° from each marker), the experimenter asked the participant to fixate the leftmost marker to indicate that he or she was ready for the trial to begin. The experimenter then initiated the trial by pushing a silent button. The fixation display was replaced by a preview display consisting of two frames to the right of fixation, each containing a picture of a real world object, or one containing a picture of a real-world object and the other containing a meaningless pattern. The participant immediately initiated a rightward horizontal eye movement to a location centered between the two frames. The distance from the initial fixation point to the location centered between the two stimuli was 20.6°. During the saccade, the preview display was replaced by the target display, consisting of the target picture in one frame and the meaningless pattern in the other. The target display remained in view until the participant responded by naming the pic-

Table 1
Example of the 12 Same-Orientation Conditions Employed in Experiment 1
When the Target Was a Picture of a Fox

	Stimuli Presented	
Condition	Preview Display	Target Display
Same position, meaningful flanker, target at top	fox bed	fox non
Same position, meaningful flanker, target at bottom	bed fox	non fox
Same position, meaningless flanker, target at top	fox non	fox non
Same position, meaningless flanker, target at bottom	non fox	non fox
Switch position, meaningful flanker, target at top	bed fox	fox non
Switch position, meaningful flanker, target at bottom	fox bed	non fox
Switch position, meaningless flanker, target at top	non fox	fox non
Switch position, meaningless flanker, target at bottom	fox non	non fox
Control, meaningful flanker, target at top	donkey bus	fox non
Control, meaningful flanker, target at bottom	bus donkey	non fox
Control, meaningless flanker, target at top	non non	fox non
Control, meaningless flanker, target at bottom	non non	non fox

Note—The first stimulus (object or pattern) of each pair represents the stimulus appearing in the top frame. Non = meaningless pattern. The switch-orientation conditions were similar, except that the left–right orientation of the target object was reversed in the target display.

ture as quickly as possible. The computer recorded the latency of the eye movement and the latency of the vocal response (timed from when the eye crossed the 0.5° boundary).

The participants were given one practice block of 12 trials and then were tested in one test block of 144 trials. The 144 trials were produced by the within-subjects factorial combination of 3 position continuity conditions (same, switch, and control) × 2 flanker conditions (meaningful and meaningless) × 2 target positions (top and bottom) × 12 target pictures. The target position factor was included for counterbalancing purposes and was not of theoretical interest. The experiment was completed in a single session that lasted less than 1 h.

Results

Mean corrected naming latencies collapsed over target position are presented in Table 2. These means exclude trials on which an anticipatory eye movement occurred (saccade latency of less than 100 msec) and trials on which the naming latency was less than 200 msec, more than 1,200 msec, or greater than three standard deviations from the mean naming latency for that subject. About 8% of the data were eliminated in total. The pattern of corrected latencies did not differ from the pattern prior to correction. Mean eye movement latency was 278 msec, averaged across conditions.

There was a reliable effect of position continuity, with mean naming latencies of 761, 800, and 829 msec in the same-position, switch-position, and control conditions, respectively [$F(2,14) = 15.96, MS_e = 2,375, p < .001$]. Planned comparisons showed that the 39-msec position-specific benefit (switch-position minus same-position conditions) was significant [$F(1,7) = 8.934, MS_e = 2,715, p < .05$], as was the 29-msec nonspecific benefit [control minus switch-position conditions; $F(1,7) = 6.376, MS_e = 2,211, p < .01$]. There was also a reliable effect of flanker type; naming latencies were 82 msec faster in the meaningless-flanker condition (756 msec) than in the meaningful-flanker condition [838 msec; $F(1,7) = 27.28, MS_e = 5,903, p < .005$]. Position-continuity and flanker type interacted reliably [$F(2,14) = 8.190, MS_e = 1,577, p < .001$]. Examining the position-specific and nonspecific components of the preview benefit separately (shown in Table 2) revealed that the position-specific component was affected by the flanker [$F(1,7) = 10.96, MS_e = 1,937, p < .05$], whereas the nonspecific component was not ($F < 1, MS_e = 1,108$). The only other reliable effect was an interaction of position and flanker type [$F(1,7) = 6.710, MS_e = 813, p < .05$], with faster naming latencies for the top versus the bottom position when there was a picture flanker (824 vs. 851 msec, respectively), but no such difference when the flanker was a nonobject (757 vs. 754 msec).

Discussion

Two main conclusions can be drawn from the results of Experiment 1. First, both position-specific and nonspecific preview benefits can be observed during the transsaccadic integration of real-world objects. Thus, the pattern of position-specific and nonspecific preview benefits that we have observed with letter stimuli in prior studies generalizes to pictures. This result suggests that both short-term episodic token representations and prestored type representations contribute to transsaccadic integration (Henderson, 1994; Henderson & Anes, 1994). We note that these results contrast with those reported by Pollatsek et al. (1990), who found only nonspecific preview benefits with line drawings of objects. The discrepancy in the results across the two studies may not be as large as first appears, however, because Pollatsek et al. did find a tendency toward position-specific benefits in their two transsaccadic viewing experiments (see Henderson, 1994, for a discussion).

Second, the position-specific and nonspecific benefits were differentially affected by the type of flanker shown in the preview display; the position-specific benefit was smaller when two potential target pictures were shown than when only a single potential target picture appeared, whereas the nonspecific benefit was relatively constant in magnitude across flanker conditions. We have reported three previous experiments (Experiments 1 and 2 of Henderson & Anes, 1994; Experiment 2 of Henderson, 1994) in which we found a similar (although only marginally reliable) interaction between preview and flanker conditions when letters were used as stimuli. We interpret this interaction as support for the dissociation of the position-specific and nonspecific benefits, as would be expected on the view that the two types of benefits reflect different underlying representational systems.

EXPERIMENT 2

In Experiment 2, we sought to determine whether the representational system that produces position-specific benefits during transsaccadic integration codes a more detailed representation of the image features present in the preview than does the system that produces the non-

Table 2
Mean Naming Latencies (in Milliseconds) and Mean Preview Benefits (in Milliseconds) as a Function of Position Continuity (Same, Switch, Control) and Flanker Type, Experiment 1

Flanker Type	Latency			Preview Benefit	
	Same	Switch	Control	Specific	Nonspecific
Picture	825	827	860	2	33
Pattern	697	772	798	75	26
M	761	800	829	39	29

specific benefits, as would be predicted by the hypothesis that the position-specific benefits are due to episodic tokens of the preview picture (Kahneman et al., 1992). In several earlier experiments, we have been unable to find support for this prediction. For example, in one study, Pollatsek et al. (1990) manipulated position continuity and the continuity of the left–right orientation of a line drawing across a saccade and found no interaction between these two factors. In a similar experiment, Henderson (1994) manipulated position continuity and the continuity of the case in which a target letter was displayed and again found no interaction. However, there are reasons to be cautious about interpreting these prior studies. In the Pollatsek et al. (1990) experiment, there was no overall effect of position continuity, whereas in the Henderson (1994) study, there was no overall effect of the case change manipulation. The lack of main effects may have impeded detection of an interaction. Furthermore, it has been suggested that letter case is not represented across saccades (McConkie & Zola, 1979; Rayner et al., 1980). Thus, the letter case manipulation may not have provided a sufficient opportunity to observe an interaction of visual features and object position.

To provide a better test of whether specific image features might be maintained by episodic tokens across saccades, in Experiment 2 we used manipulations of position and form continuity that have been shown to produce effects on transsaccadic integration. Specifically, we manipulated position as in Experiment 1 (see, also, Henderson, 1994; Henderson & Anes, 1994) and manipulated the specific visual form of the image by maintaining or changing the left–right orientation of a line drawing target from preview to target display. Given that left–right orientation can be retained and integrated across saccades (Henderson & Siefert, 1999), and given our hypothesis that specific information about object tokens is preserved by the token system, we predicted that a change in the left–right orientation of a picture should affect transsaccadic integration when the picture's position remains constant but should not influence transsaccadic integration when the position changes. In other words, the position-specific benefit should be larger when the left–right orientation is preserved than when it is reversed, but the nonspecific benefit should be unaffected by a change in left–right orientation. To test this prediction, we introduced two levels of a *left–right orientation* factor: *same orientation*, in which the same version of the pictured object was seen in the preview and target displays, and *different orientation*, in which the pictured object was changed from one left–right orientation to its mirror image from preview to target display. The orientation continuity factor allowed us to determine whether the maintenance of more detailed information than the entry-level description of an object type is maintained by the object token system.

Experiment 2 differed from Experiment 1 in five other ways. First, because we found in Experiment 1 that the position-specific benefit tended to be greatly reduced or even eliminated when a meaningful flanker was present in the preview display (see also Henderson, 1994; Henderson & Anes, 1994), and because we were interested in exploring the influence of image information on the position-specific benefit, we did not include a flanker picture in Experiment 2. Second, because we could change the display more rapidly in Experiment 2 (using a 100-Hz video refresh rate), we were able to allow participants to make a 10° rather than a 20° saccade to the preview images, thereby increasing the amount of preview information they would be able to obtain (Henderson & Siefert, 1999; Pollatsek et al., 1990). Third, in Experiment 2 all target images were shifted 11.13′ of arc to the right (away from the initial fixation point) during the saccade. The displacements were introduced to further control for the possibility that position-specific preview benefits could be due to the phosphor properties of the display monitor. Previous research has shown that such small displacements are very difficult to detect when they take place during a saccade (e.g., Bridgeman, Hendry, & Stark, 1975; Bridgeman & Stark, 1979; Henderson, 1997). Fourth, we eliminated the frames around the pictures. To the extent that object positions are allocentrically coded (Henderson, 1994; Henderson & Anes, 1994), the frames should not be necessary to produce position-specific benefits. Finally, we included a different-picture preview condition as a second control (in addition to the empty box condition) against which to examine preview benefits. A comparison of the two controls allows us to ensure that the preview benefits observed are not due to the simplicity of the preview image in the empty box control condition (Henderson & Siefert, 1999).

Method

Participants. Twelve new participants from the same pool and with the same characteristics as those used in Experiment 1 took part in Experiment 2.

Stimuli. The stimuli were two enantiomorphs of each of 10 pictures (line drawings) of common real-world objects that all had an obvious left–right orientation, taken from Snodgrass and Vanderwart (1980). These pictures were those that produced the largest left–right orientation change effect in Experiment 1 of Henderson and Siefert (1999) and depicted the basic-level concepts of a camel, cow, dog, elephant, fish, flag, football, iron, squirrel, and whistle. Each preview and target display contained a picture and a placeholder, or two placeholders, oriented vertically in the display. The placeholder was a square containing a smaller square and a plus at its center. The pictures and placeholder subtended 4.5°, on average, along the longest axis. There was 7.3° between the centers of the pictures and placeholders, and 3.9° between their nearest contours, on average. All the pictures were displaced 11.13′ of arc away from the initial fixation point between preview and target displays.

Forty target displays were used, consisting of each of the 10 pictures in each left–right orientation and in each of the two (top and bottom) frames. A placeholder occupied the potential target position that was not occupied by the preview picture. Thirty-two preview displays were generated as described for the meaningless flanker preview displays in Experiment 1. These displays were identical to the target displays. A final preview display contained two placeholders; this display served as the simple control condition.

Apparatus. The apparatus was the same as that in Experiment 1, with the following exceptions. Stimuli were displayed at a resolution of 800 × 600 pixels on a NEC flat-screen SVGA monitor operating

at 100 Hz. The contours of the stimuli appeared black (pixels off) against a gray (pixels on) background. Because of the faster video refresh rate, display changes required a maximum of 18 msec.

Procedure. The procedure was the same as that used in Experiment 1, with the following exceptions. After the eyetracker was calibrated, the participants took part in one short block of 16 practice trials and then a test block of 320 experimental trials. The 320 experimental trials were produced by the within-subjects factorial combination of 4 preview conditions (same orientation, different orientation, simple control, and different-picture control) × 2 position continuity conditions (same position and switch position) × 2 target positions (top and bottom) × 2 left–right orientations (Orientation 1 or Orientation 2) × 10 target pictures. The target position and left–right orientation conditions were included for counterbalancing purposes and were not of theoretical interest, and the position-continuity factor was dummy coded for the simple control condition (in which two placeholders were shown in the preview display). With this design, a given preview picture was equally predictive of that same picture as target, the picture of the same concept but different orientation, the picture at Orientation 1 of a different concept, and the picture at Orientation 2 of a different concept. Thus, any differences in the perceptibility of one orientation over another, in either the preview or the target display, were controlled.

Results

Mean corrected naming latencies excluded trials on which an anticipatory eye movement occurred (saccade latency of less than 100 msec) and trials on which the naming latency was less than 200 msec, more than 1,500 msec, or greater than three standard deviations from the mean naming latency for that subject. About 5% of the data were eliminated in total. The pattern of corrected latencies did not differ from the pattern prior to correction. Mean eye movement latency was 210 msec, averaged across conditions. An analysis of eye movement latencies showed that they were reliably longer when the control stimulus was presented (two placeholders) than when a picture was presented in the preview display [$F(3,33) = 7.563, MS_e = 718, p < .001$]. The eye movement latencies were 205, 204, 226, and 206 msec in the same-orientation, different-orientation, simple control, and different-picture control conditions, respectively.

The main question in this experiment was whether the position-specific benefit would be larger when left–right picture orientation was preserved. To investigate this question, the data were analyzed in a 2 (target position, top vs. bottom) × 2 (position continuity, same position vs. different position) × 3 (left–right orientation, same vs. different vs. simple control) analysis of variance (ANOVA)

that excluded the different-picture control. The naming latencies for these data, as well as the preview benefits, are presented in Table 3. The interaction between position continuity and left–right continuity was reliable [$F(2,22) = 18.25, MS_e = 740, p < .001$]. As can be seen in Table 3, the nonspecific preview benefits were unaffected by the left–right orientation change between preview and target displays (81 msec vs. 83 msec for the same-orientation and different-orientation conditions, respectively), whereas the position-specific benefits were about twice as large when the left–right orientation of the target was constant across the saccade (51 msec), as compared with when it changed (25 msec). Although the preview benefits become larger if the different-picture control is used as the baseline (see Table 3), the conclusions concerning the effects of position and left–right orientation on the preview benefit does not change. Finally, in contrast to the results of Experiment 1, there were no interactions between any of the critical factors and target position (all Fs < 1).

Discussion

The purpose of Experiment 2 was to determine whether a change to an exemplar-level visual property of a picture of a real-world object, left–right orientation, would have more of an influence on transsaccadic integration when position was maintained than when position changed across a saccade, as would be predicted by the type–token framework. We found that preview benefits were larger when left–right orientation was constant across a saccade than when left–right orientation was reversed, replicating Henderson and Siefert (1999). Most important, this additional benefit for same-orientation pictures over different-orientation pictures was apparent only when the picture remained in the same position from preview to target display. This result was predicted by the hypothesis that the position-specific benefit reflects the operation of a representational system that codes information about the specific token that was viewed, whereas the nonspecific benefit reflects the operation of a system that codes more general information about the perceptual and/or conceptual type that was viewed.

The present results contrast with those obtained by Henderson (1994), who reported that case changes to letters did not influence position-specific preview benefits any more than they influenced nonspecific preview benefits. Henderson (1994) concluded that visual features are

Table 3
Mean Naming Latencies (in Milliseconds) and Mean Preview Benefits (in Milliseconds) as a Function of Position Continuity (Same, Switch, Controls) and Left–Right Orientation Continuity, Experiment 2

Left–Right Orientation	Latency				Preview Benefit (Versus Simple Control)	
	Same	Switch	Simple Control	Different-Picture Control	Position Specific	Nonspecific
Same	594	645			51	81
			726	761		
Different	618	643			25	83
Same orientation benefit	24	−2				

no more likely to be preserved by the token system than by the type system. It now appears more likely that letter case is simply not preserved across saccades at all, so it is preserved neither by the token nor by the type system. On the other hand, given that the left–right orientation of a pictured object is preserved across saccades (Henderson & Siefert, 1999), the present results suggest that left–right orientation is preserved by the token system.

GENERAL DISCUSSION

The purpose of this study was to investigate how the information that is obtained during one eye fixation is integrated with the information that is acquired during the next fixation. Two experiments were conducted to test the hypothesis that transsaccadic object identification is supported by two representational systems, a type system in which viewed objects activate prestored representations that code generic information about object concepts and identities and an episodic token system in which viewed objects generate episodic descriptions that code more precise object details. We investigated this topic by using a transsaccadic preview paradigm in which participants were asked to identify an object following a saccade. The results of Experiment 1 demonstrated that transsaccadic integration of pictorial information is supported by both position-specific and nonspecific components. Furthermore, the position-specific component was found to be sensitive to the number of task-relevant pictures in the preview display, whereas the nonspecific component was not. Experiment 2 extended these results, showing that the position-specific benefit was sensitive to the similarity of the left–right orientation of the picture across the saccade, whereas the nonspecific component was not. This latter effect provides strong support for the notion that the position-specific benefit reflects the operation of a system that represents information about the particular token that was present in the visual field prior to the saccade.

We propose the following type–token framework for explaining transsaccadic integration. The initial preview of a visual scene prior to a saccade leads to the coding of position markers in a spatial-processing subsystem and activation of object properties in an identity-processing subsystem. Prior to a saccade, a small number of position markers will be bound to object properties in the formation of object tokens, with preference for token construction given to object(s) near the intended saccade target (Henderson, 1992b; Henderson & Hollingworth, 1999; Henderson et al., 1989; Hoffman & Subramaniam, 1995; Irwin & Andrews, 1996; McConkie & Currie, 1996; see Henderson, 1996). Following a change to the proximal stimulus owing to a saccade, preview benefits (facilitation of identification) can take place via two processes. First, new activation caused by an attended object following the saccade (i.e., in the postsaccade fixation) can combine with any residual activation in the identity-processing subsystem to facilitate identification (and any other re-

sponse supported by the representations maintained by the identity-processing subsystem). This integration of information owing to residual activation within the identity subsystem will be independent of object position and will lead to nonspecific preview benefits. Second, an episodic token will be retrieved following the change if there is spatiotemporal continuity between a currently visible object and a represented position marker in the spatial-processing subsystem. Spatiotemporal continuity appears to be based, at least in part, on position coding within an allocentric reference frame (Henderson, 1994; Henderson & Anes, 1994). In other words, episodic tokens can be tracked across a saccade by the spatial-processing subsystem via the object's position with respect to other coded object positions and, possibly, with respect to other reference frames as well (Verfaillie, De Troy, & Van Rensbergen, 1994). When spatiotemporal continuity is maintained, representations of object properties forming that token (i.e., pointed at by that position marker) will be reactivated within the identity-processing subsystem. Again, new activation caused by an object that is attended after the change can combine with activation caused by the token to facilitate performance. The integration of information owing to reactivation from the token will be dependent on the maintenance of spatiotemporal continuity and will lead to object-specific preview benefits.

Are the episodic tokens that we are discussing the same as the object files that Kahneman et al. (1992) have proposed? On the one hand, we clearly draw much of our inspiration from the object files theory, and a good deal of our data support many of the major assumptions of that theory. For example, we agree that episodic tokens are addressed by spatial position at a particular time. However, we have made additional assumptions that put us beyond the bounds of the object files theory. For example, we have proposed that the reference frame used to track the positions of objects for the purposes of transsaccadic integration is configuration centered. We are not sure whether this proposal would be consistent with the object files theory. In addition, in object files theory, the object files are the representations that are matched to long-term memory during object identification (Kahneman et al., 1992). In our view, object identification takes place entirely within the object-processing subsystem and does not involve the episodic tokens (see Henderson, 1994). The episodic tokens can affect identification only indirectly by reactivating representations within the object-processing subsystem. Because we do not feel confident claiming that the episodic tokens we have described are identical with object files, we leave this issue open for the time being.

REFERENCES

BECKER, W. (1989). Metrics. In R. H. Wurtz and M. E. Goldberg (Eds.), *The neurobiology of eye movements* (pp. 13-67). Amsterdam: Elsevier.

BREITMEYER, B. G. (1984). *Visual masking: An integrative approach.* New York: Oxford University Press.

BRIDGEMAN, B., HENDRY, D., & STARK, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, **15**, 719-722.

BRIDGEMAN, B., & STARK, L. (1979). Omnidirectional increase in threshold for image shifts during saccadic eye movements. *Perception & Psychophysics*, **25**, 241-243.

COLLEWIJN, H., ERKELENS, C. J., & STEINMAN, R. M. (1988). Binocular co-ordination of human horizontal saccadic eye movements. *Journal of Physiology*, **404**, 157-182.

DAVIDSON, M. L., FOX, M.-J., & DICK, A. O. (1973). Effect of eye movements on backward masking and perceived location. *Perception & Psychophysics*, **14**, 110-116.

DEUBEL, H., SCHNEIDER, W. X., & BRIDGEMAN, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research*, **36**, 985-996.

DRIVER, J., & KANWISHER, N. (1992). Objects, attributes, and visual attention: Which, what, and where. *Current Directions in Psychological Science*, **1**, 26-31.

FELDMAN, J. A. (1985). Four frames suffice: A provisional model of vision and space. *Behavioral & Brain Sciences*, **8**, 265-289.

GORDON, R. D., & IRWIN, D. E. (1996). What's in an object file? Evidence from priming studies. *Perception & Psychophysics*, **58**, 1260-1277.

HENDERSON, J. M. (1992a). Identifying objects across saccades: Effects of extrafoveal preview and flanker object context. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 521-530.

HENDERSON, J. M. (1992b). Visual attention and eye movement control in reading and scene perception. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 260-283). New York: Springer-Verlag.

HENDERSON, J. M. (1994). Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General*, **123**, 410-426.

HENDERSON, J. M. (1996). Visual attention and the attention-action interface. In K. Aikens (Ed.), *Perception* (Vancouver Studies in Cognitive Science, Vol. V, pp. 290-316). Oxford: Oxford University Press.

HENDERSON, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, **8**, 51-55.

HENDERSON, J. M., & ANES, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 826-839.

HENDERSON, J. M., & HOLLINGWORTH, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, **5**, 438-443.

HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1987). The effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 449-463.

HENDERSON, J. M., POLLATSEK, A., & RAYNER, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics*, **45**, 196-208.

HENDERSON, J. M., & SIEFERT, A. B. C. (1999). The influence of enantiomorphic transformation on transsaccadic object integration. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 243-255.

HOFFMAN, J. E., & SUBRAMANIAM, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, **57**, 787-795.

IRWIN, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, **23**, 420-456.

IRWIN, D. E. (1992). Perceiving an integrated visual world. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 121-142). Cambridge, MA: MIT Press.

IRWIN, D. E., & ANDREWS, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and Performance XVI: Information integration in perception and communication* (pp. 125-156). Cambridge, MA: MIT Press.

JONIDES, J., IRWIN, D. E., & YANTIS, S. (1982). Integrating visual information from successive fixations. *Science*, **215**, 192-194.

KAHNEMAN, D., TREISMAN, A., & GIBBS, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, **24**, 175-219.

MCCONKIE, G. W., & CURRIE, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 563-581.

MCCONKIE, G. W., & RAYNER, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, **17**, 578-586.

MCCONKIE, G. W., & ZOLA, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, **25**, 221-224.

O'REGAN, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, **46**, 461-488.

POLLATSEK, A., & RAYNER, K. (1992). What is integrated across fixations? In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 166-191). New York: Springer-Verlag.

POLLATSEK, A., RAYNER, K., & COLLINS, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, **113**, 426-442.

POLLATSEK, A., RAYNER, K., & HENDERSON, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 199-210.

RAYNER, K. (1978). Foveal and parafoveal cues in reading. In J. Requin (Ed.), *Attention and performance VII* (pp. 149-162). Hillsdale, NJ: Erlbaum.

RAYNER, K., MCCONKIE, G. W., & ZOLA, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, **12**, 206-226.

SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

VERFAILLIE, K., DE TROY, A., & VAN RENSBERGEN, J. (1994). Transsaccadic integration of biological motion. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 649-670.

NOTE

1. We have previously called this factor *object continuity*, following the terminology of Kahneman et al. (1992). However, because it is not possible to distinguish object and position continuity in these experiments, we use the more precise term here.