TEMPORAL PROCESSING OF ONGOING EVENT REPRESENTATIONS

by

Victoria L. Scharp

B.S., Speech-Language Pathology and Audiology, Miami University, 2000

M.A., Speech-Language Pathology, Miami University, 2002

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Communication Science and Disorders

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Victoria L. Scharp

It was defended on

July 15, 2016

and approved by

Malcolm McNeil, PhD, Chair, Department of Communication Science and Disorders

Michael Walsh Dickey, PhD, Associate Professor, Department of Communication Science and Disorders

Stephanie Huette, PhD, Assistant Professor, Department of Psychology, The University of Memphis

Dissertation Advisor: Connie A. Tompkins, PhD, Professor, Department of Communication Science and Disorders

TEMPORAL PROCESSING OF ONGOING EVENT REPRESENTATIONS

Victoria L. Scharp, PhD, CCC-SLP

University of Pittsburgh, 2016

Time cues are ubiquitous in language and the ability to interpret them is essential for understanding events during discourse comprehension. Temporal markers that signal ongoing versus completed events, like the progressive and simple past tense, prompt distinct mental event representations. However, the detailed properties of ongoing event representations remain unexplored. Drawing from both the simulation and semantic association approaches to knowledge representation, this study examines the novel prediction that ongoing events engender incremental discourse representation updating processes. Experimental sentences cued either early or late phases of an ongoing event (e.g. *Alice had* <u>*recently started/almost finished*</u> *baking a cake*). Targets in a post-sentential lexical decision task were strongly associated with either early or late event phases (e.g. EGGS/AROMA). Facilitation priming was predicted for congruent sentence-target pairs, however, priming was found for the early event phase exclusively. The results have implications for models of knowledge representation, theories of semantic priming, and discourse model updating.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

1.0 INTRODUCTION

Forming and updating mental event representations are essential pieces of the language comprehension puzzle, however, their underlying nature remains elusive. The theoretical rationale for the current study stems from two prominent event representation perspectives, simulation (e.g. Barsalou, 1999; Gallese & Lakoff, 2005; Glenberg & Kaschak, 2003; Zwaan, 2003) and semantic association (e.g. Coll-Florit & Gennari, 2011; Gennari, 2004). Simulation theories propose that primary sensory and motor networks are activated during semantic processing in a pattern similar to actually experiencing what is read or heard. The semantic association perspective suggests that specific concepts and/or attributes become linked to a given event through knowledge and associations acquired during real-world experiences. In their current forms, these perspectives are not mutually exclusive, but instead offer critical insights which shape the theoretical arguments of this study.

One key facet of event representation formation and updating relates to time. Comprehenders rely on a multitude of default frameworks and temporal cues in order to understand when events occur in relation to one another and to infer the completion status of an event. The temporal processing literature currently conceptualizes events as either in progress or completed, resulting in the treatment of event continuity as an all or nothing phenomenon. However, it is highly likely there is a dynamic nature to event model updating, specifically that in progress events engender an incremental updating process. This novel prediction is derived from unexpected and contradictory empirical findings, contributions from event representation theories, and an integration of threads from research in both the temporal and spatial processing domains.

This dissertation document begins by outlining critical concepts including event models, temporal processing cues, and the narrative timeline in Chapter 2. The evolution of research on temporal processing cues and critiques of empirical findings from this domain will be considered in the first part of Chapter 3. Then the document shifts to a more thorough examination of event continuity, the primary area of interest. More specifically, pertinent studies of grammatical aspect, lexical aspect, and event-specific cues will each be reviewed in turn. Rounding out Chapter 3, evidence from the spatial domain provides an additional thread to the theoretical rationale leading to the detailed study rationale described in Section 3.4. A more detailed consideration of the aforementioned event representation perspectives, simulation and semantic association, follows in Chapter 4. Chapter 5 includes the methods and potential results for the investigation along with the results from a pilot study conducted to assist with stimulus development. Chapter 6 summarizes the empirical findings for the primary outcome measures, control condition, and baseline condition from the multilevel analysis. Chapter 7 places the results into the larger context of the temporal processing literature and connects the findings to theories of event representation. The document concludes with a summary of the contribution to the literature, strengths and limitations of the current study, and future directions.

2.0 SITUATION MODELS

Most researchers agree that the overarching goal of discourse comprehension is to build a coherent mental representation of what is read or heard (Gernsbacher, 1990). A situation model (Kintsch & van Dijk, 1978) is the mental representation of what is happening during discourse and theoretical and empirical characteristics of situation models have been widely used to understand how comprehenders construct mental representations. A situation model does not include every word or concept that is mentioned, but instead portrays an integration of the content to facilitate the understanding of the events described. Although the notion of situation models is rather vague and underspecified, extensive investigation suggests that there are several elements that contribute to situation model formation including the specific text or discourse to be understood and a comprehender's linguistic, pragmatic, and world knowledge (Glenberg, Meyer, & Lindem, 1987). Ultimately these elements are combined to construct the representation of the situation or "mental microworld" described in the discourse (Magliano & Schleich, 2000, p. 83). As discourse comprehension proceeds, situation models are constructed and updated as additional information accumulates.

2.1 EVENT MODELS

Events are the proposed building blocks for creating and updating situation models during comprehension (Zwaan, Langston, & Graesser, 1995). A commonly used operational definition of an event is "a segment of time at a given location that is conceived by an observer to have a beginning and an end" (Zacks & Tversky, 2001, p. 17). The place on a narrative timeline and the location of an event may be specified or inferred by the comprehender depending on the surrounding context. Event models are a type of situation model where the verb is often the driving force for what types of knowledge are mentally activated (Ferretti, Kutas, & McRae, 2007). In theories of sentence processing, verbs and their morphosyntactic properties are critical ingredients that stimulate comprehension processes (e.g. MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994). For example, when an event verb is read or heard, the conceptual knowledge of a comprehender is tapped, mentally activating prototypical event properties such as agents, actions, objects/instruments, time course, and duration (Ferretti, et al., 2007). Critically, these event attributes are available very quickly during on-line sentence processing. Different types of verb-based time cues and their influence during event processing will be discussed in Chapter 2.

Events are characterized as dynamic occurrences in time. There is strong empirical evidence that comprehenders mentally activate distinctly different event attributes when an event is perceived as ongoing versus completed (e.g. Carreiras, Carriedo, Alonso, & Fernandez, 1997; Coll-Florit & Gennari, 2011; Ferretti et al., 2007; Madden, & Ferretti, 2009; Magliano & Schleich, 2000; Section 3.3). However, when an event is perceived as ongoing, it is an open question whether the mental activation of event attributes is maintained or fluctuates across the time course of the event. To date, the event processing literature conceptualizes a dichotomy between the mental activation of event attributes linked to ongoing and completed events. This leaves an empirical gap for making specific predictions about the dynamic and incremental nature of the progression of time during event processing. More specifically, the models in the literature have not addressed the potential for degrees of "ongoingness" during ongoing events which could promote a different progression or activation pattern of event attributes. One challenge to fully understanding this event continuity facet of event processing is that events and their attributes are highly variable. For example, an event can end almost as quickly as it began (e.g., opening a can of soda) or it can last for many hours (e.g., a day at the beach). How comprehenders appear to be able to easily build, update, and recall information from event models has been the focus of a considerable amount of psycholinguistic research.

Within this vast literature base, one influential theoretical framework introduced by Zwaan and colleagues (1995) accents five specific dimensions of an event model that comprehenders track during narrative level comprehension. The framework is called the Event-indexing model, and the five aspects that are cataloged by a comprehender are: time, space (location), causation, intentionality/motivation, and protagonist/object. The Event-indexing model specifies that during comprehension memory nodes are activated and then that activation pattern is calibrated based on the amount of overlap between one of the indices and new incoming events from a described situation. The original instantiation of the model focuses on explicating how quickly new information is integrated into the model if it is congruent with the established events of one of the indices.

A central finding throughout the Event-indexing model literature is that shifts or discontinuities in any of the five dimensions result in increased processing time, suggesting that integration or updating of the event model is occurring. This slower processing is typically indexed by longer reading or response times (RTs; e.g. Zwaan et al., 1995). The longer RTs correspond to probes of events attributes that are introduced before a shift and longer reading times have been found for the actual words or phrases in the sentence or narrative where the discontinuity is introduced. For example, in the time domain, readers take longer to read the phrase *an hour later* compared to the phrase *a moment later* in the following sentence from Zwaan's seminal 1996 article on how time shifts are processed:

- The professor started analyzing the data. *A moment/hour later*, her phone rang.

Studies of the Event-indexing model have addressed how world knowledge is used for interpreting causation (Singer, Halldorson, Lear, & Andrusiak, 1992) and how accessible objects are to a protagonist as narratives progress (Glenberg et al., 1987). One limitation of the model is the independent way it approaches the dimensions by not considering the potential influences a change in one index might have on another index. For example, a change in the spatial index can have a direct influence on the perceived time elapsed during the location change (e.g. walking across the room versus walking across town). Prior investigations (e.g. Anderson, Garrod, & Sanford, 1983; Speer & Zacks, 2005) have implemented a location shift and a time shift simultaneously so it is not clear if the resulting increased RTs to probes of event attributes that were introduced prior to the time shift are due to the time shift itself or to a combined time and location shift effect. While understanding the processing contributions when the different indices are combined is essential to extending the event processing literature, there are still open questions about the potential influences of the individual indices. The index of interest for the research questions in this document was time. Specifically, this study considered how time cues that signal early and late time points within an individual event influence event model construction and updating.

2.2 TIME PROCESSING

One part of the event processing picture that is not well understood is how different types of time cues influence the mental activation of concepts within event models. As outlined above, time is one of five indices that comprehenders track according to the Event-indexing model (Zwaan et al., 1995). Tests of this model have indicated that comprehenders routinely monitor both implied or specified positions of an event on the narrative timeline and the prototypical durations of everyday events (e.g. Rapp & Taylor, 2004). To more fully explore the impact of the contribution of these specific temporal properties for event processing, a closer inspection of the conceptualization of time is warranted.

2.2.1 Psychological time

Multiple conceptualizations of time have been considered in the literature. Time is a psychological construct that is unique because it not connected to a specific sense organ (Block, 2014). As humans, we are able to experience the passage of time but we cannot hold it, taste it, and so on. Researchers have grappled with what type of mechanism might be responsible for how humans perceive and track time. Early theories of time perception and tracking include the internal clock model where pulses are generated and tracked via an internal pacemaker, (Treisman, 1963) and the scalar expectancy theory that examines time estimation performance (in animals) using temporally structured reinforcement schedules (Church, 1984; Gibbon, 1997). These models are often cited and continue to be refined (Tresiman, 1993), however, it remains unclear how these types of models can account for levels of abstraction in language comprehension. In addition, in their current form, these models do not address the potential influences of attentional allocation or the complexities of real-time processing (Zakay & Block, 1994).

One prominent debate in the temporal processing literature concerns teasing apart the mechanisms that support inherent differences between how events are experienced in real-time and how events are experienced via language. Using linguistic cues, comprehenders can be transported to the past or the future and discuss events that have never been experienced first-hand. This ability to transcend the here and now perspective connects to the consideration of whether prospective and retrospective timing enact the same or different processing mechanisms. One explanation is that prospective timing taps into the premise of the internal clock while retrospective timing is linked to more general cognitive-linguistic processing that does not directly connect to conceptualization of time (Zakay & Block, 2004).

Scholars who explore questions related to different theories of time perspective often use spatial metaphors (i.e. my best years are ahead of me) due to a strong correspondence between time and space (see Section 3.3.4, i.e. Lehrer, 1990). For example, two different time-space metaphor systems have been proposed for the English language, an ego-moving and a time-moving metaphor (Genter, Imai, & Boroditsky, 2002). For the ego-moving metaphor, the perspective is that of an observer navigating forward on a timeline toward the future. This viewpoint mimics the Dynamic Mental Representation framework (see Section 2.2.2) of progressing forward on a narrative timeline. In contrast, in the time-moving metaphor, the observer remains in a constant location and the events themselves are perceived as moving from the future to the past like a conveyer belt. At least one study has found that priming participants with either an ego-moving or time-moving context results in a strong preference to interpret ambiguous statements like "Next Wednesday's meeting has been advanced by two days" (p. 556, McGlone & Harding, 1998) using the space-time metaphor perspective that was primed. Extending the possible predictions from the space-time metaphors perspectives, Walsh (2003) has proposed A Theory Of Magnitude (ATOM), a conceptual framework that links space, time, and quantity into a generalized magnitude system. This theory will be discussed next.

Walsh (2003) has proposed a magnitude mechanism in which time, space, and quantity all operate of out the same common system. One example of how this theory has been tested is a study by De Long (1981) where participants completed tasks in scaled down environments (i.e. 1/6 or 1/24 actual size) and were asked to discontinue their efforts after 30 minutes. As predicted, the scale of the environment affected the stopping times proportionately. As ATOM has been refined and tested, additional parameters have been proposed to be part of the same generalized magnitude processing mechanism including both speed and size (Bueti & Walsh, 2009). Much of the literature examining ATOM stems from animal studies, (i.e. Dehaene, Dehaene-Lambertz, & Cohen, 1998) and it is not clear how different levels of abstraction could emerge from this type of mechanism. This notion of a common magnitude framework is an intriguing one, and it could provide one explanation for the pattern of results demonstrated frequently in the temporal and spatial processing domains. Specifically, this notion is relevant to the finding that a change in one parameter (i.e. across the street versus across town) prompts a change in another (i.e. a duration estimate, Rapp & Taylor, 2004).

Temporal information is ubiquitous across languages and can be conveyed via all word classes and forms (Miller & Johnson-Laird, 1976). One way to manage this elusiveness is to use labels for segments of time (e.g. minutes, hours, years). For event processing, understanding references to time is fundamental to comprehending the context and order of events within a narrative. These references are used by comprehenders to sequence events, track event completion, and update the event or narrative timeline. Block’s (2014) approach to conceptualizing time was to propose three different aspects of psychological time. These include the following:

1) Time duration: how long an event lasts or a length of time between events
2) Time succession: the temporal order of events in a sequence
3) Time perspective: how an individual conceptualizes past, present, and future time.

Each of these aspects of time can contribute to how comprehenders incorporate time concepts and the progression of time into their mental representation of events. In the literature, each of these aspects is linked to a specific type of time cue (e.g. changing the grammatical aspect of a verb is a time cue that corresponds to duration). Empirical data specific to the temporal property of duration and its corresponding time cues will be discussed in Chapter 2 following a brief consideration of the narrative timeline.

2.2.2 Narrative timeline

Another critical factor for understanding how time is conceptualized during language comprehension is the nature of the narrative timeline. It is widely accepted that the narrative timeline for an event begins at a “here and now” reference point and progresses forward toward an endpoint (unless otherwise indicated in the discourse; Dry, 2009). This universally accepted view of the narrative timeline appears to be rooted in the psychological arrow of time (Savitt, 1996). The phrase “time’s arrow” dates back to Sir Arthur Eddington (1928, p. 34), who purports that most events and event sequences are organized and experienced in a unidirectional (i.e. forward) manner consistent with how time is perceived (e.g. the growth of a plant from a seedling to a mature plant).

In the time processing domain, the concept of directionality linked to time's arrow is included in and expanded upon in the *iconicity assumption* (Fleischman, 1990; Hopper, 1979). This is a default framework for how comprehenders interpret the narrative timeline by assuming that events described are presented in chronological order and that successive sentences describe successive events. Various aspects of the iconicity assumption have been tested, including how it relates to shifts in the narrative timeline (Zwaan, 1996), contributes to understanding event sequences (Stanfield & Zwaan, 2001), and influences actions that are described successively or simultaneously in a text (de Vega, Robertson, Glenberg, Kaschak, & Rinck, 2004). The pattern of results for these studies provides strong evidence that individuals rely on default frameworks like the iconicity assumption during event model construction and that violations of the assumption result in increased reading times or RTs. The increased processing time found in studies that test the iconicity assumption is one example of the prominent finding across the Event-indexing literature of longer processing times for discrepancies in any of the dimensions comprehenders track (Section 2.1).

One framework that incorporates the notion of time's arrow is the Dynamic Mental Representation framework proposed by Freyd (1987, 1992). While her theory has its origins in handwriting, she includes two specific assumptions related to the temporal dimension which have been applied to event processing. These assumptions are that the temporal dimension 1) cannot be extracted from the mental representation of the external world and 2) is directional. The "dynamic" piece of this theory describes her conceptualization of the direct parallel between how humans experience time during real-life events and its role in a mental representation. In this theory, because the temporal dimension mimics how real time progresses, it is directional (i.e. forward moving on the narrative timeline). This Dynamic Mental Representation theory has connections to more recent explanations of how individuals mentally represent what is read or heard, namely simulation theories (e.g. Barsalou, 1999; Gallese & Lakoff, 2005; Glenberg & Kaschak, 2003; Zwaan, 2003). Briefly, simulation theories propose that a distributed network of areas within the primary sensory and motor systems is activated during semantic processing, as if the comprehender were simulating the actual experience described. Simulation theory is discussed further in Sections 3.3.1.1 and 4.1 as it provides one possible theoretical basis for the predictions in the present study.

The narrative timeline clearly plays an essential part in event processing and individuals must be able to navigate it in order to understand when events occur in relationship to other events and the nature and progression of individual events. This dual role for the mental timeline provides some insight into its complexity. A comprehender is often required to entertain multiple timelines at once. For example, one timeline may be established for an event at the beginning of a sequence of events while another must be tracked for an intervening or overlapping event (Hinrichs, 1986). Event duration and continuity are crucial aspects of event processing and correctly interpreting them is necessary for managing one or more event timelines. Because time is a psychological phenomenon, comprehenders must rely on different types of cues (e.g. temporal adverbials, lexical or grammatical markers, and event knowledge) when building and updating timelines for their mental representations. Several different types of time cues are discussed in the next chapter, along with the applicable empirical evidence.

3.0 TIME PROCESSING CUES

Time processing cues are routinely used when event models are constructed and updated. Some types of time processing cues prompt a comprehender to establish an event boundary (e.g. Speer & Zacks, 2005; Zwaan & Radvansky, 1998). Event boundaries are used by comprehenders to segment incoming event information and to prompt an update of the current event model. Using terminology from the Structure Building Framework (Gernsbacher, 1990), new events initiate new mental substructures and then additional information either 1) is mapped onto that established structure (model) or 2) prompts a shift to a new substructure. Shifting to a new substructure often triggers an event boundary and elicits an increase in processing time. One consequence to establishing an event boundary is that event information from before the boundary may become less accessible to the comprehender. Critically, readers perceive temporal changes as event boundaries when reading about successive activities within a text (e.g. Speer & Zacks, 2005; Zwaan & Radvansky, 1998).

Some early work proposes rigid event boundaries, a result of which is that access to prior event content degrades uniformly after the boundary is established (e.g. Zacks & Tversky, 2001). However, a more recent consensus is that that readers tend to trigger *flexible* event boundaries (e.g. Rapp & Taylor, 2004). This means that triggering an event boundary does not necessarily mandate a decrease in availability of prior event attributes. Instead, some event properties may be maintained across an event boundary based on the context and background knowledge related to the event. Exactly which properties are maintained across event boundaries is an open question, but one possible factor is event continuity (i.e. whether an event is described as ongoing or completed). This factor is discussed in depth in Section 3.3.

Different types of time cues are discussed in this chapter along with the relevant empirical evidence. Section 3.1 summarizes the literature on how certain types of time cues elicit a time shift effect. This effect is reflected by longer RTs to probes of event attributes that are mentioned prior to a time shift (e.g. an hour later). However, emerging evidence will also be described that is contrary to the time shift effect. Section 3.2 discusses how event duration (e.g. the prototypical length of routine events) links to the question of what concepts may remain mentally activated or more readily available to a comprehender in a given event model. Section 3.3 provides a synopsis of the event continuity literature. Specifically, it addresses how event completion status is cued (e.g. grammatically, lexically) and the types of event attributes that are activated when an event is signaled as ongoing or completed. Limitations in the event continuity literature are considered in Section 3.3.3. Finally the rationale for the current study is introduced in Section 3.4.

3.1 TIME SHIFT EFFECT

Much of the time processing literature has focused on processing influences associated with cues that explicitly signal shifts forward in the narrative timeline (e.g. an hour later). These studies generally have found a time shift effect, manifested in longer reading times for sentences that contained these time shift cues and/or increased RTs to probes of event attributes that were described prior to the time shift cues (Anderson et al., 1983; Bestgen & Vonk, 1995, Kelter, Kaup, & Claus, 2006; Zwaan, 1996). The increased processing time has been interpreted to reflect the time it takes for a comprehender to establish an event boundary. The time shift effect initially appeared to be quite robust until a few researchers from independent labs began to report data that were inconsistent with it (Radvansky & Copeland, 2000; Ditman, Holcomb, and Kuperberg, 2008;

Weingartner & Myers, 2013). In three studies using different methods, the expected time shift effect was not found if the event could be perceived as ongoing after the time shift cue occurred. This section will outline the results, authors' interpretations, and critiques of these specific studies as a bridge to the theoretical frameworks, considerations for the study design, and research questions outlined later in this document.

Ditman et al. (2008) used event–related potentials (ERPs) to study neural activity while adults read brief narratives containing time shift cues and repeated noun phrase anaphors. The N400, a negative deflection of the waveform with a centroparietal scalp distribution, has been interpreted as a signal of increased difficulty with semantic integration (Holcomb, 1993). Characteristics of the N400 have also been used to make inferences about the ease of semantic integration into prior context at both sentence and discourse levels (e.g. van Berkum, Brown, & Hagoort, 1999). Critically, the N400 is also sensitive to anaphor resolution processes across sentences. Ditman and colleagues (2008) had 12 participants wear a 29 channel elastic cap while reading 180 two-sentence scenarios like the following:

- The maid polished the silverware. *After one second/hour/year*, the silverware sparkled brilliantly.

The first sentence set up the noun phrase which was probed (repeated) in three temporal conditions: following the continuation of the current timeline (i.e., after one second, referred to here as the no shift condition), a moderate time shift (i.e. after one hour), or a large shift forward on the timeline (i.e. after one year). A main effect of temporal condition was found for the ERPs collected at the time word (i.e. second, hour, and year) at the expected centroparietal sites. The pattern of results for the planned comparisons was also as expected: The largest N400 amplitude occurred in the *after a year* shift condition followed by the moderate shift condition and then the smallest N400 amplitude was observed in the *after one second* condition. The declining N400 amplitudes from the largest shift to the no shift condition were interpreted to reflect a decreased need for integration for each shift condition. Based on these data, the authors concluded that readers use temporal information within the first 300 ms of sentence onset to structure event models. A second set of 12 participants provided plausibility judgments for the event durations and there were no differences between the no shift and moderate shift conditions, but the large shift scenarios were judged to be implausible. The authors did not consider whether the different levels of plausibility may have shaped the results at the time word or whether absolute shift durations (e.g. year versus hour) signal similar event boundaries and updating processes.

Turning to the ERPs collected at the repeated anaphor, a distinctly different pattern emerged. Contrary to the data collected at the time word where the largest time shift forward elicited the largest N400 amplitude, the repeated noun phrase anaphor elicited a larger negativity for the no shift and moderate shift conditions. Based on planned comparisons of the N400 amplitudes, the no shift and moderate shift conditions did not significantly differ from one another, but the large shift condition N400 was significantly smaller (i.e. less negative) than those in the other conditions. The authors speculated that this result was driven by the implausibility of the largest shift scenarios due to the mismatch between the absolute duration of the shift (i.e. after one year) and the expected duration of the event described in the scenario.

Ditman and colleagues (2008) concluded that readers continued to build on the current model for the no shift and moderate shift conditions and only shifted to constructing a new event model for the scenarios in the largest shift condition of *after a year*. These results do not disentangle event plausibility influences from time shift effects, however, the authors' interpretation suggests another potential critical variable during temporal processing: Activity-defined duration as derived from world knowledge. This topic will be discussed below in the upcoming critique of the Weingartner and Myers (2013) study and later in Section 3.2.

A second study that found evidence contrary to the time shift effect was reported by Radvansky and Copeland (2010). While this study's primary research question emphasized whether and when readers update the spatial dimension of an event model, the authors targeted the temporal dimension in their second experiment. Time shifts of *a moment later* (essentially a continuation of the narrative timeline and thus referred to as the no shift condition) and *a day later* were embedded into short stories with probe words interspersed throughout each narrative. The probe words referred to objects introduced in the text prior to the shift and were used to track the availability of event attributes within current event models. Longer reading times were found for the *day later* condition sentences compared to the no shift condition as predicted. However, the time shift effect was not evident in the RT data.

Additional analyses were conducted by Radvansky and Copeland (2010) based on two different types of object probes that were embedded throughout the stimuli. These probes included references to objects associated with the protagonist which were expected to be maintained across the time shift (e.g. a credit card in a pocket) and to objects that were removed or discarded by the protagonist (e.g. a sweater that was taken off). Participants' probe responses were highly accurate for both maintained and discarded object types. As expected, there was no RT difference between the no-shift and shift conditions for the maintained objects, however there was also no difference for the discarded or removed objects. The authors' interpretation of this unexpected finding was that readers were still in the process of updating the event attributes/entities in the current event model when the probe of the removed object occurred. Because the updating process was still ongoing, the removed object was still quickly accessible to the comprehender after the time shift. Based on a careful review of the example stimulus item, however, it is plausible that the ‘removed’ objects (e.g. a phone number the protagonist gave to a woman who put it in her purse) may have continued to be relevant to the protagonist and the storyline and thus were not tagged for removal from the event model. The protagonist’s phone number still belonged to him even though he wrote it down and gave it to a friend.

There are several other study design elements and limitations to note in Radvansky and Copeland’s (2010) Experiment 2. Changes occur in multiple dimensions of the situation model (e.g. a location change from someone tanning outside in the sun to an interaction held inside a bank) in addition to shifts that range between *a moment later* and *a few weeks later*. The wide variety of time shifts coupled with shifts in other dimensions make the storylines challenging to follow and the results difficult to interpret. In addition, the referents for the probe words were mentioned more than once so the stimuli were not controlled for repeated exposure to probe words when collecting RTs. Finally, the majority of the prior literature on the time shift effect investigated a shift of exactly *an hour later* while the shift in Radvansky and Copeland’s study was *a day later*. It is unclear if the absolute duration of the time shift (e.g. an hour versus a day later) differentially impacts how likely a comprehender is to start a new event model, however, it is likely that the *day later* cue is more likely to signal a new event model than the *hour later* cue.

A third investigation with results contrary to the time shift effect was a reading study conducted by Weingartner and Myers (2013). These authors examined the processing consequences of time shifts signaled by adverbial phrases (e.g., after a moment/an hour) at the narrative level. In Experiment 1, eye fixation times were collected for both time shift phrases and the phrases immediately following them to assess for carryover effects. In Experiment 2, manual RTs were measured to probes of the attributes of events that were introduced prior to the time shift. The stimuli for this study were borrowed from Zwaan, Madden, and Whitten's (2000) experiment, one of the first to examine event continuity.

Zwaan and colleagues (2000) used stimuli that included events that could be perceived as ongoing after a time shift forward on the narrative timeline. Here is a sample stimulus item, which includes both a time shift and a verb manipulation to indicate whether the activity was ongoing or completed:

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

Zwaan et al. found increased RTs to recognition probes if the event was described as discontinued (e.g. stopped watching) for the shift condition, but no difference if the event was described as ongoing (e.g. was watching). The authors' interpretation was that readers maintain activation for relevant event information, therefore the event attributes related to the football game were maintained for the "was watching" condition. Deactivation only occurs if the event and its attributes are perceived as no longer relevant to the described situation. This conclusion is generally consistent with the predictions from both the Event-indexing model (Zwaan, et al., 1995) and the Structure Building Framework (Gernsbacher, 1990), however one open question is how quickly event information is deactivated. Recall that in the Radvansky and Copeland study (2010) described above, discarded objects that were expected to be tagged for removal from the event model were maintained in the model after the time shift. Other open questions include whether there are patterns or uniformities to deactivation processes and what types of information might be deemed relevant and continue to be easily accessed after a time shift.

Extending the methods from the Zwaan et al. (2000) study, Weingartner and Myers (2013) collected eye tracking measurements while young adults read 24 short narratives. Experiment 1 used the same stimuli as Zwaan et al. and an anaphor resolution task. The authors predicted that if a *discontinued* event is indeed deactivated following a time shift, then there should be an increase in fixation time for the anaphoric reference to that event. Conversely, the authors predicted no difference in fixation times for the no shift and shift conditions for events that were described as *ongoing*, because the events and their attributes would still be relevant to the described situation. Mean fixation times on the anaphor were collected to measure the accessibility of an event that was introduced prior to the timeframe conditions.

Both the first pass and total fixation times were reported for three regions of interest including the temporal adverbial phrase, the post adverbial region, and the anaphor. An example of the anaphor in this task was the word PAINTING in a story about an art gallery opening as a prior allusion to the implicit referent '*making additions to a portrait*.' Results were consistent with the time shift effect until the anaphor resolution analysis was conducted. Critically, the fixation data for the anaphor resolution aspect of the narratives did not reflect a time shift effect. There was no difference between the no shift and shift conditions for reading times of anaphoric references to prior event attributes. One interpretation was that establishing an event boundary does not necessarily trigger a decrease in the activation of prior event information, consistent with the flexible boundary hypothesis (Section 3.0). It is also possible that the link to a referent that was only implicit was too weak to generate the time shift effect. An alternative explanation given by the authors is that anaphor resolution and probe recognition promote different underlying processing mechanisms. In the time processing literature, probe recognition tasks are the most commonly used method when the time shift effect has emerged. To explore this possibility further, the authors made a modification to the Zwaan et al. stimuli. The sentence that included the anaphoric reference to the critical event was eliminated in the second experiment. The sample stimulus provided for Experiment 2 was:

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 condition: Late in the afternoon he *finished painting* a portrait.

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

ADVERBIAL PHRASE: After a moment/an hour he heard a siren outside.

PROBE: PAINTING

When the alternative explanation was tested, the time shift effect (shift RTs > no shift RTs) was found. The authors concluded that a probe recognition task likely induces the post-comprehension process of searching the prior representation and may cue the comprehender to more fully consider the relevance of a given probe in the prior event model (Keenan, Baillet, & Brown, 1984; O’Brien, Duffy, & Myers, 1986).

The Weingartner and Myers (2013) study adds to the emerging literature which points to a need for more refined predictions related to the time shift effect. Limitations of this work included implicit referents to the anaphor like in the original Zwaan et al. study (2000) and inconsistencies in the types of probes used for the anaphor resolution task compared to the probe recognition task. For the anaphor resolution task, the anaphor was an object (e.g. "piece" as anaphor for sculpture) while the recognition probe often referred to an activity (e.g. "painting"). The authors did not provide any predictions or discussion about how different types of event attributes (e.g. objects versus activities) that remain accessible following a time shift might vary, although this variation could influence study results and interpretation. For example, "piece" in the anaphor resolution task may not have been as salient a referent as "painting" was in the probe recognition task. It is possible that comprehenders tag event concepts differently during event model updating making it difficult to predict which event attributes remain relevant after a time shift. The authors did acknowledge that the explicit discontinuation of an event activity interrupted the protagonist's originally stated goal and perhaps this provided an additional cue to create an event boundary (see, e.g. Linderholm et al., 2004). Clearly, further work is needed to refine the predictions for when event boundaries are established, what types of event attributes may remain available or accessible after a shift forward on the narrative timeline, and the status of event completion during comprehension.

In addition to probe recognition as the primary task when a time shift effect was found, a closer inspection of the verb types in the sample stimuli from the primary studies of the time shift effect indicates that verbs and the events that were selected may have unintentionally inflated the likelihood of the effect. For example, whether an event could plausibly be perceived as ongoing after the time shift appears to strongly influence whether or not the time shift effect emerges as demonstrated in Zwaan's (1996) seminal article. Strengthening this argument, Zwaan, Madden, and Whitten (2000) continued their work by conducting a reanalysis of the items used in the original 1996 study. The three stimuli that included the word "started" paired with an action probe word (i.e. started typing/watching/talking) did not produce the time shift effect even though there was a reliable difference between the moment and hour texts across all stimuli. The importance of event continuity echoes a central claim of the *scenario model* (Anderson, et al., 1983), one of the original models proposed for time processing which will be discussed in the next section of this document. Moreover, the type of verb (i.e. activity or accomplishment) appears to influence the contents of the event model and will also be addressed in the next section. Further investigation is needed to provide a richer understanding of event continuity as a contributing phenomenon to the formation of event representations.

The remaining sections of this chapter discuss how comprehenders may perceive the different ways time cues are signaled. The literature pertaining to event continuity will be emphasized, an area that is not fully considered in Block's (2014) categories. It should be noted that the terms event duration and event continuity often have been used interchangeably in the time processing literature. In this document, event duration is defined using Block's psychological time meaning: how long an event lasts or a length of time between events. Event continuity refers to the completion status of an event (i.e. whether the event is perceived as in progress or completed). Each will be discussed in turn in the following sections.

3.2 TIME: DURATION

In the time processing literature, duration is a multifactorial concept. Event duration can be explicitly stated using a temporal adverbial (e.g. an hour later), accented using grammatical or lexical cues (e.g. imperfective), or implied via a location change or other pragmatic cue (e.g. the long journey). Event duration has a proven influence during investigations that use off-line measurements (e.g. Yap et al., 2009; Zwaan, 1996) and for studies using on-line measurements such as word-by-word reading times (Coll-Florit & Gennari, 2011) and ERP (e.g. Becker, Ferretti, & Madden-Lombardi, 2013). In addition, a few researchers have proposed that event duration is stored with other event related concepts similar to the attributes or perceptual features of objects (Barsalou & Sewell, 1985; Nottenburg & Shoben, 1980). Based on the evidence, event duration is clearly a critical element that contributes to event model formation and updating.

Anderson et al. (1983) have shown that a comprehender's background knowledge about the prototypical duration of an event is one type of cue that comprehenders use to determine the completion status of an event. In the study on which the scenario model is based, the temporal duration of the activity or scenario described in the narrative provided natural endpoints or boundaries for the beginning and end of an event model. For example, if the protagonist was watching a movie, a time shift of *ten minutes later* was clearly within the bounds of the scenario, but a time shift of *seven hours later* was clearly out of bounds of the scenario. Specifically, the authors demonstrated that general knowledge about the durations of events impacts anaphor resolution. This study spawned much of the time shift research discussed in the prior section of this document, and the few investigations pertaining to rigid versus flexible event boundaries.

The most relevant and persistent conclusion from the scenario model relates to the deictic shift hypothesis. This hypothesis states that events and their attributes that are "in" the current model are more accessible than those connected to a prior model (Zwaan et al., 2000). For example, a probe of WAITER would be responded to more quickly when the characters were still in the RESTAURANT scenario (e.g. during a 1-2 hour span) compared to when the narrative timeline has exceeded the prototypical duration of a meal (e.g. several hours later). While this model provides a starting point for identifying how event boundaries might be established with respect to typical event durations, a finer-grained approach is necessary to better understand how comprehenders process event attributes and properties within the boundaries of individual events.

3.3 EVENT CONTINUITY

As outlined in Section 2.2.2, comprehenders routinely use a narrative now point to locate and sequence events (Almeida & Shapiro, 1983; Duchan, Bruder, Hewitt, 1995; Hopper, 1979). Both verb tense (e.g. past, present, and future) and temporal adverbials (e.g. before, after) provide necessary pieces to the comprehension puzzle which facilitate the understanding of when an event(s) takes place on the narrative timeline. Verb aspect contributes by conveying the time course and the completion status of an event or activity. The completion status of an event, whether an event is perceived as ongoing or completed, appears to strongly influence the contents of a mental representation of that event.

Within the time processing literature, different types of temporal cues have been used to signal to a comprehender that an event is ongoing or completed (e.g. Carreiras et al., 1997; Coll-Florit & Gennari, 2011; Ferretti et al., 2007; Madden & Ferretti, 2009; Magliano & Schleich, 2000). These cues include grammatical and lexical verb aspect and prototypical event semantics. Each type of time cue has been shown to affect the contents of mental event representations and will be discussed in the following sections.

3.3.1 Grammatical verb aspect

Grammatical verb aspect is one type of time processing cue that highlights event information pertaining to onset, duration, and event continuity. Even subtle syntactic changes promote distinctly different mental representations (e.g. Ferretti, McRae, & Hatherell, 2001). Despite its apparent critical contribution to the formation of an event representation, there are relatively few studies which have examined the role of verb aspect in event continuity during language comprehension. The bulk of the literature addressing verb aspect emerges from linguistics (Comrie, 1976; Dowty, 1986; ter Meulen, 1995; Vendler, 1967) but cross-pollination to psycholinguistic theories of event processing is underway.

In the event continuity literature, grammatical aspect is the type of temporal cue that has been studied most frequently. One way an event is signaled as ongoing or completed is by changing the grammatical aspect of the verb (e.g. Ferretti et al., 2007; Magliano & Schleich, 2000; Yap et al., 2009). Specifically, imperfective aspect is used to indicate an event that is ongoing (e.g. was fishing) while the perfective aspect portrays the event as completed (e.g. had fished). Distinct mental event representations emerge even though both of these verbs are in the past tense. Researchers infer based on RTs and reading times that verbs and verb phrases in the imperfective (i.e. ongoing) aspect highlight attributes within the event as if it is in process. Conversely, the perfective aspect signifies a completed event where the end state of that event is emphasized as if the comprehender is envisioning a static model of the completed event. If an event is perceived as completed, longer RTs are expected to probes of attributes or objects from that event.

Digging a little deeper into how grammatical aspect influences the perception of event completion status, Moens and Steedman (1988) propose a close connection between the temporal components and causal properties of an event. In their view, there are three causal phases of an event: 1) a set of initiating conditions, 2) the actual event, and 3) the resultant states. Critically, Ferretti et al. (2007) extend this proposal by predicting specific categories of activated event attributes linked to each event phase. For instance, agents and instruments are integral to setting up an event, therefore, those attributes are strongly associated with the initiating conditions phase. The actual event (i.e. the second phase) promotes a broader set of attributes which could include the initial agents and instruments but also patients and locations. Finally, the resultant phase emphasizes the role of patients at the end of an event. Differential activation for categories or types of event attributes based on aspect connects to the two theoretical perspectives of how event knowledge shapes mental representations that will be discussed in Chapter 3.

One influential study of grammatical aspect used the categories of event attributes described above to demonstrate the connection between temporal and causal properties of events (Ferretti et al., 2007). That study built on prior results showing systematic priming effects from verbs to patients, instruments, and agents, but not to locations. The authors proposed that the expected priming effect for locations was not found because the sentences did not emphasize the correct event phase (i.e. the actual event). As previously indicated, the imperfective aspect cues the comprehender to focus on the internal structure of an event as if it were underway. When the stimuli were modified to incorporate the imperfective aspect, priming for locations emerged. This result is one example of a trend in studies of grammatical aspect, which are outlined next.

Across studies of event processing, contrasts between perfective and imperfective aspect yield systematic findings. An early series of studies by Morrow (1985, 1990) examined the availability of concepts within the mental representation while comprehenders tracked a character moving from one room to another in a house. Participants associated different locations for a given character depending on the grammatical aspect used to describe the movement. When the imperfective tense was used (e.g. was walking) the character was inferred to be “en route” between the two rooms. However, when the perfective tense was used (e.g. walked) participants consistently responded that the character was in the destination room. This pattern of results also lines up with predictions about location changes from the Event-indexing literature discussed in Section 2.1.

The type of studies conducted by Morrow (1985, 1990) highlight one possible function of verb aspect, to cue the comprehender when to update an event model as completed (i.e. perfective) or when to emphasize the ongoing event (i.e. imperfective) (Magliano & Schleich, 2000). Another possible function of verb aspect is that it may differentially “tag” information depending on whether or not the comprehender might need to have it available as the discourse proceeds (Magliano & Schleich, 2000). This is particularly relevant when considering the role of the imperfective aspect. Two studies (Carreiras et al., 1997; Schramm, 1998) link the use of the imperfective tense with an increase in mental activation and subsequent availability of target concepts at a later point when compared to either the perfective or past perfect tense. This possible function of verb aspect connects to the open questions related to what types of events and their attributes may be more or less available after an event boundary. Finally, Hopper (1979) claims that verb aspect may highlight which story elements are in the foreground or the background on the narrative timeline. Hopper’s proposal is that events described using the imperfective tense provide the background of a story while the foreground is defined by events in the perfective tense.

One seminal investigation into these possible functions of grammatical verb aspect is a narrative level study conducted by Magliano and Schleich (2000). Using a situation model framework, the authors manipulated verb aspect to address whether the imperfective/perfective distinction affects 1) the perception of event duration and 2) the availability of event information in the mental representation. In the first of four experiments, participants completed a self-paced sentence by sentence reading task of short narratives that included an introduction to an activity, critical sentences that manipulated verb aspect (perfective, imperfective), three temporally ambiguous sentences that continued the story, and a conclusion. Below is an example stimulus item provided by the authors. A critical test question (CTQ) was asked at one of four possible positions throughout the story as indicated below.

CRITICAL TEST QUESTION (CTQ): Has the baby been born yet?

INTRODUCTION: Jack's wife Betty was expecting a baby, and boy was he excited. He was planning to be her coach when she gave birth. He went to all the Lamaze classes. Every night, Jack made Betty practice her breathing. Finally, the big night was here.

ASPECT SENTENCES:

Betty was delivering their first child. (imperfective) CTQ or

Betty delivered their first child. (perfective) CTQ

POST ASPECT SENTENCES: Jack fainted dead on the spot. (CTQ) The video recorder went crashing to the ground. (CTQ) A nurse had to stop what she was doing and help Jack. (CTQ)

CONCLUSION: Jack recorded about 10 seconds of the birth of their child. He was never more embarrassed in his life.

For the critical test question, participants were asked to respond Yes/No to whether the activity was completed (incomplete/ongoing = no; completed = yes). Additional comprehension questions were included to encourage careful reading.

This series of four experiments produced multiple results, however, one of the most relevant for the current study included a contrast between short and long duration events in Experiment 2. Participants were asked to indicate (yes/no) 1) whether a verb phrase (target activity) was part of the prior narrative and 2) if the critical activity was still ongoing at four time points after the activity was introduced. Critically, for the imperfective aspect, there was a significant difference between sentence position 1 and 4 but only for the long duration events. This could be indicative of a progressive nature to the mental representation for longer duration events that are cued by the imperfective tense. This finding provides a hint toward more fine-grained incremental processing occurring for events that are perceived as ongoing, particularly if the events have a longer duration (e.g. at least two hours in this study). This experiment is one of the first to demonstrate an interaction between event duration and grammatical verb aspect.

Prior to discussing another pertinent set of results from Magliano and Schleich (2000), a brief turn is needed to the role of working memory in building and updating mental event representations. To promote discourse coherence, the mental activation level of event information appears to fluctuate with a tradeoff between what is in the foreground and background at a given point in time (Glenberg et al., 1987; Sanford & Garrod, 1981). Comprehenders use cues (e.g. grammatical aspect) to tag events and their attributes that should be maintained in working memory or removed from the current event model. It is well established that there are both limitations and individual differences in working memory capacity (e.g. Just & Carpenter, 1992). The Structure Building Framework (Gernsbacher, 1990; 1997) and the Event-indexing model (Zwaan et al., 1995) highlight working memory as a key determinant of how incoming information is tagged during comprehension processes. If the content is deemed relevant, it is mapped onto the current developing structure or incorporated into the existing event model. However, incoming information might promote a shift to a new substructure or a discontinuity in one of the event indices prompting the system to update the current event model. Shifting and updating processes can be resource consuming and often result in increased reading times or RTs. Quickly interpreting cues such as grammatical aspect becomes a critical factor for comprehension efficiency.

To test grammatical aspect as a cueing system, Magliano and Schleich (2000) used two word verb phrases (e.g. deliver baby or change tire) as probes for a recognition task while participants completed sentence level self-paced reading of the narratives from Experiment 1. Participants were asked to decide as quickly and accurately as possible whether the given verb phrase was part of the activities that had occurred in the story, so there was no explicit consideration of the temporal status of the probed activity. A computerized working memory span task (Kim, Millis, & Langston, 1999) was administered after the verb phrase recognition task. Readers were divided into high span (N=39) or low span (N=42) groups based on their performance.

There was a significant interaction for the working memory span by aspect by sentence position mixed analysis of variance (ANOVA) for both items and subjects. *Post hoc* analyses showed that working memory span was differentially related to RTs for the critical test question at sentence position 1 and 4. For high span readers, there was no difference between in progress and completed stories at sentence position 1. For sentence position 4, the high span group responded faster to the verb phrase probes for the in progress condition compared to the completed condition. This pattern suggests a higher degree of mental activation was maintained through sentence position 4 for the verb phrases in the in progress condition. The low span group did not evidence this activation advantage for the in progress stories. The investigators linked the low span group’s performance to a presumed inability to maintain activation for the in progress condition at the same level as the high span group.

This study by Magliano and Schleich (2000) is one of the very few studies examining the role of grammatical verb aspect during discourse level processing and therefore it is ubiquitously cited throughout the event processing literature. The results are generally consistent with both the Structure Building Framework and the Event-indexing model and show that comprehenders are sensitive to subtle syntactic cues such as verb aspect. This study provided some preliminary evidence that the influence of grammatical aspect remains for up to three sentences after the activity is introduced. However, there is no additional evidence from this study or in the broader literature base to corroborate this finding. Event duration was also implicated in this study via its interaction with verb aspect in the second of the series of four experiments. In particular, differences between sentence positions 1 and 4 emerged exclusively for in progress stories with long duration events.

Several study limitations merit discussion. One potentially important limitation of this study is that the authors investigated only a single verb type, activities with “duration” (p. 91), to try to provide the strongest test of verb aspect. This choice included activity verbs with a high rate of variability in duration as determined via an event duration estimation task (e.g. short event range: 10.29 seconds for scratching your nose to 32.08 minutes for packing a suitcase; long event range: 2 hours for watching a movie to 1.17 years for writing a novel). Considering the strong effect of duration reported in Experiment 2, it is likely that collapsing activities across various durations may have influenced the results for this series of experiments.

3.3.1.1 Perceptual studies

Briefly mentioned in Section 2.2.2, simulation theory is one proposal for how mental event representations are constrained during comprehension. A vast number of psycholinguistic investigations have emerged from this perspective, including studies of verb aspect. Simulation theories propose that during semantic processing, the primary sensory and motor networks are activated in a similar pattern as if the comprehender were actually experiencing what is read or heard. Zwaan (2003) is one strong proponent of simulation theory, extending his predictions from the Event-indexing model into what he calls the Immersed Experiencer Framework. A quote that exemplifies Zwaan's position is "Language is a set of cues to the comprehender to construct an experiential (perception plus action) simulation…the comprehender is an immersed experiencer of the described situation and comprehension is the vicarious experience of the described situation" (Zwaan, 2003, p. 37). More specifically, the neural traces of the perceptual experiences are activated in a systematic manner to facilitate linguistic processing. A brief outline is provided next of key studies outlining trends from the simulation theory perspective, with an emphasis on the findings from investigations of verb aspect.

Zwaan (2003) proposes that there is a specificity to perceptual activation patterns during reading or listening, calling on an economy of processing principle (Graesser, Singer, Trabasso, 1994). In essence, for comprehension processes to function efficiently, a comprehender should not activate more information than is necessary to understand a given situation. One classic comparison from his work involves the following sentences:

- The ranger saw the eagle in the SKY.
- The ranger saw the eagle in the TREE.

Simulation theories predict distinctly different perceptual traces when reading the word "eagle." For instance, the shape of the bird (wings outstretched/wings folded in) and the perceived height of the bird (sky/tree) would provide a unique set of experiential traces to the comprehender. These select traces are then used to form the appropriate (visual) representation of the event. Much of the simulation theory literature focuses on a constrained set of traces that are activated using action or motion verbs. For example, Hauk and colleagues (2004) used event-related fMRI to examine activation of brain regions when participants passively read action words with specific body part referents (e.g. lick, pick, and kick). In addition to the expected activation of typical language network regions, adjacent and overlapping brain areas along the motor strip were activated for body parts consistent with the areas that are active during actual movement. The stimuli in this study carefully controlled for psycholinguistic attributes of the words, a potential confound in other studies. Other researchers have also found this pattern of adjacent and overlapping brain area activation (e.g. Pulvermüller, Shtyrov, & Ilmoniemi, 2003) including when action words were embedded in sentences (Tettamanti, Buccino, & Saccuman, 2005). Zwaan (2003) further emphasizes the differential activation patterns for tool words in motor areas versus animal words activating predominantly visual areas (Chao & Martin, 2000).

A common finding across the simulation theory literature is a congruity effect. An assumption of this perspective is that there is immediate activation of an object's implied perceptual attributes during language comprehension (Dahan & Magnuson, 2006). Further, general knowledge is used to inform and constrain the activated experiential perceptual traces (e.g. eagle-SKY/TREE example above). Prototypical responses for the congruity effect include faster RTs to pictures that match the shape, orientation, and/or color of objects in a target sentence when compared to a mismatched picture (e.g. Zwaan & Pecher, 2012). One example is a faster RT to a vertical image of a pencil for the first sentence and a horizontal image for the second sentence in this pair:

- The pencil is in the CUP.
- The pencil is in the DRAWER.

This type of priming effect is quite robust and has been found using a multitude of tasks such as making semantic judgments referencing smaller parts of a larger object (e.g. the ATTIC/BASEMENT of a house) (Zwaan & Yaxley, 2003) and for pictures of facial expressions paired with pictures of congruent body language positions (Meeren, van Heijnsbergen, & de Gelder, 2005).

Another way these types of priming effects have been described is in terms of affordance or action compatibility effects. The term “affordance” comes from Gibson’s theory (1977) which emphasizes the potential spectrum of uses of a certain object. One example is making a sensibility judgment for protecting one’s face from the wind using a newspaper versus a matchbook (Kaschak & Glenberg, 2000). Making judgments about salient versus non-salient affordances yields faster RTs for affordance compatible stimuli. Action compatibility effects stem from other work by Glenberg and Kaschak (2002). In one study, faster RTs for sensibility judgments using a button press were found for stimuli which matched the direction of the transfer of an object (e.g. either toward or away from the participant for CLOSE THE DRAWER). The transfer effect was also found when the stimuli were extended to abstract concepts like transferring a message (Kaschak et al., 2005).

Applications of simulation theory are emerging in studies of verb aspect. One by Madden and Therriault (2009) extends the premise of an affordance or action compatibility effect to a proposed object use effect such as whether or not a target object is in use. Target sentences included

John was working/had worked on his laptop at home paired with a picture of either an open or a closed laptop. Participants completed a self-paced word-by-word sentence reading task with one of two versions (in use/not in use) of a picture replacing the target object in a sentence. The primary task was a yes/no sensibility judgement. In addition, reading times were collected for the segment of the sentence containing the picture and the two words that followed the picture.

Consistent with other findings that incorporated an imperfective/perfective contrast, a use effect emerged for both on-line sentence reading and sensibility judgments. A main effect of use was found when pictures replaced the objects, with faster response times for use-congruent pictures. Faster reading times occurred for the first word after the picture for imperfective sentences with in use pictures. This imperfective advantage continued to the measurement taken at the second word after the picture. For perfective items, the use effect was demonstrated at the picture only and did not continue for either of the following two words. This finding could be linked to inherent differences in the pictures, so the authors conducted a follow up study that confirmed that it was the context of the sentences and not the characteristics of the pictures that produced the imperfective advantage effect. The primary study was not able to separate out whether the imperfective advantage is due to sentence wrap up effects (e.g. Rayner, Kambe, & Duffy, 2000) or additional post sentence processing necessary to answer the sensibility judgements. It also does not directly address the possible temporal phases of using an object (e.g. a lit match or a burnt match) or the implied or stated progression of certain event attributes, factors that may impart their own influences on the mental representation.

3.3.1.2 Temporal phases within event models

One study that takes a rather clever approach to addressing the possibility of temporal phases within a given event incorporates the idea of attribute progression within event models.

Welke and colleagues (2014) tested the premise that events are experienced as dynamic occurrences in time by focusing on objects that begin with a particular source state that changes as the event progresses into a resulting state. The authors build on the work by McRae's group (McRae, de Sa, & Seidenberg, 1997) and Ferretti et al. (2001, 2007) which emphasizes that a set of the most frequent and/or intercorrelated semantic attributes for objects and people is activated during the processing of event verbs. In addition, Welke et al. merged predictions from a study conducted by Nuthmann and van der Meer (2005). That study is briefly considered next before returning to discussion of Welke et al.'s work.

Nuthmann and van der Meer (2005) asked whether the perceived forward *directionality* of time's psychological arrow is mentally tracked. In this study, individuals were indeed sensitive to the chronological progression of event states in a semantic relatedness judgment task. Sensitivity was indexed by faster RTs to judgments of verb states which matched future time orientation (e.g. shrinking-SMALL was responded to more quickly than shrinking-LARGE). This processing advantage was found at both stimulus onset asynchronies that were tested (i.e. 250 ms and 1000 ms). The saliency of activated event states and sensitivity to chronological changes as events progress spawned the notion of progression attributes studied by Welke and colleagues (2014).

In the Welke et al. (2014) study, participants were presented with an event verb prime, such as WASHING or CUTTING. These verbs were followed by two adjectives, presented simultaneously, that signaled different states of progression in the target event as in the following example:

- Washing (event prime) DIRTY-CLEAN or CLEAN-DIRTY
- Cutting (event prime) LONG-SHORT or SHORT-LONG

The participants were asked to respond using a manual button press if the adjectival states were interrelated. These authors propose that a *pattern completion inference mechanism* (Barsalou, 2009) is enacted when a source state (e.g. dirty) is presented such that the resulting state (e.g. clean) is likely to be predicted. This prediction appears to be unidirectional in nature, like time's arrow, as source states are not strongly associated if a resulting state is presented first (Friedman, 2003).

To test these expectations, adjectives were presented simultaneously in their chronological order (source-result), in reverse chronological order (result-source), and in an unrelated pair (i.e. starting with a source state followed by an unrelated resulting state or two unrelated states). A pattern completion effect was predicted for adjectives presented in chronological order, the same type of effect that has been demonstrated in other types of priming studies and the congruity studies within time processing in particular. This study also included two different stimulus onset asynchrony (SOA) intervals, 250 and 1,000 ms. Consistent with the McRae and colleagues priming study for locations of events (McRae, Hare, Ferretti, & Elman, 2001), a 250 ms SOA is associated with more automatic processing while a 1,000 ms SOA elicits more strategic processing.

As predicted, a priming effect was found for source and resulting states that were related to the event verb prime and there were no SOA influences. However, there was priming regardless of chronological order. This means there was not an order advantage (i.e. chronological or inverse) as long as both words were related to the event prime. The authors concluded that while events activate both source and resulting states, the temporal order is (perhaps) not part of the knowledge associated with that event. An additional interesting finding was that the source-unrelated adjective pairs produced more errors and longer RTs than the other conditions. It was proposed that the longer and more error prone responses were a result of interference from the automatic activation of the resulting state, consistent with Barsalou's (2009) pattern completion inference mechanism. Comprehenders must override the resulting state that is automatically activated in order to respond "no" to a source-unrelated adjective pair. This prediction has been substantiated in at least one other study where resulting event states were responded to faster than source states when adjective pairs were presented sequentially (versus simultaneously) (Nuthmann & van der Meer, 2005).

Another way to examine temporal phases during event processing is to consider the temporal position of the events within an activity. One study that examined early versus late time points in an event sequence was conducted by Landgraf and colleagues (2012). In this pupil monitoring and eye movement study, a typical event such as GOING TO A RESTAURANT was introduced followed by event pairs which signaled an early or late subevent such as GET MENU and PAY THE BILL. The event pairs were presented in either the correct chronological order or the inverse order just like the literature for source-result states. Overall, the predictions and results for temporal phases match those of the source-result state studies.

3.3.2 Lexical aspect

While the majority of studies addressing verb aspect have focused on grammatical manipulations, two have also incorporated a lexical aspect contrast. Both investigations explored how lexical aspect (i.e. signaling accomplishments vs. activities) interacts with grammatical aspect to constrain mental representations (Becker et al., 2013; Yap et al., 2009). Yap and colleagues determined that sentences with a perfective cue were processed more quickly and accurately with accomplishment (i.e. natural endpoints) verbs (e.g., knitting/knitted) using a sentence-picture matching task in Cantonese. In a complementary fashion, sentences with an imperfective cue were processed more quickly and accurately with activity (i.e. no natural endpoint) verbs (e.g., swimming/swum). These results are indicative of facilitation effect (i.e. faster and more accurate responses) for the imperfective + activity and perfective + accomplishment combinations.

The Yap et al. (2009) study highlights a potentially fundamental difference between accomplishment and activity verbs during event processing. Using a simulation perspective, the authors propose that the natural endpoints of *accomplishment* verbs provide a very strong perceptual contrast between the ongoing and completed stages of an event. On the other hand, the lack of natural endpoints for *activity* verbs promotes very similar mental representations for ongoing and completed events. This interpretation echoes an early unanticipated finding about the nature of ongoing event representations for accomplishment verbs reported by Madden and Zwaan (2003). The authors used sentence-picture matching with the imperfective/perfective grammatical contrast described in Section 3.3.1.1. The stimuli for three experiments included sentences like: *The man was making/made a fire* and one or two pictures of the events described. One picture depicted the event as if it were in an intermediate stage (e.g. a man putting logs into a fireplace) and the other was of an action close to or at the end of the event (e.g. a man looking at a blazing fire).

In Madden and Zwaan's (2003) first experiment, participants read the sentence and saw both pictures at the same time. Their task was to press a button on the keyboard marked LEFT or RIGHT to select the picture that best matched the sentence. In the second experiment, both versions of the sentence were followed by only one picture and participants responded yes or no whether the picture matched the sentence. Faster RTs were expected for matching pairs (i.e. imperfective + ongoing picture; perfective + completed picture) when compared to mismatching pairs. Finally for the third experiment, one of the pictures was presented before each sentence type to determine if the picture would facilitate event model formation. The dependent variable for this experiment was how long it took participants to both read the sentence and make the relatedness judgment to the picture.

The results of this series of experiments by Madden and Zwaan (2003) generally mimicked the pattern of results from other studies of grammatical verb aspect for *completed* events. When both pictures were presented, participants were more likely to choose the completed picture after reading the perfective sentence. A parallel result was found when only one picture was presented (i.e. faster RTs to completed + perfective) and when the picture was presented before the sentence.

The results, however, were not as predicted for ongoing events and the pattern of results for the imperfective sentences was unexpected. In Experiment 1, participants showed no strong preference for the ongoing picture after reading an imperfective sentence. So for sentences that cued an ongoing event, participants were equally likely to select either picture. The lack of an advantage for imperfective sentences paired with ongoing pictures continued through Experiments 2 and 3. The authors' interpretation was that imperfective and perfective sentences constrain mental representations in different ways. While the perfective constrains to the completed end state of an event, the imperfective *does not* constrain to an intermediate stage of a described event. The authors extend this interpretation to include the caveat that there could be individual differences in how comprehenders simulate ongoing stages of events. This possibility introduces a potential mismatch between what the reader was mentally representing as ongoing and the in-process stage portrayed in the image. Critically, the unexpected result for the imperfective sentences in this study opens the door for further investigation into events that are cued as ongoing.

Another investigation to include a lexical contrast was conducted by Becker and colleagues (2013). The authors used ERP to study the relationship between lexical aspect, grammatical aspect, and duration of intervening events in short narratives. Results showed faster RTs for accomplishment verb probes in the imperfective aspect (relative to perfective) for targets that followed short, but not long, intervening events. This finding is contrary to the sentence-level results of Yap et al. (2009) where facilitation was also demonstrated for activities. Becker et al. propose that the long intervening events were used as a cue to advance the timeline beyond where the advantage of grammatical aspect would be a factor. It is possible though, that it was actually the additional context provided by the narratives that prompted participants to update their event models. Another possibility is that the off-line RT measurement missed the effect. A close inspection of a sample stimulus shows that the authors chose events that varied inherently in their durations (e.g. accomplishment-to pack a lunch; activity-to exercise). Additionally, there are plausibility issues with the sample stimulus that could have contributed to the unexpected pattern of results. Additional investigations are needed to capture the relationship and influence of multiple temporal cues.

3.3.3 Limitations in the event continuity literature

A few limitations are common across the studies that incorporate verb aspect contrasts to study how time processing cues constrain mental representations. These include using only one verb type and not controlling for other known time processing cues in the stimuli. For example, Madden & Zwaan (2003) and Morrow (1985, 1990) used verbs from a single lexical category because grammatical aspect was their contrast of interest. On the other hand, Ferretti and colleagues (2007) and Magliano and Schleich (2000) included verbs from multiple lexical categories without contrasting verb classes. Multiple studies have shown that manipulating either lexical aspect or grammatical aspect can signal whether an event is ongoing or completed (e.g. Carreiras et al., 1997; Coll-Florit & Gennari, 2011). Critically, the event attributes that are active in the two types of event models are distinct. In addition to event continuity cues, event duration appears to be an influential factor on its own as it often interacts with verb aspect. Studies targeting the interaction of these time cues are beginning to emerge (e.g. Becker et al., 2013), however, the extent to which each factor influences mental event representations remains elusive.

While the traditional approaches for studying time processing effects have included judgment tasks, there are multiple drawbacks to this type of method. First, the timing of when the judgement is taken might obscure an effect that could be found in a more on-line measurement or one that occurs earlier in the processing stream. Second, judgement tasks often include a closed response set of only two choices. A forced choice paradigm is constraining in a couple of ways. If pictures are used, the images may not match the exact mental representation of the comprehender which could prompt additional processing to resolve the potentially competing interpretations. Some researchers in time processing are using more on-line measures such as event-related potentials (ERP) or eye tracking to index, for example, the processing of syntactic or semantic violations. The current investigation uses a lexical decision task to provide an implicit measure of potential semantic priming effects.

Finally, a potential gap in the current time processing literature is its treatment of event continuity as an all or none phenomenon. Within this literature base, an event is either categorized as ongoing or it is completed with concomitant processing repercussions (i.e. longer reading times or RTs) for each event category. The study outlined in this document will begin to address this gap and specifically proposes that current theories of event processing should incorporate a more temporally fine-grained approach to event updating.

3.3.4 Parallels to the spatial domain

Findings from the spatial domain point to the potential utility of examining the internal structure of ongoing event representations. Space is one of the five indices tracked by comprehenders according to the Event-indexing model (Zwaan et al., 1995; Section 2.1). There is a strong correspondence between time and space and a shift in one domain often promotes a change in the other (Rapp & Taylor, 2004). For example, if a protagonist is walking across the room or walking across town, the comprehender adapts the current event model to accommodate the time change prompted by the shift in location (e.g. longer time shift to walk across town). Zwaan (2003) also described the highly dependent relationship between time and space in his Immersed Experiencer Framework that embraces the principles of simulation theory (Section 3.3.1.1 and also discussed in Chapter 4). Casasanto and Boroditsky (2008) found recruitment of spatial metaphors is commonplace when individuals are generating non-linguistic judgments about time. Finally, Lakoff and Johnson (1980) underscore that the relationship between time and space is asymmetrical where space concepts are invoked more frequently to describe time.

Real-time processing differences for event descriptions of how actions (e.g. changes or movements in space) are conceptualized have emerged from studies that manipulate grammatical aspect, similar to the results in the temporal domain. Matlock (2011) has conducted a series of experiments where sentences written in the past progressive elicit inferences that include more action in a given time period when compared to the simple past. For example, the authors provided event descriptions written in the past progressive (e.g. *was painting* or *was planting*) and simple past tense and asked participants to provide a volume estimate for each sentence type. Past progressive sentences elicited reliably higher estimates than those written in the simple past tense, for either the total number of houses that had been painted or the total number of pine trees along a driveway (Matlock, 2011). Huette et al. used a blank visual world paradigm (i.e. individuals looking at a blank screen while listening to simple past and past progressive sentences) to assess how aspectual cues influence eye movements during event comprehension (Huette, Winter, Matlock, & Spivey, 2012). In line with the results from Matlock's work, past progressive sentences yielded fewer eye movements and longer fixations implying that more action was conceptualized.

Examining the relationship between the domains of time and space further, Anderson, Matlock, and Spivey (2013) used grammatical aspect to study temporal distance in motion descriptions. The authors take a simulation perspective and use computerized mouse-tracking (Spivey et al., 2005) to examine how the relationship between tense and grammatical aspect influences event comprehension. The comparison included how aspectual cues would influence events in the recent past versus the distant past. In this study, participants were given 16 sentences (and 16 fillers) about a person moving along a path. Sentences were written to satisfy one of four conditions:

- recent past simple past: e.g. Yesterday Paul ran to the lake.
- recent past progressive: e.g. Yesterday Paul was running to the lake.
- distant simple past: e.g. Last year Paul ran to the lake.
- distant past progressive: e.g. Last year Paul was running to the lake.

Each set of sentences was matched to a visual scene that consisted of a diagonal path that terminated in the top center of the screen with a corresponding image (e.g. a lake). A character was pictured outside the scene set apart by a black frame. Participants were asked to listen to sentences and place the character into the scene to match each sentence. Three mouse movement measurements were collected including the initial grab-click of the character, the transfer of the character into the scene, and the drop-click of the character. Prior work using computerized mouse tracking showed that past progressive motion descriptions prompted movements of longer durations. In addition to simple RT measures, computer mouse tracking measures a multitude of indices of response and motor dynamics (Spivey et al., 2005).

Results of the Anderson et al. (2013) study showed that there was an effect of how close to the destination the characters were dropped along the path. For past progressive sentences highlighting ongoing events, participants placed the character at a midway point on the path as if the character was still approaching the destination while the simple past sentences prompted placement of the character at the end of the path at the destination. There was also an additive effect of temporal context (i.e. distant versus recent past).

For movement durations (i.e. grab-click to drop-click of character) in the distant past (i.e. last year), moving the characters in response to the simple past tense sentences took less time compared to the past progressive condition. In contrast, for the recent past (i.e. yesterday) sentences, the pattern was reversed. The past progressive sentences were responded to faster than the simple past. These data resulted in an unexpected cross-over interaction of aspect and temporal context for the movement duration data. Anderson and colleagues (2013) reasoned that it is possible the distant simple past condition is ambiguous. “Last year Paul ran to the lake” could mean that this event occurred once or multiple times (i.e. iteratively) and this ambiguity may have influenced the data. The interaction between aspect and temporal context for the movement data is consistent with the facilitation effects found in the Yap et al. (2009) study and the action compatibility effects described in prior sections (Sections 3.3.1 and 3.3.2). Matlock et al. propose that it is possible that the shorter, faster movements prompted by the combinations of distant + simple past and recent + past progressive sentences were prompted by a compatibility between the timeframes represented. While the parallel to the spatial domain provides an additional thread to the rationale for the current study, there are many open questions in this area of the event processing literature as well.

3.4 CURRENT STUDY RATIONALE

Prior chapters in this document have introduced the current landscape of the event processing literature for the temporal domain. Areas of agreement include the following: Comprehenders construct and update event models to facilitate language processing, a here and now perspective is elicited when events are introduced, readers routinely track multiple indices of an event, and flexible event boundaries are established to help segment events. As emphasized throughout the document, comprehenders rely on a multitude of default frameworks and cues which are essential for locating and sequencing events on the narrative timeline. Grammatical, lexical, and pragmatic cues prompt a range of event attributes that are mentally activated as event comprehension processes are engaged. While there are pockets of consensus throughout existing theories and studies of temporal processing, many open questions remain for how the perceived passage of time influences the mental activation of concepts within event models. The rationale for the current study is detailed in the following paragraphs and knitted together from multiple sources. These threads represent several unexpected or contradictory findings in the temporal processing literature that contributed to this study’s novel predictions of how mental event representations are updated.

Event continuity has bubbled up as a prominent factor that influences event model formation. It is well established that comprehenders mentally activate distinctly different event attributes when an event is perceived as ongoing versus completed. Current event processing models treat event continuity as an all or nothing phenomenon, but certain empirical findings suggest that ongoing event representations are more incremental in nature. For example, locations were primed for imperfective sentences highlighting ongoing events but not for perfective sentences emphasizing the completed events in the Ferretti et al. (2007) study. Madden and Zwaan (2003) found that when reading an imperfective sentence about an in progress event, participants selected the completed and in progress picture options with equal frequency. It is likely that the imperfective sentences elicited a range of in progress mental event representations such that both pictures were viable choices. Magliano and Schleich (2000) found that when an ongoing event was cued, readers were more likely to update their event models to reflect a completed representation as additional time-ambiguous sentences were added. This was particularly evident for events with longer durations (i.e. at least two hours). This finding underscores the elusive role of event duration during event processing. One example of the variable impact of event duration for event processing lies in the contradictory findings between the Yap et al. (2009) and Becker et al. (2013) studies. Imperfective sentences prompted a facilitation effect for both activity and accomplishment verb types in the Yap et al. study. However, Becker et al. found faster RTs only for accomplishment verb probes for targets that followed short, but not long, intervening events.

For events of all durations, current models do not capture the probable incremental nature of the progression of time when events are cued as in progress. As outlined in Section 2.2.2, comprehenders establish and navigate timelines both to track individual events and to facilitate their understanding of when events occur in relationship to other events. Current empirical findings (e.g. Becker et al., 2013; Ferretti et al., 2007; Magliano & Schleich, 2000) support the elicitation of event specific details and attributes when processing ongoing events, but the models assume those activated attributes are static in nature. Alternatively, it is likely that comprehenders establish individualized dynamic event timelines that correspond to what Moens and Steedman (1998) called the “actual event” phase of a given event (Section 3.3.1). Therefore, when an in progress event is cued, a graded, dynamic, or fluctuating progression of event attributes occurs similar to how an event is experienced in real-time. In addition, an incremental updating interpretation also connects to the semantic association perspective because the primary finding in that work is that durative (i.e. in progress) verbs and verb phrases elicit a more diverse set of associations compared to non-durative verbs and verb phrases. This perspective will be discussed at length in Section 4.2.

Extending the proposed incremental updating interpretation further, the event model initially activated when an event is cued as in progress may undergo a transformation to more closely resemble a completed event model as it nears the end of the event-specific timeline. This novel prediction aligns with the assumptions of the Dynamic Mental Representation Theory (Freyd, 1987; 1992, Section 2.2.2), that time cannot be extracted from the mental representation and it progresses forward in a unidirectional manner. The hypothesis of a more fine-grained updating process for in progress events is also informed by the contradictory findings from the Nuthmann and van der Meer (2005) and Welke et al. (2014) and studies. Both labs explored whether readers are sensitive to chronology for the progression of event states. Nuthmann and van der Meer found facilitation effects for responses to *shrinking*-SMALL when compared to *shrinking*-LARGE, but an analogous advantage was not found for event primes and their progression states like *washing* DIRTY-CLEAN versus CLEAN-DIRTY in the Welke et al. study. Barsalou’s (2009) pattern completion inference mechanism was invoked by Welke et al. as one possible explanation for the fact that the predicted chronological facilitation effect was not found in their study. If comprehenders are using such a mechanism to facilitate comprehension processes, clearly a closer inspection is needed of how it is applied. The specific research question and potential outcomes for the current study are summarized in Chapter 4 and Appendix A. However, before that, Chapter 3 outlines a discussion of two different perspectives for how mental event representations are shaped.

4.0 EVENT REPRESENTATION THEORIES

Now that several different types of time processing cues and the rationale for the current study have been outlined, a closer inspection is needed of the theoretical rationales proposed to explain the formation of event representations. The perspectives discussed below are not inconsistent with one another. Each provides a potential explanation for the primary research question: When navigating the narrative timeline within a single ongoing event, does signaling an early time point generate a distinctly different mental representation compared to a time point signaled later in the same event?

4.1 SIMULATION PERSPECTIVE

This perspective was introduced in some detail in the discussion of perceptual studies in Section 3.3.1.1. Simulation theories embrace the notion that sensory and motor experiential traces are activated to form mental representations. The simulation perspective is underspecified and does not offer specifics related to temporal processing, however, there are inferences that can be made based on current evidence. There is strong agreement and consistent evidence that events marked as completed emphasize the end stage result of the event (e.g. Becker et al., 2013; Ferretti et al., 2007; Yap et al., 2009). So for simulation theories, completed event representations predominantly involve a more focused (i.e. smaller) and less complex combination of sensory-motor level simulations compared to those for ongoing events. Complexity in this context refers to how in depth the reenactment of the multi-modal experiences is for a given event (e.g. Matlock, Ramscar,

& Boroditsky, 2005; Zwaan, 2008). Compared to ongoing events, completed events prompt fewer multi-modal traces and therefore less in depth experiential simulations. This property of completed events results in faster reading times or RTs in empirical studies, because the only part of the event the comprehender must simulate is the end state. The predictions for simulating ongoing events are more complex and are discussed next.

4.1.1 Predictions for ongoingness

Simulation theories (e.g. Barsalou, 1999; Gallese & Lakoff, 2005; Glenberg & Kaschak, 2003; Zwaan, 2003) predict that ongoing or durative event representations engender a richer or more complex simulation than those for completed events. The longer simulation period is necessary because of the engagement of an increased number and variety of multi-modal experiential traces (i.e. perceptual, emotional, related actions). Additionally, comprehenders appear to take a vantage point from inside an ongoing event as if it is in the process of occurring. This insider view prompts the activation of experiential traces that must capture the evolution of the event. There is behavioral and electrophysiological evidence that additional characteristics of event verbs are primed when the imperfective tense is used (see Section 3.3.1). For example, when an ongoing event is perceived, lexical-level priming for locations has been found (Ferretti et al., 2007) as well as increased availability of the participants in a scene (Carreiras et al., 1997). Researchers grounded in the simulation perspective argue that emphasizing the ongoing nature of events elicits specific consideration of the motion implied in the developing situation (Anderson, Matlock, Fausey, & Spivey, 2008; Anderson et al., 2013). If this is the case, it is unclear whether the activation of the primed event attributes and experiential traces is maintained throughout the simulation or if the processing of an ongoing event is more incremental in nature.

The current study proposes that a comprehender's representation of an ongoing event is dynamic and progressive in nature. Prior studies have exclusively focused on the differences between ongoing and completed event models (e.g. Becker et al., 2013; Ferretti et al., 2007; Madden & Zwaan, 2003; Magliano & Schleich, 2000; Yap et al., 2009). It is possible, however, that there are incremental changes in mental activation that occur as an ongoing event is processed as a result of more fine-grained temporal processing and updating. This proposed extension of the simulation perspective would incorporate a distinction between early and late simulations for an ongoing event. For example, early on in the event, one set of experiential traces might be activated, but there may be a narrowing as a comprehender approaches the end of an ongoing event. Specifically, the simulation may be winding down and more closely represent the end stage completed event. The present study begins to explore the possibility of refining the simulation account of event representations and processing to include this incremental updating notion.

4.1.2 Limitations of the simulation perspective

In addition to the potential limitation in the current scope of simulation theories, as proposed by this author, there are other areas that call for additional clarification within this perspective. Primary challenges for the simulation perspective include the underspecified way it approaches both abstract concepts and scale. For example, experiential trace explanations are not currently fine-grained enough to capture the nuances of how *stood* and *waited* might be simulated (Weiskopf, 2010). In their current form, simulation approaches provide evidence of predictable motor region neural activation when a comprehender hears the word *walk*, but do not explain qualitative differences such as speed or effort (Coll-Florit & Gennari, 2011). There is recent work underway that begins to address this limitation that is motivated by clinical populations such as individuals with spinal cord injury. For example, some of this work identifies highly specific neural underpinnings of finer-grained kinematic movements such as grasping (e.g. Collinger et al., 2014; Wodlinger et al., 2015).

The concept of experiential simulations also does not currently account for the necessary time compression which occurs routinely during comprehension. Illustrations include simulating *owning a house* (Weiskopf, 2010) or *for 10 years*. A comprehender quickly compresses the act of owning or the duration of 10 years without needing to simulate the entire span described. More broadly, the bulk of the evidence in this area utilizes single sentence-level stimuli and sentence-picture verification tasks like the studies from Section 3.3.1.1. The evidence for the simulation perspective does not currently address critical elements of language comprehension including both how contextual cues might influence simulations and how dynamic comprehension processes might be represented.

4.2 SEMANTIC ASSOCIATION PERSPECTIVE

The simulation viewpoint proposes that the driving force for how mental representations are shaped is the activation of neural traces stemming from prior sensory and motor experiences. Another perspective is the knowledge-based approach championed by Sylvia Gennari (Coll-Florit & Gennari, 2011; Gennari, 2004). The knowledge-based approach is not necessarily inconsistent with simulation theory, however it proposes that the foundation of a mental event representation is linked to the diversity of semantic associations in the described situation. A series of studies applying the knowledge-based perspective by Coll-Florit and Gennari (2011) investigated how comprehenders process and represent events of different durations in an on-line manner.

Specifically, they examined what role event temporal properties may play during word-by-word sentence processing, a superior method to whole sentence reading paradigms used in the majority of other studies of event continuity. Durative (e.g. *to owe 50 euros*) and non-durative (e.g. *to lose 50 euros*) Spanish verbs and verb phrases were contrasted to determine if the expected processing differences remained when a more on-line measure was used. Durative and non-durative verbs are analogous to "ongoing" and "completed" respectively. Coll-Florit and Gennari found that durative verbs and verb phrases consistently required longer processing times compared to non-durative verbs even after controlling for plausibility judgments and the lexical properties of the verbs. The authors provided a semantic association rationale to explain the processing mechanism that could account for the durative effect.

Multiple researchers of event processing concur that it is not the lexical item that is the engine of event representation formation. Instead, situation-specific representations arise stemming from contributions from both linguistic and situational contexts (e.g. Ferretti, McRae, & Hatherell, 2001; McRae, Ferretti, & Amyote, 1997; McRae et al., 2001). Attributes become associated with a given event via the knowledge and associations acquired during real-world experiences and are referred to as event concepts. The notion of event concepts mirrors that of object concepts (Vigliocco, Vinson, Lewis, & Garrett, 2004) such that both real-world and linguistic experiences promote how the concepts are learned and stored in semantic memory. The co-occurrences of typical participants, instruments, and attributes provide the foundation for event concepts and are activated when they are processed (Ferretti et al., 2007). Coll-Florit and Gennari (2011) extend this perspective further by hypothesizing that these situation-specific representations include duration information. They propose that temporal, causal, and contingency relationships (Moens & Steedman, 1988) are critical pieces of event duration representations. The strength of the connections between the co-occurring event concepts varies as a function of experience.

Coll-Florit and Gennari (2011) propose that comprehenders are tapping into more diverse associated knowledge when reading durative (vs. non-durative) verb phrases in sentence-level and short narrative-level contexts. The diversity could stem from two possible but related sources referring to the strength of the connections between co-occurring attributes within event concepts: 1) semantic properties of verbs with strong competing interpretations and/or 2) weaker (i.e. less frequent) semantic connections similar to subordinate meanings in lexical ambiguity. The strength of the association depends on the co-occurrence frequency between the objects, events, and properties that become linked through real world experiences and the linguistic contexts in which they occur. When a durative verb is processed, a wider range of associated event attributes, world knowledge, and experiences is tapped. This broader activation necessarily increases processing time in order to resolve the competing interpretations of equally frequent semantic properties. In a similar fashion, increased processing time may be needed to allow the weaker semantic associations to be retrieved.

To corroborate this conclusion of a more diverse set of associations activated by durative verbs and verb phrases, the authors complemented their on-line measurements with an off-line free association task. The responses were categorized into one of four classifications: objects, events, states, and properties. As predicted, the durative states elicited a wide range of responses that spanned the four classifications in approximately equal proportions. Conversely, non-durative verbs and verb phrases prompted fewer associates for states and properties with a higher concentration of primarily entity and event classifications.

Another significant finding from Coll-Florit and Gennari (2011) is that increases in processing time co-varied with increases in attributed duration. Using a rating task, the authors collected data on how long the situation described in the durative and non-durative verb phrases would last (i.e. 1 = very little time through 7 = much time). There was a positive correlation between the duration ratings and the reading times of the verb phrases from the initial studies. Folding together all of the factors in this study, increases in event duration and processing times also co-varied with increases in the variability of the associates for both sets of verb phrases. Overall, the authors concluded that durative verbs and verb phrases promote a higher volume and more diverse set of semantic associates which results in increased reading times for durative versus non-durative verbs and verb phrases.

4.2.1 Semantic association approach to ongoingness

At present, the knowledge-based approach has predominantly been applied to durative and non-durative verb phrases. Applying its premise to the proposed idea of incremental updating for ongoing events, it is possible that the increase in volume and diversity of semantic associates generated when processing durative verb phrases does not occur in an all or nothing manner. The semantic association perspective reasons that the increased processing time necessary when reading durative verbs stems from competing strong interpretations and/or weaker semantic connections. It is implied that there is a resolution of these activation patterns as comprehension proceeds (i.e. resolving the competing interpretations and/or activation of a prevailing weaker interpretation). This leaves open the potential for an incremental resolution process specific to durative or ongoing event representations. In other words, the volume or type of associates available to the comprehender near the beginning of an ongoing event may narrow as comprehension proceeds. The current study begins to explore the possibility of refining the knowledge-based account of event verb processing to include an incremental updating process.

4.2.2 Limitations of the semantic association approach

The knowledge-based perspective advocated by Gennari continues to be specified, however, in its current instantiation, and as acknowledged by its proponents (e.g., Coll-Florit & Gennari, 2011), it cannot rule out the simulation approach. For the purposes of the current study, either perspective would benefit from a more specific examination of the potentially incremental nature to processing ongoing events. Critically, both simulation theory and the knowledge-based approach predict that ongoing events are multifaceted and thus promote either a longer simulation or more diverse set of semantic associations when compared to completed events. It is unclear, however, whether or not the longer simulations and/or diversity of semantic associates are maintained uniformly throughout the entire time course of an ongoing event.

4.3 SUMMARY OF THEORETICAL CONTRIBUTION

The potential results of the current study have implications for simulation theory andthe semantic association perspective. Each perspective identifies events and/or event verbs as central features to model construction and updating. Each potential application will be outlined in the following paragraphs beginning with simulation theory. If the predicted congruity effect is found, the results would add to the growing list of perceptual congruity studies by adding event phase cues as another factor that promotes faster responses for congruent items. From a simulation perspective, a congruity effect based on a time phase cue would also suggest that for ongoing events there is potentially a graded simulation occurring depending on where on the event timeline the comprehender is directed. This result is also consistent with the economy of processing principle (see Section 3.3.1.1). If the second prediction is upheld and there is a Phase x Congruity interaction, it is predicted that there would be a congruity effect for the Late phase and either no congruity effect or a smaller congruity effect for the Early phase. From the simulation perspective, this result could suggest that both the Early and Late targets are activated in the Early cue phase but that Late phase cue elicits a shorter simulation that is closer to the representation of a completed event. Finally, if the null hypothesis cannot be rejected, this result would be relevant to simulation theory by questioning whether there really is a difference between the experiential simulations prompted by an event phase cue within an ongoing event.

A congruity effect from the semantic association perspective would suggest that not only do ongoing events prompt a larger and more diverse set of attributes, but that event phase cues elicit a different subset of attributes within an ongoing event. More specifically, this finding could indicate that there are attributes that are more strongly associated with either an Early or Late event phase. If the second predicted outcome is found and there is a Congruity x Phase interaction, this finding would be consistent with the premise that ongoing events prompt a larger and more diverse set of attributes. In conjunction with this interaction prediction, if the congruity effect is found for the Late phase only, then this finding would be the first set of results to illustrate the possibility that the Late event phase representation is closer to that of a competed event. Finally, if the null hypothesis cannot be rejected, the contribution to the semantic association literature might be that there may be a similar type and volume of event attributes that are implicitly activated for an ongoing event representation regardless of the time point cued within that event. The next chapter describes the methods used in the current study.

5.0 METHODS

The research question and potential outcomes for this study are described in Appendix A. Two theoretical positions, simulation theory and semantic association, predict that ongoing events are multifaceted in nature. Ongoing events likely prompt longer simulations and/or the mental activation of a broader set of semantic associations when compared to completed events. This study is the first to examine events that are in progress and predicts that there is an incremental updating process at work. Recall that in their current instantiations, neither theoretical position rules out the other. To aid in stimulus development, a pilot study was conducted to determine the types of event concepts or associates adults link to two time points (i.e. just beginning/almost finished) within ongoing events. In this chapter, the first sections (5.1 through 5.1.5.1) describe the methods of the pilot study and its outcomes. Section 5.2 defines the criteria used to select the stimuli for the current study. The final sections outline the projected methods and potential outcomes for this study.

5.1 PILOT STUDY

The goal of the pilot study was to solicit associates or attributes that are often linked to either just beginning or being almost finished with an ongoing event (e.g. playing a board game, watching a movie). The results of the associate generation task were used in the stimulus development for the current study. Prior to this pilot study, it was unknown if ongoing events elicit a diverse set of associates when different time points within the event are cued. It was deemed likely that there would be a significant amount of overlap in associate generation between the two time points (i.e. just beginning/almost finished), but it was also possible that there would be a distinct set of concepts for each time point. Because event duration has been a critical factor in prior studies (e.g. Becker et al., 2013), it was possible that longer events would have the highest probability of producing associates unique to each time point. Fifty events spanning a range of durations (e.g. sliding down a slide, playing tennis) were included in the pilot study.

5.1.1 Participant characteristics (pilot)

Potential pilot participants met the following inclusion criteria by self-report: between 18-35 years of age, monolingual speakers of English, with no history of neurological disorders, and not currently experiencing an alcohol or drug-related dependency. Participants were excluded if they self-reported: an age of under 17 or over 35; learned or spoke more than a few words or phrases of another language as a child; have a history of a stroke, seizure, hemorrhage, brain tumor, or other type of neurological disease or condition; or indicated a drug or alcohol-related dependency. The screening questions are listed in Appendix B.

5.1.2 Participant recruitment (pilot)

Participant recruitment targeted primarily undergraduate and graduate students at the University of Pittsburgh and other regional universities. Professors sent an introductory script via email to students in their departments who expressed interest in participating in research studies. Additional participants were also recruited by word of mouth and a flyer.

5.1.3 Experimental procedure (pilot)

Both the screening questions and the experimental questions were posed via a web-based survey. Potential participants used a survey link embedded in the recruitment email and flyer that directed them to the generation study deployed via University of Pittsburgh's web-based survey tool, Qualtrics (Copyright © 2015, Provo, UT). Participants were prompted to read the introductory script and advance through the screening questions to determine eligibility (see Appendix C). The survey tool was configured to exit the survey if a potential participant answered an eligibility question in a manner that would exclude them from the study. If potential participants met all of the inclusion criteria, the survey proceeded to the experimental task instructions (see Appendix C). The Qualtrics software collected the words and phrases participants associated with either just beginning or being almost finished with an event.

5.1.4 Experimental stimuli (pilot)

Two different surveys were developed that included unique sets of 25 events (i.e. Survey A, Survey B) to keep the screening and experimental tasks to an approximate 30-45 minute commitment. Separate participants groups were recruited for each survey, however participants generated words or phrases associated with both time points for the events in the survey they were completing. Pilot stimuli were pseudorandomized so that the early and late time points for the same event did not occur within 5 items of one another. The experimental items are listed in Appendix D. Events chosen for both surveys included a range of items consisting of both short (e.g. opening an umbrella) and long durations (e.g. touring a museum). Specific event durations (e.g. 30 minutes) were not provided to avoid constraining the generative process.

5.2 SELECTION CRITERIA FOR EXPERIMENTAL STIMULI

Individuals generated between 3 and 12 single words or short phrases for both event phase conditions (Early = just beginning; Late = almost finished) for each event. While there was a considerable amount of overlap between the early and late condition associates, there were also associates that were only reported for one of the conditions. Based on this outcome, the review of the pilot data focused on identifying associates that were almost exclusively linked to only just beginning or being almost finished with an event. Events needed to have targets that met the following criteria in both conditions for inclusion in the current study: 1) concept/word reported no more than two times in the opposite condition and 2) the concept/word must have been mentioned in the first three responses. Associates were excluded as potential lexical targets if they depicted an emotional state such as happiness or anxiety because individuals often reported the opposite emotion for the same condition (i.e. putting together a puzzle elicited responses of both boring and excited). When these criteria were applied, there were some events that did not produce any potential targets and some events in which up to 78% of respondents reported a given attribute. In order to include as many potential stimuli as possible, a concept/word was included if it was reported by at least 28% (i.e., 11/40) of respondents. Example stimuli with early/late phase targets include visiting a museum – GUIDE/GIFT and baking a cake – EGGS/AROMA (see Section Appendix G for other examples).

In all, 16 events met criteria for both Early and Late lexical targets. Early targets were reported by an average of 40.06% of respondents with an average of 1.13 mentions in the opposite condition, and Late targets by 42.5% of respondents and 0.63 mentions in the opposite condition. Two independent judges conducted reliability judgements using the established criteria for each of the 16 events. There was 94% agreement between the judges (see procedures in Appendix E).

Lexical properties were calculated using the English Lexicon Project website (Balota et al., 2007). Group averages for all of the Early and Late lexical targets were compared. Lexical properties included mean RTs, word length, number of syllables, and parts of speech. Paired *t*-tests were conducted and there were no significant differences in Mean RT between Early (M = 595.87, SD = 51.96) and Late (M = 623.68, SD = 50.92) condition targets; $t(15)$ =-1.43, p = .173. Three synonym substitutions were necessary to balance the Early and Late target group averages: 1) GIFT replaced SOUVENIR for the Late target of going to a museum 2) SUCCESS replaced ACCOMPLISHED for the Late target of putting together a puzzle and 3) VEGETABLES replaced PRODUCE for the Early target of going grocery shopping. Each of the substituted words was checked to make sure that is was not reported more than two times in the opposite condition and was listed as a synonym in at least one dictionary (i.e. Merriam-Webster, 2004). Both conditions include targets with a variety of parts of speech which will be considered as a co-variate during data analysis if applicable.

Latent Semantic Analysis (LSA; Laham, 1997) was used to calculate the semantic relationship between the Early and Late targets and the time cue sentences. This analysis was completed to help interpret the predicted congruity effect. Identifying the semantic relationship of each individual word/phrase (i.e. amusement, park, amusement park) in the time cue sentences to the target words provides some evidence that the predicted congruity effect is due to the presentation of the time cue sentence as a cumulative whole. Critically, results from this analysis would strengthen the argument that a congruity effect is not driven by the semantic relationship from any individual word in the time cue sentence to the target.

The semantic relationship was calculated between each target and: the time cue (i.e. recently started/almost finished), the verb phrase, the individual content words within the time cue and verb phrases, and any other content words in the experimental sentences. Once each of these figures was calculated, the grand mean of the relationship between the target words and the content words within the time cue sentence was calculated. LSA calculates a correlation coefficient (between 0 and 1) that is used to interpret the semantic association between the different entries (i.e. words, phrases). Based on the research question, the sentence stimuli for this study should have either no relationship or a weak relationship to the target words. A more lenient criterion was set for the average of the individual sentences to maximize the number of stimuli that could be included in the study. The criterion for inclusion was ≤.5 (i.e. a moderate relationship) for the grand mean of each sentence to the target word and ≤.3 (i.e. a weak relationship) across all of the sentences in each condition (Landauer, Folotz, and Laham, 1998). The grand mean across all experimental trials for the Early targets was .19 (SD = .09) and .23 (SD = .07) for Late targets. While both averages indicate a weak relationship between the sentence-target pairs, there was a significant difference between LSA averages for the Early and Late targets in a paired samples *t*-test with alpha set to .05, $t(15) = -2.24$, $p = .04$. See Appendix I for the complete table of results.

5.2.1 Potential results and interpretations (pilot)

This pilot study and the subsequent experiment are the first to focus on ongoing events and their associates. It was unknown if participants would generate a different set of associates for two different time points within the same ongoing event. There were two possible outcomes for the types of associates prompted by early and late time points within a single ongoing event: 1) predominantly similar sets of associates/concepts or 2) a percentage of associates/concepts exclusive to each time point. In addition to different types of associations, the total volume of concepts that are linked to early and late ongoing events could vary. The next few paragraphs describe each of these outcomes and potential interpretations from both the semantic association and simulation perspectives.

If there is essentially no difference between the early and late sets of associations, then there could be a static nature to the activated associates for ongoing events. This result could possibly be used to argue against the proposed idea that ongoing event representations are incrementally updated and adapt to become more like completed representations as comprehension proceeds. However, if unique associates are not produced for each time point, it could also simply reflect the off-line, metalinguistic nature of the task. From the semantic association perspective, a no difference result could indicate that the associates that are linked to an early time point in an ongoing event are also relevant at a late time point. For simulation theory, a no difference result could reflect a consistency between the activation patterns of sensory-motor experiential traces cued by early and late markers in an ongoing representation. On the other hand, if there are concepts that are unique to early or late in the event, this could support the premise that ongoing event representations may undergo some form of transformation or updating. It could also mean that an off-line task elicits some event associations that are more strongly connected to an early versus a late temporal reference point.

The volume of associates generated at each time point offers a different kind of insight into the contents of early versus late ongoing mental event representations. Evidence from prior studies points to fewer associations for completed events. If there is an incremental updating process at work, it is possible that the late ongoing representation may more closely resemble a completed event representation.

5.2.1.1 Pilot study results

Forty complete sets of data were collected for 50 events. Table 1 reports the demographic characteristics for the pilot study participants.

Table 1. Demographic characteristics for event association generation pilot study

Characteristic	Survey A (N=40)	Survey B (N=40)
Sex	35 female	35 female
Age (years)		
Mean (SD)	22.9 (3.9)	22.5 (5.8)
Range	19-34	18-35
Education (years)		
Mean (SD)	15.0 (2.1)	14.1 (2.0)
Range	13-20	12-20

Because the events were divided into two sets simply to reduce participant burden, the data were collapsed for reporting purposes. Table 2 includes the average number of associates that were generated for 50 events at each time point. Participants provided an average of 20.74 more associates for the early versus the late time point. A paired samples *t*-test was conducted to compare the number of associates generated for early and late conditions. There was a significant difference between the early (M = 281.5, SD = 37.0) and late (M = 260.07, SD = 33) time points; $t(49)$ =4.09, p < .000, d = .58. The mean difference between the two sets of associates yields a medium effect via post hoc power analysis (Faul, Erdfelder, Buchner, & Lang, 2009). Despite the significant difference across all 50 events, participants produced and average of 23.14 more late associates than early ones for 14 of the 50 events (28%). For the 36 events that were the driving force for the significant result, the early time point prompted an average of 38.7 more associates than the late time point.

Table 2. Average number of associates generated for 50 events (N=81)

Time point	Associates Generated Mean (SD)	Range
Just beginning	281.5 (37.0)	211-363
Almost finished	260.1 (33.0)	182-329

The procedures for reviewing the responses from the event generation task are included in Appendix E. Despite an underrepresentation of males in both groups (~12% for each survey), a high percentage of the male responses exactly matched the words or phrases that were generated by females. Male participants generated an average of 31.4 attributes for the early time point and 28.3 for the late time point. This equates to an average of 11% of the total data provided for each time point. An average of 2.3 attributes did not exactly match a female participant response for both early and late time points. This equates to 7% and 8% of all the responses generated by male participants that did not match a female response. The average number of attributes generated only by male participants (i.e. 2.3 at each time point) relative to all of the data generated by both sexes is <1% for both time points. Section 5.2 discusses the results from the pilot study further and provides a rationale for the selection criteria that was used to determine lexical targets for the current investigation.

5.2.2 Rating task for experimental trials

There were 16 events that had both Early and Late lexical targets that met the criteria from the prior section. Due to the relatively low percentage of individuals who generated the targets in each condition, a rating task was completed to test how strongly associated the potential lexical targets are with different event phases. Event phases included four different time points: Before, Just Beginning, Almost Finished, and After. Three raters completed an association rating task using a

5-point scale (i.e. 1 = weak through 5 = strong) to indicate how strongly each target was associated with one of the four event phases. Trials were pseudorandomized by both event and event phase (i.e. there were at least five trials between occurrences of a given event and across events, none of the four event phases occurred more than twice in a row). All targets appeared in the task three times for a total of 96 experimental ratings. There were also 20 filler trials that included the same event phases but paired with different events and used targets intended to produce very low ratings (i.e. After: Fishing – BAIT). For the Early targets, raters indicated how strongly each word was associated with the Before, Just Beginning, and Almost Finished event phases. For the Late lexical targets, the judges rated the association for Just Beginning, Almost Finished, and After phases.

Appendix E provides the instructions for the rating task and Table 3 reports the average ratings. Average ratings were all in the predicted directions. Critically, the ratings for the experimental pairings were all greater than 4 (i.e. Early + Just Beginning; Late + Almost Finished). In addition, the paired *t*-tests for the contrasts of interest were significant including the Just beginning and Almost finished ratings for the Early ($t(15) = -33.11$, $p < .000$) and Late targets ($t(15) = -19.32$, $p < .000$).

Table 3. Relatedness judgements of early and late targets

Event Phase	Early Targets Mean (SD)	Late Targets Mean (SD)
Before	4.54 (.5)	N/A
Just beginning	4.77 (.29)	1.35 (.38)
Almost finished	1.35 (.35)	4.44 (.36)
After	N/A	4.52 (.66)
N/A = targets not given in that event phase		

5.3 STUDY PARTICIPANTS

5.3.1.1 Participant recruitment and screening

All participants were recruited from the University of Pittsburgh and surrounding areas. The primary investigator obtained written consent according to the procedures outlined by the University of Pittsburgh's Institutional Review Board before beginning any screening or experimental tasks. The same participant inclusion and exclusion criteria (Section 5.1.1) and screening questions (Appendix B) from the pilot study were used for this investigation. A total of 54 participants completed the initial phone screening. Three participants were excluded because they had learned more than a few words or phrases of a language other than English as a child.

Testing sessions began with a vision screening that was conducted after participants signed the consent form. Participants had normal or corrected to normal vision as determined by a two-step screening. First, using the experimental set up, participants read out loud to the examiner 5 single words in all capital letters with comparable characteristics to the experimental targets. Next, participants read aloud 2 sentences in the past progressive and 2 sentences in the simple past tense in the same 12 point font used in the experiment. Sentences were displayed on the same computer screen used in the experiment and participants read them with 100% accuracy. Articulation differences or errors were accepted as correct. Fifty-one participants passed the vision screening and completed the study.

5.3.1.2 Participant sample size

Based on the study design and in order to account for individual differences, a multilevel model was used for data analysis. Scherbaum and Ferreter (2009) address some of the challenges of determining sample sizes that are sufficient to maximize statistical power for multilevel models.

The sample size rule of thumb that the authors propose is N=50 and was used for this study. Additional support for this sample size stems from 1) multilevel models' use of maximum likelihood (i.e. a standard estimation procedure to optimize the characteristics of the sample), 2) repeated measures conducted for a set of items, and 3) each subject serving as his or her own control. Once the target N=50 was achieved, a post hoc power analysis was conducted to ensure the study had adequate power to detect a small effect (Hedges, 2007, see Section 6.7).

5.3.2 Participant Demographics

Table 4 reports the participant demographics. One participant completed the study but was excluded from analysis because she required a verbal cue on at least 40% of the trials to return her finger to the center button resting position on the response box between trials. The study sample was 28% male. Independent samples t-tests were conducted to test for sex differences in age and education. For age, Levene's test of equal variance was rejected, F=35.967, p < .001 so independent t-test results that corrected for heterogeneity of variance were used. There was not a significant difference in age (male: M = 25.0, SD = 5.0; female: M = 22.3, SD = 2.0; t(14.66) = 1.953, p = .07) or years of education (male: M = 15.4, SD = 1.2; female: M = 15.4, SD = 1.0; t(48) = 1.95, p = .07) between the two sexes. Independent samples t-tests were also conducted to test for differences between sexes in RTs for the experimental trials. There were no significant differences in Target RT, the primary outcome variable, with equal variances not assumed (Levene's tests for equal variances: F = 2.11, p = .004; male: M = 621.98, SD = 128.98; women: M = 620.78, SD = 136.42); t(1630.9) = .225, p = .822). In addition, there was no significant difference between the sexes for Baseline RT (male: M = 577.12, SD = 130.42; women: M = 567.93, SD = 123.11); t(3118) = 1.84.206, p = .066).

Table 4. Demographic characteristics for experimental event processing study

Characteristic	**Male**	Female
Sex	14	36
Age (years)		
Mean (SD)	25.0 (5.0)	22.3 (2.0)
Range	19-34	21-30
Education (years)		
Mean (SD)	15.4 (1.2)	15.4 (1.0)
Range	13-18	14-18

5.3.3 Experiment summary

The primary aim of this investigation was to advance knowledge of the nature of event representations by examining the likely influences of the perceived passage of time within event models. More specifically, this study is an initial step toward examining whether comprehenders engage an incremental updating process for events that are cued as ongoing. Two prominent event representation perspectives, simulation and semantic association, provide the theoretical foundation for the predictions in this study. Currently, in both theories, mental activation patterns that represent ongoing events are conceptualized as static, however, there are various empirical hints that strongly suggest a more dynamic nature. These data include unexpected and contradictory findings for ongoing events in multiple studies (i.e. Ferretti et al., 2007; Madden & Zwaan, 2003; Magliano & Schleich, 2000) and contributions pulled from the spatial event processing domain. To date, all contrasts in the event continuity literature have been between ongoing and completed event representations. This study is the first to focus specifically within ongoing events.

Results from single word priming studies (e.g. Hare et al., 2009; McRae et al., 2005) and sentence-level reading time studies (e.g. McRae, Spivey-Knowlton, & Tanenhaus, 1998; Hare et al., 2009) indicate that event knowledge exerts an immediate influence during on-line reading comprehension. Critically, the priming effects and faster reading times are not accounted for by direct lexical associations. This conclusion is derived in part from Ferretti's work manipulating grammatical aspect (Ferretti et al., 2007). In this study, direct lexical association could not account for the priming effects because grammatical aspect modulated the effect (i.e. priming effect found for locations in ongoing condition only). In addition to the work that points to the rapid influence of event knowledge during on-line comprehension, the results from Coll-Florit and Gennari (2011) provide evidence that temporal properties of event knowledge are tapped via word-by-word sentence comprehension. The study design and lexical decision task for the current study will be one of the first to include a word-by-word reading approach in lieu of whole sentence reading paradigms used in the majority of prior investigations of temporal properties (e.g. Kelter, Kaup, & Claus, 2004; Zwaan, 1996).

For the experimental task, participants read sentences using a self-paced word-by-word paradigm. Individual words in each sentence were masked except for the word in the moving window that was advanced by the participant using the spacebar. Sentences signal 1) ongoing events using the past progressive tense and 2) Early or Late temporal reference points. Immediately following each sentence, participants saw a letter string in capital letters that was right justified on the screen. All targets appeared in the same right justified position on the screen since the participants' eye gaze was to the right of the screen following the self-paced word by word reading of each sentence. Letter strings were displayed immediately after the final word in the sentence and participants decided as quickly as possible if the letter string was a real word or a nonword.

Letter strings included Early and Late lexical targets selected from the pilot study reported in Section 5.1, other real words that were unrelated to the sentence contexts, and nonwords.

Event continuation was signaled by grammatical aspect (i.e. past progressive) because this manipulation produces a robust effect. Both accuracy and RTs were collected via Yes/No manual button press. Participants were expected to be highly accurate. The RT data are meaningful because they can reflect automatic event-related priming (Hare et al., 2009) and a congruity effect was predicted. Participants were predicted to respond faster to congruent Early/Late temporal reference cues + Early/Late lexical targets and slower to incongruent pairs. In addition to the primary contrast of interest (i.e. time congruent versus time incongruent), two other conditions were included in the study and are outlined in the next sections.

Control Condition

The control condition is intended to replicate the robust ongoing versus completed event representation finding in the literature with these novel stimuli. For this condition, dubbed the Completed condition, the Early and Late lexical targets were paired with the same events in the primary task but the past progressive sentences were replaced with sentences written in the simple past tense. The expected results for the Completed condition were: 1) Early targets would yield longer RTs than the Late targets and 2) Late targets would yield faster RTs when paired with simple past sentence compared to when the Late targets are paired with the Almost Finished past progressive sentence. The prediction for the shorter RTs in the Completed condition is derived from the Moens and Steedman (1988) model of event processing (see Section 3.3.1) where the authors propose that the completed condition focuses the comprehender on the resultant state of the event. Activating an attribute that is strongly associated with the resultant phase of an event was expected to be faster than activating an attribute for an event that is in progress. The trials for this condition were folded in with the experimental trials and all trials were pseudorandomized to maximize the distance between repeated lexical targets.

Baseline Condition

If the predicted result of a congruity effect is found, it will not be clear whether it is due to facilitation or inhibitory effects or perhaps both (e.g. Posner & Snyder, 1975). The Baseline condition provides a mechanism to calculate facilitation via the difference between the related and neutral stimuli while the difference between the neutral and unrelated stimuli determines how much inhibition contributed to the result. This condition was blocked separately from the experimental and control trials and was administered after all of the experimental task blocks.

Study Design

The study employs a single group two factor design with repeated measures on conditions: event Phase or temporal reference point (Early, Late) and Congruity (Congruent, Incongruent). A secondary question and analysis contrasts phase/target Congruity (Congruent, Incongruent) for the Completed condition (i.e. simple past) as a replication of prior findings.

Experimental stimuli and filler trials

Experimental stimuli consisted of two parts: 1) one sentence that signals an early or late time point for an ongoing event and 2) an Early or Late lexical target. Time cue sentences were written using the past perfect progressive (i.e. Alice had recently started/almost finished baking a cake) to cue an in progress event representation. Early and Late lexical targets were paired with the same past perfect progressive sentence for each event for a total of 32 experimental trials (i.e. 16 congruent trials for each temporal reference point). Appendix G lists the stimuli that met the inclusion criteria (outlined in Section 5.2).

Experimental stimuli were modified to the simple past tense for a control condition and were paired with the same Early and Late targets for a given event. This condition, dubbed the Completed condition, was intended to replicate the robust ongoing versus completed finding in the literature with these novel stimuli. Based on the high ratings (i.e. average of 4.52 on a 5-point scale) between Late lexical targets and the After event phase in pilot testing, the Late targets paired with a simple past sentence should approximate the completed condition from other studies.

Filler trials were constructed to disguise the experimental trials. Four sentence types (i.e. past perfect progressive, past perfect, past progressive, and simple past) were combined with both real word and nonword letter strings. The real word filler targets did not repeat any of the experimental targets or cue a specific time point within the event. Nonwords were phonotactically plausible and generated using the English Lexicon Project website (Balota et al., 2007) to match the experimental targets on: 1) length, 2) bigram frequency by position, 3) and mean RT.

The experiment includes a total of 320 trials delivered in 10 blocks of 32 trials. Table 4 includes sample stimulus items and the total number of trials included in each condition. Twenty percent of the total trials were experimental trials and each block included 6 or 7 (i.e. ~20%) experimental trials. Fifty percent or 160 trials included a nonword letter string, and finally, this study included a 15% relatedness proportion.

To minimize the effects of repeated exposures to experimental trials, stimulus blocks were constructed to maximize the number of trials between exposures using a combination of real word and nonword filler trials. Additional considerations for task construction include: 1) each block began with at least two filler trials and ended with at least one filler, 2) there were no more than three sequential “yes” or “no” responses within each block, 3) nonwords were repeated more than once with a different preceding sentence, 4) the name of the protagonist was changed for each filler and experimental trial version, and 5) protagonist names were repeated pseudorandomly to ensure that a single protagonist was not always paired with the same type of event. After all of the experimental and control condition trials were administered, a separate testing block was given for the Baseline condition. In this condition, participants saw see a series of XXX in place of the words in the experimental sentences (see Table 4) followed by a letter string. Letter strings included the Early and Late lexical targets and the nonwords from the experimental task. The baseline task included two blocks of 32 trials.

Table 5. Stimulus examples

Stimulus Type	Stimulus Examples	Related Trials	Unrelated Trials
Experimental Condition Past perfect progressive Early/Late targets	Wendy had recently started/almost finished doing laundry. SOFTENER/FOLDING	32	32
Control Condition Simple past Early/Late targets	Laura did some laundry. SOFTENER/FOLDING	16	16
Fillers (4 types) Past perfect progressive Past perfect Past progressive Simple Past Real word/Nonword targets	 Amy had been playing tennis. Nick had played tennis. Julie was playing tennis. Ron played tennis. WATER/NAWK	 0 0 0 0	 56 56 56 56
*Baseline Condition	XXX XXXXX XX XXXXX XXXXX XXXX SOFTENER/FOLDING	0	32

*Baseline trials were administered after all of the experimental and control trials
See Appendix G for complete list of stimuli

Experimental procedure

Testing sessions were conducted with each participant individually in a quiet research lab and began with informed consent and vision screening procedures. For eligible participants, the experimental task was introduced using the script in Appendix I. E-prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2012) delivered the stimuli and recorded both accuracy and RTs for all items via button response box. All experimental testing blocks were administered prior to the baseline task.

Trials were presented visually and participants were asked to press the space bar to progress through the sentences using a self-paced word-by-word approach to more closely replicate natural comprehension processes. A moving window approach was used such that words were masked by hyphens corresponding to the number of letters for the masked words. Immediately after the last word in each the sentence was displayed, a letter string was displayed in all capital letters in the right justified screen position. Lexical decision accuracy and response time were collected for each letter string. Testing sessions lased approximately 30 minutes.

5.4 DATA ANALYSIS OUTLINE

Accuracy and RTs were collected for all trials. Participants were expected to be highly accurate for both experimental and filler trials. Accuracy percentages were calculated separately for event Phase and Congruity conditions for both Early and Late targets. It was expected that there would be no accuracy differences between event Phase and Congruity conditions. Data analysis subsequently focused on the RTs for accurate experimental trials only. RTs that were +/- >3 SD from an individual's mean RT were treated as outliers and removed from analysis. Appendix A reports the expected results and potential outcomes for the experimental question.

The main analysis used multilevel modeling (Raudenbush & Bryk, 2002; Singer & Willett, 2003; Woltman, Feldstain, MacKay, & Rocchi, 2012) with two levels (i.e. Level 1 is Items and Level 2 is Participants). The dependent variable (DV) was target RTs. This type of analysis allowed for simultaneous investigation within and between levels of predictors of the outcome of interest. This approach is superior to other types of simple linear regression techniques because it can account for variance among different levels.

Three different patterns of results were deemed possible for the experimental condition. Each of these potential outcomes assumes that explicit piloting task elicited a valid representation of the attributes that are implicitly activated at each time point. The first potential outcome was a congruity effect with faster RTs for Recently Started + Early targets and Almost Finished + Late targets and longer RTs for incongruent pairings. If there are differences between the ongoing event representations cued by Early and Late time points, a congruity effect result would indicate that there is no Phase x Congruity interaction and the size of the congruity effect between the two event phases will not be statistically different.

The second potential outcome was a Phase x Congruity interaction. This outcome hinges on the joint validity of two premises reviewed above: 1) the semantic association perspective premise that a larger and more diverse set of attributes is implicitly activated for ongoing events and 2) this author's hypothesis that a Late time point in an ongoing event elicits a mental event representation more like that of a completed event. For this outcome, a congruity effect would be predicted for the Late time point and there are two possibilities for the Early time point data. If a more diverse set of attributes is active early, both Early and Late attributes may be available in the

Early time point mental representation, so Late targets could also be primed at the Early time point. Thus, either there could be no congruity effect for the Early time point or a smaller congruity effect compared to the Late time point.

The final potential outcome was that the null hypothesis cannot be rejected. For this outcome, there would be no congruity effect for either time point. This finding might suggest that the Late time point representation is not like a completed representation and that a time point cue does not influence the attributes that are implicitly activated for ongoing event representations.

The analysis of the control, Completed condition, which was expected to replicate the robust ongoing versus completed effect, used the same multilevel model approach substituting the data from simple past sentence stimuli for those from the past progressive sentences. For this analysis, only the target Congruity condition (Congruent; Incongruent) was included because the simple past tense sentence is only linked to one time cue and indicates that the event is over (i.e. completed).

5.4.1 Proposed post hoc analyses

We proposed to analyze the data from the baseline condition if the congruity effect was found. This analysis would determine whether the congruity effect was due to facilitation of the congruent trials or inhibition of the incongruent trials or a combination of the two. A univariate ANOVA with average RTs as the dependent variable was planned for Congruent, Neutral, and Incongruent trials. If the congruity effect is due to facilitation, the Congruent condition should be faster than the other two conditions.

The pilot data suggested that there may be inherent differences in the frequency of associations that are generated for Early and Late temporal reference points for a given event. Of the 16 events included in the study, four had a higher number of attributes that were generated during the pilot study for the Late event phase compared to the Early event phase. To control for this attribute frequency bias, data from these three events could be removed and the analysis rerun to assess the potential contribution this factor may have had on the results. Based on the outcomes, potential post hoc analysis could also account for the different parts of speech of the Early and Late lexical targets using an effects coding strategy.

6.0 RESULTS

The current study examined how real world event knowledge merges with linguistic time cues to guide language comprehension. It investigated the implicit mental activation of attributes that are frequently associated with common events. Event completion status is currently conceptualized as an all or nothing phenomenon, however, the premise for this study is that ongoing event representations are not homogenous or static, but rather reflect incremental updating processes. This chapter summarizes the results of the current study. It begins with the accuracy results from the lexical decision task. Then, the bulk of the chapter focuses on the results of building the multilevel model that was used to analyze the RTs to the target words for each event Phase and for the Control Condition.

6.1 ACCURACY RESULTS

As expected, participants were highly accurate for the lexical decision task. Table 6 reports the mean accuracy, standard deviations, and range for Early and Late targets for both event phases and the control condition. Across all conditions, participants achieved a mean accuracy of 99%. Paired *t*-tests were conducted to assess for accuracy differences between Early and Late targets in the Early Phase, Late Phase, and Control Condition. Using a Bonferroni correction, the alpha level for these tests was set at .017 (i.e. .05/3). The Control Condition yielded the only significant result $t(49) = 2.69$, $p = .01$, a small numerical difference that was likely significant due to wider variance in comparison to the other trial types.

Table 6. Accuracy results for Early and Late targets for event phases and control condition

	Early Targets	Late Targets
Early Event Phase		
Mean (SD)	15.9 (.4)	15.9 (.2)
Range	14-16	15-16
Late Event Phase		
Mean (SD)	15.8 (.5)	15.9 (.4)
Range	14-16	14-16
Control Condition*		
Mean (SD)	15.9 (.3)	15.6 (.8)
Range	15-16	13-16

*significant difference, $t(49) = 2.69$, $p = .01$

6.2 IDENTIFICATION OF REACTION TIME OUTLIERS

Reaction times analyses were based on data from accurate trials only. Outliers for RTs were identified using a criterion of +/- 3 SD from the mean which was applied in two ways. First, RT outliers were removed based on the mean RT for each individual participant and then the process was repeated for each target word. In both cases, outliers were identified separately for each event phase. Identifying RT outliers by participant and target resulted in the removal of < 3% of the data. Due to the nature of the stimuli, a pairwise exclusion was then conducted resulting in a total of <6% of the data removed from the analyses. Table 7 includes the mean RTs, standard deviations, and ranges for the Early and Late targets for both event Phases for the Experimental, Control, and Baseline conditions.

Table 7. Mean RTs (in msec) with standard deviations and ranges for all conditions

	Early Targets	**Late Targets**
Early Event Phase		
Mean (SD)	601.14 (129.09)	622.14 (134.64)
Range	343-1191	384-1661
Late Event Phase		
Mean (SD)	629.53 (145.35)	632.09 (125.90)
Range	378-1465	374-1404
Control Condition		
Mean (SD)	622.91 (132.03)	620.98 (136.96)
Range	387-1475	368-1474
Baseline Condition		
Mean (SD)	565.34 (116.46)	575.78 (133.51)
Range	326-1337	275-1390

Note. msec = milliseconds; SD = standard deviation.

6.3 MULTILEVEL MODEL BUILDING

Before multilevel model building was initiated using the Stata 14.0 (StataCorp, 2015) statistical analysis package, the assumption of normality for the RT distributions was checked. Reaction time data are notoriously skewed, and each of the distributions of the RT reported in Table 7 have ratios of >|2| for both skewness and kurtosis respective to their standard errors (Dixon, Brown, Engleman, & Jennric, 1990). However, in multilevel model building, the assumptions are applied to the residuals (i.e. the observed minus the predicted values) and not the raw data. Parallel to the assumptions for typical regression analysis, the two assumptions for data distributions in multilevel modeling include normality and homogeneity of variance. Violations to either of these assumptions are corrected by using robust standard errors, so analyses that were run and are reported use robust standard errors (Raudenbush & Bryk 2002).

Reaction time data for Target RTs were submitted to multilevel modeling for analysis with items clustered in participants. Throughout the model building process, log likelihood ratios were calculated using Chi-square tests with alpha set to .05 to assess if the model fit was significantly improved by adding a specific additional component. First, the intra-class correlation (ICC) was calculated to determine if modeling a random intercept for participants increased the variance that could be accounted for in comparison to the initial model with items clustered in participants. When participants were added as a random intercept the ICC of .300 indicated that ~30% of the variance in Target RTs was attributable to the differences among individual participants. The Chi-square test was significant when the random intercept was added to the fixed effects model, χ^2 (1) = 896.48, p < .001, so a random intercept for participants was added to the model.

Starting with the random intercept model, multiple model comparisons were conducted to determine which variables significantly increased the explained variance in the model, and critically, if those variables should be added as fixed or random components. If adding a fixed component significantly improved model fit, then that component was tested again as a random component to determine if that component should also be modelled as random.

After adding a random intercept for participants to the model, model fit was evaluated when the independent variables of Event Phase and Target Type were added as fixed components. The likelihood ratio tests indicated that adding these variables significantly improved model fit, χ^2 (3) = 33.52, p < .001. To determine if modelling Event Phase and Target Type as random components improved model fit, each was added to the model separately as a random component.

There was no significant difference in model fit when either Event Phase, χ^2 (1) = 3.21, p = .07, or Target Type, χ^2 (1) = 0, p = 1.00, were modeled as random versus fixed components.

Next, fixed effect variables (i.e., Target Type, Event Phase) were evaluated through linear regression to determine whether other variables such as Baseline RT should be included in the model. While the baseline condition was originally intended to be used in a post hoc test if a congruity effect was found, adding it to the model could account for any lexical factors that may have influenced the results. Adding Baseline RT as a fixed component did significantly improve model fit, χ^2 (1) = 28.68, p < .001. When Baseline RT was subsequently added as a random component, it also significantly improved model fit, χ^2 (1) = 7.91, p = .005, so it was retained as a random component.

Another variable that was tested during the model building process was Event, representing each of the specific events investigated in this study. By their nature, events and their properties are highly diverse. The events for the current study elicited a range of attributes in the pilot study. Including Event in the model could account for some of the likely variability among the different event stimuli. Model fit was significantly improved when Event was added as a fixed component, χ^2 (15) = 153.19, p < .001 but not as a random component, χ^2 (1) = .03, p = .857.

Two other potential factors relating to the administration of the task itself were tested to determine if they should be included in the model. First, half of the participants received one order of the stimuli while the other half received a second distinct trial order. No RT differences were expected between the two counterbalanced orders, and when Order was added into the model, it did not significantly improve model fit (χ^2 (1) = .08, p = .770).

The second factor pertaining to the task administration was the possibility that participants experienced either a speeding or slowing of responses as they progressed through the 320 trial task (i.e. potentially indicative of a practice or fatigue effect). Adding Trial Number into the model as a fixed component did not significantly improve model fit, χ^2 (1) = 1.44, p = .230. Technically, testing whether Trial Number should be added as a random effect is not needed since including it as a fixed component did not improve model fit. However, there were three potential response patterns that were considered with respect to Trial Number. As the task progressed, it was possible that a subset of participants were 1) responding faster due to practice effects, 2) getting slower due to fatigue effects, or 3) were consistent responders throughout the task. Because of the likely variability in individual response patterns, Trail Number as a random component was added to the model and it did significantly improve model fit, χ^2 (1) = 18.09, p < .001.

Using notation similar to Raudenbush and Bryk (2002), the first and second levels of the mixed model are:

Equation 1. Level 1 of analytic model

$$RT_{it} = \beta_{0i} + \beta_1 EventPhase_i + \beta_2 TargetType_i + \beta_3 BaselineRT_i + \boldsymbol{\beta_{4-18}}\, \boldsymbol{Event_i} + \beta_{19}(TargetType * EventPhase)_i + \beta_{20i} TrialNumber_i + \varepsilon_{it}$$

The following components were included in the model: β_{0i} represents the intercept that has both a fixed and participant specific component, denoted by subscript i; β_1 represents the slope terms for Early and Late Event Phases (i.e. only a fixed effect component therefore no subscript); β_3 represents the fixed effect for Early and Late Targets; $\boldsymbol{\beta_{4-18}}$ represents the indicators for the event comparisons; $\boldsymbol{\beta_{19}}$ represents the interaction terms for Event Phase by Target conditions; β_{20i} represents the effect for trial number with both a fixed and random component; and ε_{it} is the level-1 residual error (i.e. the remaining unexplained variance in the outcome after controlling for all of the predictor variables).

Equation 2. Level 2 of analytic model

$$\beta_{0i} = \gamma_{00} + u_{0i}$$

$$\boldsymbol{\beta_1} = \gamma_{01} \cdots \gamma_{02}$$

$$\beta_2 = \gamma_{02}$$

$$\beta_3 = \gamma_{03}$$

$$\boldsymbol{\beta_{4-18}} = \gamma_{04} \cdots \gamma_{018}$$

$$\boldsymbol{\beta_{19}} = \gamma_{19}$$

$$\beta_{20i} = \gamma_{20} + u_{20i}$$

In this level, coefficient γ_{00} is the grand-mean (i.e. fixed effect) intercept, u_{0i} is the random intercept (i.e. the individual participants' deviation from the grand mean intercept), γ_{01} represents the fixed effect of Event Phase, γ_{02} is the fixed effect for Target Type, γ_{03} is the fixed effect for Baseline RTs for the target words, $\gamma_{04} + \cdots + \gamma_{018}$ represents the set of fixed effects for the different Events presented in the stimulus items, γ_{019} represents the fixed-effect term for the Phase by Target interaction, and $\gamma_{20} + u_{20i}$ represents the fixed-effect for Trial Number (γ_{20}) plus the individual specific deviations on the coefficient for Trial Number (u_{20i}). Combining Level 1 and Level 2 yields the composite form (i.e. both levels into one equation):

Equation 3. Composite analytic model

$$\begin{aligned} RT_{it} = {} & (\gamma_{00} + u_{0i}) + (\gamma_{01} EventPhase_i) + \gamma_{02} TargetType_i + \gamma_{03} BaselineRT_i \\ & + (\gamma_{04} Event_i + \cdots + \gamma_{018} Event_i) + (\gamma_{19} (EventPhase * TargetType)_i) \\ & + (\gamma_{20} + u_{20i}) TrialNumber_i + \varepsilon_{it} \end{aligned}$$

The composite model was fit to the data using robust standard errors to correct for violations of normality and homogeneity of variance in the data distributions (Raudenbush & Bryk, 2002) and there was a significant Phase x Congruity interaction, Mean Difference = -19.24, SE = 8.21, z = -2.34, p = .019. This interaction result is not the primary contrast of interest because in the multilevel model it compares Early Targets in the Early Phase to Late Targets in the Late Phase. To determine the result of the contrasts of interest (i.e. Early vs. Late Targets across each Event Phase) planned comparisons of the mean differences were calculated using Chi-square tests with alpha set to .05. There was a significant difference in RTs between Early and Late Targets in the Early Phase, indicative of a congruity effect, Mean Difference = 19.33, SE = 5.01, χ^2 (1) = 14.86, p < .001. However, there was no significant difference for Target Type in the Late Phase, Mean Difference = .09, SE = 6.50, χ^2 (1) = 0, p = .989. To test whether the violations of the assumptions in the data distributions were severe enough to bias the results, the model was fit without robust standard errors and the same pattern of results was found (Early Phase: Mean Difference = 19.33, SE = 5.60, χ^2 (1) = 11.93, p < .001; Late Phase: Mean Difference = .088, SE = 5.59, χ^2 (1) = 0, p = .989. The effect size for the Early Phase congruity effect was calculated based on the within participant effect from the multilevel model using Hedges' (2007) formula, $d_w = \frac{\bar{Y}_{..}^{Late} - \bar{Y}_{..}^{Early}}{\hat{\sigma}_w}$. The effect size was d = .181, a small effect according to interpretation standards set by Cohen (1988). A small effect was expected because priming effects often hinge on small differences in response times.

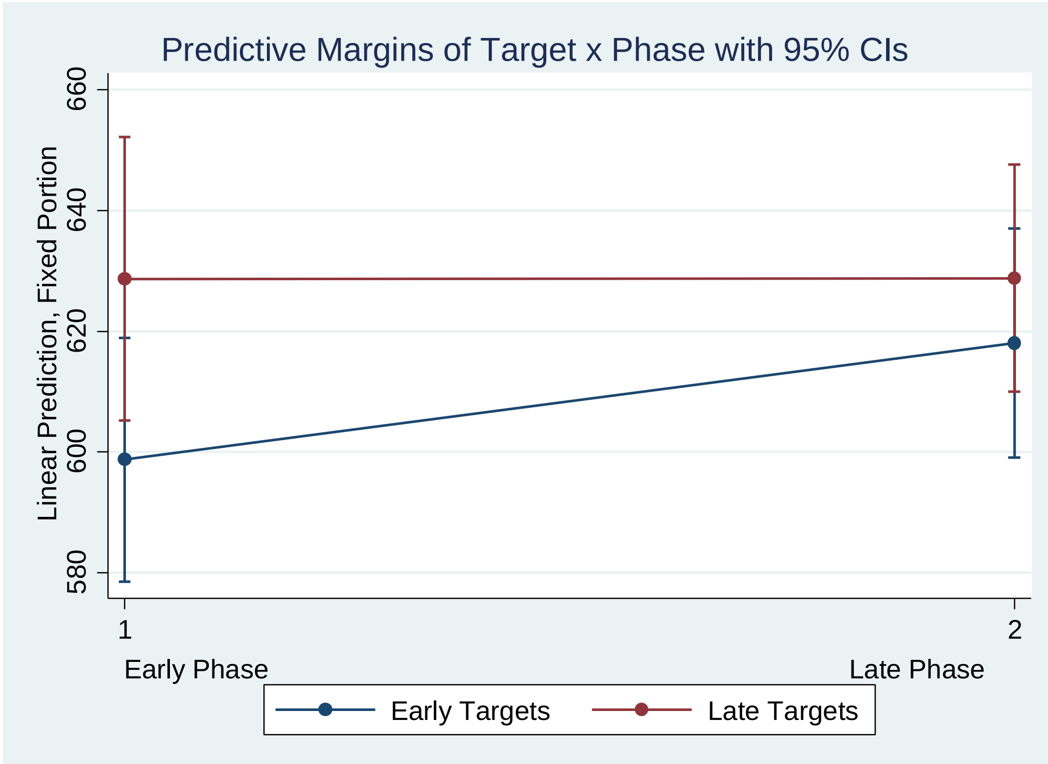
Table 8 lists the complete results from the analytic model. Event numbers were arbitrarily assigned for E-Prime programming purposes. The coefficients reported for all of the Events in the model correspond to the comparison of a given Event (i.e. Event 2 through Event 16) to Event 1.

A negative coefficient indicates a faster RT for that Event compared to Event 1. Modelling a different Event (i.e. Event 3) instead of Event 1 as the comparison Event produces different individual Event coefficients, however, the other results in the model remain the same. Figure 1 is the data plot showing the interaction between Event Phase and Target Type. A table that compares the coefficients of each individual event to the average of all the other events is included in Appendix J. The next sections link the predicted outcomes to the specific outcomes from the model.

Table 8. Analytic model results for experimental trials

Fixed Effects	*Coefficient*	*Robust Standard Error*	*Z*	*P*
Phase: γ_{01}	29.945	7.39	4.03	<0.001*
Target: γ_{02}	19.328	5.014	3.85	0.001*
Baseline RT: γ_{03}	0.077	0.026	3.01	.003*
Event 2 v. 1: γ_{04}	-26.379	9.972	-2.65	0.008*
Event 3 v.1: γ_{05}	61.909	12.990	4.77	<0.001*
Event 4 v. 1: γ_{06}	-8.642	11.648	-0.74	0.458
Event 5 v.1: γ_{07}	13.229	13.947	0.95	0.343
Event 6 v. 1: γ_{08}	-0.455	11.674	-0.04	0.969
Event 7 v.1: γ_{09}	15.759	12.999	1.21	0.225
Event 8 v. 1: γ_{010}	-19.281	11.177	-1.73	0.085
Event 9 v. 1: γ_{011}	-6.768	13.010	-0.52	0.603
Event 10 v. 1: γ_{012}	57.989	12.467	4.65	<0.001*
Event 11 v. 1: γ_{013}	-16.203	10.239	-1.58	0.144
Event 12 v. 1: γ_{014}	-23.985	10.937	-2.19	0.028*
Event 13 v. 1: γ_{015}	6.364	11.428	0.56	0.578
Event 14 v.1: γ_{016}	7.026	12.636	0.56	0.4578
Event 15 v. 1: γ_{017}	-21.395	11.003	-1.94	0.052
Event 16 v. 1: γ_{018}	-14.670	12.577	-1.17	0.243
Phase by Target: γ_{019}	-19.241	8.211	-2.34	0.019*
Trial Number: γ_{020}	-0.024	0.037	-0.65	0.516
Grand-mean: γ_{00}	557.72	17.446	31.25	<0.001*
Random Effects	*Estimate*	*Robust Standard Error*	*95% Confidence Interval*	
u_{00}: Intercept	56.657	9.693	40.517	79.228
u_{020}: Trial Number	.157	.041	.093	.261
ε_{it}: Residual	107.019	5.142	97.401	117.587

***Significant at the 0.05 probability level**



Note: CIs = confidence intervals

Figure 1. Plotted data from analytic model for experimental trials

6.3.1 Predicted Outcome 1

The first potential outcome was that there would be a congruity effect with faster RTs for time congruent attributes and/or slowing of time incongruent attributes in both event phases. This predicted outcome also specified that there would not be a Phase x Congruity interaction. When the model was fit to the data using robust standard errors, there was a significant Phase x Congruity interaction, so this predicted outcome was not found.

6.3.2 Predicted Outcome 2

The second potential outcome was that there would be a Phase x Congruity interaction. The expected pattern was a congruity effect for the Late phase and either no effect or a smaller effect for the Early phase. The actual result was the inverse of this prediction, a congruity effect was found but only for the Early phase, Mean Difference = 19.328, $p < .001$. This result indicates that the Late targets elicited a slower reaction time than the Early targets in the Early event phase. The analogous result was not found when the Early and Late targets were compared in the Late phase, Mean Difference = .09, $p = 0.989$.

6.3.3 Predicted Outcome 3

The third potential outcome was that the null hypothesis could not be rejected because the predicted congruity effect was not found in either condition. For this study, the null hypothesis was rejected but only for the Early phase.

6.4 CONTROL CONDITION RESULTS

The control condition was included in the study to replicate the robust ongoing versus completed finding in the literature. Control trials were analyzed separately from the experimental trials, and the same model building process was used. The expected results for the Completed condition were that the Early targets would yield longer RTs than the Late targets and that Late targets would yield faster RTs for the Completed Condition than in the Late Phase.

There was < 3 ms difference between the mean RTs for the Early and the Late targets in the Control condition (Early: M = 622.91 (132.03); Late: M =620.92 (136.96)), so there were no between target differences expected as a result of fitting the model to the data. The multilevel model for the Control trials included a manipulation for Target Type (i.e. Early, Late) but not for Phase because all of the sentences were written in the simple past tense. Consistent with the Experimental condition, the RT data distributions for the Control condition violate the assumptions of normality and homogeneity of variance with ratios of >|2| for both skewness and kurtosis respective to their standard errors (Dixon, Brown, Engleman, & Jennric, 1990), so robust standard errors were used in the analysis (Raudenbush & Bryk 2002). Model building began with items clustered in participants for Target RTs testing log likelihood ratios using Chi-squares with alpha set to .05. The ICC was .303 and the Chi-square test was significant when the random intercept was added to the fixed effects model, χ^2 (1) = 397.53, p < .001, so a random intercept for participants was added to the model for the Control condition.

Using the random intercept model, there was no significant difference when Target Type was added as a fixed component, χ^2 (1) = .13, p = .720, or when Baseline RT was added as a fixed component, χ^2 (1) = 3.62, p = .057. Model fit was significantly improved when Event was added as a fixed component, χ^2 (15) = 94.52, p < .001 but not as a random component, χ^2 (1) = .390, p = .534. Adding Trial Number into the model significantly improved model fit as well, as both a fixed component, χ^2 (1) = 4.20, p = .040 and a random component, χ^2 (1) = 6.32, p = .012.

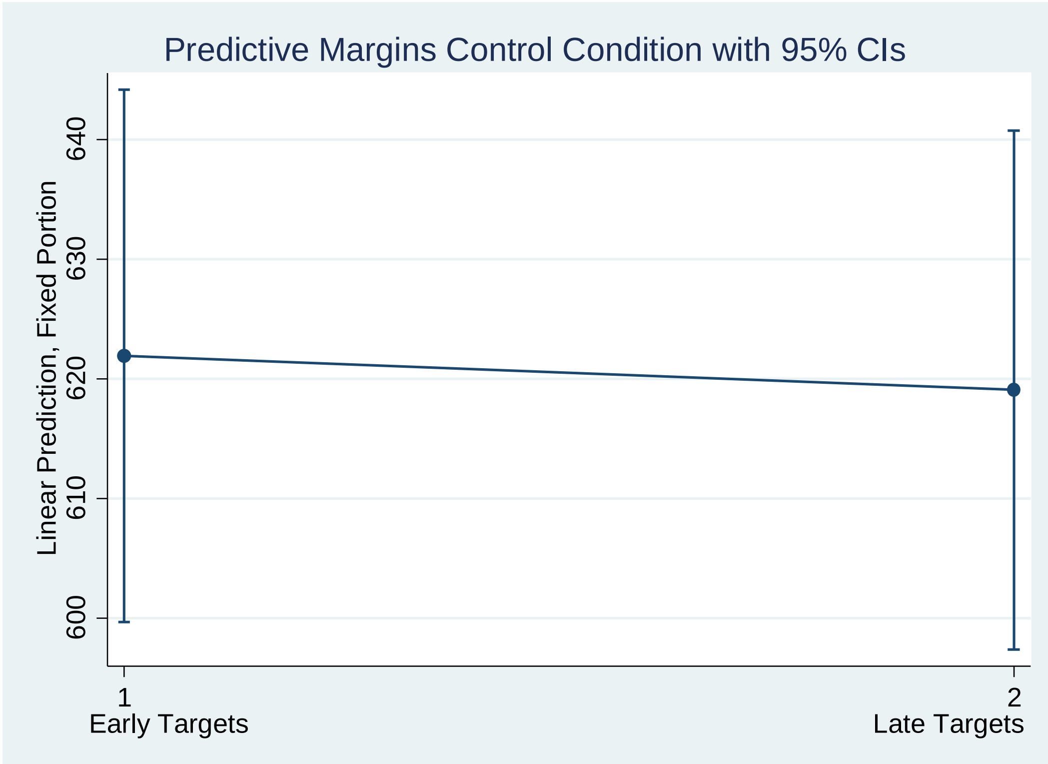
Using robust standard errors to correct for violations of normality and homogeneity of variance in the data distributions (Raudenbush & Bryk, 2000), the model was fit to the data and there was no significant difference between the RTs for the Early versus the Late targets in the

Control condition, χ^2 (1) = -2.847, p = .627. Table 9 lists the results from the analytic model for the control trials and Figure 2 is a plot of the data.

Table 9. Analytic model results for control trials

Fixed Effects	*Coefficient*	*Robust Standard Error*	*Z*	*P*
Target: γ_{01}	-2.847	5.862	-.049	.627
Event 2 v. 1: γ_{02}	-15.941	14.732	-1.08	.279
Event 3 v.1: γ_{03}	23.750	13.478	1.76	.078
Event 4 v. 1: γ_{04}	-17.926	12.146	-1.48	.140
Event 5 v.1: γ_{05}	82.497	21.880	3.77	<.001*
Event 6 v. 1: γ_{06}	18.264	14.870	1.23	.219
Event 7 v.1: γ_{07}	35.904	15.229	2.36	.018
Event 8 v. 1: γ_{08}	-14.677	13.527	-1.09	.278
Event 9 v. 1: γ_{09}	8.442	16.373	0.52	.606
Event 10 v. 1: γ_{010}	53.531	16.448	3.25	.001*
Event 11 v. 1: γ_{011}	-6.337	14.816	-.043	.669
Event 12 v. 1: γ_{012}	8.072	14.970	0.54	.590
Event 13 v. 1: γ_{013}	-16.772	14.518	-1.16	.248
Event 14 v.1: γ_{014}	37.565	16.166	2.32	.020*
Event 15 v. 1: γ_{015}	39.387	15.723	2.50	.012*
Event 16 v. 1: γ_{016}	63.844	18.283	3.49	<.001*
Trial Number: γ_{018}	.073	16.971	35.01	<.001*
Grand-mean: γ_{00}				
Random Effects	*Estimate*	*Robust Standard Error*	*95% Confidence Interval*	
u_{00}: Intercept	69.872	9.740	53.167	91.825
u_{020}: Trial Number	.155	.047	.085	.281
ε_{it}: Residual	107.452	6.786	94.943	121.611

***Significant at the 0.05 probability level**



Note: CIs = confidence intervals

Figure 2. Plotted data from analytic model for control trials

6.4.1 Across condition comparisons

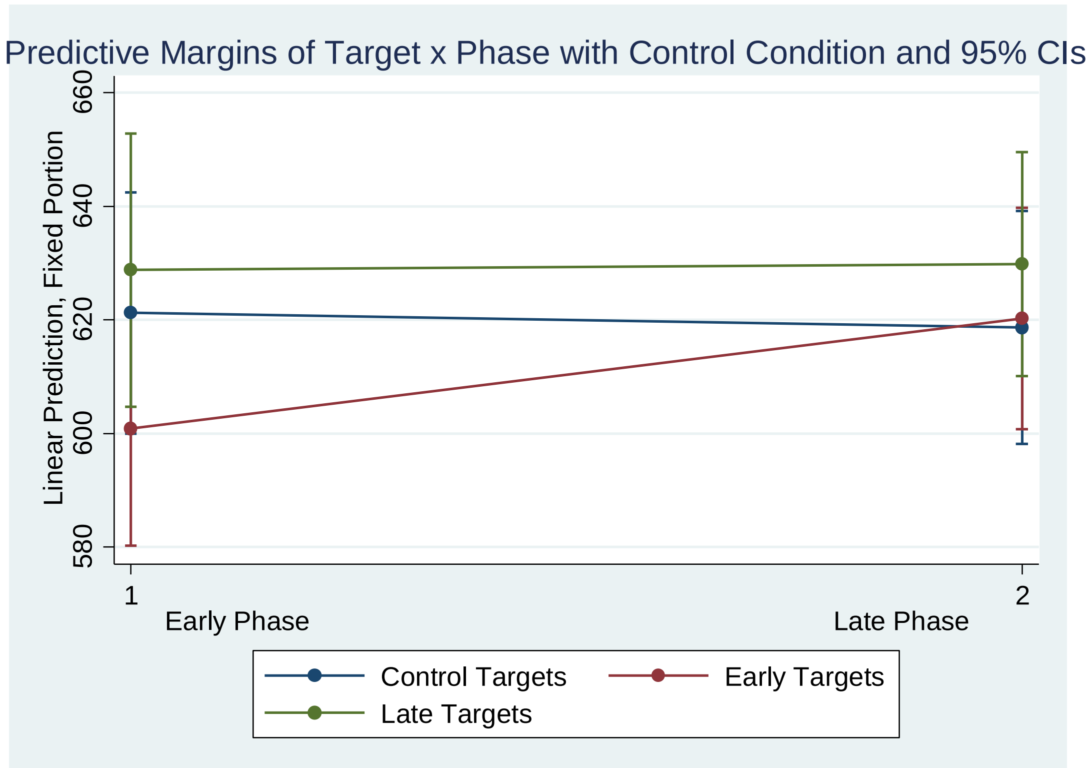
The data were added into the same multilevel mixed model to compare the ongoing conditions with the control condition. This approach was taken in lieu of a standard repeated measures ANCOVA because although the models are highly similar (i.e. both general liner models) the multilevel model can control for the random effects of Baseline RT and Trail Number. Table 10 reports the adjusted model based means, standard errors, and 95% confidence intervals for all conditions using linear combination tests (i.e. planned comparisons in multilevel models). Figure 3 is the data plot for the model. The significant congruity effect in the Early Phase (Mean Difference = 27.95, SE = 7.03, $\chi^2(1) = 15.81$, $p < .001$) and non-significant result in the Late Phase

(Mean Difference = 1.05, SE = 6.49, $\chi^2(1) = 0.03$, $p = 0.872$) and Completed Condition (Mean Difference = -2.53, SE = 5.95, $\chi^2(1) = 0.18$, $p = 0.671$) were replicated (as expected) in this model. In addition, there were two significant results when the Control Condition was compared to the ongoing conditions: 1) the Early Target RTs were significantly faster in the Early Phase compared to the Control Condition (Mean Difference = -20.40, SE = 6.80, $\chi^2(1) = 9.01$, $p = 0.003$) and 2) the Late Target RTs were significantly slower in the Late Phase compared to the Control Condition (Mean Difference = 11.13, SE = 4.07, $\chi^2(1) = 7.49$, $p = 0.006$) replicating the robust ongoing versus completed finding.

Table 10. Predicted model based means, standard errors, and 95% confidence intervals for all conditions

	Early Targets	**Late Targets**
Early Event Phase		
Predicted Mean (SE)	600.83 (10.51)	620.24 (9.95)
95% Confidence Interval	580.23, 621.42	600.73, 639.75
Late Event Phase		
Predicted Mean (SE)	628.78 (12.26)	629.82 (10.06)
95% Confidence Interval	604.74, 652.81	610.10, 649.54
Control Condition		
Predicted Mean (SE)	621.23 (10.85)*	618.70 (10.48)**
95% Confidence Interval	599.96, 642.50	598.16, 639.75

*Early targets significantly faster in Early Phase versus Control, $\chi^2(1) = 9.01$, $p = 0.003$
**Late targets in Control significant faster than in Late Phase, $\chi^2(1) = 7.49$, $p = 0.006$



Note: CIs = confidence intervals

Figure 3. Plotted data from analytic model for all conditions

6.5 POST HOC TESTING

To account for *a priori* differences for certain events based on piloting, the multilevel model that was used for the experimental trials was fit to the data without those events. An element of caution is needed when interpreting the results from models with fewer events due to a decrease in power from removing observations and a likely increase in standard errors. The first event that was dropped was sliding down a slide (Event 15) due to this event having a shorter duration than the other events (e.g., hiking). Without Event 15, the pattern of results is the same but the interaction is no longer significant, Mean Difference = -16.36, SE = 9.01, $z = -1.82$, $p = .069$, however, the interaction is not the primary contrast of interest. Critically, without Event 15, the Early Phase congruity effect remains, Mean Difference = 19.795, SE = 5.54, $\chi^2(1) = 12.76$, $p < .001$ as does the no difference result in the Late Phase, Mean Difference = 3.435, SE = 7.186, $\chi^2(1) = 0.23$, $p = .633$. For the next model, a total of four events were removed from the dataset because each of those events produced a higher number of attributes in the Late Phase vs. the Early Phase while all the other events produced the opposite pattern (i.e. a larger number of attributes generated in the Early Phase vs. the Late Phase). For this model, the events that were removed were: visiting an amusement park-Event 1, babysitting a child-Event 2, mowing the lawn-Event 11; and sliding down a slide-Event 15. Without these four events, the pattern of results is the same as the original model and the interaction is no longer significant, Mean Difference = -14.84, SE = 10.90, $z = -1.36$, $p = .173$. The Early Phase congruity effect remains, Mean Difference = 19.271, SE = 6.396, $\chi^2(1) = 9.08$, $p = .003$ with no difference in Late Phase, Mean Difference = 4.028, SE = 8.598, $\chi^2(1) = 0.27$, $p = .607$. Given these results, the primary finding of an Early Phase congruity effect appears to be a robust finding.

One other post hoc test was completed to assist with interpretation. An alternate option for examining the influence of Baseline RTs in the model is to calculate a difference value between the Target RTs and the Baseline RTs. To determine if this approach was superior to building the Baseline RTs directly into the model, the model fit was tested using the difference values instead of the original approach of autoregressing Target RTs on Baseline RTs. When the alternate approach of using the difference values was applied, the R^2 was reduced from .153 to .012 and the standard errors became inflated. While there is ongoing debate about whether or not to use difference scores in model building, the result above is consistent with a leading perspective against using difference values (e.g. Cronbach & Furby, 1970).

6.6 SUMMARY OF RESULTS

In sum, participants in this study were highly accurate across all conditions. The primary result was the inverse of the second predicted outcome: a congruity effect was found but only for the Early Phase. There was no difference between the Early vs. Late target RTs in the Control Condition. However, there were significant differences when the Control Condition was compared to the ongoing conditions. Specifically, the Late Targets were slower in the Late Phase condition compared to Control trials replicating the robust ongoing vs. completed finding in the event processing literature.

6.7 POST HOC POWER ANALYSIS

Some of the challenges for running an a priori power analysis for multilevel models were outlined in Section 5.3.1.2. The initial target sample size for this study was N=50 per Scherbaum and Ferreter (2009). Once N = 50 was achieved, a post hoc power analysis using Optimal Design Software (Raudenbush, Spybrook, Congdon, Liu, Martinez, & Bloom, 2011) was conducted to ensure the study had adequate power to detect a small effect (Hedges, 2007). The parameters for the post hoc power analysis were the following: $\alpha = .05$, $\delta = .20$, intraclass correlation (ICC) = 0.30, $R^2 = 0.13$, and n = 64. The ICC was calculated based on an unconditional model (i.e. no predictor variables) of the RT data which indicated that approximately 30% of the variance in the dataset is due to participant variability. This result underscores the use of a multilevel modelling approach versus a more basic regression analysis method. The R^2 parameter uses an ordinary least squares estimate, is based on the fixed effects model of the RT data, and resulted in approximately 13% of the variability being explained by fixed effect predictors (i.e. Event Phase, Target Type, and Events). The number of observations (i.e. n = 64) is derived from the total number of observations or data points collected for each participant. So for the current study, there was a RT for both an Early and Late time point for each of 16 events or 64 total observations. Figure 4 is the graph of the post hoc power analysis. The results from the post hoc power analysis suggested that approximately 20 subjects would be needed to identify an effect if one existed at a probability level of 0.80, so with 50 subjects the power of the current study is .994.

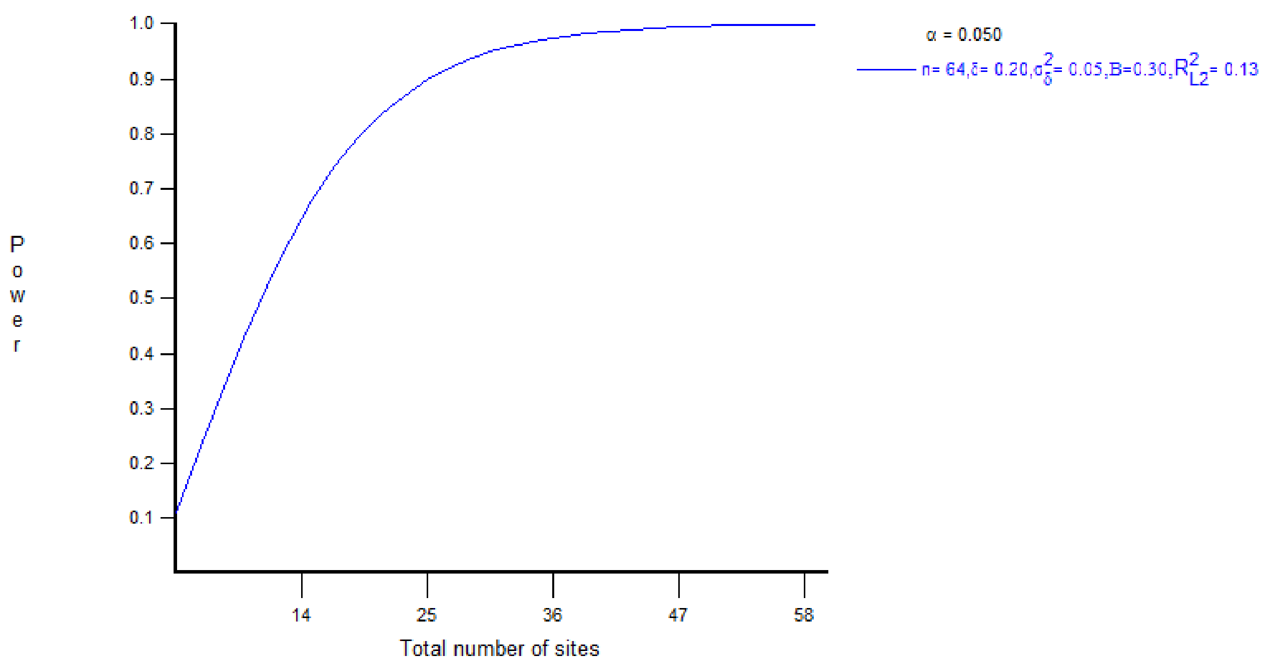


Figure 4. Post hoc power analysis

7.0 DISCUSSION

The primary focus of this investigation was to take the first step in examining whether ongoing event representations are processed in an all or nothing manner. Based on discrepancies in the current temporal processing literature and extrapolations from current theories of knowledge representations, this author proposed that it is likely that ongoing event representations engender a more fine-grained updating process during event comprehension. A pilot study was conducted to determine attributes that are associated with either just beginning or being almost finished with a common event (i.e. doing laundry). Attributes that were linked primarily to either the Early or the Late event phase were systematically chosen and used as target words in a lexical decision task.

Fifty participants used a word-by-word reading method to progress through sentence-level stimuli that cued either an Early or Late event phase within an ongoing event (see Table 5 for sample stimuli). Sentences were followed by letter strings that were real words or phonotactically plausible nonwords. Participants used a button box to respond as quickly and accurately as possible whether the letter string was a real word or a nonword. Both accuracy and RTs were collected for each target. The experimental task incorporated sentences written in the simple past tense as a control condition as well as filler trials to disguise the task. Following the experimental task, baseline RTs to the target words were collected in a separate block of trials using XXXs instead of words for the sentence stimuli.

Participants were highly accurate on the lexical decision task. Valid RTs (i.e. RTs for correct items and RTs within 3 SD of both the individual participant's mean and the target's mean RT) were analyzed using multilevel modeling. The primary finding in this study was a congruity effect in the Early event phase. The expected congruity effect in the Late phase was not found. This chapter will consider these results relative to the temporal processing literature as well as the implications for theories of event representation. The final sections of this chapter address how these results connect to the larger theoretical picture of event processing, strengths and limitations of the study, and future directions.

7.1 IMPICATIONS FOR THEORIES OF EVENT REPRESENTATION

This author's primary argument is that comprehenders engage in dynamic model updating for ongoing events contrary to the static approach proposed by current models. The theoretical rationale for this study derived from contributions from simulation theories and the semantic association perspective. In their current forms, neither perspective can account for the Early Phase congruity effect found in this study. There is a growing body of evidence that at least some temporal properties are part of event knowledge representations, however, it remains elusive what those temporal properties are. Both simulation and semantic association theories make temporally-based predictions related to event continuity (i.e. distinct mental event representations for ongoing versus completed events). Response time and reading time differences resulting from small grammatical manipulations (i.e. Ferretti et al., 2007) illustrate that there are distinct event models generated by event continuity contrasts. This section will begin by placing the findings of the current study into the landscape of temporal processing studies followed by a discussion of how the results could influence theories of event representation.

Researchers have used a variety of approaches to study temporal processing and more specifically event continuity. One of the primary studies of the processing differences between ongoing and completed events was conducted by Madden and Zwaan (2003, see Section 3.3.2). In that study, imperfective and perfective sentences were paired with black and white drawings of an event that depicted either and intermediate stage or a concluding stage of an event such as building a fire. Participants were more likely to select the completed picture for the perfective sentences but were equally likely to select either picture for the imperfective sentences. The authors concluded that the imperfective (i.e. ongoing) tense may not constrain to the specific pictures of the intermediate or completed stages of the described event, and therefore, both event phase drawings could be plausible representations of the ongoing event. The current study's congruity effect result in the Early Phase extends Madden and Zwaan's primary finding. Both studies hint that there could be a more dynamic nature to ongoing event representations. However, if replicated, the congruity effect in this dissertation would provide evidence that Early Phase time cues may help comprehenders constrain their ongoing mental representations.

In addition to event continuity, a handful of studies have examined temporal order as a different way to explore temporal processing. One example is a study by Welke et al. (2014) which contrasted source-result states of events (i.e. DIRTY-CLEAN vs. CLEAN-DIRTY). The authors found a no difference result for the order of source-result pairs (i.e. chronological or inverse order) and concluded that events activate both source and resultant states and that the temporal order may not be part of the knowledge that is associated with the event. This conclusion is more consistent with the implicit activation of an unordered set of attributes than the type of event progression that is predicted by simulation theory. Similar results were found when Landgraf et al. (2012) compared relatedness judgment RTs to subevent pairs such as GET MENU and PAY THE BILL. The results of this dissertation study are inconsistent with the findings from these two studies and suggest instead that temporal order might be part of the event representation. One explanation for the difference could be that the additional context provided by the event phase cue in the sentence level stimuli from the current study signals to the comprehender that temporal order should be part of the event representation. An alternative explanation is that the lexical decision task revealed a type of temporal order effect that the relatedness judgments were not sensitive enough to detect at the single word and short phrase levels. Specifically related to the no difference finding in the Landgraf study, it could also be that the subevents in their stimuli did not tap different event phases per the Moens and Steedman (1998) framework.

One final study that examined whether temporal sequencing is encoded in semantic memory was conducted by McRae and colleagues (Khalkhali, Wammes, and McRae, 2012). In this study, three subevents of a given event were responded to the fastest in a relatedness decision task if they were presented in the real world sequence (i.e. MARINATE-GRILL-CHEW). This result underscores the finding from the current study that temporal order may indeed be encoded as part of the knowledge associated with a given event. It could be that the subevents in the Khalkhali et al. study (2012) and the attributes from the current study more closely match the event phases proposed by Moens and Steedman (1998) than studies that did not find a temporal order difference. The next sections discuss the results of the current study in light of current event processing models.

7.1.1 Simulation theory

Simulation theories propose that sensory and motor experiential traces are activated to form mental representations. Researchers from this camp claim that during event processing, the experiential traces are mentally activated in a pattern that is similar to what occurs when comprehender is actually experiencing the event. The bulk of the evidence from this perspective is consistent with perceptual congruity or action compatibility effects (i.e. The pencil in the DRAWER/CUP paired with a picture of a pencil in a horizontal or vertical position). The prediction from this perspective that is most relevant for the current study is that ongoing events prompt the mental activation of a larger number of multi-modal experiential traces compared to completed events. Consequently, simulating only the end stage result of a completed event has typically prompted faster reading times or shorter RTs to event attributes in contrast to ongoing events. Simulation theories suggest that a comprehender takes an internal vantage point to an ongoing event as if that event is in progress, but whether there is an incremental nature to the processing of ongoing events has not been considered. The leading critique of this perspective is that it is underspecified. One example of this problem in the temporal processing domain is that simulation theories do not make any specific predictions beyond an ongoing versus completed dichotomy for event processing.

For the purposes of the current study, predictions beyond the ongoing versus completed dichotomy were inferred from current literature. It was argued that it is possible that ongoing events prompt incremental changes in mental activation as a result of more fine-grained temporal processing and updating. Incorporating this proposed distinction between early and late simulations for an ongoing event would extend the simulation perspective and increase its specificity. For example, the prediction for the current study was that the experiential traces that are mentally activated early on in an ongoing event representation may narrow as a comprehender approaches the end of an ongoing event. Specifically, the simulation may be winding down and more closely represent the end stage completed event. Based on this prediction, the expected pattern of results was a congruity effect: Early Targets prompting faster RTs in the Early Phase compared to the Late Phase and the opposite prediction for the Late Targets.

The current study used time cues (i.e. recently started/almost finished) to examine the predicted congruity effects within ongoing events. The effect was found, but only for the Early Phase. This result indicates that there was a clear advantage for the Early Targets when they were paired with the Early Phase cue that was not evident for the Late Targets in the Late Phase. This Early Phase facilitation result points to: 1) the activation of a distinct cluster of attributes at the beginning of an event simulation and 2) the possibility that this cluster of attributes might become less relevant as the simulation progresses. Instead of the originally proposed narrowing of event attributes when a comprehender was prompted by a Late Phase cue, there could be different additional attributes that become activated as the simulation progresses. It is unclear whether this Early Phase congruity effect represents a true facilitation effect or a slowing for the responses in the incongruent condition.

The Early Phase congruity effect is consistent with two prominent event processing concepts discussed in Chapter 3. The first is the economy of processing principle (see Section 3.3.1.1; Graesser, Singer, Trabasso, 1994) which was derived from the simulation point of view. According to this principle, comprehenders activate only what is needed to understand a given situation (i.e. differential experiential trace activation for the shape of the eagle for: The eagle is in the TREE/SKY). For the current study, it is possible that comprehenders are selectively activating only the experiential traces that are necessary to fulfill the beginning stages of the simulation. Based on the longer RTs for the Early Targets with the Late Phase cue, it is possible that those attributes become less relevant as the simulation progresses.

The second prominent event processing concept that the Early Phase congruity effect finding underscores is the Moens and Steedman (1988; see Section 3.1.1) proposal that there are three phases of an event: 1) a set of initiating conditions, 2) the actual event, and 3) the resultant states. The Early Phase cue might have facilitated the attributes that are connected to the initiating conditions event phase. The results of the strength of association rating task that was completed during piloting are consistent with this conclusion. The Early Targets were strongly associated with both the Before and the Just Beginning event phases. So, it is possible that the Early Targets were consistent with the set of initiating conditions for a given event and were primed by the Early Phase cue.

The no difference result for the Late Phase could indicate that both sets of attributes are relevant toward the end of the ongoing simulation. However, this result could also be driven by a processing mismatch between the Early Targets and the Late Phase cue. The results of the pilot study support the mismatch interpretation based on the low ratings between the Almost Finished Phase and the Early Targets. Irrespective of the potential interpretations of the no difference result, it does not appear that the Late Phase cue facilitated the processing of the Late Targets. In fact, based on the adjusted model based means from Table 10, the Late Targets were responded to ~10 ms faster in the Early Phase cue condition compared to the Late Phase condition.

Turning to the Completed Condition, the simulation perspective and Moens and Steedman (1988) predict that completed events focus the comprehender on the end stage result of an event. This emphasis on the end of the event prompts the activation of fewer experiential traces or a shorter simulation compared to the ongoing conditions. The first predicted result for the Completed Condition was that the Early targets would yield longer RTs than the Late targets. This prediction was derived from the idea that the Early Targets might no longer be relevant in completed event representation and was supported by the low ratings in the Almost Finished event phase. This predicted result was not found as there was no difference between the RTs for the Early and Late Targets in the Completed Condition. The second expected result and the primary contrast of interest for replicating the robust ongoing versus completed finding in the literature was that the Late targets would yield faster RTs for the Completed Condition than in the Late Phase. As predicted, there was a significant difference for the Late Targets with faster RTs for the Completed Condition compared to the Late Phase (Mean Difference = 11.13, SE = 4.07, $\chi^2(1) = 7.49$, p = 0.006).

In their current form, simulation theories would seem to imply that there is a progressive nature to ongoing event representations, but this premise is untested in that literature. If the Early Phase congruity effect result from this study is replicated, it could be argued that ongoing event representations are not homogenous and that there could be a temporal progression to simulations of ongoing events. The congruity effect finding could also prompt a necessary addition to the simulation perspective. Specifically, the models would need to reflect that time based cues can differentially influence the experiential traces that are activated during ongoing event representations.

The type of congruity effect found in the current study also begins to address the under-specification critique of the simulation perspective. In their current form, simulation theories do not readily account for different levels of abstraction such as how the abstract concepts like lucky and boredom emerge from perceptually driven traces. More specifically for the current finding, there is no proposed mechanism that modulates the scale of a simulation or for how simulations capture the necessary compression of the event timeline (i.e. for ten years, own a house). The congruity effect result in the current study could add a level of abstraction to the current set of perceptually driven effects that dominate the literature. At present, the bulk of the literature that is consistent with the simulation perspective is driven by stimuli like: The pencil is in the CUP/DRAWER with a picture of a pencil in a horizontal or vertical position. Simulation theories argue that these visually driven stimuli selectively activate the perceptual traces that match the object's orientation which drives the congruity effect.

This perceptual trace driven argument for congruity effects does not hold for time cues because they are not directly linked to a specific sense organ. It is possible that subtle grammatical distinctions and event phase cues are used to signal a progressive nature to ongoing event representations. If this is true, it would provide a much needed example of how one component of the abstract concept of time is integrated into and influences event simulations. If replicated, the primary finding for the current study could support the previously implied incremental nature to ongoing event representations, and argue that the activation of a set of experiential traces is potentially not maintained throughout the entire simulation. This argument of differential experiential activation throughout ongoing event simulations would advance the broader conceptualization of how time processing cues shape event processing and more specifically how ongoing event representations might be simulated.

7.1.2 Semantic association theory

The semantic association perspective proposes that the foundation of a mental event representation is linked to the diversity of semantic associations in the described situation (Coll-Florit & Gennari, 2011). For this camp, durative verbs (i.e. ongoing) prompt longer reading times than non-durative verbs (i.e. completed). The authors claim that the longer RTs are representative of comprehenders tapping more diverse associated knowledge when reading durative verbs. The increased processing time reflected in the longer RTs has been linked to either resolving potentially competing interpretations of co-occurring semantic associations or the increased processing time that is necessary to allow for weaker semantic associations to be retrieved.

The semantic association perspective continues to be specified, but in its current form it does not make any specific predications about the potential for more fine-grained processing during an ongoing event. However, it can be inferred from the theory that as comprehension proceeds, there is a resolution of the attribute activation patterns (i.e. resolving competing semantic associations or activation of the weaker interpretation). Focusing on the predicted mean RTs between the Late Phase and the Control Condition (see Table 10), there was a significant difference for the Late Targets in the predicted direction (i.e. longer RTs for Late Phase/ongoing versus Control/completed). In addition, while not significant, the RTs for the Early Targets are in the same direction (i.e. longer RTs for Late Phase/ongoing versus Control/completed). The congruity effect in the Early Phase is suggestive of a time cue helping to pinpoint relevant associations in the early stages of an event and that there might be different subsets of associations that are activated throughout an ongoing event.

The goal of the current study was to examine how event representations unfold within events that are in progress. In its current form, the semantic association perspective only predicts differences between ongoing and completed representations (i.e. durative versus non durative verbs). In addition, these event continuity differences are reflected in longer reading times for the durative or ongoing verb phrases compared to the non-durative or competed condition. Reading time differences were not expected for the experimental trials in the current study because all of sentences were written in the past perfect progressive tense to signal an ongoing event representation. Instead of examining reading times, this study tapped into likely semantic priming processes for attributes that were strongly associated with a given event phase. The critical finding from the current study is that the internal context of the event appears to make a difference in how salient or available the event attributes are at different points in the ongoing event. It is possible that the larger volume and/or diversity of event attributes that are predicted for the ongoing event verbs are not maintained uniformly throughout the entire time course of the ongoing event. If the results of the current study are replicated, the semantic association perspective would need to be updated to reflect a difference among the attributes that are implicitly activated across event phases. The potential classification of those differences will be considered next.

The leading premise of the sematic association perspective is that ongoing events elicit a larger and more diverse set of associations compared to their completed counterparts. If ongoing event representations do elicit a larger volume and diversity of event attributes, it is currently not clear which element (i.e. volume or diversity) might induce a stronger influence. It is possible that the volume versus diversity question will be a difficult one to tackle due to the inherent variability across the temporal aspect of events such as duration as well as other event properties (e.g., familiarity). Studies like the attribute generation pilot study could help to begin to explore this piece of semantic association theory.

One final argument to consider relative to the semantic association perspective is that the ambiguity of ongoing event representations themselves could be a primary factor in the longer reading times in the Gennari studies (2011). The past (perfect) progressive tense might be setting up an expectation that comprehenders should anticipate some form of additional context. For example, reading a sentence like *Alice was baking/recently started baking a cake* could imply that another event was happening at the same time, that the reader should anticipate additional details about that event, or that the comprehender should expect a subsequent event to be introduced.

7.2 IMPLICATIONS OF CONGRUITY EFFECT

Empirical findings strongly suggest that at least some temporal properties of events influence knowledge representation. Examples include the robust ongoing versus completed finding throughout the event processing literature, the grammatical and lexical studies reviewed in Section 3.3.1 (e.g. Becker et al., 2013; Ferretti et al., 2007; Magliano & Schleich, 2000; Yap et al., 2009), and the argument for a durative versus non-durative distinction that drives the semantic association perspective (Coll-Florit & Gennari, 2011). If replicated, the Early Phase congruity effect finding in the current study could add to the growing list of studies that highlight the importance of temporal properties and their potential influences on event representations. It is possible that an Early event phase cue may help to constrain the number or type of attributes that are connected to a given event. Based on the preliminary findings in this dissertation study, it does not appear that a Late event phase cue provides the same advantage.

Despite the lack of a congruity effect in the Late Phase, the Late Targets did elicit longer RTs in the Late Phase versus the Control Condition. This finding supports the event continuity contrast prediction from both the simulation and semantic association perspectives. It is possible that despite their high associative ratings to the Almost Finished and After event phases, the Late Targets are not representative of activated attributes toward the end of an ongoing event representation. It is noteworthy though that the replicated ongoing versus completed effect does not support this possibility. If the Late Targets are representative of attributes in the Late Phase, then perhaps the Late Phase cue does not help comprehenders to: 1) truncate or focus their simulation toward the end of the ongoing event or 2) decrease or narrow the number of activated attributes. It is possible that, given a Late Phase time cue, comprehenders still need to simulate at least part of the earlier stages of an ongoing event or activate the attributes that are also associated with the portion of the ongoing event that occurs prior to the time cue. An alternate possibility could be that the faster RTs for both sets of Targets (i.e. Early and Late) in the Early Phase compared to the Late Phase are a reflection of the unfolding simulation or the pattern of activated attributes within ongoing event representations.

7.3 POTENTIAL CLINICAL IMPLICATIONS

There are multiple potential clinical implications for this line of research. One leading possibility for a clinical application is exploring if individuals with discourse comprehension deficits have a decreased ability to integrate temporal processing cues. Based on the primary result from this dissertation, it appears that non-brain-damaged adults may benefit from Early event phase cues. If this result is replicated, this finding could be tested with older adults and eventually with adults with brain damage who demonstrate discourse processing deficits. It will be imperative to compare the results of the young and older non-brain-damaged adults prior to any application in the clinical domain because there is limited but emerging evidence that older adults may be more effective at integrating their event knowledge compared to younger adults after the expected differences in baseline RT between the two groups are considered (Brod, Werkle-Bergner, & Shing, 2013; Prull, Gabrieli, Bunge, 2000).

Two distinct patient populations, adults with either right hemisphere brain damage (RHD) or aphasia, frequently exhibit sentence and discourse level processing deficits (i.e. Blake, 2011; Tompkins, Klepousniotou & Gibbs Scott, 2012; Ferrill, Love, & Shapiro, 2015; Tompkins, 1995). Empirical evidence from individuals with discourse comprehension deficits as a consequence of RHD indicates that some adults benefit from strong contextual cues (i.e. Blake, 2009; Blake, Tompkins, Scharp, Meigh, & Wambaugh, 2015). If a decreased ability to integrate time cues can be identified, it could be that providing redundant temporal cues may help to mitigate a possible impairment in the time processing domain. In the aphasia literature, the clinical profile of some adults includes a specific difficulty integrating a variety of grammatical cues (i.e. Charles et al., 2013; Dick, Bates, Wulfeck, Utman, Dronkers, & Gernsbacher, 2001). Current aphasia batteries are deficient in providing accurate and comprehensive language impairment profiles for a range of verb forms and their properties which can make addressing syntactic deficits in individuals with aphasia particularly difficult (i.e. Cho-Reyes & Thompson, 2012). Continued study is needed to fully address the potential impact of syntactic deficits in adults with aphasia and it could be that examining the possible benefits of incorporating temporal processing cues could help to address this need. Finally, adults with aphasia often benefit from additional context or demonstrate increased comprehension given discourse level versus sentence level stimuli (Ferrill, Love, & Shapiro, 2015). This positive impact on comprehension via providing additional or redundant contextual cues points to the possibility that providing temporal processing cues may be a fruitful avenue to explore with patient populations.

7.4 INDIVIDUAL DIFFERENCES

As discussed in Section 2.1, events and their properties are highly variable (e.g., duration, agents, objects, and locations). In addition to the multitude of differences between events, it is a given that there are person-level differences linked to individual events. Specifically, each individual’s prior knowledge, familiarity, and experiences could influence the subset of experiential traces and/or attributes that are implicitly mentally activated when an adult reads (or hears) a sentence or story about a particular event. Taking the customization of event knowledge a step further, there could also be influences of how recently or frequently a comprehender experiences an event. For example, the mental representations are likely quite different between an adult who went finishing one time as a child compared to an avid fisherman or someone who went fishing yesterday. Event knowledge is also likely to be strongly influenced by what someone does for a living or enjoys as a hobby (i.e. a baker or gardener). One option for trying to even out likely event-level and person-level differences in order to study the potential influences of temporal processing cues could be to provide stimuli that highlight specific attributes and link them to different event phases. Another possibility could be to design a familiarity questionnaire for the events used in a stimulus set to use as a potential covariate during data analysis.

In addition to the types of individual differences described above, there are also potential differences in how participants approached the lexical decision task. In total, there were 320 trials and all required the same moving window word-by-word reading followed by a lexical decision. Trials were grouped into sets of 32, and the software was programmed to offer participants a break after each set. Participants were instructed to read each sentence carefully and then to make the word or nonword judgement. Because there were not comprehension questions presented throughout the task, it is possible that some participants were quickly progressing through the sentence without carefully reading it to get to the lexical decision portion. This strategy of quickly advancing to the lexical decision in order to advance through and finish the task in the shortest amount of time possible could negatively influence the results. This approach to the task would mean that the results reflect a pure lexical decision only without the potential priming influence of the sentences and the time cues.

It is likely that the majority of the participants completed the task as instructed and were reading the sentences prior to making the lexical decision. This inference is based on the result of a strong congruity effect for the Early Phase cue only. Because the targets were the same across conditions and the experimental, control, and filler sentences were counterbalanced and delivered within the same blocks, if participants were not influenced by the sentences, then a no difference result would be predicted between conditions. However, it could be argued that the lack of a difference between the Early versus Late Target RTs in the Control Condition might be indicative of a pure lexical decision approach to the experimental task, where participants were not carefully reading the sentences.

One way to ascertain if participants were quickly progressing through the sentences without reading them is to analyze the reading time data between the different sentence types. In the original Gennari (2004) study, ongoing sentences resulted in longer reading times than their completed counterparts. The critical region for this dissertation study is the verb phrase itself: *baking a cake/baked a cake*. Based on the Gennari result, longer reading times would be predicted for the critical region of all of the ongoing sentences (i.e. both time cue sentences) compared to the critical region for the completed sentences (after accounting for the slight difference in the number of characters between the verb phrases). The two time points within the ongoing condition for this study differ from the original Gennari (2004; Coll-Florit & Gennari, 2011) studies where only an ongoing versus completed contrast was used. This author predicts that reading times for the critical regions for the Early Phase cue sentences would be longer than the Late Phase cue sentences, and that the reading times for the critical region in both time cue sentences would be longer than the parallel critical region in the completed condition.

7.5 STRENGTHS AND LIMITATIONS

This study was innovative because it was the first to examine the effects of signaling different time points within a single ongoing event representation. It was an initial step toward determining whether there is a fine-grained processing mechanism triggered by temporal processing cues during mental event representation updating. Both simulation theory and the semantic association perspective predict that ongoing events are more dynamic in nature than completed events. However, it is unclear whether the mental activation of either the experiential simulation or the diversity of semantic associations is maintained uniformly throughout the time course of an ongoing event. From the constraint-based perspective, it is an open question whether temporal cues are part of the common knowledge that is used to constrain the implicit mental activation during event processing. This study examined two time points near the two extremes of an individual event timeline as a first test of the hypothesis that there is an incremental nature to ongoing event representations. The primary result was that comprehenders were sensitive to time cues within ongoing events, but the congruity effect was limited to Early Phase time cues.

This study addressed several limitations that are commonplace in the temporal processing literature. First, it incorporated word-by-word sentence processing versus the whole sentence reading paradigms used in the majority of other studies of event continuity. Second, this study used lexical decision instead of relatedness judgement to a probe word or a picture. A lexical decision task is more implicit than a relatedness judgment task. In addition, relatedness judgement tasks are limited in that the pictures that are meant to depict a specific event phase may not exactly match the mental representation of the comprehender resulting in interference. The application of a multilevel analysis approach, the inclusion of a control condition to replicate prior findings, and adding a baseline condition were each elements of this study’s design that have been rarely considered in the temporal processing literature. Finally, the set of stimuli used in this study were tightly controlled and culled from a pilot study that included both an attribute generation and a rating task in order to increase experimental precision.

In addition to the strengths stated above, this study had several limitations. First, the lack of a "neutral" condition limited the interpretation of the congruity effect that was found in the Early Phase. It is unclear whether the Event Phase cues that were paired with the Early Targets reflect a true facilitation in the Early Phase or perhaps a slowing in the Late Phase. Adding in a "neutral" condition that pairs the target words with past perfect progressive sentences without a time cue is one way test the potential influences of event phase cues. Adding another condition that includes target words that are strongly linked to both event phases might also assist with interpretation. For example, in the pilot study, multiple events elicited the same word with approximately equal frequency for both the early and late ongoing event cues (i.e. HELMET for riding a bike). Including a block of trials where there is no predicted RT difference to the target words between any of the event phases could provide additional justification for the expected congruity effects in the experimental conditions.

Focusing on the target words used in the study, it is possible that the event attribute generation task did not elicit an accurate array of the typical attributes that are commonly associated with a given event or event phase. For the event attributes that were selected, an average of 40% of individuals generated those features. Perhaps dividing up the piloting task differently (i.e. >2 sets of individuals generating the attributes or a smaller group of events or event phases per group to reduce task duration) or multiple rounds of piloting (i.e. subsequent surveys using the events with the largest attribute overlap in a given condition) might provide additional target options with a higher average response percentage. Iterative rounds of piloting might also reveal potential target words for events that only had a target word surface for one but not both of the event phases. Another potential limitation was conducting the rating task with only three raters. Perhaps using a larger pool of raters might have generated a more precise set of ratings for how strongly a potential target word was associated with specified event phase.

Another study limitation is that the target words that were selected for the current study had essentially parallel ratings (i.e. strong association) with both the "before" and "just beginning" phases and the "almost finished" and "after" event phases. Perhaps using additional raters with multiple potential target options could enhance target selection by further specifying the attributes that are linked to a given event phase or reveal if there is a distinction for some target words between other event phases (i.e. between Almost Finished and After). Finally, this study examined only two time points within an ongoing event. It is possible that probing other time points on the event timeline would yield a different pattern of results.

7.6 FUTURE DIRECTIONS

The ability to quickly integrate time cues is a critical piece of the discourse comprehension puzzle. The overarching question for this dissertation study was whether there are degrees of ongoingness in mental event representations. The primary result of a congruity effect in the Early Phase for ongoing events, if replicated, could support the idea that ongoing event representations are graded and dynamic in nature. Two time points near the extremes of an ongoing event were selected to elicit a strong contrast within in progress events as a first test of this idea. Probing additional time points on the timeline within individual events would be necessary to continue to address this question.

One challenge that surfaces when considering different methods for probing other time points is identifying a cue that focuses a comprehender on a specific place in the event timeline. This challenge is relevant not only for studies of within event processing but also for examining boundaries between events as outlined briefly in Section 3.3. There is consensus that readers construct *flexible* event boundaries (e.g. Rapp & Taylor, 2004) between events, but it is unclear what event properties might permeate those boundaries during processing. This is where the importance of event continuity bubbles up in the literature. It could be that comprehenders use time cues like event phase markers to help divide up the event processing pie, however, person-level differences and event-level variability (i.e. duration) add to the complexity of exploring event boundaries and the potential maintenance of event properties both within and across event boundaries. In addition, the specific time cue provided may not exactly match the boundaries that comprehenders implicitly use during event comprehension which could prompt inflated processing times for resolving competing interpretations.

If multiple points within an ongoing event could be isolated, it is likely that there will be a large amount of overlap in the attributes that are linked to those time points. For example, in the pilot study for this investigation many of the same potential targets were generated for both Recently Started and Almost Finished event phases. Additionally, one interpretation of the no difference result for the response times between Early and Late targets for the Late Phase could be that both sets of attributes are relevant in the Late event phase even though the Early Targets were strongly linked (only) to the Early Phase in a rating task. The lack of a clear demarcation for the types of event properties that are linked to a given event phase leaves open the possibility that that there is a fluctuation in the strength of the mental activation among the different attributes that are strongly linked to different events and event phases. Perhaps this is an entry point for further study of the potential for levels of ongoingess within events that are in progress. If there is a fluctuation in the mental activation for specific attributes within ongoing events, this finding would underscore the progressive or dynamic nature to processing in progress events. Taking this line of inquiry a step further, the trend of faster RTs for both sets of Targets in the Early Phase and slower RTs in the Late Phase suggests that there could be a type of transformation occurring as comprehenders navigate the timeline of an individual event that is cued as in progress. Probing additional points within ongoing event representations could help to illuminate the validity and implications of this notion.

Another facet of event processing that adds to the complexity of examining the properties of events and their boundaries is that events often overlap or co-occur with other events. Comprehenders must be able to navigate multiple event timelines in order to understand the nature and progression of individual events and when events occur in relationship to other events (Hinrichs, 1986). Perhaps the need to entertain multiple event timelines is where flexible event boundaries and the maintenance of activation for certain event properties are most relevant, especially given the likelihood of co-occurring events sharing at least some overlapping properties. The need to maintain multiple event properties and timelines provides a natural connection to the high span vs. low span working memory finding from Magliano & Schleich (2000). In their study, only the high span readers maintained activation of the in progress events for up to three additional sentences when those events had a long duration. Measuring working memory abilities and addressing them as an individual difference variable in future studies may be fruitful for interpreting person-level differences for processing time cues.

Based on the current findings, one area that appears to warrant further study is the three event phase distinctions proposed by Moens and Steedman (1988). The Targets that were selected for the Early Phase were strongly linked to both the Before and Recently Started event phases during the piloting rating task. Moens and Steedman propose that the first event phase is a set of initiating conditions, so it could be that the Early Targets were representative of the set of initiating conditions rather than the actual event (i.e. the second phase). The same is true for the resultant phase (i.e. the third phase). The Late Phase targets were strongly associated with both the Almost Finished and After phases in the rating task. It is unclear from the current findings if an Almost Finished time cue focuses the comprehender toward the end of the actual event phase or the resultant phase. Therefore, an extension of the current study could include either identifying additional event-target pairs with a range of association ratings or adding alternative event phase cues (i.e. before, during, after).

The results of this dissertation have generated multiple directions for future research in the temporal processing domain. One area that is ripe for future study relates to the prediction from Ferretti and colleagues (2007) that there are different types of event attributes that are linked to the three event phase framework of Moens and Steedman (2000). The attributes that Ferretti's group propose are connected to the first event phase or set of initiating conditions are agents and instruments, while the second phase emphasizes patients and locations, and the resultant phase focuses the comprehender on the patients in the end-stage of a given event. Adding additional weight to exploring the different types of event attributes that could be linked to a range of event phases is the fact that Gennari's group (2011) also specifies that ongoing event representations emphasize a larger and more diverse set of attributes compared to completed events. It could be that an answer to the open question of what type of attributes are maintained throughout the processing of an ongoing event (or across event boundaries) relies on examining a range of attribute types at different points on the event timeline.

Reconciling the leading perspectives in event processing will be a challenging task. There is a considerable amount of overlap in the positions with supporting empirical evidence among the theories. A central premise for both simulation theories and the semantic association perspective is that comprehenders quickly activate stored event-specific knowledge that gives rise to semantic interpretations and incorporates linguistic markers. Critically for the current study, comprehenders appear to be sensitive to temporal processing cues including both small grammatical distinctions, and now, perhaps event phase cues. Leading event processing theories consider event verbs to be the driving force for the types of knowledge that are tapped during event model formation and updating (Ferretti, Kutas, & McRae, 2007; Zwaan, 2003). Time-based cues are used to help shape event representations that likely include the implicit activation of multimodal experiential traces and/or an array of attributes that are associated with a given event.

In their current forms, simulation theory and the semantic association perspective cannot rule out one another. One challenge to reconciling these perspectives is the lack of a proposed mechanism for online language processing for simulation theories. It is unclear how the rather abstract perfective vs. imperfective linguistic distinction arises from perceptually driven experiential traces. Simulation theories do not articulate how the proposed event specific simulations that are rooted in sensory-motor properties enable quick integration of different types of linguistic markers with a given verb phrase and/or other context from the surrounding sentence(s).

On the other hand, the semantic association perspective proposes a mechanism similar to those offered for resolving lexical ambiguities (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994). This probabilistic mechanism takes more of a computational modeling approach that is not limited to consideration of sensory-motor properties. Instead, scholars like Gennari (2011) assert that regularities from real world experiences merge with linguistic experience to propel “word-to-world mappings” (pg. 67). The primary finding that Gennari points to for evidence of world-to-word mappings is that event duration ratings correlate with contextual diversity (i.e. corpus analysis, see Section 4.2). Pinpointing event duration, despite its complexities (see Section 3.2), as a central feature in how event representations are structured and updated offers a framework for testable predictions. For example, specifying and contrasting events of different durations (i.e. ran for 1 mile vs. 12 miles) might provide a window into the role event duration could play in structuring event representations. One caution to this approach parallels a challenge in the time shift literature from Section 3.1, specifically, defining how the boundaries are defined. The hour later vs. minute later distinction from the time shift studies is an arbitrary one, and further study of event duration will need to find firmer footing in order to advance the theoretical underpinnings of event representation models.

The current study is the beginning of a line of research that stretches beyond the standard semantic and syntactic manipulations in the temporal processing literature, and provides a natural application to understanding the nature of strengths and weaknesses in clinical populations with discourse comprehension deficits. Future work will continue to be theoretically motivated by pitting predictions from different event processing perspectives against one another while targeting the open questions related to specifying the nature of temporal event processing. One specific direction for this research will be to explore the possibility of degrees of ongoingess in mental event representations for events that are in progress by probing multiple points along the ongoing event timeline. In conclusion, this study controlled for several limitations of prior investigations, furthered the understanding of the contribution of temporal processing cues for ongoing event representations, and yielded multiple avenues for future study.

APPENDIX A

RESEARCH QUESTION, HYPOTHESES, POTENTIAL OUTCOMES, AND INTERPRETATIONS

<u>Research Question:</u> At either an Early or Late temporal reference point for an ongoing event, do time congruent attributes prompt faster RTs than time incongruent attributes?

Assumption: The explicit piloting task generated a valid representation of the attributes that are implicitly activated for an ongoing event representation at each time point.

Hypotheses	Potential Outcomes	Potential Interpretations
H_0: When an Early or Late temporal reference point is cued for ongoing events, there is no difference between RTs for time congruent versus time incongruent event attributes. H_1: When an Early or Late temporal reference point is cued for ongoing events, time congruent attributes prompt faster RTs than time incongruent attributes	Potential Outcome #1 • Congruity effect EARLY time cue + EARLY target RTs faster than EARLY time cue + LATE target LATE time cue + LATE target RTs faster than LATE time cue + EARLY target • Highly accurate responses for both early and late lexical targets	Potential Outcome #1 • Facilitation for time congruent attributes and/or slowing of time incongruent attributes • No interaction between time point and congruity • Congruity effects for Early and Late time points are not statistically different
	Potential Outcome #2 • Congruity x Phase interaction	Potential Outcome #2

	• Congruity effect for Late time point AND either: 1) No congruity effect for Early time point 2) Smaller congruity effect for Early time point Potential Outcome #3 • Cannot reject the null hypothesis • No congruity effect for phase congruent attributes and no difference in RTs between early and late lexical targets for either Phase	• Semantic association perspective is valid: a larger and more diverse set of attributes implicitly activated for ongoing events • PI's hypothesis is valid: a late time point event representation similar to completed event representation Potential Outcome #3: • The Early Phase event attributes that are implicitly activated for ongoing event representations are also relevant for Late Phase representations • Simulation perspective: There is no difference between the experiential simulations within a single ongoing event • Semantic association perspective: Similar type and volume of attributes generated for ongoing event representations regardless of time point within that event

APPENDIX B

PILOT STUDY SCREENING QUESTIONS

1. What is your age?

    [choices provided: individual years listed for 18-35; other]

    [IF OTHER: survey will end]

2. What is your gender?

    [choices provided: Male; Female]

3. What is the highest level of education you have completed?

    [choices provided: Less than High School, High School/GED, Some College, 2-year degree, 4-year degree, Master’s Degree, Doctoral Degree, Professional Degree (JD, MD)]

4. When you were learning to speak as a child, did you learn any language other than English?"

    [choices provided: Yes; No]

    [IF YES]

    Did you speak or understand more than a few phrases at home?

    [IF YES: survey will end]

5. Have you ever had any condition that affects your brain such as: a stroke, seizure, hemorrhage, brain tumor, or other type of neurological disease or condition?

    [choices provided: Yes; No]

[IF YES: survey will end]

6. Do you currently have an alcohol- or drug-related dependency?

    [choices provided: Yes; No]

    [IF YES: survey will end]

APPENDIX C

PILOT STUDY INSTRUCTIONS

In this task you will type a list of words, phrases, attributes, or attributes that you associate with a common event. Sometimes you will be asked to list words that are associated with the beginning of an event and sometimes you will be asked to list words that you would think of when an event is almost finished. There are no right or wrong answers for this task. You will be given at least 90 seconds to generate as many words or short phrases as you can for each event. When you are finished with one event, click the continue button to proceed to the next event. This task includes approximately 40 events.

For example, responses for playing a football game might include words or phrases:

Just beginning: playing a football game

1. pads
2. put on gear
3. helmet
4. kick off
5. huddle
6. first down
7. team
8. whistle
9. clean uniform

10. sit on the bench

OR

Almost finished: playing a football game

1. tired
2. sweaty
3. clock
4. leave the stadium
5. win
6. lose
7. overtime
8. dirty uniform
9. celebrate
10. disappoint
11. locker room

APPENDIX D

PILOT STUDY EXPERIMENTAL STIMULI

Survey A

1 Just beginning: playing tennis

2 Just beginning: talking on the phone

3 Just beginning: watching a TV show

4 Almost finished: lighting candles

5 Just beginning: doing laundry

6 Almost finished: taking a hike

7 Almost finished: reading a novel

8 Just beginning: playing golf

9 Just beginning: touring a museum

10 Almost finished: putting together a puzzle

11 Almost finished: swimming

12 Almost finished: sliding down a slide

13 Just beginning: playing a board game

14 Just beginning: taking an exam

15 Almost finished: swinging a baseball bat

16 Almost finished: taking a road trip

17 Almost finished: grocery shopping

18 Just beginning: reading a novel

19 Almost finished: playing a board game

20 Almost finished: mowing the lawn

21 Almost finished: playing tennis

22 Just beginning: taking a road trip

23 Almost finished: taking an exam

24 Almost finished: playing golf

25 Just beginning: running a marathon

26 Almost finished: riding a bike

27 Just beginning: going swimming

28 Almost finished: going bowling

29 Just beginning: mowing the lawn

30 Almost finished: watching a TV show

31 Almost finished: cleaning a bathroom

32 Just beginning: pulling weeds

33 Almost finished: running a marathon

34 Just beginning: sliding down a slide

35 Just beginning: pitching a tent

36 Almost finished: doing laundry

37 Just beginning: cleaning a bathroom

38 Just beginning: babysitting a child

39 Almost finished: touring a museum

40 Almost finished: talking on the phone

41 Almost finished: pulling weeds

42 Just beginning: putting together a puzzle

43 Just beginning: lighting candles

44 Just beginning: swinging a baseball bat

45 Almost finished: babysitting a child

46 Just beginning: going bowling

47 Just beginning: grocery shopping

48 Just beginning: taking a hike

49 Almost finished: pitching a tent

50 Just beginning: riding a bike

Survey B

1 Almost finished: watching a movie

2 Just beginning: attending a meeting

3 Just beginning: going running

4 Almost finished: painting a room

5 Just beginning: playing poker

6 Almost finished: kicking a ball

7 Just beginning: baking a cake

8 Just beginning: rearranging furniture

9 Almost finished: checking email

10 Just beginning: watching a movie

11 Almost finished: going out to dinner

12 Just beginning: studying for a test

13 Just beginning: walking a dog

14 Almost finished: building a fire

15 Just beginning: opening an umbrella

16 Almost finished: pouring water into glasses

17 Almost finished: rearranging furniture

18 Just beginning: planting a garden

19 Almost finished: washing dishes

20 Almost finished: playing baseball

21 Just beginning: painting a room

22 Just beginning: taking a bath

23 Just beginning: folding towels

24 Almost finished: going to the gym

25 Just beginning: going to an amusement park

26 Almost finished: attending a meeting

27 Almost finished: going running

28 Just beginning: swinging a golf club

29 Just beginning: washing dishes

30 Almost finished: folding towels

31 Just beginning: kicking a ball

32 Just beginning: going fishing

33 Almost finished: opening an umbrella

34 Almost finished: arranging flowers

35 Almost finished: taking a bath

36 Just beginning: going to the gym

37 Just beginning: building a fire

38 Almost finished: going to an amusement park

39 Just beginning: pouring water into glasses

40 Just beginning: going out to dinner

41 Almost finished: playing poker

42 Almost finished: walking a dog

43 Just beginning: playing baseball

44 Almost finished: going fishing

45 Almost finished: swinging a golf club

46 Just beginning: arranging flowers

47 Almost finished: planting a garden

48 Almost finished: studying for a test

49 Just beginning: checking email

50 Almost finished: baking a cake

APPENDIX E

PROCEDURES FOR REVIEWING PILOT STUDY RESPONSES

To illustrate the process of attribute review, examples from the attributes that were generated for 'playing a board game' are included throughout this description of procedures. First, Early and Late responses were reviewed separately for each event with the primary objective of identifying responses that were reported by at least eight (i.e. 20%) individuals. Attribute/concept responses were grouped and counted as the same concept if they were: 1) the same word in its plural form (i.e. player, players), 2) a different form of the same word (i.e. laugh, laughing, and laughter), 3) a synonym or a closely-related word as determined by Merriam-Webster Dictionary (2004) (i.e. rules, directions, instructions), or 4) the exact word that was an essential part of a phrase but did not change the meaning of the phrase (i.e. 'pick pieces' included in the count for 'pieces,' 'family time' included in the count for 'family'). Reliability judgements were completed by two raters who achieved 94% agreement and discrepancies were resolved via discussion.

Once the concepts that were mentioned by at least 20% of respondents were identified, two other factors were considered. The first was the number of times the word/concept was reported in the other condition and the second was checking that the response was included in the first three responses in the group data. If attributes were repeated more than two times in the other condition or were not reported at all in the first three responses, the concept was removed from further consideration. The number of mentions in the opposite condition was the most limiting factor, but it was also given the most weight because the research question examines the potential differences in RTs between Early and Late targets. Attributes that were equally likely in both conditions or were reported by more than 5% of participants in the opposite condition do not provide a valid test of the question. If more than one option emerged, the concept with the highest reporting percentage was selected. Exclusionary factors were if the concept reflected an emotional state (i.e. excited/excitement, happy, sad) or was part of a phrase that could not be easily consolidated into a single word/concept (i.e. set up, game night). The concepts/words for the event of playing a board game are listed below:

1) At least 20% of individuals reported the attributes:
    a. Early: competition/competitive, dice, family, friends, fun, cards, (pick) pieces, players, (read the) rules, set up, and win(ner/ning)
    b. Late: competition/competitive, family, fun, friends, put away (pieces), laugh(ing/ter), lose(r), and win(ner/ning)
2) Remaining options with no more than two mentions in the opposite condition:
    a. Early: cards (2 mentions in Late) and rules (2 mentions in Late)
    b. Late: loser (2 mentions in Early)
3) Remaining options after checking concept for inclusion in the initial three responses:
    a. Early: cards and rules
    b. Late: loser
4) "Rules" chosen with 53% of participants versus 23% reporting "cards"

APPENDIX F

RATING TASK INSTRUCTIONS

For this task you will use a 5-point scale to rate how strongly you associate a word with a particular event phase.

1 = the word is not associated or is very weakly associated with the event phase

5 = the word is very strongly associated with the event phase

Event phases:

- Before an event begins
- Just beginning or early in an event that is in progress
- Almost finished or late in an event that is in progress
- After an event is over

For example, if the event was going to the movies, you might use the rating scale as follows:

Event: Going to a movie

Word: **PARKING**

	Low/weak				High/strong
Phase: *Before* the event starts	1	2	3	4	**<u>5</u>**
Phase: *Almost finished* with the event	**<u>1</u>**	2	3	4	5
Phase: *After* the event is over	1	2	3	**<u>4</u>**	5

In this example, you park your car before entering the theater and before the movie begins so you might rate the word PARKING at a 4 or 5 for the Before phase since PARKING is strongly associated with something that happens before the event begins. For the Almost Finished phase, you might rate the word PARKING as a 1 or 2 since that word is *not* strongly associated with actually watching the movie. Finally, for After the event, you might rate PARKING as a 4 or 5 since once the movie is over you will be going to the parking lot to leave the theater.

Here are a few more examples for the same event with different words. Your job is to rate how strongly each word is associated with the event phases provided. There are three parts to each example. The first part gives the event phase, the second part states the event, and the last part is a word in all CAPS that you will rate on a 1 to 5 scale for how strongly it is associated with the event phase.

Just beginning: Going to a movie

PREVIEWS

Low/Weak

High/Strong

1 2 3 **4** 5

The PREVIEWS occur very early on when you go to the movies, so you might rate that word quite high for the Just Beginning event phase like a 4 or 5.

Almost finished: Going to a movie

PREVIEWS

Low/weak

High/Strong

1 2 3 4 5

The PREVIEWS are over by the time the event is almost finished, so you might answer with a low rating of 1 or 2 for how strongly associated the word PREVIEWS is for the Almost Finished event phase.

Before: Going to a movie

CREDITS

Low/weak

High/Strong

1 2 3 4 5

You might rate the word CREDITS with a 1 or 2 or as weakly associated with the Before event phase since they have not happened yet.

Almost finished: Going to a movie

CREDITS

Low/weak

High/Strong

1 2 3 4 **5**

CREDITS are strongly associated with being Almost Finished going to a movie, so you might choose a 4 or 5.

Now it is your turn. Please read each event phase, event or activity, and WORD. Then, rate each WORD on a 1 to 5 scale based on how strongly you associate that word with the event phase for the event listed. You may be asked to rate more than one word for a given event, so please read each ***event phase*** carefully before you indicate your rating. Thank you!

APPENDIX G

STUDY STIMULI WITH EARLY AND LATE TARGETS

	Sentence with temporal reference point	Early Target	Late Target
1	Rick had recently started/almost finished his visit at an amusement park.	LINES	TIRED
2	Jane had recently started/almost finished her babysitting job that day.	GAMES	PAID
3	Joe had recently started/almost finished riding his bike one morning.	PADS	SWEATY
4	Tim had recently started/almost finished going bowling one evening.	NAMES	WINNING
5	Alice had recently started/almost finished baking a cake one night.	EGGS	AROMA
6	Jack had recently started/almost finished eating his dinner.	HUNGRY	FULL
7	Julie had recently started/almost finished playing a board game.	RULES	LOSING
8	Mary had recently started/almost finished her grocery shopping trip.	VEGETABLES	CASHIER
9	Robert had recently started/almost finished going hiking one morning.	BOOTS	DIRTY
10	Linda had recently started/almost finished doing laundry one Saturday.	SOFTENER	FOLDING
11	Brad had recently started/almost finished mowing the lawn.	GAS	SMELL
12	Josh had recently started/almost finished touring a museum.	GUIDE	GIFT
13	Sara had recently started/almost finished doing a puzzle.	CORNER	SUCCESS
14	Josh had recently started/almost finished taking a road trip.	MAP	MEMORIES
15	Emily had recently started/almost finished sliding down a slide.	CLIMB	FAST
16	Zack had recently started/almost finished watching a television show one night.	COUCH	SUSPENSE

APPENDIX H

STUDY INSTRUCTIONS

Examiner: You are going to read some sentences one word at a time by pressing the spacebar. After each sentence you will see a series of capital letters. Your job is to decide as quickly and accurately as possible if the letters make up a real word or not. If the letters make a real word, you'll push YES [gesture to response box] and if they don't you'll press NO. Let's do some practice.

[60 practice trials in 3 blocks of 20 with appropriate feedback]

(Aschersleben, Bachmann, & Musseler, 1999)

Examiner: Great, let's move onto the real thing. Now you will read sets of about 30 sentences and do the same thing. Push YES if the letters make a real word and NO if they don't. Respond as quickly and accurately as you can. The computer will offer you a break after each set of about 30 sentences. Ready?

[Block 1 of 32 trials]

[Continue administration of additional blocks that contain a pseudorandomization of experimental, filler, and control trials]

[After all experimental blocks are administrated, continue to Baseline Task], Examiner: This time, you will see a series of XXX instead of words. After the series of XXX there will be a series of capital letters just like before. Your task is the same: to decide as quickly and accurately as you can if the letters make a real word [gesture to YES button] or if they don't [gesture to NO button].

[20 practice trials with appropriate feedback (Aschersleben, Bachmann, & Musseler, 1999)]

Examiner: Great, let's move onto the real thing. Now you will read sets of about 30 sentences and do the same thing. Push YES if the letters make a real word and NO if they don't. Respond as quickly and accurately as you can. The computer will offer you a break after each set of about 30 sentences. Ready?

[then 2 blocks of 32 trials with a pause between blocks to offer a break if needed]

APPENDIX I

LATENT SEMANTIC ANALYSIS

Latent Semantic Analysis (Laham, 1997) was completed to determine the relationship between early and late targets to each content word for all time cue sentences (i.e. LINES-recently; LINES-started; LINES-recently started; LINES-visit; LINES-amusement; LINES-park; LINES-amusement park).

Early Targets	**Average LSA for Content Words for Early Time Cue Sentences**
LINES	0.08
GAMES	0.12
EGGS	0.17
NAMES	0.16
SOFTENER	0.09
HUNGRY	0.39
VEGETABLES	0.18
BOOTS	0.29
GAS	0.08
RULES	0.20
CORNER	0.30
PADS	0.12
CLIMB	0.32
MAP	0.12
GUIDE	0.24
COUCH	0.21
Grand Mean: .19	

Note: LSA = Latent Semantic Analysis

Late Targets	**Average LSA for Content Words for Late Time Cue Sentences**
TIRED	0.31
PAID	0.19
AROMA	0.22
WINNING	0.22
FOLDING	0.22
FULL	0.39
CASHIER	0.16
DIRTY	0.26
SMELL	0.18
LOSING	0.27
SUCCESS	0.24
SWEATY	0.17
FAST	0.34
MEMORIES	0.21
GIFT	0.19
SUSPENSE	0.18
Grand Mean: .23	

Note: LSA = Latent Semantic Analysis

Early Targets LSA grand mean (SD)	.19 (.09)*
Late Targets LSA grand mean (SD)	.23 (.07)

Note: LSA = Latent Semantic Analysis; SD = standard deviation
*Significant difference at .05; $t(15) = -2.24$, $p = .04$

APPENDIX J

EVENT BY EVENT COMPARISON TABLE

To determine the influences of the individual events in the primary multilevel model for experimental trials, the RTs from each of the 16 events were compared to the average RTs across all the other events for each event phase. The results of these comparisons are reported in the table below with robust standard errors and *p* values with alpha set to .05. For a point of reference, the pooled data for both phases was also included in the table. If the coefficient is negative, then the target for that event was responded to faster than the average of the targets from all the other events.

For the Early Phase, 5 of the 16 events were significantly different from the average of all the other events. The event that had targets that were responded to faster than the average of all the other event targets were babysitting (GAMES) and touring a museum (MAP). There were three events that had targets that were responded to more slowly than the average of the other events. The events and targets that prompted slower RTs included: riding a bike (SWEATY), doing laundry (FOLDING), and taking a road trip (MEMORIES).

For the Late Phase, 8 events different significantly from the overall event average. There were 5 event targets that prompted faster responses and 3 events that were responded to more slowly in comparison to the average of the other events in the Late Phase. The events that had targets that were responded to more quickly included: babysitting (PAID), touring a museum (GIFT), hiking (DIRTY), mowing the lawn (SMELL), and sliding down a slide (FAST).

Conversely, riding a bike (SWEATY), playing a game (LOSING), and doing laundry (FOLDING), elicited slower responses.

Event	**Both Phases** Coefficient (SE)	**Early Event Phase** Coefficient (SE)	**Late Event Phase** Coefficient (SE)
1. Amusement park LINES/TIRED	-2.19 (8.86) $p = .805$	-13.86 (12.47) $p = .266$	8.75 (12.77) $p = .493$
2. Babysitting GAMES/PAID	-29.31** (7.55) $p < .001$	-27.26* (9.69) $p = .005$	-30.81* (10.29) $p = .003$
3. Riding a bike PADS/SWEATY	64.05** (11.71) $p < .001$	32.24** (12.94) $p = .013$	94.33** (15.84) $p < .001$
4. Going bowling NAMES/WINNING	-10.09 (7.51) $p = .248$	-4.54 (11.32) $p = .688$	-15.32 (10.84) $p = .158$
5. Baking a cake EGGS/AROMA	13.60 (10.56) $p = .198$	12.43 (13.28) $p = .349$	18.07 (14.23) $p = .204$
6. Eating dinner HUNGRY/FULL	-2.08 (8.34) $p = .803$	-5.21 (11.31) $p = .645$	0.82 (12.07) $p = .946$
7. Playing a game RULES/LOSING	17.73* (8.69) $p = .041$	11.13 (12.11) $p = .358$	26.18* (12.45) $p = .036$
	-21.99*	-19.54	-27.89*

8. Grocery shopping VEGETABLES/ CASHIER	(9.14) *p* = .372	(12.71) *p* = .124	(9.89) *p* = .005
9. Taking a hike BOOTS/DIRTY	-8.16 (9.14) *p* = .372	4.61 (13.77) *p* = .738	-21.88* (10.11) *p* = .030
10. Doing laundry SORTING/ FOLDING	60.74** (9.02) *p* < .001	50.04** (12.89) *p* < .001	70.08** (12.08) *p* < .001
11. Mowing the lawn GAS/SMELL	-18.60* (6.42) *p* = .004	-15.25 (11.14) *p* = .171	-21.89* (8.74) *p* = .012
12. Touring a museum GUIDE/GIFT	-26.22** (7.93) *p* < .001	-19.85* (9.59) *p* = .038	-31.84* (11.13) *p* = .004
13. Doing a puzzle CORNER/ SUCCESS	4.78 (7.35) *p* = .515	7.07 (9.64) *p* = .463	1.87 (10.48) *p* = .858
14. Taking a road trip MAP/MEMORIES	7.76 (9.43) *p* = .410	28.51* (12.59) *p* = .024	-11.72 (10.34) *p* = .257
15. Sliding down a slide CLIMB/FAST	-23.60* (7.95) *p* = .003	-18.85 (10.47) *p* = .072	-28.10* (11.43) *p* = .014
16. Watching a TV show COUCH/ SUSPENSE	-16.53 (8.59) *p* = .054	-17.89 (10.79) *p* = .097	-16.13 (11.63) *p* = .166

BIBLIOGRAPHY

Almeida, M. J., & Shapiro, S. C. (1983, May). Reasoning about the temporal structure of narrative texts. In *Proceedings of the Fifth Annual Meeting of the Cognitive Science Society* (pp. 5-9).

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *The Quarterly Journal of Experimental Psychology Section A*, *35*(3), 427-440. doi:10.1080/14640748308402480

Anderson, S. E., Matlock, T., Fausey, C. M., & Spivey, M. J. (2008). On the path to understanding the on-line processing of grammatical aspect. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 143-148). Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson, S. E., Matlock, T., & Spivey, M. (2013). Grammatical aspect and temporal distance in motion descriptions. *Frontiers in Psychology*, *4*(337), 1-9.

Aschersleben, G., Bachmann, T., & Müsseler, J. (1999). *Cognitive contributions to the perception of spatial and temporal events*. Elsevier Science.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B.,… & Treiman, R. (2007). The English Lexicon Project. Behavior Research Methods, 39, 445-459. Retrieved January 26, 2016 from http://elexicon.wustl.edu/

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577-660.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281-1289.

Barsalou, L. W., & Sewell, D. R. (1985). Contrasting the representation of scripts and categories. *Journal of Memory and Language*, *24*(6), 646-665.

Becker, R. B., Ferretti, T. R., & Madden-Lombardi, C. J. (2013). Grammatical aspect, lexical aspect, and event duration constrain the availability of events in narratives. *Cognition*, *129*(2), 212-220.

Bestgen, Y., & Vonk, W. (1995). The role of temporal segmentation markers in discourse processing. *Discourse Processes*, *19*(3), 385-406. doi:10.1080/01638539509544924

Bestgen, Y., & Vonk, W. (2000). Temporal adverbials as segmentation markers in discourse comprehension. *Journal of Memory and Language, 42*(1), 74-87. doi:10.1006/jmla.1999.2670

Blake, M. L. (2009). Inferencing processes after right hemisphere brain damage: Effects of contextual bias. *Journal of Speech, Language and Hearing Research, 52*(2), 373–384.

Blake, M. L. (2011). Cognitive Communication Disorders Associated with Right Hemisphere Brain Damage (pp. 119-168). In M.L Kimbarow (Ed.) *Cognitive Communication Disorders*. Plural Publishers.

Blake, M. L., Tompkins, C. A., Scharp, V. L., Meigh, K. M., & Wambaugh, J. (2015). Contextual Constraint Treatment for coarse coding deficit in adults with right hemisphere brain damage: Generalisation to narrative discourse comprehension. *Neuropsychological rehabilitation, 25*(1), 15-52.

Block, R. A. (Ed.). (2014). *Cognitive models of psychological time*. Psychology Press.

Brod, G., Werkle-Bergner, M., & Shing, Y. L. (2013). The influence of prior knowledge on memory: a developmental cognitive neuroscience perspective. *Frontiers in behavioral neuroscience*, *7*, 139.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1525), 1831-1840.

Carreiras, M., Carriedo, N., Alonso, M., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory & Cognition*, *25*(4), 438-446. doi:10.3758/BF03201120

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579-593.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*(4), 478-484.

Charles D., Olm C., Powers J., Ash, S., Irwin, D. J.… Groosman, A., (2013). Grammatical comprehension deficits in non-fluent/agrammatic primary progressive aphasia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *85*(3):249-256.

Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production and comprehension in aphasia: Northwestern Assessment of Verbs and Sentences (NAVS). *Aphasiology, 26*(10), 1250-1277.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Coll-Florit, M., & Gennari, S. P. (2011). Time in language: Event duration in language comprehension. *Cognitive Psychology*, *62*(1), 41-79. doi:10.1016/j.cogpsych.2010.09.002

Collinger, J. L., Vinjamuri, R., Degenhart, A. D., Weber, D. J., Sudre, G. P., Boninger, M. L., ... & Wang, W. (2014). Motor-related brain activity during action observation: a neural substrate for electrocorticographic brain-computer interfaces after spinal cord injury. *Frontiers in Integrative Neuroscience*, *8*.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems* (Vol. 2). Cambridge University Press.

Cronbach, L. J., & Furby, L. (1970). How we should measure" change": Or should we? *Psychological Bulletin*, *74*(1), 68.

Dahan, D., & Magnuson, J. S. (2006). Spoken word recognition. *Handbook of Psycholinguistics*, *2*, 249-284.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 21*(8), 355-361.

DeLong, A. J. (1981). Phenomenological space-time: toward an experiential relativity. *Science, 213*(4508), 681-683.

de Vega, M., Robertson, D. A., Glenberg, A. M., Kaschak, M. P., & Rinck, M. (2004). On doing two things at once: Temporal constraints on actions in language comprehension. *Memory & Cognition, 32*(7), 1033–1043.

Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review, 108*(4), 759.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2008). Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. *Psychonomic Bulletin & Review*, *15*(4), 750-756. doi:10.3758/PBR.15.4.750

Dixon, W. J., Brown, M. B., Engelman, L., & Jennrich, R. I. (1990). BMDP statistical software manual. Berkeley, CA: University of California Press.

Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics?. *Linguistics and Philosophy*, *9*(1), 37-61.

Dry, H. (2009). Sentence aspect and the movement of narrative time. *Text - Interdisciplinary Journal for the Study of Discourse*, *1*(3), pp. 233-240. doi:10.1515/text.1.1981.1.3.233

Duchan, J. F., Bruder, G. A., & Hewitt, L. E. (1995). Deixis in narrative. *A cognitive science perspective*. (pp.159-189). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Eddington, S. A. (1928). *The Running-down of the Universe*. Wm. Benton.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 182-196. doi:10.1037/0278-7393.33.1.182

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, *44*(4), 516-547.

Ferrill, M., Love, T., & Shapiro, L. P. (2015). The Use of Local Anticipatory Processing Cues During Sentence Comprehension In Aphasia. *Language and Communicative Disorders*, 107.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*(6), 825-850.

Fleischman, S. (1990). Tense and narrativity. Austin, TX: University of Texas Press

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, *94*(4), 427.

Freyd, J. J. (1992). *Dynamic representations guiding adaptive behavior* (pp. 309-323). Springer: Netherlands.

Friedman, W. J. (2003). Arrows of time in early childhood. *Child Development*, *74*(1), 155-167.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3-4), 455-479.

Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies, American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers, 36*, 432-443.

Gennari, S. P. (2004). Temporal references and temporal relations in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 877.

Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: Evidence for two systems in processing space→ time metaphors. *Language and Cognitive Processes, 17*(5), 537-565.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, *23*(3), 265-304.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279.

Gibbon, J., & Allan, L. (1984). *Timing and time perception*. New York Academy of Sciences; New York, NY: New York Academy of Sciences, 1984.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558-565.

Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. *Psychology of Learning and Motivation*, *43*, 93-126.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69-83. doi:10.1016/0749-596X(87)90063-5

Gibson, J. J. (1977). The theory of affordances. Hilldale, USA.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371.

Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language*, *31*(4), 507-524. doi:10.1016/0749-596X(92)90026-T

Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, *111*(2), 151-167.

Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping*, *21*(3), 191-201.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*(4), 341-370.

Hinrichs, E. (1986). Temporal anaphora in discourses of English. *Linguistics and Philosophy*, *9*(1), 63-82.

Hinrichs, E. (1986). Temporal anaphora in discourses of English. *Linguistics and Philosophy*, *9*(1), 63-82.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*(1), 47-61.

Hopper, P.J. (1979): Aspect and foregrounding in discourse. *Syntax and Semantics*, *12*, 213-241.

Huette, S., Winter, B., Matlock, T., & Spivey, M. (2012). Processing motion implied in language: eye-movement differences during aspect comprehension. *Cognitive Processing*, *13*(1), 193-197.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review*, *99*(1), 122.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, *43*(3), 508-529.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*(3), B79-B89.

Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, *23*, 115-126.

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, and Cognition*, *30*(2), 451-464.

Khalkhali, S., Wammes, J., & McRae, K. (2012). Integrating words that refer to typical sequences of events. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *66*(2), 106.

Kim, H., Millis, K. K., & Langston, W. (1999). Computerizing the reading span test. Unpublished manuscript.

Kintsch, W. & Van Dijk, T. A., (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363-394. doi:10.1037/0033-295X.85.5.363

Laham, D. (1997). Latent semantic analysis approaches to categorization. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society* (p. 979).

Lakoff, G., & Johnson, M. (1980). Conceptual metaphor in everyday language. *The Journal of Philosophy*, *77*(8), 453-486.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25,* 259-284.

Landgraf, S., Raisig, S., & van der Meer, E. (2012). Discerning Temporal Expectancy Effects in Script Processing: Evidence from Pupillary and Eye Movement Recordings. *Journal of the International Neuropsychological Society, 18*(02), 351-360.

Lehrer, A. (1990). Polysemy, conventionality, and the structure of the lexicon. *Cognitive Linguistics (includes Cognitive Linguistic Bibliography), 1*(2), 207-246.

Linderholm, T., Gernsbacher, M. A., van den Broek, P., Neninde, L., Robertson, R. R., & Sundermier, B. (2004). Suppression of story character goals during reading. *Discourse Processes, 37*(1), 67-78.

Madden, C. J., & Ferretti, T. R. (2009). Verb aspect and the mental representation of situations. In P. Li & W. Klein (Eds), *The expression of time in language*. Berlin: Mouton de Gruyter.

Madden, C. J., & Therriault, D. J. (2009). Verb aspect and perceptual simulations. *The Quarterly Journal of Experimental Psychology*, *62*(7), 1294-1303.

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, *31*(5), 663-672.

Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes*, *29*(2), 83-112. doi:10.1207/S15326950dp2902_1

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676.

Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science, 189*(4198), 226-228.

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). On the experiential link between spatial and temporal language. *Cognitive Science*, *29*(4), 655-664.

Matlock, T. (2011). The conceptual motivation of aspect. *Motivation in Grammar and the Lexicon*, *27,* 133-147.

McGlone, M. S., & Harding, J. L. (1998). Back (or forward?) to the future: The role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1211.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), B33-B42.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition, 33*(7), 1174-1184.

McRae, K., Hare, M., Ferretti, T., & Elman, J. L. (2001). Activating verbs from typical agents, patients, instruments, and locations via event schemas. In *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 617-622). Erlbaum Mahwah, NJ.

McRae, K., Ferretti, T. R., & Liane Amyote, T. R. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*(2-3), 137-176.

McRae, K., & Matsuki, K. (2013). Constraint-based models of sentence processing. *Sentence Processing*, 51-77.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, *38*(3), 283-312.

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16518-16523.

Merriam-Webster Inc. (2004). *Merriam-Webster's collegiate dictionary*. Merriam-Webster.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception*. Belknap Press.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, *14*(2), 15-28.

Morrow, D. G. (1985). Prepositions and verb aspect in narrative understanding. *Journal of Memory and Language*, *24*(4), 390-404.

Morrow, D. G. (1990). Spatial models, prepositions, and verb-aspect markers. *Discourse Processes*, *13*(4), 441-469.

Nottenburg, G., & Shoben, E. J. (1980). Scripts as linear orders. *Journal of Experimental Social Psychology*, *16*(4), 329-347.

Nuthmann, A., & Van Der Meer, E. (2005). Time's arrow and pupillary response. *Psychophysiology*, *42*(3), 306-317.

O'Brien, E. J., Duffy, S. A., & Myers, J. L. (1986). Anaphoric inference during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 346-352.

Posner, M. I. & Snyder, C. R. (1975). Attention and Cognitive Control. In Robert L. Solso (ed.), *Information Processing and Cognition*: The Loyola Symposium. Lawrence Erlbaum

Prull, M. W., Gabrieli, J. D., & Bunge, S. A. (2000). Age-related changes in memory: A cognitive neuroscience perspective. In: Craik F. I. M., Salthouse T. A., editors. *The Handbook of Aging and Cognition, 2*. Erlbaum; Mahwah, NJ: 2000. p. 91.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2003). Spatiotemporal dynamics of neural language processing: an MEG study using minimum-norm current estimates. *Neuroimage, 20*(2), 1020-1025.

Qualtrics Copyright © 2015 (Version 5976342) [Computer Software]. Provo, UT

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 210-216. doi:10.1037/a0017258

Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 608-620. doi:10.1002/wcs.133

Rapp, D. N., & Taylor, H. A. (2004). Interactive dimensions in the construction of mental representations for text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(5), 988.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology*, *55A*, 1339-1362

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., & Bloom, H. (2011). Optimal design software for multi-level and longitudinal research (Version 3.01) [Software]. Available from www.wtgrantfoundation.org.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology: Section A, 53*(4), 1061-1080.

Sanford, A. J., & Garrod, S. C. (1981).*Understanding written language: Explorations of comprehension beyond the sentence*. New York: Wiley.

Savitt, S. F. (1996). The direction of time. *The British Journal for the Philosophy of Science*, *47*(3), 347-370.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347-367.

Schneider, W., Eschman, A., and Zuccolotto, A. (2012). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.

Schramm, A. H. (1998). Aspect and causal inferences: The role of temporal coherence in a process model of inference generation in text comprehension. Unpublished doctoral dissertation, University of Minnesota

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language, 31*(4), 507-524.

Singer, J. D., & Willett, J. B. (2003). Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language, 53*(1), 125-140. doi:10.1016/j.jml.2005.02.009

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(29), 10393-10398.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.

StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In P. Eimas & J. Miller (Eds.), *Handbook in perception and cognition, Volume 11: Speech, language, and communication*. (pp. 217-262). Academic Press.

ter Meulen, A. G. B. (1995). *Representing time in natural language: The dynamic interpretation of tense and aspect*. Cambridge, MA: MIT Press

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., ... & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273-281.

Tompkins, C. A., Klepousniotou, E., Gibbs Scott, A. (2012). Nature and assessment of right hemisphere disorders. In I. Papathanasiou, P. Coppens, & C. Potagas (Eds.), Aphasia and related neurogenic communication disorders, pp. 297-343. Sudbury, MA: Jones & Bartlett.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the" internal clock". *Psychological Monographs: General and Applied*, *77*(13), 1.

Treisman, M. (1993) On the structure of the temporal sensory system. *Psychologica Belgica*. *33*(2), 271-283.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285-318.

Van Berkum, J. J., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*(2), 147-182.

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422-488.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483-488.

Weingartner, K. M., & Myers, J. L. (2013). Effects of changes in narrative time on eye movements and recognition responses. *Journal of Cognitive Psychology*, *25*(3), 283-298. doi:10.1080/20445911.2013.768232

Weiskopf, D. A. (2010). Embodied cognition and linguistic comprehension. *Studies In History and Philosophy of Science Part A*, *41*(3), 294-304.

Welke, T., Raisig, S., Nowack, K., Schaadt, G., Hagendorf, H., & van der Meer, E. (2014). Cooking from cold to hot: goal-directedness in simulation and language. *Cognitive Linguistics*, *25*(4), 559-581.

Wodlinger, B., Downey, J. E., Tyler-Kabara, E. C., Schwartz, A. B., Boninger, M. L., & Collinger, J. L. (2015). Ten-dimensional anthropomorphic arm control in a human brain−

machine interface: difficulties, solutions, and limitations. *Journal of Neural Engineering, 12*(1), 016-011.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology, 8*(1), 52-69.

Yap, F. H., Chu, P.C., Yiu, E. S., Wong, S. F., & Kwan, S. W.,…Shirai, Y. (2009). Aspectual asymmetries in the mental representation of events: Role of lexical and grammatical aspect. *Memory & Cognition*, *37*(5), 587-595. doi:10.3758/MC.37.5.587

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3-21. doi:10.1037/0033-2909.127.1.3

Zakay, D., & Block, R. A. (2004). Prospective and retrospective duration judgments: An executive-control perspective. Acta neurobiologiae experimentalis, 64(3), 319-328.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1196-1207. doi:10.1037/0278-7393.22.5.1196

Zwaan, R. A. (2003). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, *44*, 35-62.

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*(5), 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*(6), 1022-1028.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS ONE 7*(12): e51382.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*(4), 954-958.