BRIEF REPORTS

Time travel through language: Temporal shifts rapidly decrease information accessibility during reading

TALI DITMAN AND PHILLIP J. HOLCOMB
Tufts University, Medford, Massachusetts

AND

GINA R. KUPERBERG
Tufts University, Medford, Massachusetts
and Massachusetts General Hospital, Charlestown, Massachusetts

The present study dissociated the immediate neural costs from the subsequent neural consequences of integrating time shifts into our mental representations of events. Event-related potentials were recorded as participants read scenarios that included words referring to short temporal shifts (e.g., *after one second*), moderate temporal shifts (e.g., *after one hour*), or long temporal shifts (e.g., *after one year*). These words were followed by repeated noun-phrase anaphors, which are preferred as referents for information no longer in attentional focus. The N400 was measured as an index of online conceptual integration. As the discourse unfolded, the N400 was larger for long- (e.g., *year*) than for short- (e.g., *second*) shift words. For the anaphor, the N400 was modulated in the opposite direction. Thus, the introduction of a temporal discontinuity leads to immediate neural integration costs, as well as to decreased accessibility of earlier information.

Time is essential for structuring, understanding, and predicting real-world events. Although the tracking of time as one reads appears to be effortless and instantaneous, it is far more complex than it may seem. In the present study, we employed event-related potentials (ERPs) to investigate immediate neural activity when readers encountered a temporal shift and to examine the structure of the resulting discourse representation, or situation model (Johnson-Laird, 1983; Zwaan & Radvansky, 1998).

Consistent with an event-indexing model that posits that readers track information along multiple dimensions (Zwaan, Langston, & Graesser, 1995; for neuroimaging evidence, see Speer, Zacks, & Reynolds, 2007), behavioral research has demonstrated that temporal information is continually tracked and integrated during reading (Speer & Zacks, 2005; Zwaan, 1996). Readers perceive temporal references as event boundaries, especially when there is an implied temporal shift (*an hour later*), as opposed to a continuity (*a moment later*) (Speer & Zacks, 2005). Readers also spend more time reading sentences with temporal discontinuities than they do sentences without temporal shifts; this may occur in an all-or-none fashion (Zwaan, 1996), regardless of the plausibility of the scenario. However, there are also conditions under which

shifts are processed in a gradient fashion (Kelter, Kaup, & Claus, 2004) or under which plot-driven expectations and temporal expectations interact (Rapp & Gerrig, 2002).

There are potentially two reasons why temporal shifts can increase processing times: (1) increased processing cost in shifting time and updating the situation model itself and (2) increased cost in accessing information that is presented prior to the temporal shift (cf. Ericsson & Kintsch, 1995; Zwaan & Radvansky, 1998). Evidence for the latter comes from studies using probe-verification tasks, as well as from studies examining the ease of processing an *anaphor*—a word or words that refer to a previously presented word or words (an *antecedent*; Anderson, Garrod, & Sanford, 1983; Speer & Zacks, 2005). Readers take longer to respond to a probe when it is presented after a temporal shift than when it is presented after a continuity (Zwaan, 1996) and take longer to resolve an anaphor referring to an antecedent presented before a temporal shift than one whose antecedent is presented before a continuity (Speer & Zacks, 2005). However, when Speer and Zacks compared reading times for sentences with and without anaphors, they found similar reading slowdowns following the temporal shift, regardless of anaphor presence. Thus, increased reading times may have been due to

---

T. Ditman, tali.ditman@tufts.edu

the cognitive effort associated with shifting the situation model as a result of the temporal shift, rather than with accessing information preceding the shift.

One way to disentangle the influences of cognitive effort due to a temporal shift from the accessibility of prior information is to employ repeated noun-phrase (NP) anaphors. These anaphors normally serve the function of referring to an antecedent that is no longer in a reader's focus of attention. Readers are slower to read repeated NP anaphors (e.g., *John*) that refer to antecedents within their focus of attention (e.g., following *John went to the store*) than to read repeated NP anaphors referring to nonfocused antecedents (e.g., following *John and Mary went to the store*; Gordon & Scearce, 1995)—an effect termed the *repeated-name penalty*. Thus far, this effect has been explored by using syntactic cues to manipulate the focus of attention both behaviorally (Gordon & Scearce, 1995) and electrophysiologically (Swaab, Camblin, & Gordon, 2004). In the present experiment, we manipulated antecedent prominence using semantic cues (i.e., temporal shifts in discourse) rather than syntactic cues (e.g., single vs. conjoined subjects), in order to determine whether an analogous cost in processing repeated NP anaphors referring to nonfocused antecedents would be observed—henceforth referred to as a *semantic repeated-name penalty*.

We used ERPs and the semantic repeated-name penalty to assess the accessibility of information presented prior to a temporal shift. Because ERPs provide millisecond resolution, they are an excellent tool for examining dynamic fluctuations in neural activity associated with tracking temporal information during reading. We examined amplitude changes in an ERP component that gauges the ease of semantically integrating information into a preceding context—the N400 (Kutas & Hillyard, 1984). The N400 is

a negative deflection in the waveform with a centroparietal scalp distribution, peaking approximately 400 msec after word onset, with larger negative-going amplitudes indicating increased integration difficulty (Holcomb, 1993). It is sensitive to semantic relationships between words (Bentin, McCarthy, & Wood, 1985; Rugg, 1984), stored knowledge within semantic memory (Federmeier & Kutas, 1999), real-world knowledge (Hagoort, Hald, Bastiaansen, & Petersson, 2004), and sentence- and discourse-level contexts (Kutas & Hillyard, 1984; van Berkum, Brown, & Hagoort, 1999). Importantly, the N400 is modulated by anaphor-resolution processes across sentence boundaries (Ditman, Holcomb, & Kuperberg, 2007) and is sensitive to the repeated-name penalty, with larger N400 amplitudes evoked for repeated names referring to focused antecedents than for those referring to nonfocused ones (Swaab et al., 2004).

To examine whether readers are immediately sensitive to temporal discontinuities within discourse and whether a temporal shift quickly decreases the accessibility of preceding information, in the present experiment we monitored dynamic levels of neural activity as scenarios unfolded (see Figure 1 and the Appendix). In each scenario, the first sentence introduced an event (e.g., *Kelly scolded the child*). The second sentence began with a temporal clause that indicated short, moderate, or long shifts in the discourse (e.g., *After one second/hour/year*), followed by a repeated NP anaphor referring to an antecedent presented as the final word in the first sentence (*child*), followed by a predicate (*whimpered to his mother*). We examined ERPs at the time word (*second/hour/year*) and at the anaphor (*the child*).

Examining N400 modulation at the time word allowed us to determine whether readers immediately integrate temporal information into their situation model. Following
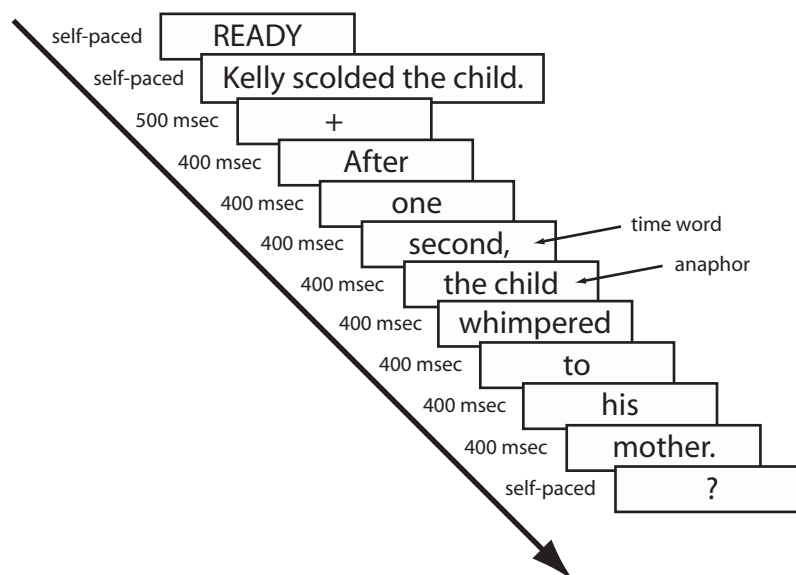

Figure 1. A typical experimental trial. The amount of time that sentences and words were on the computer screen is referenced on the left; the interstimulus interval (ISI) was kept constant at 100 msec, with the exception of the time between the sentence-final word and the question mark, which had an ISI of 700 msec.

previous findings (Zwaan, 1996), we predicted that time words indicating both moderate and long shifts (e.g., *hour/year*) would evoke larger amplitude N400s than would time words indicating a short temporal shift (e.g., *second*).

Examining N400 modulation at the repeated NP anaphor allowed us to determine whether such shifts rapidly decrease accessibility of preceding information. If they do, it should be harder to integrate a repeated NP anaphor—resulting in a larger amplitude N400—after a short shift than after moderate and long shifts (e.g., in the second sentence, it should be harder to integrate *the child* preceded by *After one second* than when it is preceded by *After one hour/year*).

METHOD

Participants

Twenty-four right-handed native English speakers (15 of whom were male, with a mean age of 19.2 years) participated for monetary compensation. Informed consent was obtained in accordance with Tufts Human Subjects Research Committee guidelines.

Stimuli

A total of 180 scenarios were constructed (following that in Figure 1). There were three temporal conditions—short shift, moderate shift, and long shift—which differed only with respect to the time word in the second sentence. Short- and moderate-shift temporal conditions were equally plausible, and long-shift temporal conditions were implausible (as confirmed by ratings; see below). The sentences were randomized across the three conditions in a Latin square design with 60 scenarios per condition. In addition, 60 filler scenarios were constructed, with 30 scenarios that were temporally implausible with short or moderate time shifts and 30 long-shift scenarios that were plausible (e.g., *Frank applied for the citizenship. After one second/year, the citizenship was granted by the government*). The fillers ensured that participants were reading the full scenario, rather than making decisions following the time word (e.g., determining that the scenario was implausible after reading *year*). On average, long-shift words had the highest word frequency ($M = 423.95$), moderate-shift words were least frequent ($M = 87.25$), and short-shift words fell in between ($M = 351.55$; Kučera & Francis, 1967).

Plausibility ratings were collected ($n = 12$) to ensure that short- and moderate-shift scenarios were equally plausible and that both were more plausible than long-shift scenarios. After they gave informed consent, participants were instructed to read each scenario for comprehension and to rate the real-world likelihood (on a scale of 1–6) of the described events. One-way repeated measures ANOVAs by participants (F_1) and by items (F_2) confirmed significant differences in ratings across the temporal conditions [$F_1(2,22) = 21.73$, $p < .001$; $F_2(2,358) = 53.36$, $p < .001$], and planned comparisons confirmed that long-shift scenarios ($M = 3.25$, $SD = 0.85$) were less plausible than short- ($M = 4.08$, $SD = 0.65$) and moderate- ($M = 4.32$, $SD = 0.61$) shift scenarios (all $ps < .01$). Short- and moderate-shift scenarios did not differ significantly from one another ($p > .10$) in the participant analysis, but they did in the item analysis ($p < .05$). In the filler scenarios, the short- and moderate-shift scenarios ($M = 1.65$, $SD = 1.03$) were rated as less plausible than their long-shift counterparts ($M = 5.21$, $SD = 1.44$) [$t_1(11) = 24.07$, $p < .001$; $t_2(59) = 26.97$, $p < .001$].

Procedure

For each scenario, participants saw the complete first sentence until they pressed a button to advance. The second sentence, preceded by a fixation cross, was then presented word by word in the center of the screen, except that the anaphor was presented with its definite article (e.g., *the child*). At the question mark, participants decided whether, given the event described in the first sentence, the event in the second sentence was plausible.

Recording procedure. Participants wore an elastic cap (Electro-Cap International) with 29 tin electrodes (Figure 2). Electrodes were located in the standard International 10–20 System locations, as well as in additional sites over the hemispheres. The electrode montage in Figure 2 depicts the placement of electrodes, with five sites along the midline (FPz, Fz, Cz, Pz, and Oz), three sites along the medial column on the left (FC1, C3, and CP1) and the right (FC2, C4, and CP2) hemispheres, four left- (F3, FC5, CP5, and P3) and right- (F4, FC6, CP6, and P4) hemisphere lateral sites, and five sites along the left (FP1, F7, T3, T5, and O1) and right (FP2, F8, T4, T6, and O2) peripheral column. Electrodes placed below the left eye and lateral to the right eye monitored vertical and horizontal eye movements, respectively. Electrodes were referenced to the left mastoid, and an electrode placed on the right mastoid monitored differential activity.

The electroencephalogram was amplified by an SA Bioamplifier, using a bandpass of 0.01–40 Hz and continuously sampled at a rate of 200 Hz. Impedances were kept below 10 kΩ for the eyes and below 3 kΩ for other sites. ERPs were averaged offline for each electrode site for each condition. Trials contaminated with eye artifact (exceeding 50 μV) or amplifier blockage (exceeding 140 msec of nearly identical values) were excluded.

Data Analysis

Behavioral data. Participants' acceptability judgments were analyzed by examining the percentage of scenarios classified as plausible.

ERP data. All analyses of time words and anaphors were conducted on mean amplitude values using 100 msec of activity preceding word onset as a baseline. Analyses were conducted at 0–250 msec and 300–500 msec (corresponding to the N400) following word onset. The 0- to 250-msec time window was examined to ensure that deviations in the N400 could not be accounted for by earlier differences; in the event of differences in this window, we employed a post-stimulus baseline (100–250 msec post-stimulus-onset; cf. Hagoort, 2003) and a prestimulus baseline.

We included all sites in a systematic, comprehensive columnar pattern of analysis, which has the advantage of yielding statistical information about differences in the distribution of effects along the anterior–posterior axis of the scalp (AP distribution) at the midline electrode column and across the two hemispheres (hemisphere) at the medial, lateral, and peripheral electrode columns. Specifically, repeated measures ANOVAs were performed for each column, with temporal condition (short shift, moderate shift, long shift), AP distribution (with the number of levels depending on the number of electrode sites in each column), and, for the three lateral electrode columns, hemisphere (left, right) as variables (see Figure 2). A Greenhouse–Geisser correction was applied to analyses with more than one degree of freedom in the numerator. We report original degrees of freedom with corrected p values. Significant interactions were examined further with simple effects tests and planned comparisons.

RESULTS

Behavioral Data

Participants were more likely to judge short- ($M = 55.6\%$, $SD = 11.9$) and moderate- ($M = 59.5\%$, $SD = 12.5$) shift scenarios as plausible than they were long-shift scenarios ($M = 18.7\%$, $SD = 11.9$). One-way repeated measures ANOVAs by participants and by items confirmed differences [$F_1(2,46) = 197.99$, $p < .001$; $F_2(2,358) = 231.27$, $p < .001$]. Planned comparisons revealed that long-shift scenarios were least likely to be judged plausible (all pairwise comparisons, $ps < .001$), whereas short- and moderate-shift scenarios did not differ ($ps > .05$). Short- and moderate-shift filler scenarios were more likely to be
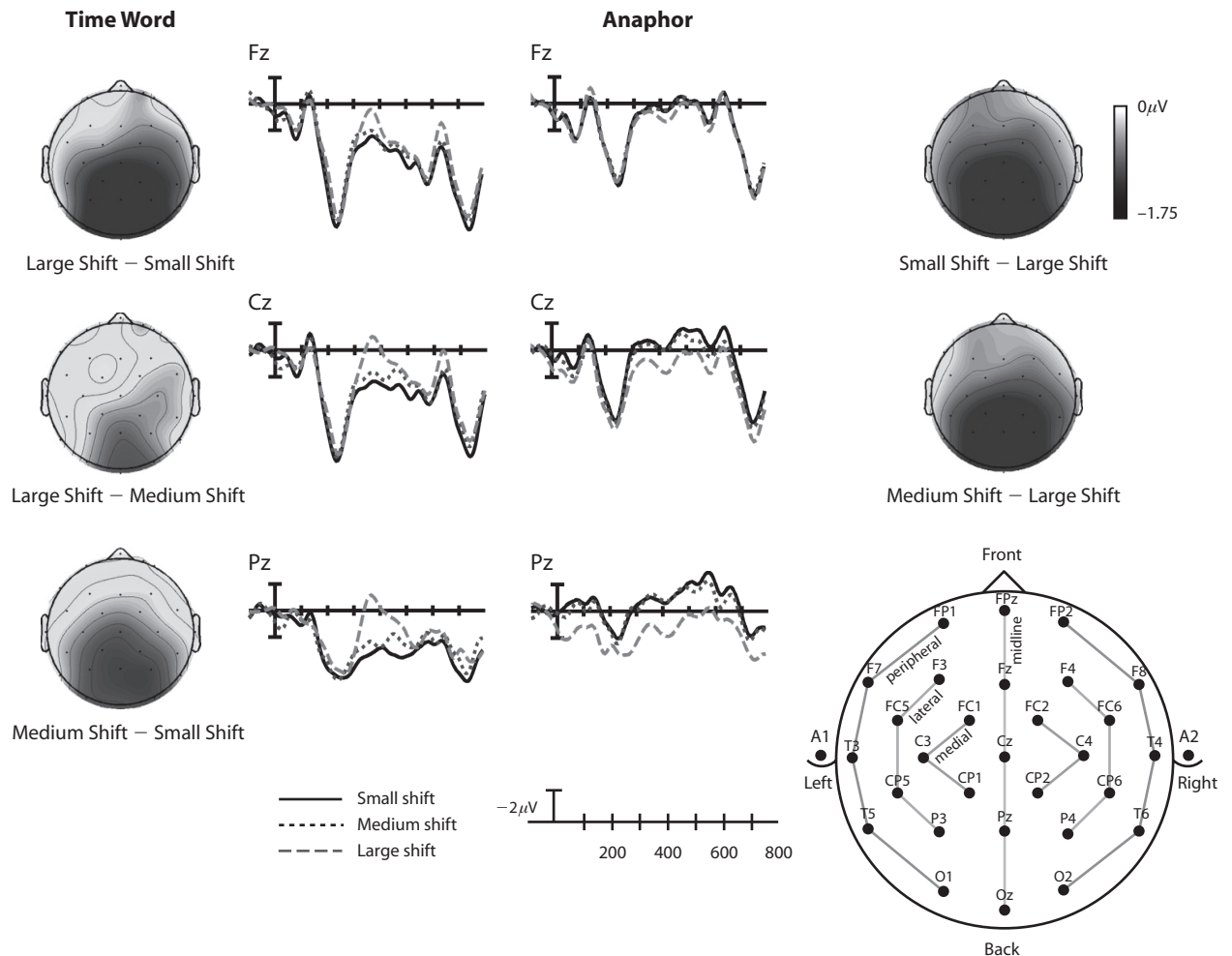
Figure 2. Wave components and voltage maps depicting neural activity across the scalp 450 msec post-word-onset for the time word and for the anaphor.

judged plausible ($M = 87.64$, $SD = 7.71$) than long-shift filler scenarios ($M = 6.25$, $SD = 5.67$) [$t_1(23) = 43.51$, $p < .001$; $t_2(59) = 32.07$, $p < .001$].

ERP Data

Artifact contamination led to rejecting 13.9% of the time-word trials [short, $M_{number} = 52.25$ (range, 40–60); moderate, $M_{number} = 59.96$ (range, 40–58); long, $M_{number} = 51.79$ (range, 44–58)] and 11.9% of the anaphor trials [short, $M_{number} = 53.29$ (range: 44–60); moderate, $M_{number} = 52.17$ (range, 39–57); long, $M_{number} = 53.08$ (range, 46–58)]. The artifact rejection rate did not differ by condition, as determined by one-way repeated measures ANOVAs at the time word and the anaphor ($ps > .10$). (See Figure 2.)

The ERP analyses reported below are on all trials. However, in order to ensure that results were not solely due to differences in participants' classifications during the ERP experiment, we repeated analyses using only short-, moderate-, and long-shift scenarios that were classified as implausible. We chose implausible re-

sponses because few long-shift scenarios had been classified as plausible.

ERPs at the Time Word

Early components (0–250 msec). No main effects of, or interactions with, temporal condition were observed at any column (all $ps > .05$). (See Table 1 for statistical analyses.)

N400 (300–500 msec). Main effects of temporal condition showed that the N400 amplitude evoked by time words across the three conditions differed from one another, and temporal condition × AP distribution interactions indicated that this effect varied across the scalp. Simple effects tests demonstrated a main effect of temporal condition at centroparietal sites but no effect at anterior sites, consistent with the scalp distribution of the N400. Planned comparisons at centroparietal sites showed that long-shift time words evoked the largest amplitude N400, followed by time words indicating moderate shifts, and the smallest amplitude N400 was evoked by short-shift time words. (See Table 1.)

Table 1
Overall ANOVAs Examining N400 Amplitude at the Time Word

Column	Main Effect of Temporal Condition	Temporal Condition × AP Distribution Interaction	Effect of Condition at Each Site
Midline	$F(2,46) = 14.44$***	$F(8,184) = 8.61$***	Fz~
			Cz***: 1 < 2*, 1 < 3***, 2 < 3*
			Pz***: 1 < 2*, 1 < 3***, 2 < 3***
			Oz***: 1 < 2~, 1 < 3***, 2 < 3***
Medial	$F(2,46) = 10.32$**	$F(4,92) = 7.37$**	FC1 and FC2**: 1 < 3**, 2 < 3~
			C3 and C4***: 1 < 2~, 1 < 3***, 2 < 3**
			CP1 and CP2***: 1 < 2*, 1 < 3***, 2 < 3**
Lateral	$F(2,46) = 14.85$***	$F(6,138) = 5.92$**	F3 and F4*: 1 < 3*
			FC5 and FC6**: 1 < 3**, 2 < 3*
			CP5 and CP6***: 1 < 2*, 1 < 3***, 2 < 3**
			P3 and P4***: 1 < 2*, 1 < 3***, 2 < 3**
Peripheral	$F(2,46) = 12.31$***	$F(8,184) = 5.70$**	T3 and T4**: 1 < 3**, 2 < 3*
			T5 and T6***: 1 < 3***, 2 < 3**
			O1 and O2***: 1 < 2*, 1 < 3***, 2 < 3***

Note—1, short shift; 2, moderate shift; 3, long shift; all differences are expressed in mean amplitude (greater values indicate greater negativity).   ~$p < .10$.   *$p < .05$.   **$p < .01$.   ***$p < .001$.

ERPs at the Anaphor

Early components (0–250 msec). Differences were observed, demonstrating a larger negativity for anaphors following short and moderate shifts than for those following long shifts. (See Table 2 for statistical analyses.)

N400 (300–500 msec). At the anaphor, we observed a very different N400 amplitude pattern from that elicited by the time word. Using a prestimulus baseline, a larger negativity was evoked by the short and moderate shifts than by the long shifts. This effect was also largest at cen-

troparietal sites, as demonstrated by temporal condition × AP distribution interactions. Planned comparisons at these sites demonstrated that, whereas the N400 for short and moderate shifts did not differ, the N400 amplitude for long shifts was significantly lower than that for short or moderate shifts. These findings remained robust, even after equating earlier differences by using a poststimulus baseline and a prestimulus baseline.[1] The same pattern of results was observed after equating responses (using only responses classified as implausible).[2] (See Table 2.)

Table 2
Overall ANOVAs at the Anaphor Using a Prestimulus Baseline

Column	Main Effect of Temporal Condition	Temporal Condition × AP Distribution Interaction	Effect of Condition at Each Site
		Early Component Amplitude	
Midline	$F(2,46) = 4.87$*	$F(8,184) = 8.26$***	Cz**: 1 > 3**, 2 > 3**
			Pz***: 1 > 3***, 2 > 3***
			Oz***: 1 > 3***, 2 > 3***
Medial	$F(2,46) = 3.38$*	$F(4,92) = 10.04$***	FC1 and FC2~
			C3 and C4**: 1 > 3**, 2 > 3***
			CP1 and CP2***: 1 > 3**, 2 > 3***
Lateral	$F(2,46) = 3.61$*	$F(6,138) = 8.81$**	CP5 and CP6***: 1 > 3**, 2 > 3***
			P3 and P4***: 1 > 3***, 2 > 3***
Peripheral	$F(2,46) = 4.24$*	$F(8,184) = 7.12$**	T3 and T4*: 2 > 3*
			T5 and T6***: 1 > 3***, 2 > 3***
			O1 and O2***: 1 > 3***, 2 > 3***
		N400 Amplitude	
Midline	$F(2,46) = 12.53$***	$F(8,184) = 5.12$**	Cz**: 1 > 3**, 2 > 3**
			Pz***: 1 > 3***, 2 > 3***
			Oz***: 1 > 3***, 2 > 3***
Medial	$F(2,46) = 10.47$**	$F(4,92) = 7.53$**	FC1 and FC2~
			C3 and C4**: 1 > 3**, 2 > 3***
			CP1 and CP2***: 1 > 3**, 2 > 3***
Lateral	$F(2,46) = 9.91$**	$F(6,138) = 5.22$**	CP5 and CP6***: 1 > 3**, 2 > 3***
			P3 and P4***: 1 > 3***, 2 > 3***
Peripheral	$F(2,46) = 9.68$***	$F(8,184) = 3.63$*	T3 and T4*: 2 > 3*
			T5 and T6***: 1 > 3***, 2 > 3***
			O1 and O2***: 1 > 3***, 2 > 3***

Note—1, short shift; 2, moderate shift; 3, long shift; all differences are expressed in mean amplitude (greater values indicate greater negativity).   ~$p < .10$.   *$p < .05$.   **$p < .01$.   ***$p < .001$.

DISCUSSION

These findings demonstrate that readers use temporal information immediately—within 300 msec—in order to structure a developing mental model. At the time word, the largest N400 amplitude was evoked by long temporal shifts, followed by moderate shifts, and the smallest N400 amplitude was evoked by short temporal shifts.[3] However, immediately following the time word, the N400 was larger for repeated NP anaphors that referred to information preceding both short and moderate shifts than for repeated NP anaphors referring to information preceding long shifts.

Discourse scenarios with long temporal shifts were generally less plausible than those with short or moderate temporal shifts. Thus, the larger N400 for time words indicating long temporal shifts than for those indicating short or moderate ones may have been driven by the incongruity between the time word and the expected duration of the event described, rather than by the cost of integrating the temporal shift per se. Therefore, of particular interest was the N400 evoked by moderate-shift time words, which was significantly larger than that for short ones; however, they were rated as equally plausible. This finding suggests that the absolute duration of the temporal shift can lead to an immediate cost in processing discourse and is consistent with previous behavioral reports of increased reading times for clauses following time words indicating temporal shifts, regardless of sentence plausibility (Zwaan, 1996). We have extended these findings by demonstrating an immediate neural cost in integrating the temporal shift at the time word itself.

In contrast, the N400 observed for the repeated NP anaphor was larger following both short- and moderate-shift than it was following long-shift time words. We interpret this as reflecting increased difficulty in accessing the antecedent, as a result of a semantic repeated-name penalty (cf. Sanders & Gernsbacher, 2004). Specifically, a long shift made it more appropriate to use a repeated NP anaphor to refer to the antecedent. With these results, we have extended previous findings of the repeated-name penalty to contexts in which attention is shifted using semantic, rather than syntactic, cues (Swaab et al., 2004). Most important, we have demonstrated that a long temporal shift immediately led to the creation of a new situation model, decreasing the accessibility of preceding information. As discussed above, note that the creation of this new situation model may have been driven by the semantic implausibility of the time shift relative to the expected duration of its preceding event, rather than by the absolute duration of the time shift. This may explain why there was no decrease in the N400 amplitude between repeated NP anaphors following moderate and short temporal shifts: Readers may not have constructed a new model in the moderate-shift scenarios, given that the described events could plausibly occur within the noted time frame.[4]

In sum, the present study demonstrates that as a discourse model is built up during reading, we appear to "travel through time," with the introduction of a temporal discontinuity leading to an immediate increased cost in neural integration as well as to decreased accessibility of earlier information.

REFERENCES

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology*, **35A**, 427-440.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography & Clinical Neurophysiology*, **60**, 343-355.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). The contributions of lexico-semantic and discourse information to the resolution of ambiguous categorical anaphors. *Language & Cognitive Processes*, **22**, 793-827.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, **102**, 211-245.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory & Language*, **41**, 469-495.

Gordon, P. C., & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory & Cognition*, **23**, 313-323.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, **15**, 883-899.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, **304**, 438-441.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, **30**, 47-61.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.

Kelter, S., Kaup, B., & Claus, B. (2004). Representing a described sequence of events: A dynamic view of narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 451-464.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, **307**, 161-163.

Rapp, D. N., & Gerrig, R. J. (2002). Readers' reality-driven and plot-driven analyses in narrative comprehension. *Memory & Cognition*, **30**, 779-788.

Rugg, M. D. (1984). Event-related potentials in phonological matching tasks. *Brain & Language*, **23**, 225-240.

Sanders, T. J. M., & Gernsbacher, M. A. (2004). Accessibility in text and discourse processing. *Discourse Processes*, **37**, 79-89.

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory & Language*, **53**, 125-140.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, **18**, 449-455.

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience*, **16**, 715-726.

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early

referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory & Language*, **41**, 147-182.

ZWAAN, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1196-1207.

ZWAAN, R. A., LANGSTON, M. C., & GRAESSER, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, **6**, 292-297.

ZWAAN, R. A., MADDEN, C. J., & WHITTEN, S. N. (2000). The presence of an event in the narrated situation affects its availability to the comprehender. *Memory & Cognition*, **28**, 1022-1028.

ZWAAN, R. A., & RADVANSKY, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, **123**, 162-185.

NOTES

1. Using a poststimulus baseline, there were main effects of temporal condition [midline, $F(2,46) = 4.59$, $p < .05$; medial, $F(2,46) = 5.28$, $p < .01$; lateral, $F(2,46) = 3.95$, $p < .05$; peripheral, $F(2,46) = 2.31$, $p > .10$]. Using a prestimulus baseline (100 msec of activity preceding *after*), there were temporal condition × AP distribution interactions [midline, $F(8,184) = 2.94$, $p < .05$; medial, $F(4,92) = 2.73$, $p < .07$; lateral, $F(6,138) = 3.92$, $p < .05$; peripheral, $F(8,184) = 0.82$, $p > .10$]. Importantly, no early differences were observed using the poststimulus or prestimulus baseline (all $ps > .10$).

2. N400 differences between the conditions were observed both at the time word (main effect of temporal condition, all $Fs > 5.76$, all $ps < .01$; temporal condition × AP distribution interaction, all $Fs > 2.93$, all $ps \leq .05$) and at the anaphor (main effect of temporal condition, all $Fs > 6.65$, all $ps < .01$; temporal condition × AP distribution interaction, all $Fs > 2.68$, all $ps < .06$).

3. It is unlikely that ERPs were modulated by word frequency differences, since these would predict the smallest, rather than largest, N400 amplitude for long-shift words.

4. Along with the plausibility of the moderate-shift scenarios, the use of the word *after* may have decreased the likelihood of shifting to a new situation model, since *after* may be interpreted as describing continuous events (Zwaan, Madden, & Whitten, 2000). We thank Rolf Zwaan for this interpretation.

APPENDIX
Examples of Temporal Scenarios

1. The maid polished the silverware. After one second/hour/year, the silverware sparkled brilliantly.
2. Sue pet the cat. After one second/hour/year, the cat purred happily.
3. The man proposed to the woman. After one second/hour/year, the woman accepted his proposal.
4. The play depressed the spectator. After one second/hour/year, the spectator cried uncontrollably.
5. The mother sang the lullaby. After one second/hour/year, the lullaby soothed the baby.
6. The movie amused the man. After one second/hour/year, the man laughed loudly.
7. The singers angered the crowd. After one second/hour/year, the crowd booed wildly.
8. The programmer encrypted the software. After one second/hour/year, the software was secure from hacking.
9. The professor discussed the exam. After one second/hour/year, the exam was distributed to the students.
10. Adam waxed the table. After one second/hour/year, the table glistened like new.