Published in final edited form as:

J Cogn Psychol (Hove). 2013 May 1; 25(3): 283–298. doi:10.1080/20445911.2013.768232.

Effects of changes in narrative time on eye movements and recognition responses

Kristin M. Weingartner and
Hofstra University

Jerome L. Myers
University of Massachusetts, Amherst

Abstract

In two experiments we examined how temporal aspects of narrative events influence comprehension. In Experiment 1 participants read paragraphs in which a critical event was followed by a phrase that signaled a time shift (*After an hour* versus *After a moment*). Consistent with earlier findings (e.g., Zwaan, 1996), fixation durations were longer on the phrase that signaled a larger time shift. However, there was no evidence that a larger time shift affected the accessibility of event information in Experiment 1, when the dependent measure was ease of anaphor comprehension, or in Experiment 2, when a recognition probe task was used. Although the discontinuation of an event (Maurice *stopped* versus *was painting*) did not affect anaphor reading times, it did lead to longer recognition times for the event. These results indicate that at least some event aspects remain accessible following a change in time and that the dependent measure can have a critical impact on the conclusions.

Keywords

Reading; anaphor comprehension; eye movements; time shifts; temporal updating

An important goal of research in text comprehension is to understand how readers incorporate changes that occur in the evolving situation into their ongoing mental representation of the text. In the two experiments reported here we examined how readers integrate changes with respect to the temporal aspects of the described situation. Specifically, we used eye tracking methodology in Experiment 1, and recognition memory probes in Experiment 2, to investigate further two reported consequences of temporal shifts signaled by adverbial phrases in narratives: (1) an increase in processing time with larger shifts forward in time (e.g. *An hour later* versus *A moment later*); and (2) a reduced accessibility of previously presented information with larger time shifts. Consider the following excerpt of a sample stimulus from Zwaan (1996):

(1) Maurice was in an excellent mood. He was shaking hands and beaming. *A moment later / An hour later / A day later*, he turned very pale.

Using materials such as these, several investigators reported that, compared to adverbial phrases that denote little or no passage of time (*A moment later*), those that do (*An hour later*, *A day later*) were associated with elevated reading times for the sentences or clauses in which they were contained (Anderson, Garrod, & Sanford, 1983; Radvansky & Copeland, 2010; Radvansky, Copeland, Berish, & Dijkstra, 2003; Speer & Zacks, 2005; Zwaan, 1996),

and in one study (Ditman, Holcomb, & Kuperburg, 2008) they led to immediate neural integration costs.

In addition to increasing processing load, there is evidence that time shifts lead to a reduction in the accessibility of information that precedes them. Zwaan (1996) used sequences of events similar to (1), embedded in lengthier paragraphs. Immediately after the sentence that contained a temporal adverbial phrase, recognition response times to an event from the preceding sentence (*beaming*) were longer when there was a time shift of an hour or a day compared to when there was relatively no shift in time. Importantly, this occurred despite the fact that the physical distance between where the critical event was mentioned and where the test probe appeared was held constant. Thus, a shift forward in time in the *described situation* appears to have a critical impact on the accessibility of information. This finding, along with the processing load effects described previously, provides support for the claim that readers perceive temporal changes as event boundaries between consecutive episodes of activity in the text (Speer & Zacks, 2005; Zwaan & Radvansky, 1998).

The findings from a number of other investigations (Anderson et al., 1983; Claus & Kelter, 2006; Ditman et al., 2008; Kelter, Kaup, & Claus, 2004; Radvansky et al., 2003; Rapp & Taylor, 2004; Speer & Zacks, 2005; Zwaan, Madden, & Whitten, 2000) are consistent with the conclusion that a change in time alters the accessibility of events and other features of the described situation. However, whether a change in time is sufficient to affect the accessibility of event information presented earlier may depend upon a number of other factors. One such factor appears to be whether or not changes in narrative time are accompanied by location shifts. For example, Rapp and Taylor (2004) found that the passage of narrative time affected response times to probes of material read earlier when accompanied by location shifts, but not otherwise (for a related finding, see Levine and Klin, 2001).

A second factor that is particularly relevant to the experiments reported here is whether or not an activity can reasonably be assumed to still be ongoing when a time shift is introduced. For example, if the protagonist is viewing a movie, *after six hours* marks a discontinuation of the activity whereas if the protagonist has gone for a day of skiing, the same phrase does not necessarily indicate that the activity has stopped. In the movie example, information that is temporally bound to that activity should be out of focus and thus more difficult to retrieve following *after six hours* than if only an hour had passed. On the other hand, in the skiing example prior information may still be in focus and therefore readily accessible even following *after six hours*.

Zwaan et al. (2000) tested this proposal directly with a recognition probe task. They used sentence pairs similar to those in (2a), where it was plausible that the event mentioned in the first sentence (e.g., *watching a football game*) would still be going on after an hour and the second sentence introduced a shift in time of one hour or objectively no shift in time. Zwaan et al. also manipulated how the events themselves were introduced; they were either described as ongoing (*was watching*, as in 2a) or were explicitly discontinued (*stopped watching*, as in 2b).

(2a) Wayne *was watching* a football game. *After an hour / An hour later / A moment later*, he made a sandwich

(2b) Wayne *stopped watching* a football game. *After an hour / An hour later / A moment later*, he made a sandwich

As Zwaan et al. (2000) had predicted, recognition response times to the events (*watching*) were elevated in the Shift conditions when they had been discontinued (*stopped watching*) but were similar for the Shift and No-shift conditions when they had been described as

ongoing (*was watching*). These results are consistent with models that hold that prior information is kept accessible when it is relevant to the described situation, but is deactivated when it is not (see the Event-Indexing Model, Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998 and the Structure Building Framework, Gernsbacher, 1990).

A third variable that is potentially relevant for conclusions about the effects of temporal changes is how the activation of information is measured. Whereas in several investigations a secondary task such as responding to a recognition probe was used (e.g., Kelter et al., 2004, Experiment 2; Radvansky et al., 2003; Rapp & Taylor, 2004; Zwaan, 1996; Zwaan et al., 2000), in a smaller number of studies the dependent measure was ease of processing an anaphoric reference to an event and thus more closely resembled natural reading (e.g., Anderson et al., 1983; Kelter et al., 2004, Experiments 1 and 3; Speer & Zacks, 2005). Consistent with the conclusions from the recognition probe task, there is some evidence that a change in time affects the comprehension of anaphoric references to event information. For example, Kelter et al. (2004) found that reading times for a sentence containing an anaphoric reference to event information were elevated when the text signalled a larger change in time. However, the change in time was accompanied by a shift in location from the antecedent event. Thus, it is not clear whether information was found to be less accessible following a larger change in time because of the additional passage of time alone, or if it was the additional passage of time spent away from the original location that was critical.

Even assuming that a change in location is not a critical factor in the effects, the evidence is only mixed that changes in narrative time influence the accessibility of information when anaphor processing is the dependent measure. For example, Anderson et al. (1983) failed to find a significant difference between reading times for sentences that contained a pronominal reference to scenario-dependent characters (*projectionist* in a movie scenario) as a function of the length of the time shift (e.g., *Seven hours later* versus *Ten minutes later*). Furthermore, in Ditman et al. (2008), although event information was found to be less accessible following a large time shift (*one year* versus *one second*), in contrast to some other findings (e.g., Zwaan, 1996; Zwaan et al., 2000), event information appeared to remain accessible following a moderate time shift (*one hour*).

Another issue in interpreting the effects of temporal changes is related to the suggestion (Speer & Zacks, 2005) that longer anaphor reading times following a time shift might reflect a general carryover in processing load, rather than a reduction in the accessibility of the antecedent information. Speer and Zacks (2005) found that following a moderate time shift (*An hour later* versus *A moment later*), reading times were similarly elevated for sentences that did or did not contain an anaphoric reference to a preceding object. Importantly, this raises the question of whether a similar carryover in processing load might have been responsible for some of the findings from other studies.

In view of the issues raised, the goal of the current research was to further investigate the effects of manipulations of narrative time during online comprehension. To accomplish this we used a design similar to that used by Zwaan et al. (2000), such that a critical event was explicitly discontinued or ongoing (e.g., Maurice *stopped* versus *was* making additions to an unfinished sculpture) and was followed by an adverbial phrase that denoted a time shift of *one hour* or *one moment* (see the sample passages in Table 1). In Experiment 1 the event information was referred to with an anaphoric reference (e.g., the *piece*). Participants' eye movements were measured while they read the passages and reading times were examined on the adverbial phrase, the anaphor, and the region in between. This enables us to disentangle a potential carryover in processing load from a difference in memory

accessibility. In Experiment 2 a recognition probe task was used with passages similar to those used in Experiment 1. This allows us to compare our eye tracking results with those obtained with recognition responses, the measure of activation that has most frequently been used in previous investigations of changes in narrative time.

If changes in narrative time signal event boundaries between consecutive episodes of activity in the text, as earlier findings suggest, these boundaries should take time to be integrated with readers' ongoing mental representation. Thus, in Experiment 1 processing times on the adverbial phrase should be longer when there is a time shift of *one hour* compared to when there is relatively no time shift (*one moment*), at least for discontinued events. If event boundaries lead to the deactivation of information that is no longer relevant to the described situation, then anaphor reading times (Experiment 1) and recognition probe response times (Experiment 2) should be longer following a time shift when the events are explicitly discontinued but not when they are described as ongoing.

Experiment 1

The main purpose of Experiment 1 was to examine the effects of temporal discontinuities and time shifts during on-line comprehension. Passages similar to the two examples in Table 1 were used. Based on previous findings, a time shift should lead to an increased processing load, which would be indicated by longer fixation times on the adverbial phrase in the Hour condition than in the Moment condition. If time shifts serve a general role, signalling that there is going to be (or might be) a shift from one episode of activity to another, they should lead to an increased processing load regardless of whether the preceding event is discontinued or ongoing. However, if the predicted increase in processing load instead reflects the additional effort required to initiate a new substructure (Gernsbacher, 1990) or new event node (Zwaan et al., 1995), time shifts should lead to a larger processing load in the Discontinued condition than in the Ongoing condition.

Mean fixation times on the anaphor were used as a measure of the accessibility of the critical event antecedent. If a discontinued event is deactivated following a time shift, an anaphoric reference to that event should take more time to resolve. If this is the case, for the Discontinued conditions fixation times in the anaphor region should be longer in the Hour condition than in the Moment condition. In contrast, for the Ongoing conditions, fixation times in the anaphor region should be similar for the Hour and Moment conditions.

An additional goal of Experiment 1 was to determine whether the distance between an anaphor and its event antecedent moderates the effects of temporal changes on anaphor processing. To accomplish this the anaphor was separated from its event antecedent by an average of either 17 words (Experiment 1a) or 9 words (Experiment 1b). If the greater distance between the anaphor and its antecedent in Experiment 1a results in a general reduction in the accessibility of the antecedent, its retrieval time might be elevated regardless of whether a temporal change occurs. If this is the case, then the predicted effects of the time shift and continuity variables on anaphor fixation times should be weaker in Experiment 1a than in Experiment 1b. Addressing this question might shed light on some of the apparent inconsistencies in the literature regarding the effects of temporal changes on ease of anaphor processing.

Method

Participants—One-hundred-and-eleven Native English speaking students from the University of Massachusetts community participated (52 participants in Experiment 1a; 59 participants in Experiment 1b) in exchange for \$8 or 2 extra credit points in a psychology course. All participants had normal or corrected-to-normal vision. The data from 23

participants were eliminated because their response accuracy was less than 70% for basic comprehension questions that followed the passages (four participants in Experiment 1a; three participants in Experiment 1b) or because blinks accompanied by large calibration errors occurred throughout the experiment (four participants in Experiment 1a; 12 participants in Experiment 1b).¹ Therefore, the analyses in Experiments 1a and 1b were each based on 44 participants.

Materials and Design

Experiment 1a: There were a total of 24 experimental passages similar in structure to the sample passage in Table 1. The Introduction of each passage foregrounded one or two protagonists. Following this, the Critical Event section contained a description of an event that was referenced by an anaphor that appeared later in the passage. This description conveyed that the event was either discontinued (Discontinued conditions) or ongoing (Ongoing conditions). To help obscure the manipulation, a variety of linguistic expressions were used to describe the critical event (e.g., Maurice *was / stopped* making additions to a sculpture; Joe *was / finished* reading a magazine; Eve *started / finished* working on a charcoal drawing). The critical event description was immediately followed by a temporal adverbial phrase that signaled a time shift of one hour (*After an hour*) or one moment (*After a moment*). The text that appeared after the adverbial phrase was identical across all four conditions and contained a description of another event that was unrelated to the critical event (Maurice decided to get a cup of coffee). This information provided a natural continuation of the passage and never contained an explicit change in physical location. Following this information the critical event was referred to by a category noun-phrase anaphor (e.g., the *piece*) that did not overlap lexically with the event-antecedent (*sculpture*).² The text that appeared between the description of the critical event and the anaphor consisted of approximately 17 words (a range of 12 to 23 words). The anaphors were nouns that represented the critical events, or objects or characters that were associated with the critical events, but not with any of the earlier information in the passage. Finally, one to two subsequent sentences provided a sensible conclusion of the passage.

In addition to the 24 experimental passages there were 24 filler passages. Neither of the temporal adverbials from the experimental passages appeared in any of the filler passages. Four lists were constructed. Each list contained 48 passages: Six passages in each of the four experimental conditions and the 24 filler passages. The experimental passages were rotated within the four conditions in each list by a Latin square design. Eleven participants were randomly assigned to each list, and for each participant the experimental and filler passages were presented in a random order. A Yes/No comprehension question followed 25% of the passages to encourage participants to read carefully and to obscure the purpose of the experiment. Participants responded to each question by pressing a corresponding Yes or No button on a response box. The number of correct “Yes” and “No” responses was equated across the questions. For the experimental passages, these questions did not target any information related to the critical event. Participants had little difficulty answering the questions, with an average accuracy of 91% in Experiment 1a.

Experiment 1b: Modified versions of the 24 experimental passages from Experiment 1a were used (see the sample passage in Table 1). The most significant change to the passages from Experiment 1a involved a reduction in the amount of text that appeared between the description of the critical event and the anaphor, from approximately 17 words (a range of

¹Data files were excluded for which calibration errors occurred within a given analysis region (or on the fixation prior to entering or after leaving the region) on 50% or more of the trials within a single condition.

²In two of the passages used in Experiment 1a there was some lexical overlap between the event-antecedent and its anaphor. This was not the case for any of the antecedent-anaphor pairs in Experiment 1b.

12 to 23 words) to approximately 9 words (a range of 6 to 13 words). In this regard, the distance between the anaphor and the event in Experiment 1b was comparable to that between the probe word and the event in Zwaan et al.'s (2000) materials. A number of other modifications were made to accommodate this change and also to ensure that the category noun-phrase anaphors did not overlap lexically with their antecedents. As a consequence of these modifications the antecedent-anaphor pairs that were used in Experiment 1a were replaced for six of the passages.

Modified versions of the 24 filler passages from Experiment 1a were used. All other aspects of the materials and design were identical to those used in Experiment 1a.³ Participants had little difficulty answering the Yes/No comprehension questions in Experiment 1b, with an average accuracy of 93%.

Procedure and Apparatus—In Experiments 1a and 1b the participants' eye movements were monitored by an SR Research Ltd. Eye-Link II eye tracking system. The eyetracker is a video-based tracking system with two cameras (one for each eye) and two infra-red LEDs that illuminate each eye. The cameras sample pupil location at a rate of 250 Hz (a temporal resolution of 4 ms). The spatial resolution of the eyetracking system is less than one-half a degree of visual angle. Although viewing was binocular, only data from the right eye were used for analysis. The passages were presented in lowercase letters (except when capitals were appropriate) on a 19-in. Princeton LCD color monitor. Prior to the experiment a calibration routine was performed, consisting of a 9-point horizontal calibration grid. Each trial started with a fixation box on the left-hand side of the monitor, the location of which coincided with the location of the first letter of the first word in the passage. Participants' gaze in the fixation box triggered the appearance of the entire passage on a single display. The participants were instructed to read for comprehension so that they would be able to answer questions that would appear after some of the passages. Participants pressed a button on a response box to signal that they had finished reading a passage. Prior to each trial, a drift correction was performed to compensate for minor head movements. The 9-point calibration routine was repeated when deemed necessary by the experimenter. The experiment began with two practice passages, which were immediately followed by the 48 experimental and filler passages.

Analyses—Prior to all of the analyses, paragraphs with blinks in the regions of interest were removed. Following this, fixations less than 80 ms in duration, and within one character of the previous or subsequent fixation, were incorporated into the neighboring fixation. Remaining fixations of less than 40 ms were removed if they were within three characters of the nearest fixation. These procedures led to the removal of 4.7 % of the data in Experiment 1a and 3.2% of the data in Experiment 1b.

Fixation times were analyzed for three regions: the temporal adverbial phrase (*After an hour*, *After a moment*), the post-adverbial region, consisting of one to four words, and the anaphor (*piece* and *painting* in the sample passages from Experiments 1a and 1b, respectively). For each region of interest, we report one measure that reflects early processing, first pass time (the combined fixation time in a region prior to a saccade out of that region, also referred to as gaze duration for measurements of single words), and one measure that reflects later processing, total fixation time (the combined fixation time in a region, including re-fixations). Analyses were conducted with both variability due to subjects (F_1) and variability due to items (F_2) in the error term. An alpha level of .05 was used for all statistical analyses reported throughout this paper. Furthermore, for all analyses,

³Due to a file labeling error in Experiment 1b, 12 participants were assigned to one of the counterbalancing lists and ten were assigned to a second list (eleven participants were assigned to each of the other two lists).

counterbalancing list was included as a factor in the ANOVAs for the items analyses and group membership (i.e., one group of participants was assigned to each list) was included as a factor in the ANOVAs for the subjects analyses. Given the lack of their theoretical relevance, effects for the list and group variables are not reported.

Results

Adverbial Region—This region includes the three words that comprised the temporal adverbial phrase for the Hour and Moment conditions (*After an hour* and *After a moment*, respectively). Means and standard errors for first pass and total times for Experiments 1a and 1b appear in Table 2. To minimize the effect of the difference in length between the Hour and Moment conditions, fixation times were divided by the number of characters in the adverbial region.

Experiment 1a: Participants spent significantly more time reading the adverbial region in the Hour condition than in the Moment condition. The first pass Hour and Moment condition means are 33.53 and 29.19 ms per character, respectively, $F_1(1,40) = 24.32$, $MSE = 34.06$, $p < .01$; $F_2(1,20) = 44.10$, $MSE = 11.56$, $p < .01$. The total time means for the Hour and Moment conditions are 37.08 and 31.96 ms per character, $F_1(1,40) = 41.92$, $MSE = 27.47$, $p < .01$; $F_2(1,20) = 42.52$, $MSE = 16.21$, $p < .01$.

The fixation times in the adverbial region were longer when the protagonist's activity was discontinued than when the activity was ongoing. The mean first pass times for the Discontinued and Ongoing conditions are 32.01 and 30.71 ms per character, respectively. The difference is significant only in the items analysis; $F_1(1,40) = 2.47$, $MSE = 30.20$, $p = .12$; $F_2(1,20) = 5.29$, $MSE = 9.58$, $p = .03$. The mean total times for the Discontinued and Ongoing conditions are 35.49 and 33.56 ms per character. The difference is significant in both the subjects and items analyses; $F_1(1,40) = 4.39$, $MSE = 37.34$, $p = .04$; $F_2(1,20) = 8.65$, $MSE = 10.93$, $p < .01$.

None of the interaction tests were significant. For the first pass measure, $F_1(1,40) = .66$, $MSE = 30.90$, $p = .42$; $F_2(1,20) = 1.62$, $MSE = 10.10$, $p = .22$; for the total time measure, $F_1(1,40) = 1.50$, $MSE = 45.02$, $p = .23$; $F_2(1,20) = 2.82$, $MSE = 16.38$, $p = .11$.

Experiment 1b: Consistent with the pattern from Experiment 1a, participants spent more time reading the adverbial region in the Hour condition than in the Moment condition. For the first pass measure, the difference between the Hour and Moment conditions is significant, 30.71 versus 27.08 ms per character, $F_1(1,40) = 24.48$, $MSE = 23.64$, $p < .01$; $F_2(1,20) = 43.51$, $MSE = 7.84$, $p < .01$. The total time means for the Hour and Moment conditions are 32.19 and 30.59 ms per character. The difference is not significant in the subjects analysis and is only marginally significant in the items analysis; $F_1(1,40) = 2.07$, $MSE = 54.12$, $p = .16$; $F_2(1,20) = 3.31$, $MSE = 23.94$, $p = .08$.

As in Experiment 1a, fixation times in the adverbial region were longer when the protagonist discontinued an activity than when the activity was ongoing. Mean first pass times for the Discontinued and Ongoing conditions are 29.97 and 27.82 ms per character, respectively. The difference is significant; $F_1(1,40) = 11.90$, $MSE = 17.12$, $p < .01$; $F_2(1,20) = 8.25$, $MSE = 13.50$, $p < .01$. The mean total times for the Discontinued and Ongoing conditions are 32.29 and 30.49 ms per character. The difference is significant in the subjects analysis and is marginally significant in the items analysis; $F_1(1,40) = 4.65$, $MSE = 30.29$, $p = .04$; $F_2(1,20) = 4.09$, $MSE = 16.53$, $p = .06$.

None of the interaction tests were significant. For the first pass measure, both F s were less than one; for the total time measure, $F_1(1,40) = .97$, $MSE = 38.47$, $p = .33$; $F_2(1,20) = 1.50$, $MSE = 19.12$, $p = .24$.

Combined Analyses: Because the two experiments had the same design and number of participants, for each of the three regions we did further analyses combining the two data sets with experiment as a between-participants factor. For all of the dependent measures, the mean fixation times in Experiment 1a were significantly longer (or marginally so) than those in Experiment 1b. This main effect of experiment is most likely due to differences in reading speed among the participants across the two experiments. For the post-adverbial region, it is likely that the length difference also played a role (the post-adverbial region was slightly larger in Experiment 1a). Given the lack of its theoretical relevance, we do not devote additional attention to this main effect of experiment. Unless stated otherwise, none of the possible interactions involving experiment were significant in any of the combined analyses for the three regions.

Averaging over the two experiments, both first pass and total times in the adverbial region were significantly longer in the Hour than in the Moment condition: for first pass fixation time, $F_1(1,80) = 48.37$, $MSE = 28.85$, $p < .01$; $F_2(1,40) = 86.85$, $MSE = 9.70$, $p < .01$; for total fixation time, $F_1(1,80) = 24.27$, $MSE = 40.80$, $p < .01$; $F_2(1,40) = 30.78$, $MSE = 20.07$, $p < .01$. Continuity was also a significant factor in the combined analysis; fixation times in the adverbial region were longer in the Discontinued condition than in the Ongoing condition: for first pass fixation time, $F_1(1,80) = 11.10$, $MSE = 23.66$, $p < .01$; $F_2(1,40) = 13.53$, $MSE = 11.54$, $p < .01$; for total fixation time, $F_1(1,80) = 9.00$, $MSE = 33.81$, $p < .01$; $F_2(1,40) = 11.72$, $MSE = 13.73$, $p < .01$. The F s for the time shift $\times$ continuity interaction were less than one in both the first pass and total time analyses.

There was a significant interaction between the time shift variable and experiment in total fixation times on the adverbial region, $F_1(1,80) = 6.66$, $MSE = 40.80$, $p = .01$; $F_2(1,40) = 7.50$, $MSE = 20.07$, $p = .01$. The nature of this interaction indicates that the difference between the Hour and Moment conditions is larger in Experiment 1a than in Experiment 1b.

Post-Adverbial Region—This region includes the two to four words (Experiment 1a) or one to two words (Experiment 1b) that appeared after the adverbial region.⁴ Analyses of these data allow us to test whether the differences observed in the adverbial region carry over beyond that region.

Experiment 1a: There was no evidence that either of the variables affected reading times on the post-adverbial region. For both the first pass and total time measures all F s were less than one in the analyses for the main effect of the time shift variable and in the analyses for the time shift $\times$ continuity interaction. For the continuity variable, both F s were less than one for the first pass measure; for the total time measure, $F_1(1,40) = 2.45$, $MSE = 2697$, $p = .13$; $F_2(1,20) = 1.67$, $MSE = 2414$, $p = .21$.

Experiment 1b: First pass reading times in the post-adverbial region were similar for the Hour and Moment conditions, $F_1(1,40) = .83$, $MSE = 3376$, $p = .37$; $F_2(1,20) = 1.11$, $MSE = 1756$, $p = .31$. Total fixation times in the post-adverbial region were actually *shorter* in the Hour condition (308.35 ms) than in the Moment condition (330.48 ms); $F_1(1,40) = 4.19$,

⁴In both Experiments 1a and 1b the passages were constructed with the constraint that none of the critical regions (adverbial, post-adverbial, anaphor) appear at either of the extremes on a line of text. To satisfy this constraint, it was necessary to reduce the length of the post-adverbial region slightly in Experiment 1b because all three critical regions often appeared in a single line of text in that experiment.

$MSE = 5122$, $p = .05$; $F_2(1,20) = 8.57$, $MSE = 1276$, $p < .01$. The main effect of the continuity variable did not approach significance for the first pass measure (both F s were less than one) or the total fixation measure, $F_1(1,40) = 1.92$, $MSE = 3851$, $p = .17$; $F_2(1,20) = 1.16$, $MSE = 2261$, $p = .29$. There is some indication in Table 2 that fixation times in the post-adverbial region are longer in the Discontinued-Moment condition than in the other three conditions, particularly in total fixation times. However, the time shift $\times$ continuity interaction was not significant for the first pass measure ($F_1(1,40) = 1.78$, $MSE = 3011$, $p = .19$; $F_2(1,20) = 1.18$, $MSE = 1859$, $p = .29$) and was significant only in the subjects analysis for the total time measure; $F_1(1,40) = 5.99$, $MSE = 3614$, $p = .02$; $F_2(1,20) = 3.24$, $MSE = 2632$, $p = .09$.

Combined Analyses: For the combined analyses in the post-adverbial region, there were no reliable effects for the first pass measure; all F s were less than one. For the total time fixation measure there was no effect of the time shift variable ($F_1(1,80) = 1.69$, $MSE = 5252$, $p = .20$; $F_2(1,40) = 2.32$, $MSE = 1887$, $p = .14$). The main effect of the continuity variable was significant only in the subjects analysis; $F_1(1,80) = 4.27$, $MSE = 3274$, $p = .04$; $F_2(1,40) = 2.82$, $MSE = 2337$, $p = .10$. Consistent with analyses from Experiment 1b, with the combined analyses there is some suggestion that total fixation times in the post-adverbial region are longer in the Discontinued-Moment condition than in the other three conditions. However, this interaction trend was only marginally significant, $F_1(1,80) = 3.84$, $MSE = 4273$, $p = .053$; $F_2(1,40) = 3.21$, $MSE = 2364$, $p = .08$.⁵

Anaphor Region—This region includes the noun from the category noun-phrase anaphor (e.g., *piece* and *painting* in the sample passages for Experiments 1a and 1b, respectively). As can be seen in Table 2, the mean first pass and total fixation times in the anaphor region are similar across the four conditions. None of the analyses for the time shift variable approached significance; *Experiment 1a*: all F s < 1 ; *Experiment 1b*, first pass: $F_1(1,40) = .04$, $MSE = 1399$, $p = .85$; $F_2(1,20) = 1.18$, $MSE = 564.78$, $p = .29$; total time: both F s < 1 ; *Combined analyses*, first pass: $F_1(1,80) = .24$, $MSE = 1778$, $p = .63$; $F_2(1,40) = 1.53$, $MSE = 807.81$, $p = .22$; total time: $F_1(1,80) = .06$, $MSE = 3173$, $p = .81$; $F_2(1,40) = 1.10$, $MSE = 1380$, $p = .30$. Similarly, none of the analyses for the continuity variable approached significance; *Experiment 1a*, first pass: $F_1(1,40) = .96$, $MSE = 1631$, $p = .33$; $F_2(1,20) = 1.43$, $MSE = 733.71$, $p = .25$; total time: $F_1(1,40) = .28$, $MSE = 4263$, $p = .60$; $F_2(1,20) = 1.01$, $MSE = 939.13$, $p = .33$; *Experiment 1b*, first pass: $F_1(1,40) = 1.10$, $MSE = 1082$, $p = .30$; $F_2(1,20) = .26$, $MSE = 453.77$, $p = .62$; total time: $F_1(1,40) = 1.74$, $MSE = 1668$, $p = .20$; $F_2(1,20) = 1.31$, $MSE = 862.88$, $p = .27$; *Combined analyses*: all F s < 1 . Finally, none of the analyses for the time shift $\times$ continuity interaction were significant; *Experiment 1a*, first pass: $F_1(1,40) = 1.09$, $MSE = 2387$, $p = .30$; $F_2(1,20) = 1.46$, $MSE = 629.45$, $p = .24$; total time: both F s < 1 ; *Experiment 1b*: all F s < 1 ; *Combined analyses*, first pass: $F_1(1,80) = 1.91$, $MSE = 1707$, $p = .17$; $F_2(1,40) = 1.35$, $MSE = 768.18$, $p = .25$; total time: $F_1(1,80) = 1.40$, $MSE = 3212$, $p = .24$; $F_2(1,40) = .91$, $MSE = 1189$, $p = .35$.

Of particular relevance to our secondary goal of determining whether the distance between the anaphor and its antecedent moderates the effects of temporal changes on anaphor processing, for the combined analyses, none of the interactions involving the experiment factor were significant (all p s $> .16$).

⁵As another test of whether the differences observed in the adverbial region carry over beyond that region, the duration of the first fixation after leaving the adverbial region (i.e., spillover times) was analyzed for both experiments. Only in Experiment 1b, spillover times were 9.52 ms *shorter* in the hour conditions than in the moment conditions, and then only when tested against subject variability, $F_1(1,40) = 4.72$, $MSE = 3971$, $p < .05$; $F_2(1,20) = 2.87$, $MSE = 679$, $p = .11$. This reversal from the fixation patterns observed on the adverbial region makes no sense under any account we can construct and is most likely a product of chance variation.

Supplemental Analyses: To further explore the potential effects of temporal changes on anaphor resolution we examined the probability of skipping the anaphor region. Skipping rates for this region did not differ significantly across conditions, although they were marginally lower in the Hour condition (13.29 %) than in the Moment condition (18.08 %) in Experiment 1b; $F_1(1,40) = 3.94$, $MSE = 255.00$, $p = .054$; $F_2(1,20) = 3.28$, $MSE = 171.19$, $p = .09$. A similar trend was observed in the combined analyses; skipping rates were 15.51 % and 18.73 % in the Hour and Moment conditions, respectively; $F_1(1,80) = 3.24$, $MSE = 281.06$, $p = .08$; $F_2(1,40) = 3.21$, $MSE = 181.29$, $p = .08$. There were no other significant effects in any of the analyses for the individual experiments or in the combined analyses (all F s < 1). For the combined analyses, none of the interactions involving the experiment factor were significant (all F s < 1).

Given that none of the analyses for the fixation time data in the anaphor region approached significance and that the effect of the time shift variable on skipping rates was only marginal, we conducted post hoc power analyses using the program *G*Power* (Faul, Erdfelder, Lang, & Buchner, 2007) to confirm that the design of Experiment 1 had enough power to detect main effects of the time shift and continuity variables and their interaction. For both the subjects and items analyses, Experiment 1 had a power greater than .93 to reject the null (two-tailed test, $\alpha = .05$) for a medium effect size (.25) as defined by Cohen's (1988) standards.

Discussion

The patterns of fixation times for the adverbial region from Experiments 1a and 1b lead to two straightforward conclusions. First, reading times are longer for adverbial phrases that signal a larger time shift (*After an hour* vs. *After a moment*). Thus, consistent with other findings (Anderson et al., 1983; Radvansky & Copeland, 2010; Radvansky et al., 2003; Speer & Zacks, 2005; Zwaan, 1996), the current results from the eye tracking methodology indicate that time shifts lead to increases in processing load. The second conclusion is based on the finding that compared to when an event is described as ongoing, reading times are longer for adverbial phrases that follow discontinued events. In other words, the discontinuation of an event leads to a processing load on a subsequent adverbial phrase. Although this processing load might have emerged on the adverbial phrase itself, an additional possibility is that it reflects a carryover in processing from the event description.⁶ Regardless of which of these alternatives is true, the main conclusion is the same: the explicit discontinuation of an event leads to a processing load.

Because there was no evidence of an interaction between the time shift and continuity variables we can conclude that time shifts and event discontinuations have additive effects on the processing load of a temporal phrase. Interestingly, there is some evidence that situational discontinuities from different dimensions (time, causation, goal, character) have additive effects on reading times (Zwaan, Radvansky, Hilliard, & Curiel, 1998). The current results suggest that this conclusion generalizes to circumstances where multiple changes occur within the same dimension.

Turning to the findings for the post-adverbial region, the only fully significant effect was in Experiment 1b, where total fixation times were *shorter* in the hour conditions than in the moment conditions. Given this, we can be confident that the processing load observed on the adverbial phrase itself did not carry over to the subsequent text.

⁶Although an analysis of fixation times on the event descriptions themselves might address this question, these descriptions consisted of several words that were often distributed across two lines of text. Given the inherent variability in fixation times for words that appear at the beginning and at the end of a line of text, the outcome of such an analysis would be difficult to interpret.

For the anaphor region, there were no significant differences among the condition means for the first pass or total fixation time analyses. This finding is particularly striking in Experiment 1b, where the anaphor was separated from its antecedent by only nine words. If in Experiment 1a the lengthier material that appeared between the critical event antecedent and the anaphor led to a general reduction in the accessibility of the antecedent, delaying its reinstatement across all of the conditions, this should have been less likely in Experiment 1b, where the distance was substantially smaller.

Although the anaphor region was skipped slightly less often in the Hour condition than in the Moment condition in Experiment 1b, this difference was not significant in Experiment 1a and was only marginally significant in Experiment 1b and in the combined analyses. Furthermore, there was no hint of an effect of the continuity variable or an interaction between it and the time shift variable in any of the analyses for the anaphor region. Thus, based on the findings from Experiment 1, it is reasonable to conclude that event information remains accessible following a time shift, even when the events are discontinued.

Given that time shifts have been shown to lead to the deactivation of event information elsewhere (e.g., Anderson et al., 1983; Ditman et al., 2008; Radvansky et al., 2003; Zwaan, 1996; Zwaan et al., 2000), we next consider possible explanations for why this did not occur in Experiments 1a and 1b. Perhaps the simplest possibility is that in both experiments, participants failed to reinstate the antecedent. This is extremely unlikely. Using a variety of measures, many published studies have reported evidence of reinstatement of categorical antecedents that were separated from the anaphor by at least as much material as in our experiments (e.g., Dell, McKoon, & Ratcliff, 1983; McKoon & Ratcliff, 1980; O'Brien, Duffy, & Myers, 1986; Wiley, Mason, & Myers, 2001).

One obvious difference between those studies reporting the deactivation of event information preceding a temporal shift and Experiment 1 is in the way activation was measured. The conclusions from Zwaan et al. (2000) and several other reports of deactivation were based on recognition probe response times; that is, participants were presented with a probe word and had to indicate whether that word had appeared in the text. Why might a recognition probe task be more sensitive to the effects of temporal changes than ease of anaphor comprehension? One possibility is based on evidence (Keenan, Baillet, & Brown, 1984; O'Brien et al., 1986) that reconstructive processes may underlie recognition responses. According to this view, a post-comprehension probe task might induce readers to retrospectively consider an event's relevance to the situation described by the preceding text in ways that an anaphoric reference does not. The purpose of Experiment 2 was to address whether the discrepancy between the results from Experiment 1 and those from some earlier investigations of temporal changes might be due in part to the nature of the dependent measure used.

Experiment 2

The design of Experiment 2 was similar to that of Experiment 1 except that the anaphor sentence was eliminated and a recognition probe task was included, where the probe word represented the critical event. If the effects of temporal changes on the accessibility of event information are task-dependent as suggested previously, then in contrast to the lack of a reading time difference on the anaphor region in Experiment 1, we should observe an effect of our variables on recognition probe response times.

Method

Participants—Thirty-six Native English speaking students from the Hofstra University community participated as part of a course requirement. The data from one participant who

had an accuracy score of 53% for the comprehension questions was eliminated and the data from three additional participants were eliminated because they incorrectly responded “No” to the recognition probe on 50% or more of the trials within a condition. Therefore, the analyses were based on 32 participants.

Materials and Design—Modified versions of the 24 experimental passages from Experiment 1 were used. A sample passage appears in Table 3. The most significant change involved the removal of the sentence containing the anaphoric reference to the event information. For each passage the description of the critical event was immediately followed by the sentence containing the temporal adverbial phrase, which was in turn followed by a recognition probe word. The average length of the adverbial sentence was 7 words (a range of 6 to 8 words). In this regard, the distance between the probe word and the event in Experiment 2 is comparable to the distance in Zwaan et al.’s (2000) materials as well as the distance between the anaphor and the event in Experiment 1b. Similar to Zwaan et al.’s materials, each probe word was an explicit repetition of the verb from the critical event in the preceding passage (*painting* in the sample passage). For several of the passages additional changes were made to those used in Experiment 1 to avoid the use of “light” verbs that have little meaning on their own as the probe word (e.g., *making*).

In addition to the 24 experimental passages there were 36 filler passages: the 24 filler passages from Experiment 1 and 12 new filler passages. Neither of the temporal adverbials from the experimental passages appeared in any of the filler passages. Each filler passage was followed by a recognition probe word. For ten of the filler passages the probe task required a correct “Yes” response. For the remaining 26 fillers, although the probe word did not appear in the passage, it was related to the storyline of the passage. This was intended to prevent participants from making a rapid judgment based on mere familiarity. To help obscure the experimental manipulations, the probed information always appeared in the beginning of the filler passages. Furthermore, for both the “Yes” and “No” filler passages, one-half of the probe words were nouns and the other half were verbs. All other aspects of the design were identical to those from Experiment 1. Participants had little difficulty answering the comprehension questions, with an average accuracy of 91%.

Procedure and Apparatus—The experimental and filler passages were presented with E-Prime (version 1.1, Psychology Software Tools, Inc., 2002). Participants were instructed to read each passage for comprehension at their own pace. Each trial began with the presentation of the word READY in the center of the screen. Participants were instructed to press the advance button when they were ready to start the next trial. Similar to the procedure used in Zwaan et al. (2000), each press of the advance button caused the current line of text to be erased and the next to be presented. Immediately after the final line of the passage (which was always the temporal adverbial sentence), pressing the advance button caused the final line to be replaced by the recognition probe word in capital letters. Participants were instructed to indicate, as quickly and accurately as possible, whether the probe word had appeared in the preceding passage by pressing a “Yes” or a “No” button on the response box. The experiment began with four practice passages, which were immediately followed by the 60 experimental and filler passages.

Results

The accuracy scores for the recognition probe task ranged between 91% and 94% and did not differ across conditions (all $ps > .41$). The mean response times for correct probe responses appear in Table 4. Prior to the analyses response times greater than 3500 ms were removed. This led to the removal of less than 1% of the data. The outcome of the analyses revealed that response times were significantly longer for probe words that represented

discontinued events (1311 ms) than those that represented ongoing events (1221 ms), $F_1(1,28) = 8.32$, $MSE = 31157$, $p < .01$; $F_2(1,20) = 6.91$, $MSE = 41231$, $p = .02$. There was no evidence that the shift variable affected response times (both $F_s < 1$) or that it interacted with the continuity variable, $F_1(1,28) = 1.88$, $MSE = 22250$, $p = .18$; $F_2(1,20) = 1.01$, $MSE = 28505$, $p = .60$. (Insert Table 4 about here)

Discussion

In contrast to the results from Experiments 1a and 1b, where temporal changes did not significantly affect fixation times in the anaphor region, in Experiment 2, recognition probe response times were significantly longer for discontinued events than for ongoing events. In other words, the continuity of an event was influential when the measure of its activation was recognition probe response times but not ease of anaphor comprehension. Thus, the combined pattern of results from Experiments 1 and 2 suggests that with investigations of the effects of temporal changes on event accessibility, the dependent measure can have a critical impact on the conclusions.

Although the expected interaction between the continuity and time shift variables was not observed in Experiment 2, it is interesting to note that there is a suggestive trend that is similar to the pattern reported in Zwaan et al. (2000, Experiment 1). That is, in both experiments the mean response time was longer in the Discontinued-Hour condition than in the Discontinued-Moment condition, which was in turn longer than the mean response times in either of the Ongoing conditions (and in both experiments, the mean response times for the latter two conditions were similar). It is also noteworthy that in Zwaan et al. (2000) neither the main effect of the time shift variable nor the interaction between the variables was fully significant. Taking these points into consideration, it appears that although recognition probe response times are sensitive to whether an event is described as discontinued or ongoing, the nature of the interaction between the continuity variable and that of the time shift variable is somewhat elusive.

General Discussion

In the current experiments we examined how changes in narrative time influence reading comprehension. In Experiments 1a and 1b a change in time, marked by a time shift or the discontinuation of an event, led to longer fixations on a temporal adverbial phrase. This finding adds to a growing body of evidence (Anderson et al., 1983; Radvansky & Copeland, 2010; Radvansky et al., 2003; Speer & Zacks, 2005; Zwaan, 1996) that temporal changes lead to increases in processing load. This conclusion is consistent with the general claim that readers perceive temporal changes as event boundaries between consecutive episodes of activity in the text. For example, according to the Event Segmentation Theory (Zacks, Speer, & Reynolds, 2009; Zacks, Speer, Swallow, Braver, & Reynolds, 2007) readers maintain an “event model” in working memory that guides the perception of incoming information. To the extent that incoming text information is predictable based upon the most recently constructed event model, that model remains active. However, when the incoming text signals an unexpected change (e.g., a change in time or location), an event boundary is perceived and the model in working memory is updated. According to the Event Segmentation Theory, elevated reading times at event boundaries reflect these updating processes.

The failure to observe an effect of changes in narrative time on anaphor resolution in Experiment 1 indicates that event boundaries do not necessarily lead to the deactivation of prior information. This conclusion is at odds with the Event Segmentation Theory and other accounts (e.g., Event-Indexing Model, Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998 and Structure Building Framework, Gernsbacher, 1990) that assume that

information presented prior to an event boundary is more difficult to retrieve following that boundary. However, it is noteworthy that the general pattern obtained in Experiment 1—a difference in processing load on an adverbial phrase combined with the lack of a difference in the accessibility of information—has been reported in two other recent investigations (Ditman et al., 2008; Radvansky & Copeland, 2010). In Radvansky and Copeland's (2010) study, while time shifts (e.g., *a day later* versus *a moment later*) led to longer sentence reading times, they did not affect the accessibility of previously mentioned objects, even when the objects were no longer relevant to the described situation after the time shift. In an ERP investigation, Ditman et al. (2008) observed a larger N400 amplitude when readers processed an adverbial phrase that denoted a moderate time shift (*one hour*) compared to one that denoted objectively no shift in time (*one second*). In contrast, for these two conditions there was no difference in the N400 amplitude when readers processed an anaphoric reference to information that preceded the adverbial phrase, which suggests that the time shift of one hour did not affect the activation of information. Because there were several methodological differences between each of these investigations and Experiment 1, it is risky to generate detailed conclusions based on the similarities in the findings. The important point is that despite the methodological differences, the results from three independent investigations demonstrate that a shift in time can lead to a processing load without deactivating preceding information. The results from Experiment 1 indicate that this conclusion can be generalized to circumstances where the preceding events are explicitly discontinued.

In Experiment 2 recognition response times were longer for probes that represented discontinued events compared to those that represented ongoing events. This result is consistent with theoretical accounts that assume that information presented prior to an event boundary is more difficult to retrieve following that boundary. However, in light of the findings from Experiment 1, the influence of an event boundary on the event's activation appears to depend in part on the nature of the task. What might account for this task difference? We have already described one explanation, which is anchored in the claim (Keenan et al., 1984; O'Brien et al., 1986) that reconstructive processes may underlie recognition probe responses. Given this, the post-comprehension probe task might have induced readers to retrospectively consider an event's relevance to the situation described by the preceding text in ways that an anaphoric reference does not.⁷ This conclusion is consistent with the claim that offline and online tasks encourage readers to focus on different features of the text (Rapp & Mensink, 2011).

A second feature that might have contributed to the task difference observed in these experiments is related to the manner in which the text was presented. In contrast to Experiment 1, where the entire paragraph was present on the screen, in Experiment 2, the text was presented one line at a time. This could have induced processes that do not occur under ordinary circumstances. More generally, the combination of the line-by-line text presentation and the probe task could have invited strategic processes that are absent during natural reading.

Another notable methodological difference between Experiments 1 and 2 is that whereas the anaphors in Experiment 1 were nouns that often referred to objects associated with the events (e.g., the *piece* in Table 1), the probes used in Experiment 2 (and in Zwaan et al., 2000) were always action words that represented the events themselves (e.g., *painting* in

⁷Although this conclusion is plausible for the recognition probe task, we note that it is not clear whether it would generalize to comparisons with all of the other measures of accessibility that have been used in the literature, such as naming, sentence completion, and question accuracy (e.g., Ferretti, Kutas, & McRae, 2007; Ferretti, Rohde, Kehler, & Crutchley, 2009; Magliano & Schleich, 2000). While it is reasonable to assume that these tasks might induce post-comprehension processes similar to those that are a concern with the recognition probe task, additional research is needed to be fully confident of this.

Table 2). This leaves open the possibility that although discontinued events are deactivated, and regardless of whether this is a general phenomenon or one that is linked to time shifts as concluded by Zwaan et al. (2000), some *features* of the events remain accessible (see Ferretti, Kutas, & McRae, 2007 for a related proposal about the role of verb aspect on the activation of events and their associated features). Consistent with this general claim, in Radvansky and Copeland's (2010) recognition probe investigation, a large time shift (*a day later*) led to the deactivation of probed activities but not objects that should have been removed from the foreground. However, in some other recognition probe investigations, temporal changes were found to deactivate objects (e.g., Kelter et al., 2004; Speer & Zacks, 2005). Clearly additional experimentation is needed to fully examine this proposal.

In the Introduction we suggested that changes in narrative time might interact with a number of text variables to affect comprehension. To work toward isolating the effects of changes in narrative time, they were never accompanied by location shifts in the current experiments. However, our manipulation of narrative time was not completely isolated from changes in other dimensions. In the passages used in the current experiments, discontinuing an event always interrupted a character's goal plan (e.g., Maurice stopped working toward his goal of completing the art work). It remains to be determined whether discontinued events would still be perceived as event boundaries when goal plans are maintained. In the context of narrative film, there is some evidence (Magliano & Zacks, 2011) that when a change in space and/or time occurs across an editing break in a film, viewers are more likely to perceive an event boundary when the portrayed actions are not part of a continuous goal plan than when they *are* part of a continuous goal plan. In light of these findings, it is reasonable to hypothesize that a discontinued event in a narrative text would no longer be perceived as an event boundary when there is continuity in a character's goal plan.

In summary, the results of the experiments reported here lead to two main conclusions about how changes in narrative time affect text comprehension. First, the fixation patterns on the temporal adverbial phrase contribute to a growing body of evidence that a change in time signals an episodic boundary and that it takes time for readers to incorporate such a boundary into their mental representation of the text. Because the reading process is more natural with the eye movement method than with that involved in pressing a button to bring on each sentence, or with that involved in making a post-comprehension judgment, the results from Experiment 1 increase our confidence that these conclusions generalize to online comprehension. Second, the results on the anaphor region in Experiments 1a and 1b add to a smaller body of recent evidence (e.g., Ditman et al., 2008; Radvansky & Copeland, 2010) that at least some event information remains accessible following a moderate time shift, even when the event is no longer ongoing. Clearly more research is needed to understand the accessibility of event information following changes in narrative time. Although we can conclude there are methodological differences, we need to formulate and test hypotheses about the underlying processes that contribute to these differences. Furthermore, although we have demonstrated effects of temporal discontinuities during natural reading with the use of eye movement measures, we need to disentangle those effects from possible effects of goal discontinuities. In summary, despite the results of these experiments and the many others cited on the effects of changes in narrative time, there is still much to be done.

Acknowledgments

This research was supported by National Institute of Mental Health Training Grant MH16745 awarded to the University of Massachusetts, Amherst. Portions of these data were presented at the 46th Annual Meeting of the Psychonomic Society. We would like to thank Kathryn Chace, Matthew Doiron, Kristine Kivel, Meghan Pedersen, and Donald Wood for their assistance with data collection.

References

- Anderson AA, Garrod SC, Sanford AJ. The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*. 1983; 35(3-A):427–440.
- Claus B, Kelter S. Comprehending narratives containing flashbacks: Evidence for temporally organized representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2006; 32(5):1031–1044.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd ed.. Erlbaum; Hillsdale, NJ: 1988.
- Cowles H, Garnham A. Antecedent focus and conceptual distance effects in category noun-phrase anaphora. *Language and Cognitive Processes*. 2005; 20(6):725–750.
- Dell GS, McKoon G, Ratcliff R. The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning & Verbal Behavior*. 1983; 22(1):121–132.
- Ditman T, Holcomb PJ, Kuperburg GR. Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. *Psychonomic Bulletin & Review*. 2008; 15(4): 750–756. [PubMed: 18792500]
- Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*. 2007; 39:175–191. [PubMed: 17695343]
- Ferretti TR, Kutas M, McRae K. Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2007; 33(1):182–196.
- Ferretti TR, Rohde H, Kehler A, Crutchley M. Verb aspect, event structure, and coreferential processing. *Journal of Memory and Language*. 2009; 61(2):191–205. [PubMed: 22690039]
- Gernsbacher, M. *Language comprehension as structure building*. Lawrence Erlbaum Associates, Inc.; Hillsdale, NJ England: 1990.
- Keenan JM, Baillet SD, Brown P. The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning & Verbal Behavior*. 1984; 23(2):115–126.
- Kelter S, Kaup B, Claus B. Representing a Described Sequence of Events: A Dynamic View of Narrative Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30(2):451–464.
- Levine WH, Klin CM. Tracking of spatial information in narratives. *Memory & Cognition*. 2001; 29(2):327–335.
- Magliano JP, Schleich MC. Verb aspect and situation models. *Discourse Processes*. 2000; 29(2):83–112.
- Magliano JP, Zacks JM. The impact of continuity editing in narrative film on event segmentation. *Cognitive Science*. 2011; 35(8):1489–1517. [PubMed: 21972849]
- McKoon G, Ratcliff R. The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning & Verbal Behavior*. 1980; 19(6):668–682.
- O'Brien EJ, Duffy SA, Myers JL. Anaphoric inference during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1986; 12(3):346–352.
- Radvansky GA, Copeland DE. Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2010; 36(1):210–216.
- Radvansky GA, Copeland DE, Berish DE, Dijkstra K. Aging and situation model updating. *Aging, Neuropsychology, and Cognition*. 2003; 10(2):158–166.
- Rapp, DN.; Mensink, MC. Focusing effects from online and offline reading tasks. In: McCrudden, MT.; Magliano, JP.; Schraw, G., editors. *Text relevance and learning from text*. IAP Information Age Publishing; Charlotte, NC US: 2011. p. 141-164.
- Rapp DN, Taylor HA. Interactive Dimensions in the Construction of Mental Representations for Text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30(5):988–1001.
- Speer NK, Zacks JM. Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language*. 2005; 53(1):125–140.

- Wiley J, Mason RA, Myers JL. Accessibility of potential referents following categorical anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27(5):1238–1249.
- Zacks JM, Speer NK, Reynolds JR. Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*. 2009; 138(2):307–327. [PubMed: 19397386]
- Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR. Event perception: A mind-brain perspective. *Psychological Bulletin*. 2007; 133(2):273–293. [PubMed: 17338600]
- Zwaan RA. Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1996; 22(5):1196–1207.
- Zwaan RA, Langston MC, Graesser AC. The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*. 1995; 6(5):292–297.
- Zwaan RA, Madden CJ, Whitten SN. The presence of an event in the narrated situation affects its availability to the comprehender. *Memory & Cognition*. 2000; 28(6):1022–1028.
- Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. *Psychological Bulletin*. 1998; 123(2):162–185. [PubMed: 9522683]
- Zwaan RA, Radvansky GA, Hilliard AE, Curiel JM. Constructing multidimensional situation models during reading. *Scientific Studies of Reading*. 1998; 2(3):199–220.

Table 1
Sample Passage from Experiment 1a

INTRODUCTION: Maurice was getting ready for the grand opening of his new art gallery. The opening day was quickly approaching and he still had a lot to get done. First he called the caterers. He needed to confirm the cocktail menu. Then he went to do some work in his studio.

CRITICAL EVENT

Discontinued conditions: By noontime he stopped making additions to an unfinished sculpture.

Ongoing conditions: By noontime he was making additions to an unfinished sculpture.

ADVERBIAL PHRASE

Hour conditions: *After an hour*, he decided to get a cup of coffee.

Moment conditions: *After a moment*, he decided to get a cup of coffee.

ANAPHOR: He thought that he could probably finish the *piece* after his break.

CONCLUSION: Maurice finally felt as though he was close to being ready for the opening day.

Sample Passage from Experiment 1b

INTRODUCTION: Maurice was getting ready for the grand opening of his new art gallery. The opening day was tomorrow and he still had much to get done. First he called the caterers to confirm the cocktail menu. Then he walked over to his studio to get some work done.

CRITICAL EVENT

Discontinued conditions: By noon, he had finished making additions to a portrait of a peasant girl.

Ongoing conditions: By noon, he was making additions to a portrait of a peasant girl.

ADVERBIAL PHRASE

Hour conditions: After an hour, he looked over

Moment conditions: After a moment, he looked over

ANAPHOR: the *painting* to see if there were any flaws.

CONCLUSION: Then, after one final touch, Maurice was done. He finally felt as though he was close to being ready for the opening.

Note: regions of interest appear in italics

Table 2
Mean (and Standard Error) First Pass and Total Time Durations as a Function of Region
for Experiments 1a and 1b

Experiment 1a				
Measure	Region			
		Adverbial	Post-Adverbial	Anaphor
First Pass	DH	33.84 (1.43)	347.96 (14.60)	247.24 (9.29)
	DM	30.18 (1.32)	347.01 (15.64)	236.29 (8.69)
	OH	33.22 (1.18)	344.77 (13.96)	233.60 (6.44)
	OM	28.20 (1.25)	345.55 (13.44)	238.00 (6.86)
Total Time	DH	37.43 (1.30)	393.53 (14.71)	275.82 (12.84)
	DM	33.55 (1.31)	396.64 (16.86)	263.05 (11.03)
	OH	36.73 (1.30)	386.42 (13.98)	262.54 (9.88)
	OM	30.38 (1.27)	379.27 (14.68)	265.97 (9.86)

Experiment 1b				
Measure	Region			
		Adverbial	Post-Adverbial	Anaphor
First Pass	DH	31.92 (1.25)	286.28 (10.51)	219.53 (7.02)
	DM	27.72 (1.13)	304.43 (11.63)	213.48 (5.84)
	OH	29.29 (0.93)	294.71 (11.73)	219.37 (5.97)
	OM	26.16 (1.06)	289.29 (12.78)	223.32 (5.71)
Total Time	DH	33.42 (1.32)	302.51 (10.71)	235.00 (7.81)
	DM	30.78 (1.53)	345.12 (16.97)	230.12 (7.86)
	OH	30.71 (0.94)	312.70 (12.45)	236.61 (9.41)
	OM	30.02 (1.27)	310.28 (14.71)	243.99 (8.40)

Note: Fixation times for the Adverbial region are reported in milliseconds per character. Fixation times for the Post-adverbial and the Anaphor regions are reported in milliseconds. DH = *Discontinued-Hour*; DM = *Discontinued-Moment*; OH = *Ongoing-Hour*; and OM = *Ongoing-Moment*.

Table 3
Sample Passage from Experiment 2

INTRODUCTION: Maurice was nervous about the opening of his new art gallery.

The big day was quickly approaching and he had a lot to do.

First he called the caterers to confirm the cocktail menu.

Then he went to do some work in his studio.

CRITICAL EVENT

Discontinued conditions: Late in the afternoon he finished painting a portrait.

Ongoing conditions: Late in the afternoon he was painting a portrait.

ADVERBIAL PHRASE

Hour conditions: *After an hour*, he heard a siren outside.

Moment conditions: *After a moment*, he heard a siren outside.

PROBE: PAINTING

Table 4
Mean Correct Recognition Probe Response Times (in milliseconds) for Experiment 2

	Mean	Standard Error
DH	1338.68	62.03
DM	1282.65	58.30
OH	1212.57	49.10
OM	1228.79	49.09

Note: DH = *Discontinued-Hour*; DM = *Discontinued-Moment*; OH = *Ongoing-Hour*; and OM = *Ongoing-Moment*.