Establishing object-state representation in language comprehension: Evidence from picture verification, eye-tracking and ERPs

Xin Kang

PhD

University of York

Psychology

June 2015

ABSTRACT

In a set of behavioral, eye tracking, and ERP experiments, this thesis explored when and how object-state representation is established, maintained, and retrieved in language comprehension. We firstly examined whether different object-state representations could be established under two contrasting linguistic contexts (e.g., no change – "*choose the ice cream*" vs. change – "*drop the ice cream*"). Our findings showed that when linguistic context was provided, the representation that matched the consequences of described events was verified faster than the one that mismatched the expected outcome. Then, we studied the time course of establishing object-state representations with the visual world paradigm. Our results suggested that: a) the difference in looks towards the depicted versions of the situationally appropriate target object (an intact vs. a dropped ice cream) often manifested at the reference to the object but not prior to it; b) eye movements were primarily driven by semantic overlap between the visual display and the described object-state representation. Moreover, we found ERP evidence that was consistent with the need to keep track of, and retrieve object-state representations from episodic memory. We conclude that object-state representations were activated and retrieved during language processing. The work reported in this thesis highlights the need to take account of dynamics of event representation to capture the interplay between general semantic knowledge about objects and the episodic knowledge introduced by the sentential context in language comprehension.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost I want to express my deepest gratitude to my supervisor Professor Gerry T.M. Altmann for his tremendous support throughout my Ph.D. journey. His enthusiasm in research was contagious and motivational for me. He provided excellent guidance, caring, patience, and a relaxing atmosphere for my growth and progress as an apprentice and researcher. He also guided me through the process of writing this thesis, cheered me up and stood by me through good times and bad.

I would like to thank Dr Gitte Joergensen. When I started my PhD, she just finished hers and worked as a postdoc with Gerry. It has been a great pleasure to work closely with her. She is like a big sister for me. She looked after me and was always there for me. Discussion with her contributed many ideas in this thesis. Many thanks go also to Dr Silvia Gennari and Dr Shireley-Ann Ruschermeryer in my Thesis Advisory Panel. It has been a challenge for me to articulate ideas, but their encouragement and insightful comments helped me to clear my thoughts. Shireley-Ann also helped with the design of the ERP experiment and ensured access to her lab for conducting the experiment.

Further thanks go to Professor Sven Mattys who kindly agreed to be my second supervisor after Gerry moved to the US in the fourth year. Many thanks go also to Psycholinguistic Research Group, Early Career Forum, the Department of Psychology, and the Wentworth Welfare Team who have contributed immensely to my professional and personal time at York. In particular, I have benefited greatly from help by Markus van Anderson for analyzing the ERP data. Thanks go also to Gina Humphreys, Myrthe Faber, Elaine Tham, Rosa Kwok, Xiaoqian Yan, Soojin Soon, Adele Goman, Emma Holms, Clare Sutherland, Lily Law, Evangeline Tsao, Ling Ge, Jiaying Yu, and many others for their friendship. Besides, love from my family has been the encouragement for me to pursue the present research.

Finally, I would also like to acknowledge the financial support of an Overseas Research Scholarship awarded by the University of York.

DECLARATION

This thesis contains original work completed by the author under the supervision of Professor Gerry T.M. Altmann. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. The data reported in chapters 2, 3, 4, and 5 have been presented at the following conferences:

Conference talks:

Kang, X., Joergensen, G., Altmann, G. (2014).Representing conflicting states of a single object requires identity binding: Evidence from the visual world paradigm. *Talk presentedat the Experimental Psychology Society Meeting,* Kent, UK.

Poster presentation:

Kang, X., & Altmann, G. (2015). Event-related brain potentials reflect processing of object-state representations in language comprehension. *Poster Presentation at the Seventh Annual Meeting of the Society for the Neurobiology of Language,* Chicago, US.

Kang, X., Joergensen, G., Altmann, G. (2014).Representing multiple instantiations of an object: Effects of visual and linguistic context on real-time event processing. *Poster Presentation at CUNY Annual Sentence Processing Conference*, Columbus, Ohio, US.

Kang, X., Joergensen, G., Altmann, G. (2013). From altered states to altered eye movements: Object state-change in the visual world. *Poster Presentation at the Architecture and Mechanism of Language Processing (AMLaP)*, Marseille, France.

Kang, X., Altmann, G. (2012). Penny or egg: Accessibility of object representations in the visual world. *Poster Presentation at the Public Engagement Conference*, York, UK.

CHAPTER 1

Introduction

1.1 Overview

In this thesis, we shall explore how the language comprehension system keeps track of the changes that are entailed by the events described by the unfolding language. Specifically, we shall focus on the representation and selection of representations of object state, during (and following) descriptions of events in which objects experience changes of state. Such changes are often entailed by action verbs (e.g., break the window). Nonetheless, the consequence of such action is not necessarily determined by the lexical semantics of the verb alone, but by semantic knowledge associated with the objects acted upon (e.g. stomp on the penny vs. stomp the egg). Here, we shall explore how and when the comprehension system keeps track of the 'before' and 'after' of event representations.

In this chapter, theoretical background and foundational studies on the construction of mental representations in language comprehension are reviewed. We firstly review theories from philosophy, lexical semantics, and situation models to provide a theoretical foundation for event representation in language. Then, we discuss how events are linked and whether perceptual properties of objects are involved in event representations. Finally, we introduce the studies in the thesis and the structure of the following chapters.

1.2 Conceptualization of events in language

In philosophy, there is little consensus on the ontology of events, but entities within a specific spatiotemporal framework are considered to be essential for identifying event structure. Immanuel Kant (1781/1961) proposed that our ability to conceive of and understand events comes from experience, but this is also structured in some basic primitive elements, such as space, time, and causality. Davidson (1980) further stated that spatial and temporal dimensions individuate events, while causality helps to determine how events relate to each other. Barwise and Perry (1981) distinguished between two types of events: states-of-affairs and courses-of-events. A state-of-affairs is not dynamic and exists on a single spatiotemporal location, which is similar to a snapshot. A course-of-events, however, unfolds over time and space, which is linked by the uniformity of entities.

Despite the differences in definition, we can see from the above philosophical views that events are often associated with time and space, which are known to provide a spatial and temporal framework for events that occur within that space (e.g. a room) and time (e.g. a day) (Wyer & Radvansky, 1999). These ideas form the foundation for theories of event structure and event representation.

Following this line of theoretical framework, theories of mental models (Johnson-Laird, 1983) and situation models (van Dijk & Kintsch, 1983) propose that spatiotemporal and causal relations are also involved in language comprehension (Zwaan & Radvansky, 1998). An individual event representation is a building block of an integrated situation model. Readers keep track of the relations between events on different indexes, such as space, time, and protagonists. A new integrated model will be constructed when an event index that links the current model and a future model needs to be updated (Zwaan, Radvasky, & Witten, 2002). In this way, we are able to interpret continuous description of events that may be conveyed in an order and manner that is different from our everyday experience. We are also able to understand events that are not necessarily restricted to objects and people within our proximity, but to refer to any event in the past, in the future, or even in an imaginary world which is not limited by our existence in the current space and time.

By contrast, lexical semantics theories propose that lexical entries capture the underlying cognitive structure of events (Rappaport Hovav & Levin, 1998). Verbs contain not only phonological, orthographic, and semantic information, but also a conceptual structure of events that is projected to the syntactic structure of utterances (Goldberg, 1995; Jackendoff, 1990; Levin, 1993; Slobin. 1991, 1996; Talmy, 1975). Learning verbs to describe events requires abstracting the action. Verbs like *break* and *crack* indicate a specific type of damage resulting from a forceful impact and have the same event structure. As the endpoint of such events is often lexicalized as part of verb meaning, the consequences of the action cannot be changed in subsequent linguistic context.

In the following two subsections, empirical evidence of how events are represented in language comprehension from the lexical semantics and situation models perspectives is provided. Then, the potential role of objects in forming consistent memory and language comprehension is discussed.

1.2.1 Events and verb semantics

According to Vendler (1967), verbs are the means by which language describes events. There are four types of events, that are conceptualized in verbs and verb phrases, including states (e.g*., have, believe*), activities (e.g., *swim, run*), accomplishments (e.g., *paint a picture, deliver a parcel*), and achievements (e.g*., reach, find*). Events that are described by state and activities verbs have no predicted end point. By contrast, events that are described by accomplishments and achievements describe a change of state that extends over a period of time.

Moreover, verbs may also lexicalize the perspective of events differently and influence event perception. Papafragou, Hulbert, Trueswell (2008) asked participants to view motion events while preparing for verbal description or memorizing the events. They recruited native speakers of Greek who typically describe motion events with verbs about the path of events (e.g., approach) and native speakers of English who usually use verbs about the manner of events (e.g., slide). They found significant cross-language differences: Participants focused on the event components typically encoded in their native language in the verb description task but not in the memory task. The results indicate that speakers' native language influences how events are perceived in the visual scene.

Nevertheless, associating events with verb semantics does not always give rise to the appropriate event representation, because sometimes the outcome of events is not determined by a single verb in isolation but derived from the entire verb phrase, or even the entire utterance or discourse. For example, *stomp on*, may or may not indicate damage to the objects in (1).

(1) a. Joe will stomp on the *penny*. – no change
b. Joe will stomp on the *egg*. – change

The propositional representation of the *‘stomp on’* action for the two sentences would be the same, despite the nouns indicating two distinctive consequences. In this case, lexical semantics theories tell us little about how we integrate the meanings of words into an event representation for the utterance.

In addition, lexical semantics theories assume that there is a one-to-one mapping between a word entry and its conceptual meaning, but in fact, properties of entities can be changed as language unfolds. For example, if you hear or read “*The girl will stomp on the egg. And then, she will look down at the egg”*, for the same noun phrase ‘the egg’ you will know its first reference (possibly an intact egg) is different in its physical state from the second reference (possibly a broken egg) due to the ‘stomp on’ action. Thus, a static representation of the event based on verb semantics alone cannot capture the dynamic change of ‘the egg’ form before to after the action.

1.2.2 Events and situation models

An alternative account of event conceptualization, however, suggests that language could be seen as a set of instructions to construct a mental representation of the described situation rather than the text itself. Mental Models (Johnson-Laird, 1983) and Situation Models (van Dijk & Kintsch, 1983) assume that representations derived from language and perceptual-motor experiences share most of their properties. When we construct a situation model, we bring together our experiential knowledge with linguistic descriptions (Zwaan & Radvansky, 1998). As a ‘penny’ and an ‘egg’ differ in their experienced physical properties (eggs are fragile but pennies are not), we predict different outcomes for the entities.

In narrative comprehension, described situations have shown to have effects on how readers remember events. For example, Bransford, Barclay, and Franks (1972) had participants listen to sentences such as (2a) and (2b) and then (2c) and (2d) for a recognition test.

(2) a. Three turtles rested *on* a floating log, and a fish swam beneath *them*.

b. Three turtles rested *beside* a floating log, and a fish swam beneath *them*.

c. Three turtles rested *on* a floating log, and a fish swam beneath *it*.

d. Three turtles rested *beside* a floating log, and a fish swam beneath *it*.

Essentially, (2a) and (2c) describe the same situation in which three turtles resting on a log with a fish swimming beneath them (and the log). (2b) and (2d) describe different situations because the fish was under the turtles in (2b) and under the log in (2d). Results showed that participants who had heard (2a) frequently mistook (2c) as the answer, while people who had heard (2b) rarely confused it with (2d). These findings showed that we understand narratives by constructing representations of events described in the text rather than simply remembering the text itself.

Situation models are considered as dynamic representations rather than static snapshots of events. A situation model is said to encode features of the current event across multiple dimensions, such as location (e.g., Glenberg, Meyer, & Lindem, 1987, Radvansky, 2005; Radvansky & Copeland, 2006; Radvansky & Copeland, 2010), time (Radvansky, Zwaan, Federico, & Franklin, 1998; Speer & Zacks, 2005; Zwaan, 1996), goals, agents, and objects. As Zwaan, Radvasky, & Whitten (2002) pointed out, readers track the relations between events on at least five indexes: time, space, causality, intentionality, and protagonist. Therefore, they construct, update, and retrieve the situation models based on these dimensions. When a change happens in any dimension, the current model will be updated to integrate the most recent information and deactivate irrelevant information.

In line with situation models, Zacks, et al. (2007) proposed the Event Segmentation Models that changes in the environment would introduce a transient period of event model updating and resettling in working memory, known as an event boundary. Encountering this event boundary, such as temporal (e.g., before, an hour later) or location shifts (e.g., from one room to another) introduces a need to update one's current situation modes, and this updating process is effortful. Sentences describing this boundary were thus read more slowly than other clauses in narrative comprehension (Speer, Zacks, & Reynolds, 2007), but objects that are associated with the event boundary becomes more prominent than nonbounary objects (Swallow, Zacks, & Abrams, 2009). Previously encountered objects, however, were less accessible after crossing event boundaries. Retrieval interference, which occurs when there are multiple models in memory and only a single model is required to be retrieved, is attributed to this decreased accessibility (Radvansky, 2012).

For example, Glenberg, Meyer, and Lindem (1987) showed that changes of the location influence the accessibility of entities in language comprehension. In this study, participants were asked to read a story and judge if an object (e.g., *sweatshirt*) had been mentioned earlier in the story. They found that participants had shorter response time when the target object was spatially associated with the agent (e.g., *he put on his sweatshirt and went jogging*) than dissociated (e.g., *he took off his sweatshirt and went jogging*). They argued that when a target object was detached from an agent in a new location, it was deactivated from the situation model.

Similarly, a temporal change encountered during reading is also shown to influence the accessibility of information prior to the time shift. Zwaan (1996) asked participants to read a sentence that either included a short time shift (e.g., *Teresa walked onto the stage. A moment later, she collapsed.*), or a long time shift (e.g., *Teresa walked onto the stage. An hour later, she collapsed.*). Results showed that information from the previously narrated event was less accessible when it was followed by "*an hour later*" than "*a moment later*", suggesting a long narrative time shift created a large temporal distance between events. Similarly, Speer & Zacks (2005) asked participants to read narratives that contained continuous description of a single protagonist's everyday activity. Either "*a moment later*…" or "*an hour later*" was used to indicate a change in narrative time. They found that when a temporal change is encountered during reading, it is perceived as event boundaries and readers are slower and less able to accurately retrieve prior information.

There is limited evidence on whether this kind of shift of object states would influence language comprehension in a similar way as the spatial or temporal shifts, as most empirical evidence is based on spatial or temporal shifts alone (e.g., Glenberg, et al., 1987; Radvansky, 2009; Radvansky & Copeland, 2006, 2012; Zwaan, 1996) or the combination of multiple dimensions (Zacks, Speer, and Reynolds, 2009). Nonetheless, events occur when objects change or interact (Miller & Johnson-Lair, 1976). Objects are also concrete and can be re-identified, but a given event can only be experienced once. For example, a window is not an event, but a window being broken is. Due to such events, the window can be shifted to a new form of existence in time (e.g., break the window). In the next section, we review empirical studies that have shown that perceptual features of objects are activated in event representation. We propose that object properties are likely to be updated in current situation models as language unfolds.

1.3 Perceptual properties of objects are activated in language comprehension

Associated with each object are its perceptual properties, such as colour, size, orientation, and motion direction. In the real world, we rely on these features to determine an object's identity or category if these features are diagnostic for the decisions (Lee and Chun, 2001; Palmeri and Tarr, 2008). Barsalou (1999) argued that even these features are schematic in the brain, and that perceptual simulations are involved in human cognitive processes, including language comprehension. In line with this idea, previous have shown that perceptual features can be activated by object names (Allopenna et al., 1998; Cooper, 1974), achromatic line drawings of objects (Martin et al., 1995), semantically related words (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2005, 2007; Yee & Sedivy, 2006), and linguistic context (Ferguson, Tresh, & Leblond, 2013; Stanfied &Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002; Zwaan, Madden, Yaxley, & Aveyard, 2004; Zwaan & Taylor, 2006).

Moreover, the activation of object representation is shown to be context dependent. Words are usually not encountered in isolation but in the context of other words. For example, the word "chair" has at least two distinct unrelated meanings, one referring to a piece of furniture for one person to sit on (e.g., "*He sank back into his chair*") and the other to the position of being in charge of a meeting or committee (e.g., “*Address your questions to the chair, please*”). A word like “fish” has multiple related meanings. It may refer to an animal that lives in the water (e.g., “*I caught three huge fish this morning*”) or the action of trying to catch the animal (e.g., “*My grandpa really loves to fish*”).

Importantly, there are even subtler distinction of meanings. Stanfield and Zwaan (2001) provided evidence on the activation of an object’s different orientations depending on the linguistic context. In this study, they used the sentence-picture verification task. Participants were asked to read sentences like "*He hammered the nail into the wall*" (indicating a horizontal orientation) and "*He hammered the nail into the floor*" (indicating a vertical orientation) and to judge if a nail in the subsequently presented picture (either being horizontal or vertical) was mentioned in the sentence. They found that participants reacted faster to the probe picture when the orientation in the picture matched with the description than when they mismatched.

Using the same task, Zwaan, Stanfield, and Yaxley (2002) illustrated that the shape of the described objects could be activated in language comprehension. In this study, participants read sentences like "*The eagle was in the nest*" (implying an eagle with folded wings) or "*The eagle was in the sky*" (implying an eagle with outstretched wings) and verified whether the item depicted in a subsequent picture was mentioned in the sentences. Consistent with Stanfied and Zwaan (2001), participants were faster to respond to the matched picture than the mismatched one.

Related, Zwaan, Madden, Yaxley, and Aveyard (2004) showed that the direction of motion of objects could also be activated. In this study, participants listened to recorded sentences that described the motion of an object toward or away from the participants (e.g., *The pitcher hurled the softball to you*). Then, two pictures of the object (e.g., the ball) were presented with the second one smaller or bigger than the first one, indicating motion direction of the ball. Participants were required to judge whether the two objects were the same or not. Results showed that they responded faster when the depicted objects matched the implied motion direction of the recorded sentences.

Perhaps, the activated perceptual properties were part of an integrated representation of the object. As Lee and Chun (2001) noted, a single feature of an object is not sufficient to characterize it. Instead, as each object is comprised of multiple features (e.g., a long, red, vertical line), an integrated aggregation of features is likely to be involved in object representation rather than individual features. Palmeri and Tarr (2008) further argued that we know more about objects than just their individual features like shape, colour, or size. We rely on these features to determine an object's identity or category if these features are diagnostic for the decisions.

However, these studies assume that the object would remain unchanged in its location and physical state. The fact that objects around us often change locations and properties has been largely neglected in the language comprehension literature. The study by Altmann and Kamide (2009) demonstrated that when an object is associated with multiple locations (before and after being moved), our attention can be directed to its future location despite its current location on a concurrently presented visual scene. The results suggest that an object can be associated with different locations in memory and these locations compete in subsequent language comprehension. Recently, Kukona, Altmann and Kamide (2014) provide further evidence on the existence of multiple representations of the same object as it is described to move locations. When a location is not referred to in the linguistic context, we are able to retrieve it due to its association with the target object, which may form an important link in the temporal-spatial conceptual framework. These findings are consistent with abstract representations of objects in visual memory, with location-based perceptual representations of objects serving as episodic memory traces (Hoover & Richardson, 2008).

Nonetheless, objects may have distinct representations at the same location when it is described to experience a substantial change of state. For example, when you hear "*The teenager smashed the back window of his neighbour's car*", you would understand window would be most likely in pieces after being smashed, but was intact before this event. Location information is not particularly helpful in this case, as the window may be either intact or broken in the same location. In a more complex scenario, you may buy a new vase from IKEA and take it home. You place it on the table. The next day, your cat accidently knocks it over and breaks the vase. In this scenario, the vase is associated with multiple locations and physical states across time. It also interacts with at least two different protagonists, you and your cat. When you tell this to a friend, how do they achieve successful understanding of this story?

Recent work has demonstrated that the change of object state is likely to activate multiple states of the same object and these representations may compete during language comprehension. Using functional magnetic resonance imaging (fMRI) technique, Hindy, Altmann, Kalenik, and Thompson-Schill (2012) conduted two experiments with two sets of stimuli. In the first experiment, the degree of change was manipulated by using different verbs (e.g., *The woman* *will drop/choose the ice cream. And then, she will look at the ice cream*), but by using different nouns in the second experiment (e.g., *The girl will stomp on the egg/penny. And then, she will look down at the egg/penny*). In both experiments, they revealed that there was a correlation between the degree of change (that was entailed by the description of what happened to the object) and BOLD responses in the left ventral lateral prefrontal cortex (that is sensitive to word-colour interference in the stroop task). This pattern suggests competitive effects between representations of the same object increase with the dissimilarity between object states before and after change. Understanding of the described scenario thus comes at a cost of resolving this competition between object representations.

This is consistent with theories of event boundaries (Radvansky, 2012; Zacks, et al., 2007) that retrieval interference occurs when a single model is required to be retrieved after crossing event boundaries due to temporal shifts (e.g., *but first/and then*). Thus, to understand stories involving description of changes, we may need to overcome retrieval interference of competing situation models. Yet, due to the less than ideal temporal resolution of fMRI technique, it remains unknown at what point during language processing the interference arose and solved. For example, on reading "*The woman dropped the ice cream. And then, she looked at the ice cream sadly*", when do we change the representation of the ice cream she was perhaps holding, to a representation of the '*ice cream*' dropped onto the floor or other surface? When we subsequently read that '*she looked at the ice cream'*, do we at that point construct the appropriate representation of the ice cream (dropped onto a surface) or was that representation already constructed by the time we finished reading the first sentence? In other words, when do we establish and perhaps shift between the one representation and the other? If we did not construct it then, how would we know at the end of the second sentence which is the correct "interpretation" of the ice cream that is intended by the writer?

In an attempt to identify at what point in sentence comprehension this conflict may reveal, Kalenik (2012) asked participants to read these sentences in a self-paced reading task by using the same set of stimuli in Hindy et al. (2012). The hypothesis was that conflict resolution would be indicated by a slower reading times for a segment of the sentence. Results showed that reading times of the object name was longer when the object was described to experience a substantial change of state than no change. In Experiment 1, this difference was revealed at the second mention of the object (e.g., *The woman will choose/drop* <u>*the ice cream*</u>*. And then, she will look at the ice cream*), but in Experiment 2 the difference was at the first mention of the object (e.g., *The girl will stomp on* <u>*the penny/egg*</u>*. And then, she will look down at the penny/egg*). The results have demonstrated that reading times of subsequent segment was likely to be influenced by the degree of change of the target subject, but there was no consistent evidence on when the influence was manifested. This discrepancy may come from the difference in experimental stimuli in the two experiments. Perhaps, using more temporal sensitive measures, such as eye tracking and evet related potentials (ERP) method would provide a clearer picture of the time course in language comprehension.

In sum, existing literature on sentence comprehension suggests that perceptual properties of objects are activated in language processing. These properties may be linked to a particular location or physical state. When this location or state changes, multiple representations of a target object can be activated. However, it is not clear at what point in language processing object representations are updated when temporal shifts are involved in subsequent language comprehension. In the next section, the current research is introduced, which attempts to explore the behavioural and neurophysiological evidence for keeping track of object-state representations in language comprehension as the target object experiences a change of physical properties brought about by external forces.

1.4 Current research

So far, there is limited evidence on whether object-state representations can be activated when an object is described as experiencing a change of state while the location remains the same. The time course of the activation of object-state representations and behavioural correlates of such sensitivity remains unknown. If there are indeed multiple representations of object-state, we shall find an object is associated with different states at different times relative to the narrated events, can be activated and switched between during online language comprehension, especially how temporal shifts (but first/and then) influence the retrieval of an object representation.

Specifically, we ask three research questions. First, are we aware of the different states of the same object as the language describing a state-changing event unfolds? Second, what is the time course with which we keep track of the states of the object? That is, at what time during language comprehension do we know which state is intended to retrieve after encountering temporal shifts (but first/and then). Third, what is the cognitive mechanism that enables us to keep track of object-state representation?

The paradigms that we use in this research include the picture verification task (Stanfield & Zwaan, 2001), eye-tracking (Tanenhaus et al., 1995; Altamnn & Kamide, 1999), and event-related potentials (Luck, 2005), which will be introduced in more details in the following chapters. With the picture verification task, Experiments 1, 2, and 3 explore whether are we are sensitive to the states of the described target object. With eye-tracking measures, Experiments 4 to 8 examine the time course of keeping track of object-state representations in real time language comprehension in the visual world. Moreover, Experiment 9 explores ERPs evidence of activating object-state representations as the object was described to experience a change of state event.

1.5 Outline of the thesis

Chapter 2 explores whether object-state representations can be constructed in language comprehension. The picture verification paradigm is used to explore whether match/mismatch effect can be replicated when object features is manipulated by the degree of change. Chapter 3 examines the time course of constructing and updating object representations by using the visual world paradigm of eye-tracking. Chapter 4 focuses on the underlying reason for eye movements towards the appropriate object-state in language comprehension. Chapter 5 provides ERP evidence for keeping track of object-state representation during the comprehension process. Chapter 6 summarizes the findings and discusses methodological implications and theoretical contributions.

CHAPTER 2

What does “the ice cream” look like? Conflicting object-state representations in language comprehension

2.1 Overview

The first three experiments reported in this thesis investigate whether detailed mental representations of objects can be activated in different linguistic contexts, and whether these established representations could be modified over the course of a sentence as new information is provided. For instance, we may activate a representation of what the ice cream would be like when we hear “*The woman will drop the ice cream”* (e.g., a dropped ice cream), but this representation may be modified if its original intact state was implied in the subsequent sentence “*But first, she will look at the ice cream*” (e.g., an upright ice cream).

The idea that language comprehension involves activating detailed mental representations (see Zwaan & Radvansky, 1998 for a review) has been supported by empirical evidence. For example, Stanfield and Zwaan (2001) asked participants to read sentences like “*The carpenter hammered the nail into the wall/floor*” and verified whether a pictured nail was mentioned in the sentence. Critically, the nail in the picture either matched or mismatched the implied orientation. Although the nail’s orientation was irrelevant to the task, participants reacted faster to the pictured nail (e.g., a vertically oriented nail) that was compatible with its implied orientation (“*The carpenter hammered the* *nail into the floor*") than the picture showing a nail with an incompatible orientation ("*The carpenter hammered the nail into the floor*"). Using a similar task, subsequent studies found that semantic other perceptual properties of objects can also be activated, including shape (Zwaan et al., 2002), motion direction (Zwaan, et al., 2004), and visibility (Yaxley & Zwaan, 2006) of objects that were implied in the linguistic context. Response latency is a particularly important measure in these studies because response accuracy is usually high. Recently, Zwaan and Pecher (2012) replicated the above-mentioned lab-based results via Amazon's Mechanical Turk, showing shorter verification times for objects with perceptual properties that matched the linguistic description than mismatched. These studies demonstrated that language comprehension involves activation of contextually appropriate detailed mental representation of described objects.

Nevertheless, target objects in Zwaan and colleagues' studies (e.g., *the nail*) were often associated with a particular location (e.g., *on the wall/floor*) that encoded spatial properties of events (e.g., Morrow & Clark, 1989; Radvansky & Copeland, 2006, 2010). Thus, the target object may be activated in one location but may not be accessible in another location. For example, a study by Glenberg et al. (1987) ask participants read short stories with the target object either spatially associated (e.g., "*after doing a few warm-up exercises, John put on his sweatshirt and went jogging*") or disassociated with the protagonist ("*after doing a few warm-up exercises John took off his sweatshirt and went jogging*"). Then they were asked to make a recognition response to the target word ("*sweatshirt*"). They found that it took longer for participants to respond to the target word after reading the spatially dissociated sentences, compared with the spatially associated sentences. This shows that spatial location shifts influenced the accessibility of the target object.

However, in a different scenario, the target object may stay at the same location, but it may be physically changed. For example, the nail may not be successfully pounded into the wall but bent half way. The sweatshirt may be torn by a tree branch when the protagonist is running across a forest. In other words, the same location would be possibly associated with different physical states of the target object. In this case, do we construct contextually appropriate representations for this object before and after being changed? Another issue that has not been explored in previous studies is whether these representations can be retrieved in subsequent sentences and whether backward and forward temporal shifts (but first/and then) influence the retrieval process differently. In the remainder of this chapter, and indeed this thesis, we explore these questions further.

In this chapter, we aim to examine whether object representation can be activated, maintained, and retrieved when location information is not provided. Data were collected via Mechanical Turk with the sentence-picture verification paradigm that has been used by Zwaan and colleagues. On each trial, participants read a word (e.g., *ice cream*) or a sentence and then indicate whether a subsequent picture is mentioned in the text. By using pictures as probe, what the object name refers to is depicted visually and might be closer to the nature of our underlying representation of this object. In this way, any observed effects should not be attributed to the imageability of the object, but the construction of mental representations. The durations that it took participants to read the text and respond to the picture were both measured. Accuracy rates of responding the probe pictures were also calculated.

Typically, people are faster and more accurately to make judgments when the visually depicted picture is consistent with linguistically described information (e.g., Stanfield & Zwaan, 2001; Zwaan, Madden, Yaxley, & Aveyard, 2004; Zwaan, Stanfield, and Yaxley, 2002; Yaxley & Zwaan, 2006). Therefore, we predicted that participants would react faster to a probe picture when it matched the described state of the target object than it mismatched the state (e.g., shorter reaction times to a flatten ice cream when they read "*drop the ice cream*" than "*choose the ice cream*").

In all three experiments, experimental stimuli were adapted from sentences used in the Hindy et al. (2012) study and modified to accommodate the requirements of the picture-verification paradigm. We used the same participant recruitment and exclusion criteria and analyzed the data with the same trimming procedure as specified in Zwaan and Pecher (2012). First, we asked participants about their native language at the end of the experiment as part of a demographic questionnaire. We excluded the data from the small number of participants who were nonnative speakers of English. We also specified that participants should not have done other experiments from our lab before. If there were someone who had taken part in more than one experiment, we would exclude him or her from data analysis. Second, we eliminated the last-run participants of the longer list to create equal-length lists by following Zwaan & Pecher (2012) to ensure that each item's mean is calculated over the same number of observations. Third, all RTs that were from incorrect answers were removed. Besides, RTs that were faster than 300 ms and slower than

3000 ms were excluded. Then, median response times (RTs) of picture probes were calculated across subjects and items. Average RTs were thus calculated based on these median scores.

In Experiment 1, a set of objects (e.g., *ice cream*) were presented in words and were visually depicted in a subsequently presented probe picture. In each trial, participants read an object name and judged whether a probe picture showed the correct object. Each object was depicted in an intact or changed form, corresponding to the object which had undergone a 'minimal/no' or a 'substantial' changes of state, respectively. The intact state (e.g., an upright ice cream) was showing what an object would naturally occur in, while the changed state (e.g., a dropped ice cream) appeared to be altered from the intact state by external forces. In this way, we investigated whether the intact state or the changed state was closer to the prototype of the object.

Experiment 2 investigated whether the pattern observed in Experiment 1 could be modulated by the linguistic context. The stimuli in Experiment 2 consisted of single sentences that described either a substantial change of state (e.g., *The woman has dropped the ice cream*) or no change (e.g., *The woman has chosen the ice cream*). We expected to see different reaction times to the same picture depending on the sentences, such as faster reaction times to a dropped ice cream in '*drop'* than '*choose'* and vice versa, as indicated by the match/mismatch effect (e.g., Stanfield & Zwaan, 2001).

Experiment 3 extended the stimuli to two sentences (e.g., *The woman chose/dropped the ice cream. But first/And then, she looked at the ice cream*). In this way, we aimed to examine which state (if multiple states have existed) was treated as the final representation of the target object after temporal shifts (but first/and then). We also investigated whether accessibility of object representation would be influenced by the time gap between sentence reading and probe recognition. The picture probe was thus presented either after a short delay (250 ms) or a long delay (1250 ms). We expected to see match/mismatch effect that has been demonstrated in previous studies by Zwaan and colleages (e.g., Stanfield & Zwaan, 2001), e.g., shorter response times were predicted when participants responded to a dropped ice cream in "*drop, and then"* than "*choose, and then*". Besides, following Hindy et al. (2012) we expected to see the degree of change that the object underwent influencing the constructed object-state representation(s). Specifically, we predicted that a substantial change of an object would lead to competition between the conflicting object states ('before' and 'after' states). Thus, we predicted that there would be longer reaction times to the probe picture when the first-sentence contained the "*drop*" than the "*choose*". Nonetheless, the retrieval of object representation may be subject to the time that is available for constructing situation models before probe recognition.

2.2 Experiment 1

Experiment 1 intended to investigate the activation an object's representation after participants had read an object name (e.g., *ice cream*), which was measured by means of reaction times to a picture probe. At first blush, object state was not relevant as long as both pictured versions of the object could be correctly identified as what the object name referred to. Rosch (1975) suggested, however, good exemplars of a category could be identified faster than poor exemplars, because they closely match a prototype of this category.

A dropped ice cream may be less typical of ice creams than are the ones that we eat (and which are intact, either in a cone or a tub), in which case perhaps the *unusual* image of a dropped ice cream may contribute to the time it takes to process a sentence that refers to one. Thus, in Experiment 1 object names (e.g., *ice cream*) were presented without giving any context so as to get the typicality of the intact state and changed state as an object's representation.

2.2.1 Method

Participants. 118 participants were recruited via Amazon's Mechanical Turk system (54 female, mean age 36, range 19-64 years old). They received \$1.5 for their participation in the experiment, which lasted approximately 20 min[1]. There was one nonnative speaker of English in the sample. With the exclusion of this participant, the sample included 117 native speakers of English.

Materials. 32 stimuli were selected from the original 120 sets from Experiment 1 in Hindy et al (2012). Each stimulus from this set was consisted of two sentences (e.g., *The woman will choose/drop the ice cream. But first/And then, she will look at the ice cream*). For Experiment 1, only nouns in the selected stimuli were kept for the current experiment (e.g., *ice cream*). A picture that depicted either an intact state or a changed state of the named object was presented after each experimental item. As all experimental items required "yes" responses, 32 foil items that required "no" answers were added. All words referred to concrete objects such as "*ice cream*", "*acorn*", etc. All pictured objects were scaled to approximately three inches. The pictures were taken from clipart.com and were edited with Adobe Illustrator for optimal similarity in perceptual features, as shown in Figure 2.1.

[1] The median hourly rate on Mechanical Turk is \$1.38 (Horton & Chilton, 2010).

(a)

(b)

Figure 2.1 Experiment 1: Example pictures showing the intact state (left) and the changed state (right) of an ice cream. In each trial, an object name was mentioned and then followed by a picture like *a* or *b*. Participants were required to verify if the item depicted in the picture was the one that the object name referred to.

Design. A within-subjects design was used. Two lists of stimuli were created to counterbalance items and conditions. Each participant saw either an intact or changed version of the same target object after reading the object name.

Procedure. The stimuli were presented online using the Qualtrics Research Suite. Participants first completed a practice session with 4 warm-up trials. Then, 32 experimental trials and 32 foil trials were presented in a random sequence. Each trial started with a left justified and vertically centered fixation cross for 1000 ms, immediately followed by a written word, centered at the same location as the fixation cross. Participants pressed the Space bar when they had understood the word. After the key press, a fixation cross appeared for 500 ms, immediately followed by a picture. Participants were asked to judge if the item depicted in the picture had been mentioned in the text that was presented before. They responded by pressing the c-key for “no” responses and the m-key for “yes” responses. The next trial started 500 ms after the response. After completing all trials, participants answered questions about their age, gender, and native languages.

2.2.2 Results

Four participants had accuracy scores <80%, which was three times less than the standard deviation from the mean response accuracy (*M*=96.63%, *SD*=1.27). Data from these participants were removed. Besides, to equate the number of participants per list by following Zwaan & Pecher (2012) to ensure that each item's mean is calculated over the same number of observations, the last-run three participants were removed so that each list had the same number of participants. Thus, data analysis involved 110 participants, with 55 persons per list. Mean verification accuracy rates and latencies are shown in Figure 2.2.

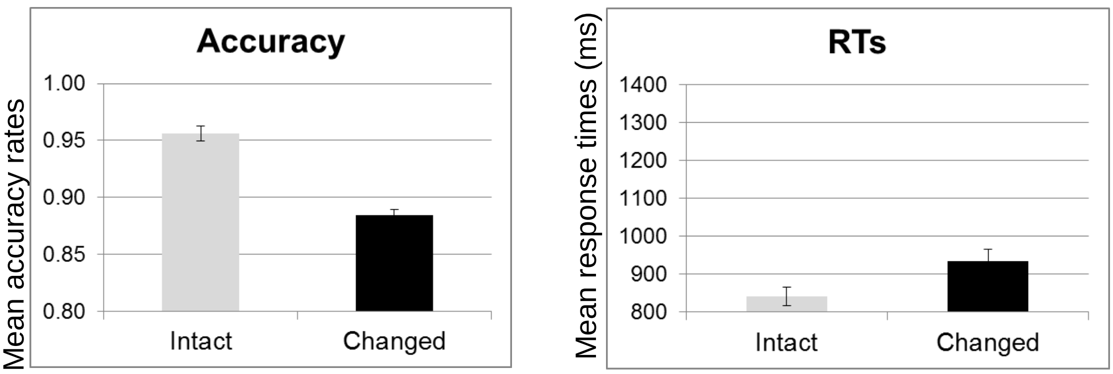


Figure 2.2 Experiment 1: Mean accuracy rates (left) and response times (right) of correct responses in the picture verification task of Experiment 1.

A paired-sample t-test showed that participants responded more accurately to the pictures depicting the intact state than the change state (significant by subjects but not by items), $t_1(109) = 8.22$, $p < .001$, $t_2(31) = 1.51$, $p = .140$. They also responded faster to the intact state than the changed state (significant both by subjects and by items), $t_1(109) = - 5.75$, $p < .001$, $t_2(31) = - 2.59$, $p = .015$. Therefore, Experiment 1 showed that verification times for the intact state was shorter than the changed state. This result suggests that the intact state of object may be the default or typical object-state representation in mind when no contextual information was provided.

2.3 Experiment 2

Experiment 2 investigated whether constructing object-state representation could be modulated by linguistic contexts. Is the representation of the *ice cream* in its intact (pre-dropped) state more, or less, accessible depending on whether one reads, *"The woman has dropped the ice cream*" or "*The woman has chosen the ice cream*"? In both cases, we presume that the initial state of the ice cream is *intact* supported by the evidence from Experiment 1 – so we are therefore asking whether mentioning a change of state event (in the "*drop*" case) makes the intact initial state (hereafter referred to as "intact state") less accessible but the changed end state (hereafter referred to as "changed state") more accessible. In this experiment, participants read sentences such as "*The woman has dropped the ice cream*", which describes a change of state that happened to an ice cream, or "*The woman has chosen the ice cream*", which describes a minimal or no change of state that occurred on it[2]. We measured how long it took participants to finish reading each of the two versions of sentences (*change vs. no change*). We also measured how accurate and how long it took participants to respond to the two versions of the picture probes (*intact state vs. changed state*).

Following Kalenik's (2012) result (see Section 2.1), we predicted that it would take longer for participants to read the sentence involving a change of state (e.g., *The woman has dropped the ice cream. And then/But first, she looked at the ice cream*) than no change (e.g., *The woman has chosen the ice cream. And then/But first, she looked at the ice cream*). Besides, by following

[2] These sentences were rated in a separate study to determine the degree of change that the object was deemed to have undergone, showing significantly higher degree of change in the '*drop*' condition than the '*choose*' condition. $t(119) = 27.63$, $p < .001$ (120 sets of stimuli, Hindy et al, 2012)

Zwaan and colleagues, we predicted that the intact state of the ice cream would be more accessible (higher accuracy rates, faster verification times) when no change was described, but the changed state would be more accessible when a change of state event was described.

2.3.1 Method

Participants. 211 participants were recruited via Amazon's Mechanical Turk system (104 female, mean age 34, range 18-69 years old). They received $1.50 for their participation in the experiment, which lasted approximately 30 minutes. There were 8 nonnative speakers of English in our sample. With the exclusion of these participants, our sample included 203 native speakers of English.

Materials. Experimental stimuli consisted of 32 pairs of sentences and corresponding picture probes. The pictures were the same in Experiment 1, showing the target object either in its intact state or changed state, as illustrated in Figure 2.1. Example sentence stimuli are shown in 1a and 1b[3].

1a) *The woman has <u>chosen</u> the ice cream.*

1b) *The woman has <u>dropped</u> the ice cream*.

The sentences described an action involving either no change (e.g., *choose*) or a change of state (e.g., *drop*) on the target object (e.g., *ice cream*). All experimental items required "yes" responses. 32 foil sentences were added. These trials were similar to the experimental trials in length, but were followed by an unrelated picture, thus requiring "no" responses.

---

[3] We also ran an experiment with the sentence stimuli in the future tense (e.g., *the woman will choose/drop the ice cream*). The results were consistent with Experiment 2. See Experiment A1 in Appendix III.

Design. A factorial design was used with 2 (Event type: No Change vs. Change) by 2 (Picture type: Intact vs. Changed). Four lists were created, including 32 sentence-picture experimental pairs in each version. Items and conditions were counterbalanced across lists. Each participant saw only one of the lists.

Procedure. The procedure was identical to that of Experiment 1. The sentence was presented left-aligned to the screen.

2.3.2 Results

4 participants seemed to confuse the response keys as they had between 0% and 3% correct responses. Another 5 participants had accuracy scores <70%, which was three times less than the standard deviation from the mean response accuracy; data from these participants were also removed. The removal of these 9 participants yielded unequal numbers of participants across lists. To equate the number of participants per list, the last-run participants of three lists were removed so that each list had the same number of participants as the shortest list (N=44). This means that the data analysis involved 176 participants.

Mean reading times of sentences, accuracy rates, and verification times for picture probes (intact state, changed state) in both event types (change, no change) were calculated across participants, as shown in Table 2.1.

Table 2.1 Experiment 2: *Mean reading times, accuracy rates and response latencies that were calculated based on median scores across subjects in Experiment 2*

Type of events	No Change		Change	
Reading time (ms)	2351		2297	
Type of pictures	Intact state	Changed state	Intact state	Changed state
Accuracy rate (%)	97	88	94	97
Response latency (ms)	952	1153	991	1065

A paired-sample t-test showed that there was no difference in reading times between the two versions of sentences, $t_1(175) = .56$, $p = .756$, $t_2(31) = 1.94$, $p = .062$. For picture verification times a two-way repeated measures ANOVA revealed no main effect of event, $F_1(1, 175) = 1.93$, $p = .167$, $\eta_p^2 = .01$, $F_2(1, 31) = 2.09$, $p = .159$, $\eta_p^2 = .06$, but a significant main effect of picture (the intact state being verified faster than the changed state), $F_1(1, 175) = 71.52$, $p < .001$, $\eta_p^2 = .29$, $F_2(1, 31) = 4.51$, $p = .042$, $\eta_p^2 = .13$, replicating the finding in Experiment 1. Importantly, the interaction between event and picture was also significant, $F_1(1, 175) = 14.06$, $p < .001$, $\eta_p^2 = .07$, $F_2(1, 31) = 11.56$, $p = .002$, $\eta_p^2 = .27$, indicating the influence of linguistic context on the construction of object-state representations. Paired-sample t-tests confirmed that the intact state was verified faster, $t_1(175) = 2.12$, $p < .001$, $t_2(31) = 2.64$, $p = .013$, and more accurately under the context of no change than change, $t_1(175) = 3.65$, $p < .001$, $t_2(31) = 2.43$, $p = .021$. By contrast, the changed state was verified faster, $t_1(175) = 3.01$, $p = .003$, $t_2(31) = 2.60$, $p = .014$, and more accurately under the context of change than no change, $t_1(175) = 10.38$, $p < .001$, $t_2(31) = 2.48$, $p = .019$.

In sum, in Experiment 2 there was no difference between reading times of sentences depending on the type of events. For picture verification, overall participants reacted faster to the intact state than the changed state. However, verifications were also influenced by the described changes. Compared with no change (e.g., *choose*), sentences involving a change of state (e.g., *drop*) lead to shorter verification times to the dropped ice cream but longer times to the intact ice cream, and vice versa. This indicates that the changed state was more accessible when changes were described to happen on the object than no change. These findings are consistent with previous findings by Zwaan and his colleagues (e.g., Stanfield & Zwaan, 2001; Zwaan & Pecher, 2012) that contextually appropriate perceptual information about described objects were activated in language comprehension. In particular, the description of a change of state event activated perceptual features of the target object in its changed state, showing a facilitation effect of events on the accessibility of object state.

However, it remains unclear in Experiment 2 whether the differential accessibility of object states would be maintained in subsequent language comprehension – e.g. if the object were to be referred to again in a subsequent sentence. To address these issues, more data were collected via Mechanical Turk, as reported in the following section. In this experiment, it aimed to investigate whether object-state representation could be maintained in subsequent sentences. It also examined whether the time gap between sentence reading and picture probe presentation would influence the pattern of picture verification times.

2.4 Experiment 3

Experiment 3 intended to further investigate the extent to which the accessibility of an object representation is modulated in linguistic context and whether the amount of time that is available for comprehending the text prior to verification influences response latencies. As suggested by Hindy et al. (2012) and Kalenik (2012), conflict of object representations may be introduced into language comprehension when the object was described to experience a substantial change. To investigate investigate whether subsequent language comprehension is influenced by the degree of change, a second sentence was added. Besides, two types of temporal shifts of events were added as event boundaries. Thus, the stimuli included two sentences. The first sentences were the same as the ones used in Experiment 2, describing an event involving either no change or a change of state on a target object (e.g., *The woman chose/dropped the ice cream*). The second sentences referred to the object again and started with either a backward temporal shift (e.g., *But first, she looked at the ice cream*) or a forward temporal shift (e.g., *And then, she looked at the ice cream*).

Moreover, picture probe was presented either after a short delay (250 ms) or a long delay (1250 ms) so as to examine whether the time available for comprehension would influence retrieval of object representations. This manipulation was motivated by findings from the negation literature of sentence comprehension. Kaup and Zwaan (2003) examined the accessibility of a colour term (e.g., *pink*) that was either presented in an affirmative situation (e.g., *Sam was wished that Laura was not wearing her pink dress*) or a negative situation (e.g., *Sam was relieved that Laura was not wearing her pink dress*). Importantly, as shown in aforementioned examples the critical clause was the same for both situations despite the difference in sentence meanings due to the main verbs ("wished" vs. "relieved"). Participants were asked to indicate whether a probe word (e.g., "pink") was mentioned or not in the sentences. When there was a short delay (500 ms) between sentence reading and probe recognition, no difference of response latencies between these two situations were found, suggesting "pink" was not accessible in both situations. However, after a long delay (1500 ms), difference in response times of the two situations were found and "pink" was more accessible in the affirmative situation than the negative situation, suggesting the influence of described situations rather than the text itself.

Subsequently, with a sentence-picture verification paradigm, Lüdtke, Friedrich, De Filipis, & Kaup (2008) asked participants to read affirmative or negative sentences followed by a matching or mismatching pictures while response latencies and event-related potentials were measured during reading and picture verification. When the delay was short (250 ms) verification latencies and ERPs evoked by the picture showed a priming effect regardless of the linguistic context as being affirmative or negative. In contrast, only when the delay was long (1500 ms), they observed main effects of linguistic context in addition to the priming effect, suggesting longer processing time was required to process more complex structure, such as negation so as to establish the situation models.

. The procedure was the same as in Experiment 1 and Experiment 2. Upon finishing reading the sentences, a probe picture was presented and participants were asked to judge if the depicted object was mentioned in the sentences that they had just read. However, unlike in Experiment 1 and Experiment 2, the probe picture was presented to participants either 250 ms or 1250 ms after they finished reading the sentences. We predicted that: a) there would be match/mismatch effects between linguistic contexts and picture probes. Linguistic contexts, including degree of change and temporal shifts would influence the construction of object representations; b) the delay between sentence reading and picture presentation would influence verification times. In the short delay (250 ms), there would be no main effect of linguistic contexts due to the limited time to process the sentence meaning. Whereas in the long delay (1250 ms), object-state representation would be modulated by language contexts as specified in a).

2.4.1 Method

Participants. 404 participants were recruited via Amazon's Mechanical Turk system (219 female, mean age 35, range 18-69 years old). They received $2.00 for their participation in the experiment, which lasted about 40 min. There were nine nonnative speakers of English and one participant who had taken part in Experiment 2. With the exclusion of these 10 participants, our sample included 394 native speakers of English.

Materials. Experimental stimuli consisted of 32 sets of linguistic stimuli adapted from Experiment 2 by adding a second sentence for each stimulus. The first sentence described a target object (e.g. *ice cream*) undergoing either no change (e.g., *choose*) or change of state (e.g., *drop*) and the second sentence referred to the objects again. Half of the stimuli began with a backward temporal shift (*but first*) and the other half with a forward temporal shift (*and then*). The pictures were the same as in Experiment 1 and Experiment 2, which either depicted a target object in its intact state or the changed state. Because all experimental items required a "yes" response, 32 foil items were added. The foil trials and the experimental trials were identical in sentence structure and the number of words in each sentence. Yet, the foil sentences were followed by an unrelated picture, and thus required a "no" response. The time delay between sentence reading and picture presentation was manipulated. The picture was shown to participants either 250 ms or 1250 ms after reading the second sentence. As for foil trials, half of them had a delay of probe pictures for 500 ms and the other half for 1750 ms.

Design. A mixed design was used in this experiment. With-subject factors included Event (*No Change vs. Change*), Temporal shift (*But first vs. And then*), and Picture (*Intact state vs. Changed state*). Half of the stimuli in the “but first” condition were followed by a picture probes in 250 ms, in one list but the same stimuli were presented in the “and then” condition in another list and followed by a picture probe 1250 ms later. Thus, within 8 lists, the time delay was treated as a between subject factor. In each list, there were 32 sentence-picture experimental pairs and 32 foil pairs.

Procedure. The procedure was similar to that of Experiments 1 and 2. The stimuli were presented online using the Qualtrics Research Suite. A trial started with a left justified and vertically centered fixation cross for 1000 ms, immediately followed by the first sentence (e.g. *The woman chose the ice cream*). Participants pressed the Space bar when they had understood the sentence. Following the key press, the second sentence appeared on the screen (e.g., *But first, she looked at the ice cream*). They pressed the Space bar again when they finish reading the second sentence and a fixations-cross appeared on the screen for either 250ms or 1250 ms before the probe pictured was presented. The foil trials were presented by following the same procedure, but the time delay between sentence reading and picture verification was either 500 ms or 1750 ms. Upon seeing the picture probe, participants verified whether the object depicted in the picture was mentioned in the sentences they had just read by pressing the c-key for “no” responses and the m-key for “yes” responses. Their answers and verification times were recorded. The next trial started 500 ms after the response.

2.4.2 Results

7 participants appeared to confuse the response keys and reversed their “yes/no” response mappings, as they had between 0% and 4% correct responses. Another 23 participants had accuracy scores <70%, which was three times less than the standard deviation from the mean response accuracy. Data from these participants were removed. The removal of them yielded unequal numbers of participants across lists. To balance the lists, 20 participants from 7 lists who were the last participants were removed to match the lists with the one containing the least number of participants (N=43). Thus, the data analysis of Experiment 3 involved 344 participants.

2.4.2.1 Sentence reading times

Table 2.2 Experiment 3: *Mean reading times that were calculated based on median scores across subjects in Experiments 3*

Type of events	No Change		Change	
S1 (ms)	1081		1092	
Temporal shifts	But first	And then	But first	And then
S2 (ms)	1387	1277	1394	1300

Table 2.2 shows average reading times for the first and second sentences. For the first sentence (S1), there was no difference between reading times, $t_1(343)=.80$, $p = .567$, $t_2(31)=.51$, $p = .613$. For the second sentence (S2), repeated measures ANOVAs showed that there was no main effect of Event on reading times, $F_1(1, 343) = 1.27$, $p = .261$, $\eta_p^2 = .004$, $F_2(1, 31) = 2.22$, $p = .146$, $\eta_p^2 = .01$ but a main effect of Temporal shift by subjects (longer reading times

for sentences with “but first” than “and then”), $F_1(1, 343) = 42.04$, $p < .001$, $\eta_p^2 = .11$, but not by items, $F_2(1, 31) = 1.84$, $p = .185$, $\eta_p^2 = .01$. No interaction between Temporal shift and Event was found, $F_1(1, 343) = .28$, $p = .594$, $\eta_p^2 = .001$, $F_2(1, 31) = 0.45$, $p = .833$, $\eta_p^2 = .001$. Thus, unlike in Kalenik (2012), no difference is reading time between the two event types was revealed. This discrepancy may come from the difference in stimuli presentation. In our experiment, participants read in a sentence-by-sentence way, but phrase-by-phrase in Kalenik (2012).

2.4.2.2 Picture verification times

Mean accuracy rates and response times for the picture verification task are shown in Table 2.3 and 2.4.

Table 2.3 Experiment 3: Mean accuracy rates and response latencies that were calculated based on median scores across subjects in Experiments 3 (250 ms)

Temporal shifts	But first				And then			
Type of events	No Change		Change		No Change		Change	
Type of pictures	Intact state	Changed state	Intact state	Changed state	Intact state	Changed state	Intact state	Changed state
Accuracy rate (%)	97	96	97	96	97	95	94	98
Response time (ms)	950	977	940	989	880	1068	929	1047

Statistical analyses were conducted separately for the short time gap (250 ms) and (1500 ms) separately. When there was a 250 ms delay between sentence reading and picture verification, a repeated-measures three-way ANOVA with Event, Picture, and Temporal shift as within-subjects factors showed that there was a main effect of the type of Picture on verification times (shorter response times for the intact state than the changed state), $F_1(1, 85) = 8.46$, $p = .005$, $\eta_p^2 = .09$, $F_2(1, 15) = 11.16$, $p = .004$, $\eta_p^2 = .43$, but there were no main effects of Temporal shift, $F_1(1, 85) = .17$, $p = .683$, $\eta_p^2 = .002$, $F_2(1, 15) = .04$, $p = .842$, $\eta_p^2 = .003$, nor Event, $F_1(1, 85) = .37$, $p = .546$, $\eta_p^2 = .004$, $F_2(1, 15) = .41$, $p = .534$, $\eta_p^2 = .03$. No interactions between the three factors were found, $F_1(1, 85) = 1.15$, $p = .286$, $\eta_p^2 = .01$, $F_2(1, 15) = .71$, $p = .413$, $\eta_p^2 = .05$. These results suggest that when the time delay was short (250 ms) participants reacted faster to the picture depicting an intact state than a changed state regardless of the degree of change and temporal shifts. This pattern has been seen in Experiment 1 when the intact state was treated as the default object-state representation. Thus, when there was a short delay, participant did not incorporate the meaning of sentences in comprehension but probably focused on the object name and hence activated the typical representation of the object.

Table 2.4 Experiment 3: Mean accuracy rates and response latencies that were calculated based on median scores across subjects in Experiments 3 (1250 ms)

Temporal shifts	But first				And then			
Type of events	No Change		Change		No Change		Change	
Type of pictures	Intact state	Changed state	Intact state	Changed state	Intact state	Changed state	Intact state	Changed state
Accuracy rate (%)	99	94	97	97	97	97	94	98
Response time (ms)	926	956	964	979	865	1065	898	969

By contrast, when there was a 1250 ms delay between sentence reading and picture verification, repeated-measures three-way ANOVAs with Event, Picture, and Temporal shift as within-subjects factors showed that there were no main effects of Temporal shift, $F_1(1, 85) = .01$, $p = .929$, $\eta_p^2 = .00$, $F_2(1, 15) = .22$, $p = .648$, $\eta_p^2 = .01$, or Event, $F_1(1, 85) = .26$, $p = .612$, $\eta_p^2 = .003$, $F_2(1, 15) = 1.94$, $p = .184$, $\eta_p^2 = .11$, but there was a marginal main effect of Picture (shorter response times for the intact state than the changed state), $F_1(1, 85) = 3.14$, $p = .080$, $\eta_p^2 = .04$, $F_2(1, 15) = 7.50$, $p = .015$, $\eta_p^2 = .33$. Besides, there was a significant interaction between Temporal shift and Event by subjects, but not by items, $F_1(1, 85) = 6.58$, $p = .012$, $\eta_p^2 = .07$, $F_2(1, 15) = 1.33$, $p = .268$, $\eta_p^2 = .08$. Besides, there was a significant three-way interaction between Picture, Temporal shift, and Event by subjects, $F_1(1, 85) = 5.68$, $p = .019$, $\eta_p^2 = .06$, but not by items analysis, $F_2(1, 15) = 2.94$, $p = .107$, $\eta_p^2 = .16$. Mean verification times and accuracy rates of picture probe after a long time delay (1250 ms) were shown in Table 2.4. Paired t-tests showed that verification times for the changed state in the context of "*drop, and then*" were significantly faster than "*choose, and then*" by subjects, $t_1(1,85) = 2.53$, $p = .010$, but not by items, $t_2(1,85) = 1.96$, $p = .069$, suggesting the changed state induced by the action in the first sentence was likely to be maintained in the second sentence. No difference of probe recognition was found for the "but first" condition.

In sum, Experiment 3 was conducted to investigate whether the accessibility of object-state representation could be retrieved and maintained in subsequent linguistic context and whether verification was influenced by the amount of time that is available for comprehending prior to picture presentation. Reading times for read sentences containing "but first" were longer than "and then". The results may reflect that greater retrieval interference and cognitive effort was required to process sentences violating the temporal order of events in real life (Münte, Schiltz, and Kutas, 1998). The lack of match/mismatch effect in the “but first” sentences suggested that longer time gap than 1250 ms may be needed to update situation models. By contrast, there was a match/mismatch effect for “and then” sentences but only when there was a long delay (1250 ms). When there was a short delay between sentence reading and probe recognition, participants were susceptible to typicality effect, showing a preference for the intact state to the changed state regardless of the type of linguistic contexts. This finding is consistent with Lüdtke et al. (2008) that a short delay between reading comprehension and picture presentation only allows for a word-picture priming effect. Instead, a longer delay is required to complete constructing situation models and retrieve object-state representation, particularly for more complex linguistic structure.

2.5 General discussion

With the picture-sentence verification paradigm, three experiments have investigated whether the accessibility of object state could be modulated by linguistic context. Results have shown that when only object names were presented, the intact state was more accessible than the changed state (Experiment 1). The difference between the accessibility suggests that the intact state may be a more typical object-state representation than the changed state. However, accessibility was also shown to be influenced by linguistic context (Experiment 2). When contextual information was available, the changed state became more accessible if the object was described to undergo a change of state event. Besides, we found that object-state representation could be maintained across a subsequent sentence, although the somewhat limited evidence for this is related to available comprehension time between the sentence and probe (Experiment 3). After a short delay between sentence reading and picture verification, there was no change in accessibility as a function of state-change until sufficient time had elapsed for a marginal change after a longer delay, showing shorter latencies for the intact state than the changed state. After a long delay, however, participants tended to respond faster to the changed state that was required to be maintain the representation after a change of state events, suggesting the accessibility was also influence by the content of the described situation in sentences.

Our findings are not consistent with the amodal proposition account for which the perceptual features of objects should not be relevant for the accessibility of object-state representation. However, semantic mapping between the verbs (e.g., *choose vs. drop*) and nouns seems to be able to account for the results. As shown by Altmann and Kamide (1999), verbs can be used to restrict the choice of appropriate objects in the visual scene. For example, when participants heard "*The boy will eat …*", they were more likely to look at something edible in the scene (e.g., cake) even before the object was mentioned. Perhaps, the verbs in the stimuli helped participants to identify the appropriate state of the object. For example, "*choose*" is usually associated with an intact object, while "*drop*" is often associated with a "dropped" object or an object that could be dropped. When a "flattened/dropped" ice cream was presented to participants, they might associate it with the "dropping" action, but not necessarily the 'choosing'. In this case, the findings in Experiment 2 may simply indicate that there was closer link between the 'dropping' event and a dropped ice cream compared with the "*choosing"* event. However, as shown in Experiment 3, the delay between sentence reading and picture presentation was crucial for the facilitation effects between a change of state event and a pictured changed state to appear. This indicates that the findings in Experiment 2 are possibly not just the result of the mapping between verb meaning and a depicted state. Instead, understanding what was described in the sentences is important for constructing an appropriate object-state representation under the linguistic contexts.

Our findings are in line with the situation model theories, which state that language comprehension involves activating situated object features (Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002, Yaxley & Zwaan, 2006). Experiment 1 showed that the intact state would be the default object representation. Experiment 2 replicated and extended the finding of Zwaan and colleagues that the match/mismatch effects could appear when object change was described. Experiment 3 further showed that after crossing a temporal shift situation models were updated when there was enough time to understand the content of the linguistic description. However, as the picture verification happened after participants finished reading the sentences, it was not clear how mental representation was constructed, maintained, and retrieved during online language comprehension. Besides, it is not clear at what point during language processing participants have decided on the “”appropriate” object representation. Perhaps, pragmatics context may have facilitated inference of consequences of actions (Niuewland & Kuperberg, 2008; Tian, Breheny, & Ferguson, 2010). Due to difference in semantic properties of the verbs (choose vs. drop) and temporal shifts (but first vs. and then) there are different expectations of the consequences. By the time before the noun phrase is mentioned, we may have already started to infer the results.

In the next three chapters, I will report findings of experiments using eye-tracking and event-related potentials paradigms, which allowed us to explore when linguistic input would influence the accessibility of objects' state during online language comprehension. Given the statistically weak results which moderate confidence in the generalizability of the interpretation, these other studies may also shed light on when and how object-state representation is established during language comprehension.

2.6 Conclusion

In this chapter, findings from three experiments with the picture-sentence verification paradigm were reported. The intact state was shown to be a typical object-state representation compared with the changed state. However, when contextual information was provided, object-state representation was influenced by the content of described situations, showing match/mismatch effect between linguistic context and object state. Besides, object-state representation was also shown to be affected by the amount of time allowing for language comprehension. When the described event has not yet been integrated into sentence interpretation, verification decisions appear to be modulated only by the typicality of object-state representation.

CHAPTER 3

Tracking object-state representation during language comprehension: Evidence from eye movements

3.1 Overview

With the picture-sentence verification paradigm, the three experiments in Chapter 2 have demonstrated that an intact state of an object is likely to be a typical object representation, but mental representation can also be context-dependent. Verification latencies tended to be shorter for a picture depicting an intact state than a changed state of the object, suggesting the intact state is a more typical representation of the object than the changed state. Importantly, information from linguistic contexts appeared to modulate object-state representation. For example, verification latencies for the changed state (e.g., *a dropped ice cream*) tended to be shorter when the object was described to have experienced a substantial change (e.g., *drop the ice cream*) than no change (e.g., *choose the ice cream*). Nevertheless, sufficient amount of time seemed to be needed so as to integrate the meaning of a described change into the discourse; otherwise verification latencies were shorter for the intact state than the changed state. Therefore, it seems that when an object is described as undergoing changes from one state to another, we may establish multiple mental representations for the object. For instance, on hearing "*The woman will drop the ice cream*", we may activate its changed state after being dropped, but we may keep its intact state on hearing "*The woman will choose the ice cream*".

However, since probe pictures were presented *after* sentence reading, it is not clear from Chapter 2 whether object-state representation can be constructed as part of the language comprehension process.

The goal of the two experiments reported in this chapter was to examine the pattern and time course of keeping track of object-state representation during language comprehension. An eye-tracking method, known as the visual world paradigm was used for data collection due to its sensitivity to online sentence processing. It has been shown that incoming spoken language cues are able to drive eye movements around a current static scene (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003) or on a blank screen (Altmann, 2004). Eye movements can be directed to where information is introduced (Altmann & Kamide, 2009; Hoover & Richardson, 2008), to an item that is anticipated to be mentioned (Altmann & Kamide, 1999; Boland, 1995; Kamide, Altmann, & Haywood, 2003; Knoeferle, Crocker, Scheepers, & Pickering, 2005), or to semantically-related items (e.g., Huettig & Altmann, 2005, 2011; Yee & Sedivy, 2006).

A typical experimental trial with the visual world paradigm involves asking participants to listen to recorded auditory stimuli while looking at visual displays that either include real objects (e.g., Tanenhaus et al., 1995) or clipart items (e.g., Altmann & Kamide, 1999) as their eye movements are recorded. Participants may also be instructed to perform a simple task (e.g. "pick up the candy") (e.g., Allopenna, et al., 1998; Tanenhaus, et al., 1995) or just "look and listen" (e.g., Altmann & Kamide, 2007). Experiments reported in this chapter and the next chapter adopted the “look and listen” version of the visual world paradigm. One advantage of using the "look and listen" task is to avoid interference of task demands. When participants are asked to complete a task, they may fixate the target objects *because* they are asked to touch or move them rather than as part of a natural process in language comprehension. Besides, the "look and listen" task does not require participants to utilize conscious response strategies that may affect how language is processed. Therefore, the "look and listen" task examines online language processing without interrupting it and provides a continuous implicit record of cognitive processes as language unfolds over time.

In this chapter and in all other following eye-tracking experiments, two measures of eye movements were reported: fixations and saccades (Henderson & Ferreira, 2004). Fixations refer to the situation when gaze position is held relatively still on a single location, while saccades refer to the situation when the eyes are moving fast from one fixation location to another. Thus, fixations reflect the extent to which an object holds attention and does so in the face of competing demands on attention, but saccades are a measure of shifts in attention (see Altmann & Kamide, 2007 for further discussion). Both measures are particularly important for the purpose of this thesis. Fixations were used to describe static patterns in the data and reflect the existing internal representations of the target items. In addition to fixations, detailed statistical analyses on proportions of saccadic eye-movements were also reported, which reflected the dynamic shift of attention from one fixation point to another, as language unfolded.

Raw measurements of our eye-tracking data are time points of fixations on a region of interests (ROI). Since onset and offset of critical words in auditory stimuli were marked, we were able to figure out whether a fixation appeared at a given time point. Thus, for each trial of linguistic stimuli and a given time window we also knew whether there was a fixation on a particular ROI. Fixations were thus calculated by sampling at specific, linguistically-motivated points in time, the proportion of trials in which participants were looking at the ROI at that moment in time. To compare across ROIs, we calculated the percentage of fixations on each region. For by-subjects analysis, we calculated the percentage of fixations for each participant by dividing the total number of trials with a fixation and the number of trials per condition. Similarly, for by-item analysis, the percentage fixations on a ROI were calculated by dividing the number of participants with a fixation and the total number of participants in each condition. Besides, we calculated the number of trials on which at least one saccade was launched at the region of interest within a particular time window. Thus, we were able to know the number of saccades on each ROIs within a certain time window as specified in the linguistic stimuli. Percentages of saccades have also been calculated both by subjects and by items just as percentages of fixations. In this way, we achieved the dynamic mapping of the unfolding language and percentage of eye movements on the visual display. Thus, we were able to compare whether two regions of interest differ in their likelihood of attracting eye movements during each of a set of critical time windows.

In Experiment 4, eye movements towards two versions of the target object were compared (See Figure 3.1). while linguistic stimuli "*The woman will choose/drop the ice cream*. *But first/And then, she will look at the ice cream"* unfolded. In the first sentence, the verbs were different between two conditions. In the Change condition, the verb indicated a substantial change of state (e.g., “*drop*”), while in the No Change condition the verb described a minimal change or even no change of state (e.g., “*choose*”). The second sentence referred to the target object again and began with either a backward temporal shift (*but first*) or a forward temporal shift (*and then*). These stimuli were selected and modified based on three criteria: the intact state and changed state of the target object can be visually presented; the Change condition was rated to indicate a greater change of state happening on the object than the No Change condition. No difference was found between imageablility and ambiguity between Change and No Change condition.

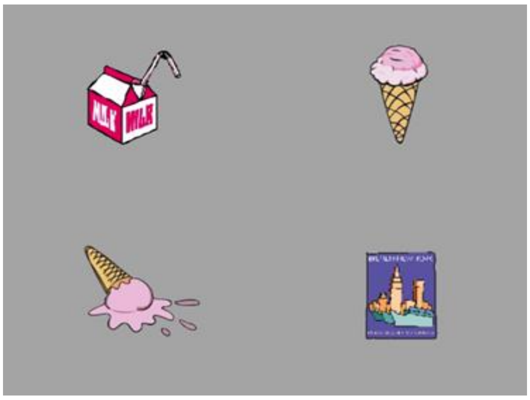


Figure 3.1 Experiment 4: Example of an experimental visual stimulus used in Experiment 4. The target object (e.g., the ice cream) was depicted in the intact state and the depicted state. Two unrelated distractors were also included. The corresponding auditory stimuli were “*The woman will choose/drop the ice cream. But first/ And then, she will look at the ice cream*”.

We hypothesized that participants would look at a particular version of the target object in the visual scene depending on the linguistic context. Based on findings from Chapter 2 and in line with the literature, we predicted that the item that matched the intended object-state representation would attract more eye movements than when they mismatched. For example, participants might prefer to look at the changed state (e.g., the dropped ice cream) when the object was described to experience a substantial change of state (the condition in which the woman was described as dropping the ice cream). By contrast, when only minimal or no change of state was mentioned (i.e. the "*choose*" condition), the changed state was not relevant to the event, and thus in this case participants might look less at it.

Experiment 5 intended to replicate the findings of Experiment 4. Auditory stimuli in Experiment 5 had a similar structure as in Experiment 4, with the first sentence describing either a substantial or a minimal change of state on an object and the second sentence referring to the object again. For example, "*The girl will stomp on the egg/penny. But first/And then, she will look down at the egg/penny"*. Unlike Experiment 4, the degree of change in Experiment 5 was manipulated by using two different target objects that experienced the same action. Whereas one object (e.g., *the egg*) would experience a substantial change of state if "stomped on", the other object (e.g., *the penny*) would not. We expected to see a similar pattern and time course of eye movements in Experiment 5 as in Experiment 4: participants would prefer to look at a particular depicted version of the object depending on the language context.

In sum, the two experiments reported in this chapter aimed to explore when a particular object-state representation was established in real-time language processing. Two conflicting object states were presented on the visual scene together with two distractors. Eye movements around the visual scene were expected to be modulated by described events.

3.2 Experiment 4

Experiment 4 aimed to investigate the time course of constructing object-state representations. The target object was described as undergoing an event involving either a change of state or no change on the object. With the "look and listen" version of the visual world paradigm, participants were presented recorded sentences via loud speakers while their eye movements around a visual scene were recorded. In this way, we were able to map participants' eye movements onto real-time language processing. We predicted that there would be differences in participants' eye movements depending on the language context. Specifically, we would see more looks towards a particular object state when it was indicated as the appropriate object-state representation. For example, there might be more looks at the dropped ice cream in "*drop*" than "*choose*".

3.2.1 Method

Participants. 64 students (42 female, mean age 21, range 18-25 years old) from the University of York participated in this study. They received either half an hour course credit or two pounds for their contribution. All participants were native speakers of British English and had normal or corrected to normal vision.

Materials. The auditory stimuli were created based on the stimuli used in Experiment 3. The first sentence of each stimulus indicated either a change of state or no change (*choose vs. drop*) on the object, while the second sentence involved either a backward temporal shift (*but first*) or a forward temporal shift (*and then*) when the object was mentioned again. In addition to the 32 sets of sentence stimuli in Experiment 3, four sets of new stimuli were added. The stimuli were all in the future tense, as shown in the examples below:

a) The woman will choose the ice cream. But first, she will look at the ice cream.

b) The woman will choose the ice cream. And then, she will look at the ice cream.

c) The woman will drop the ice cream. But first, she will look at the ice cream.

d) The woman will drop the ice cream. And then, she will look at the ice cream.

There were 36 foil items in total, which had the same sentence structure as the experimental stimuli. Half of the foil items mentioned a backward temporal shift ("*but first")* while the other half a forward temporal shift ("*and then*"). Different from experimental stimuli, foil sentences did not always refer to the same object in the first and the second sentences. One third of the foil items described a distractor on the scene in the first sentence, but referred to an object that was not depicted in the second sentence. One third referred to a depicted distractor in the second sentence, but a non-depicted object in the first sentence. Besides, one third mentioned both depicted distractors in the first and second sentences respectively. All experimental and foil sentences were recorded by a male native speaker of British English at the sampling rate of 44,100 Hz and 16-bit sound resolution in a sound proof booth. Identical words in auditory stimuli within the same set were spliced and re-used.

Figure 3.2 Experiment 4: Example of a foil visual stimulus used in Experiment 4. The corresponding auditory stimulus was "*The boy will open the book. And then, he will pick up the pen*"

36 experimental pictures (e.g., Figure 3.1) and 36 foil pictures (e.g., Figure 3.2) were created using commercially available Clipart. In each experimental picture, there were two states of the same object, including an intact state (e.g., the upright ice cream), a changed state (e.g., the inverted/dropped ice cream) and two distractors (e.g., *the milk carton and the magazine*). The locations of the four depicted items were rotated and counterbalanced across experimental stimuli.

Procedure. The experiment was conducted on an Eyelink II Head-mounted eye-tracker, which sampled at 250 Hz from the right eye, but viewing was binocular. Participants previewed the visual display for a second before hearing a description about the objects in the scene. They were instructed to look at the visual display freely while listening to the sentences. Visual stimuli were presented at 800 x 600 pixels resolution at the centre of a computer screen that was about 60 cm in front of each participant, while audio files were played via two loudspeakers placed at each side of a computer screen during the experiment. Prior to every six trials, a nine-point calibration procedure and later a validation procedure was performed. Between each trial, a single black dot was shown in the centre of the screen and was used to correct calibration drift. Participants were told when they finished half of the experiment and were encouraged to take a short break, if they would like, before carrying on. Each trial lasted about 9 seconds and the total length of the experiment was about 25 minutes.

3.2.2 Results

Data were analyzed with the following procedure: First, four interest areas were marked on the visual scene, including the intact state, the change state, and two distractors, as shown in Figure 3.3.

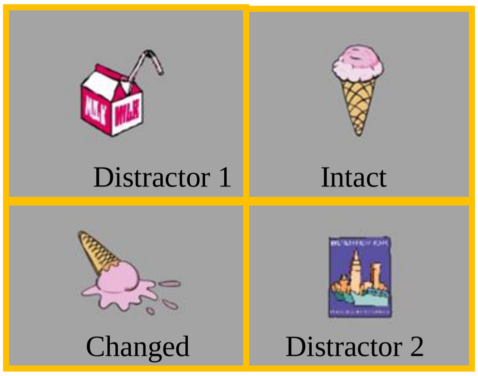


Figure 3.3 Experiment 4: Example of interest areas on visual stimuli in Experiment 4.

As the eye-tracker recorded x-y co-ordinates of participants' eye movements, this output was mapped onto the marked areas. As critical words and phrases in each sentence were marked from word onsets to offsets in the speech files, such as *"choose/drop", "the", "ice cream"*, looks to the different interest areas were thus synchronized with the auditory stimuli. In this way, we calculated both fixations on and saccades towards a particular visual region as language unfolded[4].

Subject and item means of eye movements were entered into separate statistical analyses. The data entered into the statistical analysis in this experiment and all the following eye-tracking experiments in this thesis were arcsine-transformed proportions[5]. The presentation of the result in the following

[4] See Table a2.1 in Appendix II for percentage of trials with saccades and fixations on the depicted items on the visual display when a critical word/phrase was.

[5] Identical ANOVAs were also conducted on original proportions. Arcsine transformation was done to adjust the proportion between 0 to 1. Unless noted, statistically significant effects were consistent between the transformed and untransformed data.

section is divided into two parts. First, a general description of participants' overall fixation patterns is provided, which gives an indication of eye movements as language unfolded. Second, to get a more detailed description of data, particularly how eye movements changed over time, statistical analyses of eye movements during time windows of critical words and phrases were reported, including the verb in the first sentence (e.g., "*choose/drop"*), the first reference to the target object (e.g., "*the ice cream"*), temporal shifts (*but first/and then*), the verb in the second sentence (e.g., "look at"), and the second reference to the target object (e.g., "*the ice cream*").

3.2.2.1 Fixations

Figure 3.4 illustrates eye movements over time by showing the mean proportion of fixations on the depicted items on the visual scene during the presentation of auditory stimuli. It presents an overview of eye-movements to all items on the visual scene.

Overall, there were more fixations on the two depicted states of the target object than distractors throughout the experiment. In the first sentence, at the onset of the verb there were more looks to the dropped ice cream than the upright ice cream regardless of the verbs. Then, there were steadily increasing looks towards the upright ice cream when the verb indicated a minimal change of state (e.g., "*choose*"), but the dropped ice cream when the verb indicated a substantial change of state (e.g., "*drop*"). This pattern became more obvious at the end of the first sentence when they had heard "*the woman will choose/drop the ice cream*", suggesting the influence of the described events on eye movements towards the depicted states on the visual scene.

In the second sentence, at the onset of the second sentence eye movements pattern of the first sentence was maintained. There seems to be more fixations on the upright ice cream than the dropped ice cream when no change was described in the first sentence (Figure 3.4a), but more fixations on the dropped ice cream than the upright ice cream when a change on the object was mentioned (Figure 3.4b). This pattern continued to the onset of the verb with an interaction between the described change and the depicted object state. However, eye movements in the "*drop, but first*" condition started to change after the verb onset, with increasing looks towards the upright ice cream, but decreasing looks towards the dropped ice cream. As shown in Figure 3.4b, at the offset of the second sentence, there were more fixations on the upright ice cream than the dropped ice cream, while patterns for other conditions remained more or less unchanged.

To summarize, proportions of fixations revealed that the linguistic context seemed to influence participants' eye movements on the visual scene. For instance, in the first sentence when a change of state event was described, they tended to look at the dropped ice cream, but when a backward temporal shift was mentioned, they tended to look at the upright ice cream. In the second sentence, the established object-state representation was kept unchanged untill the linguistic context indicated that a different object-state representation should be retrieved.

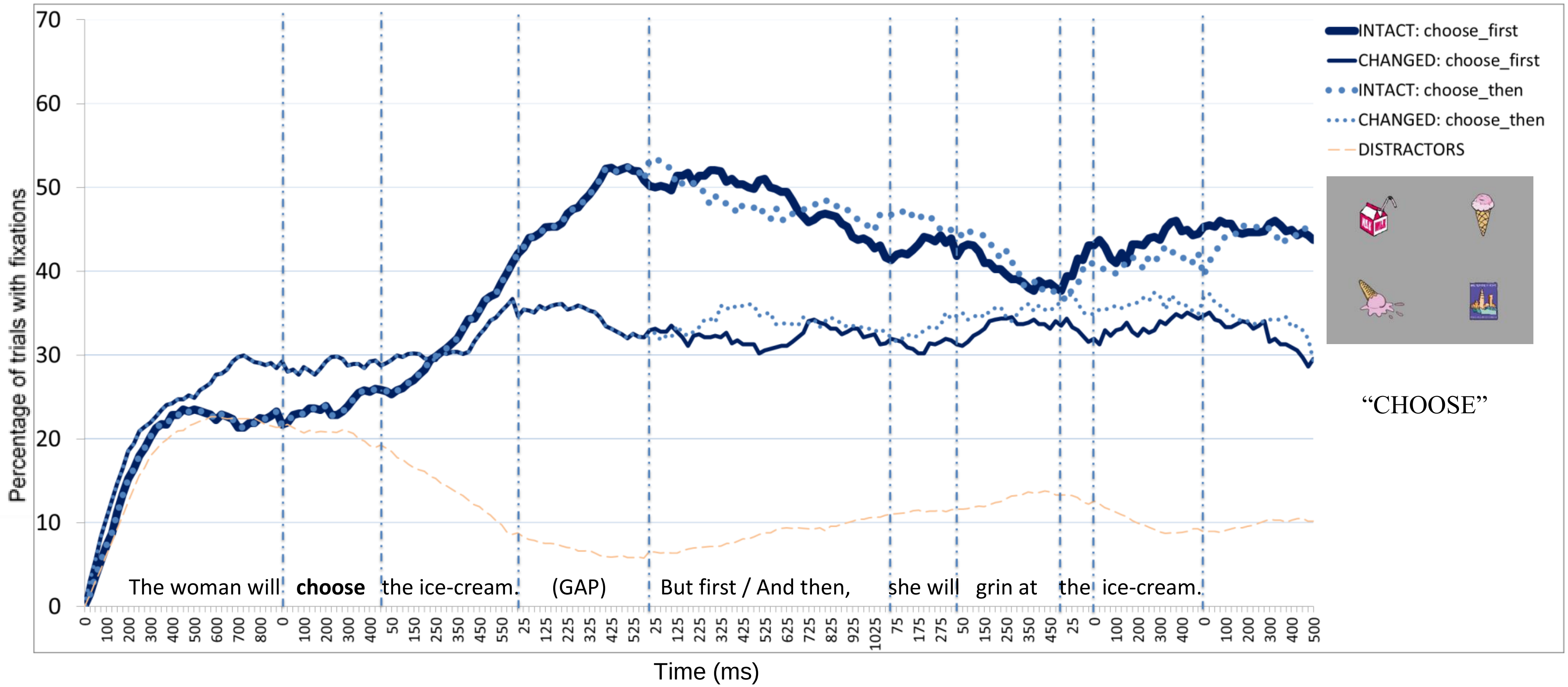


Figure 3.4a Experiment 4*:* Percentage of trials with fixations on the two versions of the target object (INTACT, CHANGED) and distractors on the visual display when *no change* (e.g., "*choose*") was mentioned in the first sentence. The x-axis shows the average time windows of critical phrases and words in the utterance. The y-axis shows mean proportion of trials with fixations that were calculated every 25 ms sequentially from the synchronization point on a trial-by-trial basis across subjects. The straight lines indicate fixations on the two versions of the target object in the "*but first*" condition, while the dotted line represent fixations in the "*and then*" condition. As the temporal shift (BUT FIRST/AND THEN) was mentioned in the second sentence, fixations during the first sentence were averaged between "but first" and "and then" conditions. The dashed line reveals averaged inspections of the distractors.

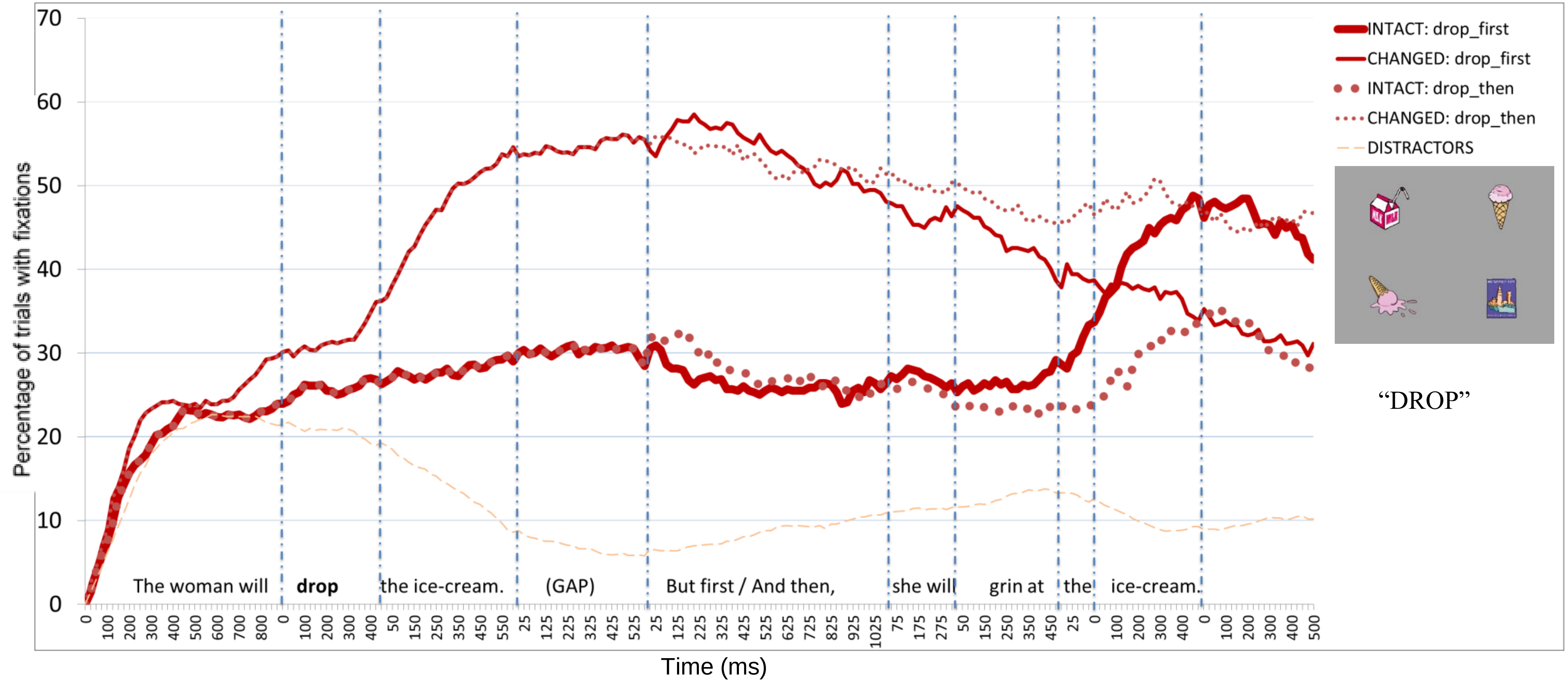


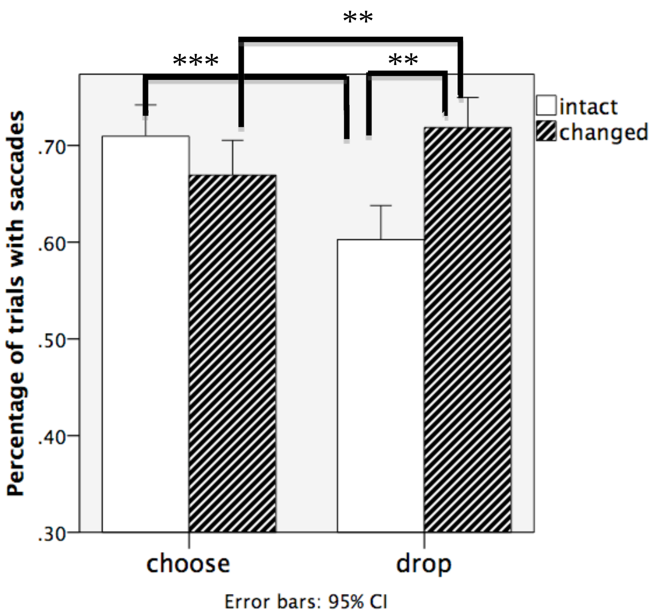
Figure 3.4b Experiment 4*:* Percentage of trials with fixations on the two versions of the target object (INTACT, CHANGED) and distractors on the visual display when *a change of state* (e.g., "*drop*") was mentioned in the first sentence. The x-axis shows the average time windows of critical phrases and words in the utterance. The y-axis shows mean proportion of trials with fixations that were calculated every 25 ms sequentially from the synchronization point on a trial-by-trial basis across subjects. The straight lines indicate fixations on the two versions of the target object in the "*but first*" condition, while the dotted line represent fixations in the "*and then*" condition. As the temporal shift (BUT FIRST/AND THEN) was mentioned in the second sentence, fixations during the first sentence were averaged between "but first" and "and then" conditions. The dashed line reveals averaged inspections of the distractors.

3.2.2.2 Saccades

"*The woman will* *choose/drop* *the ice cream*"

(i) *During "choose/drop":* A two-way repeated measures ANOVA revealed that there was no main effect of the depicted state (*intact, changed*), $F_1(1, 63) = .71$, $p = .402$, $\eta_p^2 = .01$, $F_2(1,35) = .21$, $p = .653$, $\eta_p^2 = .01$, nor main effect of the degree of change (*choose, drop*), $F_1(1,63) = .14$, $p = .139$, $\eta_p^2 = .02$, $F_2(1,35) = 2.74$, $p = .107$, $\eta_p^2 = .07$. No interaction was found between the depicted state and the degree of change by subjects, $F_1(1,63) = 1.66$, $p = .200$, $\eta_p^2 = .01$, but a significant interaction was found by items, $F_2(1,35) = 8.80$, $p = .005$, $\eta_p^2 = .20$. Planned comparisons revealed that there were more saccades towards the dropped ice cream than the upright ice cream during "*drop"* than "*choose"*, $t_1(63) = 2.08$, $p = .040$, $t_2(35) = 3.09$, $p = .004$, showing increasing eye movements towards the dropped ice cream if a change of state event was mentioned. There were more looks to the upright ice cream after "*choose*" than after "*drop*" by subjects, $t_1(63) = 4.51$, $p < .001$, but not with by item analysis, $t_2(35) = .06$, $p = .956$. Hence, there tended to be increasing looks towards the dropped ice cream when a change was described in the auditory stimuli.

(ii) *During "the ice cream"*: Figure 3.5 shows mean proportions of saccades towards the two depicted object states that were averaged for each condition. A two-way repeated measures ANOVA showed that there was no main effect of the degree of change on saccades towards the depicted state, $F_1(1,63) = 3.49$, $p = .064$, $\eta_p^2 = .03$, $F_2(1,35) = .001$, $p = .981$, $\eta_p^2 = .00$, but there was a significant main effect of the depicted state by subjects but not by items (more saccades towards the dropped ice cream than the upright one), $F_1(1,63) = 4.78$, $p = .030$, $\eta_p^2 = .04$, $F_2(1,35) = .08$, $p = .783$, $\eta_p^2 = .002$. A significant interaction between the degree of change and the depicted state was found both by subjects and by items, $F_1(1,63) = 20.00$, $p < .001$, $\eta_p^2 = .14$, $F_2(1,35) = 14.50$, $p = .001$, $\eta_p^2 = .29$. Planned comparisons confirmed that participants looked more at the upright ice cream in the *"choose"* condition (*M*= .71, *SD* = .19) than the "*drop"* condition (*M* = .60, *SD* = .20, $t_1(63) = 4.51$, $p < .001$, $t_2(35) = 3.45$, $p = .001$, but they looked more at the dropped ice cream in the "*drop"* condition (*M*= .72, *SD* = .18) than in the "*choose"* condition (*M* = .67, *SD* = .21, $t_1(63) = 2.16$, $p = .043$, $t_2(35) = 2.09$, $p = .044$).



Note: * *p*< .05; ** *p*<.01, *** *p*<.001

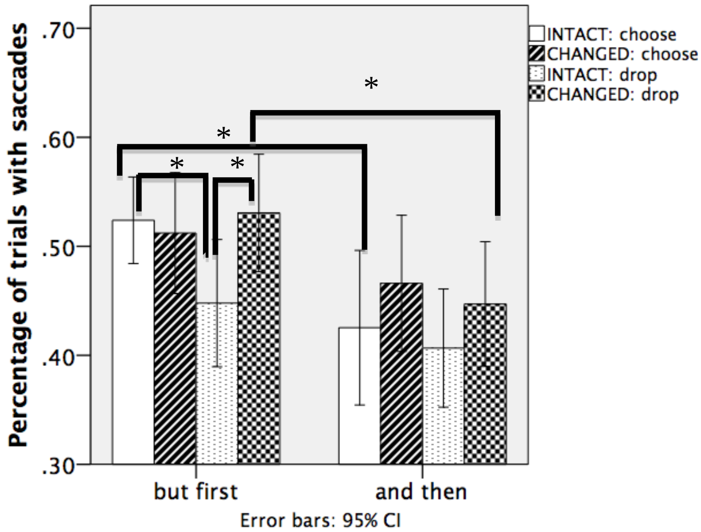
Figure 3.5 Experiment 4: Percentage of trials with saccades launched towards the intact state and the changed state between the onset and offset of the noun phrase (e.g. "*the ice cream*") (*Mean duration* = 603 ms) in the first sentence. The labels on the x-axis refer to the two described events in the first sentence, with "*choose*" representing the condition of No Change and "drop" the condition of Change of state. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

In sum, data from the first sentence showed that eye movements towards the depicted object states tended to be influenced by what was indicated as the appropriate object-state representation in the linguistic context in real time as language unfolded (e.g., a changed state for a change of state event but not for no change). This finding was reflected in the fixation data. It was also consistent with Experiment 2, where shorter picture verification times were found for pictures depicting an object state that matched the linguistic description than that mismatched.

"*But first/And then, she will look at the ice cream*"

(iii) *During "But first/And then":* A three-way repeated measures ANOVA showed that there was a significant main effect of the temporal shift (*but first vs. and then*) on saccades towards the depicted object state, with more saccades launched in the "*but first*" than in the "*and then*" condition (See Figure 3.6), $F_1(1, 63) = 11.44$, $p = .001$, $\eta_p^2 = .15$, $F_2(1, 35) = 10.37$, $p = .003$, $\eta_p^2 = .23$. There was no main effect of the degree of change (*choose vs. drop*), $F_1(1, 63) = 1.61$, $p = .210$, $\eta_p^2 = .03$, $F_2(1, 35) = .16$, $p = .689$, $\eta_p^2 = .01$, nor the depicted state (*intact vs. changed*), $F_1(1, 63) = 3.71$, $p = .059$, $\eta_p^2 = .06$, $F_2(1, 35) = .01$, $p = .982$, $\eta_p^2 = .00$. No interactions were found. Nonetheless, to evluate whether there would an immediate influence of time shifts on eye movements, follow-up paired-sample t-tests compared saccades between the two versions of the target object. If there was such influence, the upright ice cream was expected to attract more saccades than the dropped one when "but first" was mentioned. Contrary to this prediction, results showed that the pattern for the "*but first*" condition was largely similar to what has been

seen when the target noun phrase was mentioned in the first sentence, with more saccades towards the changed one than the upright one (see Figure 3.6).

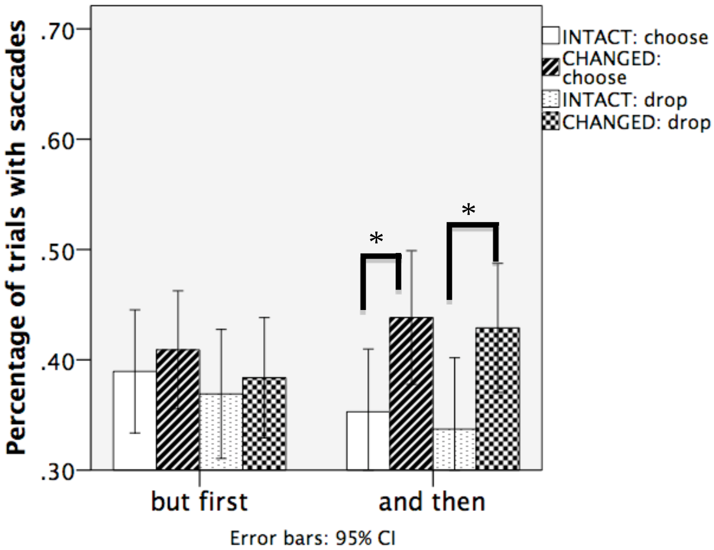


Note: * *p*< .05; ** *p*<.01, *** *p*<.001

Figure 3.6 Experiment 4: Percentage of trials with saccades towards the intact state and the changed state between the onset and offset of temporal shifts (e.g., "*but first/and then*") in the second sentence (M = 619 ms). The labels on the x-axis refer to temporal shifts at the beginning of the second sentence. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

(iv) *During "look at":* To further examine whether eye movements could be adjusted prior to the target noun phrase, saccadic eye movements during the verb region were analysed. A three-way repeated measures ANOVA showed that there was no significant main effect of the temporal shift (*but first vs. and then*) on saccades towards the depicted object state, $F_1(1, 63) = .01$, $p = .928$, $\eta_p^2 = .00$, $F_2(1, 35) = 1.26$, $p = .270$, $\eta_p^2 = .04$, nor main effect of the degree of change (*choose vs. drop*), $F_1(1, 63) = 1.00$, $p = .322$, $\eta_p^2 = .02$, $F_2(1, 35) = .56$, $p = .461$, $\eta_p^2 = .02$, but there was a main effect of the depicted state (*intact vs. changed*) by subjects (more looks to the dropped ice cream than the upright one), $F_1(1, 63) = 7.20$, $p = .009$,

η_p^2 = .10, but not by items, $F_2(1, 35) = .01$, p = .996, η_p^2 = .00. Nonetheless, a significant interaction between the temporal shift and the state was found, $F_1(1, 63) = 4.35$, p= .041, η_p^2 = .07, $F_2(1, 35) = 4.93$, p = .033, η_p^2 = .12. Planned comparisons have shown that there were more saccades toward the dropped ice creams by subjects but not by items, $t_1(63) = 2.63$, p = .011, $t_2(35) = 1.52$, p = .138, suggesting there tended to be increasing saccades towards the changed state in "*and then*" despite exisiting fixations.

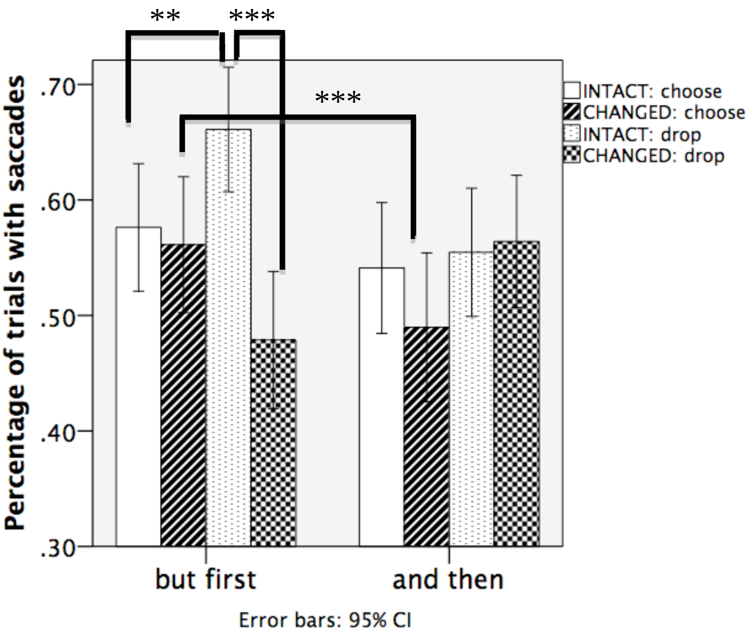


Note: * p< .05; ** p<.01, *** p<.001

Figure 3.7 Experiment 4: Percentage of trials with saccades towards the intact state and the changed state between the onset and offset of verbs (e.g., "*look at*") (M = 507 ms) in the second sentence. The labels on the x-axis refer to temporal shifts at the beginning of the second sentence. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

(v) *During "the ice cream"*: Figure 3.8 illustrates mean proportions of saccades towards the two depicted object states that were averaged for each condition across subjects. A three-way repeated measures ANOVA showed that there was a significant main effect of the depicted state with more saccades towards the intact state than the changed state overall,

$F_1(1, 63) = 11.23$, $p = .001$, $\eta_p^2 = .15$, $F_2(1, 35) = 12.67$, $p = .001$, $\eta_p^2 = .27$. There was no significant main effect of the temporal shift (*but first vs. and then*), $F_1(1, 63) = 3.38$, $p = .071$, $\eta_p^2 = .05$, $F_2(1, 35) = 1.83$, $p = .185$, $\eta_p^2 = .05$, nor the degree of change (*choose vs. drop*), $F_1(1, 63) =1.47$, $p = .230$, $\eta_p^2 = .02$, $F_2(1, 35) = 2.21$, $p = .147$, $\eta_p^2 = .06$. Nevertheless, there was a significant interaction between the temporal shift and the depicted state, $F_1(1, 63) = 4.78$, $p = .032$, $\eta_p^2 =07$, $F_2(1, 35) = 7.76$, $p = .009$, $\eta_p^2 = .18$. A significant three way interaction was also found, $F_1(1, 63) = 9.47$, $p = .003$, $\eta_p^2 = .13$, $F_2(1, 35) = 8.48$, $p = .006$, $\eta_p^2 = .20$.



Note: * *p*< .05; ** *p*<.01, *** *p*<.001

Figure 3.8 Experiment 4: Percentage of trials with saccades towards the intact state and the changed state between the onset and offset of the noun phrase (e.g, “*the ice cream*”) (*M* = 565 ms) in the second sentence. The labels on the x-axis refer to temporal shifts at the beginning of the second sentence. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

To confirm the source of the interaction described above, we conducted follow-up pairwise comparisons. No differences in saccades after the "*choose*" sentences. In "*drop*", however, there were more looks towards the upright ice cream after "*but first*" than after "*and then*", $t_1(63) = 2.31$, $p = .024$, $t_2(35) = 2.88$, $p = .007$, while there were conversely fewer looks to the dropped ice cream after "*but first*" than after "*and then*", $t_1(63) = 2.55$, $p = .013$, $t_2(35) = 1.94$, $p = .061$, suggesting increasing looks to the upright ice cream were driven by the degree of change ("*drop*") and a backward temporal shift ("*but first*").

In sum, data from the second sentence illustrate that participants tended to retrieve an object-state representation that fitted in the current linguistic description. When a substantial change was described to happen on the object in the first sentence, there were increasing looks towards the intact state but decreasing of looks at the changed state, probably due to retrieval of an intact object-state representation as required. No such difference was found in eye movements when no change was described. Importantly, in the second sentence this effect was not manifested until the object name was explicitly mentioned.

3.2.3 Discussion

In Experiment 4, participants' eye movements were recorded while they listened to auditory sentences about a target object (e.g., an ice cream) and viewed a visual scene that illustrated an upright ice cream and a dropped ice cream along with two distractors. We have obtained eye movement patterns that were manifested during the unfolding of the sentences. Results showed that eye movements were influenced by the difference between the two depicted versions of the ice cream and moderated by linguistic description. Initially, there were more fixations on the dropped ice cream than the intact ice cream even before verbs were mentioned. Later, at the first reference to the target object, participants tended to look at the upright ice cream when no change was described, but the dropped one when a change of state was mentioned. This pattern was consistent with findings from the sentence-picture verification paradigm that also showed an interaction between object state and degree of change on picture verification response latencies.

At the second reference to the target object, participants tended to look at a particular object state depending on whether a backward temporal shift (*but first*) or a forward temporal shift (*and then*) was indicated. In the "*but first*" condition, there were increasing looks at the upright ice cream but decreasing looks at the dropped ice cream when a change of state was described in the first sentence. By comparison, in the "*and then*" condition, there tended to be more looks at the dropped ice cream when a change of state than no change was described in the first sentence. Interestingly, it was during the second reference to the target object when participants moved their eyes to a relevant object state but not immediately after the temporal shift was mentioned ("*but first/and then*"). Thus, although participants' eye movements on the static visual scene were influenced by temporal shifts, such difference of eye movements towards the appropriate version of the target object was manifested at the final reference to it, but not before.

The question is whether the lack of anticipatory eye movements towards a particular object state in Experiment 4 is resulted from a lack of appropriate contextual constraint. On the one hand, it is possible that the verbs in the first sentence could not constrain what kind of objects would follow because there were four different objects depicted on the scene. Similarly, in the second sentence, participants could favour (and switch to) the upright ice cream as soon as they heard "*but first*", since they were clearly required to shift back in time and the current object-state representation would not be relevant. What would be referred to next could only be drawn from the remaining three items unless the dropped ice cream was not 'bound' to the same identity as the ice cream that was dropped (i.e. it was just another object, unrelated to the one in the story, in which case the referential set would remain at four, and would not reduce down to three). Perhaps, the lack of anticipatory eye movements towards a specific object state came from: (a) limited processing time and cognitive resources available in online language comprehension, since establishing object-state representation required enough amount of time allowing for integrating sentence meaning, as shown in Experiment 3; (b) nothing mentioned to pull attention away from the "*ice cream*" that they've already been looking at, or (c) they are not anticipating with any great certainty that the "*ice cream*" will be mentioned again because the verbs in the second sentence (e.g., "*look at*" is non-selecting the "ice cream", but there might be earlier looks to the intact ice cream in "*the woman will drop the ice cream but first she will <u>lick</u>…*").

Taken together, with the visual world paradigm, Experiment 4 demonstrated that eye movements on the visual scene were primarily drawn to the two conflicting states of a target object. The effect was mostly manifested at the explicit reference to the object itself and modulated by the degree of change and temporal shifts that were described in auditory stimuli.

3.3 Experiment 5

Experiment 5 aimed to replicate and extend the findings of Experiment 4. The paradigm was identical to Experiment 4. Different from Experiment 4, the auditory stimuli in Experiment 5 indicated the degree of change in the first sentence by using two contrasting objects (*stomp on the penny vs. stomp on the egg*) instead of two different actions (*choose the ice cream vs. drop the ice cream*). Thus, participants would not be able to get complete information of the described event until a particular object was mentioned. In this way, the degree of change was manipulated by the affordance of the objects, but not the actions alone.

3.3.1 Method

Participants. 64 students (38 female, mean age 22, range 18-26 years old) from the University of York participated in this study. They received either half an hour course credit or two pounds for their contribution. All participants were native speakers of British English and had normal or corrected to normal vision. None of the participants were in any of the previous experiments.

Materials. There are 36 sets of auditory stimuli. The first sentence indicated either a change of state or no change by using the same verb with two different nouns (*stomp on the egg vs. stomp on the penny*), while the second sentence involved either a backward temporal shift (*but first*) or forward shift (*and then*). The stimuli were all in future tense, as shown in the examples below:

a) The girl will stomp on the penny. But first, she will look down at the penny.

b) The girl will stomp on the penny. And then, she will look down at the penny.

c) The girl will stomp on the egg. But first, she will look down at the egg.

d) The girl will stomp on the egg. And then, she will look down at the egg.

The foil stimuli were the same as the ones used in Experiment 4. All sentences were recorded by a male native speaker of British English at the sample rate of 44,100 Hz and 16-bit sound resolution in a sound proof booth. Identical words in the recorded stimuli of each set were spliced and re-used to exclude effects of any subtle difference in auditory cues, such as intonation.

36 pairs of quadrant experimental pictures and were created by commercially available Clipart packages, as shown in Figure 3.9a and Figure 3.9b. In each pair of quadrant pictures, the target objects were placed in the same locations. For the no change condition ("*stomp on the penny"*), two identical version of the object were presented together with two distractors (the oven and gloves). For the change condition ("*stomp on the egg"*), the intact state and the changed state of the object were depicted along with two distractors. The foil pictures were the same as the ones used in Experiment 4.

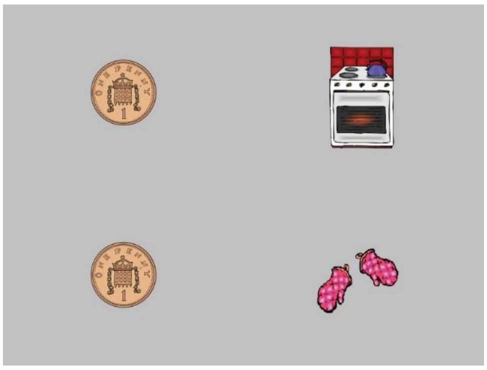


Figure. 3.9a Experiment 5: Example of an experimental visual stimulus. The corresponding auditory stimuli were "*The girl will stomp on the penny. But first/And then, she will look down at the penny".*

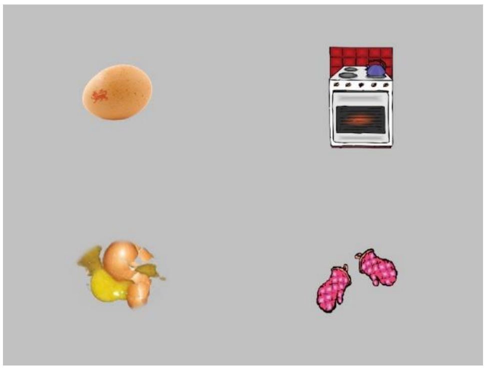


Figure 3.9b Experiment 5: Example of an experimental visual stimulus. The corresponding auditory stimuli were "*The girl will stomp on the egg. But first/And then, she will look down at the egg*".

Procedure. The procedure was identical to Experiment 4.

3.3.2 Results

Interest areas on the visual scene were marked with quadrants around depicted objects. Participants' eye movements on each region were synchronized to the onset and offset of critical words and phrases in the auditory stimuli, including *"stomp on", "the", penny/egg", "but first/and then", "look down at", "the", and "penny/egg"*. The presentation of the result is divided into two sections. First, overall fixation patterns are described, which gives a general indication of eye movements as language unfolded. Second, analyses within time windows are provided by examining the percentage of trials with saccades towards the two versions of each target object (*egg vs. penny*) over time in time windows of the verb in the first sentence (*stomp*), the first reference to the target object (*the penny/egg*), temporal shifts (*but first/and then*), and the final reference to the target object (*the penny/egg*)[6].

3.3.2.1 Fixations

Figure 3.10a and Figure 3.10b give a detailed illustration of eye movements over time by showing the mean proportion of fixations on the two versions of each target object on the visual scene during the presentation of trials. Although the two versions of the "PENNY" are the same, they are marked as "Intact" and "Changed" by following the locations of the "EGG" (i.e. the "changed" *penny* takes the same location as the "changed" *egg*, and vice versa).

[6] See Table a2.2 in Appendix II for mean saccades and fixations during the critical time windows on the visual display as the utterance unfolded.

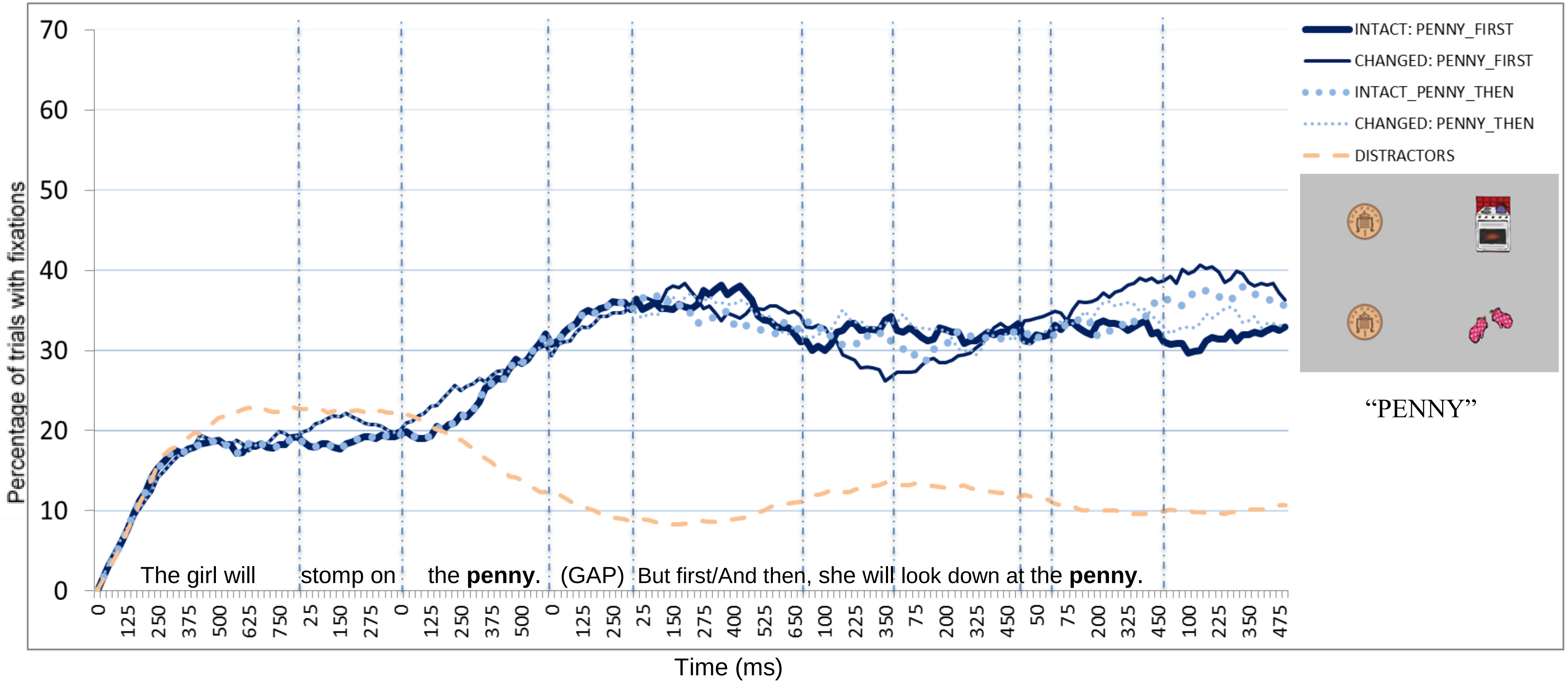


Figure 3.10a Experiment 5: Percentage of trials with fixations on the visual display as the auditory stimuli unfolded. This graph shows conditions in which the target object underwent no change of state (e.g., STOMP ON THE PENNY). Fixations were calculated every 25 ms sequentially from the onset critical words and phrases, as indicated in the x-axis. The y-axis shows mean proportions of fixations that were calculated on a trial-by-trial basis and averaged across subjects. The straight lines are fixations on the two different version of the target object when "BUT FIRST" was mentioned at the beginning of the second sentence, while the dotted lines showing fixations when "AND THEN" was mentioned. Although both versions of the "PENNY" are the identical, they are marked as "CHANGED" and "INTACT" to match the locations of the two versions of the "EGG". The dashed line shows averaged percentage of fixations on the two distractors.

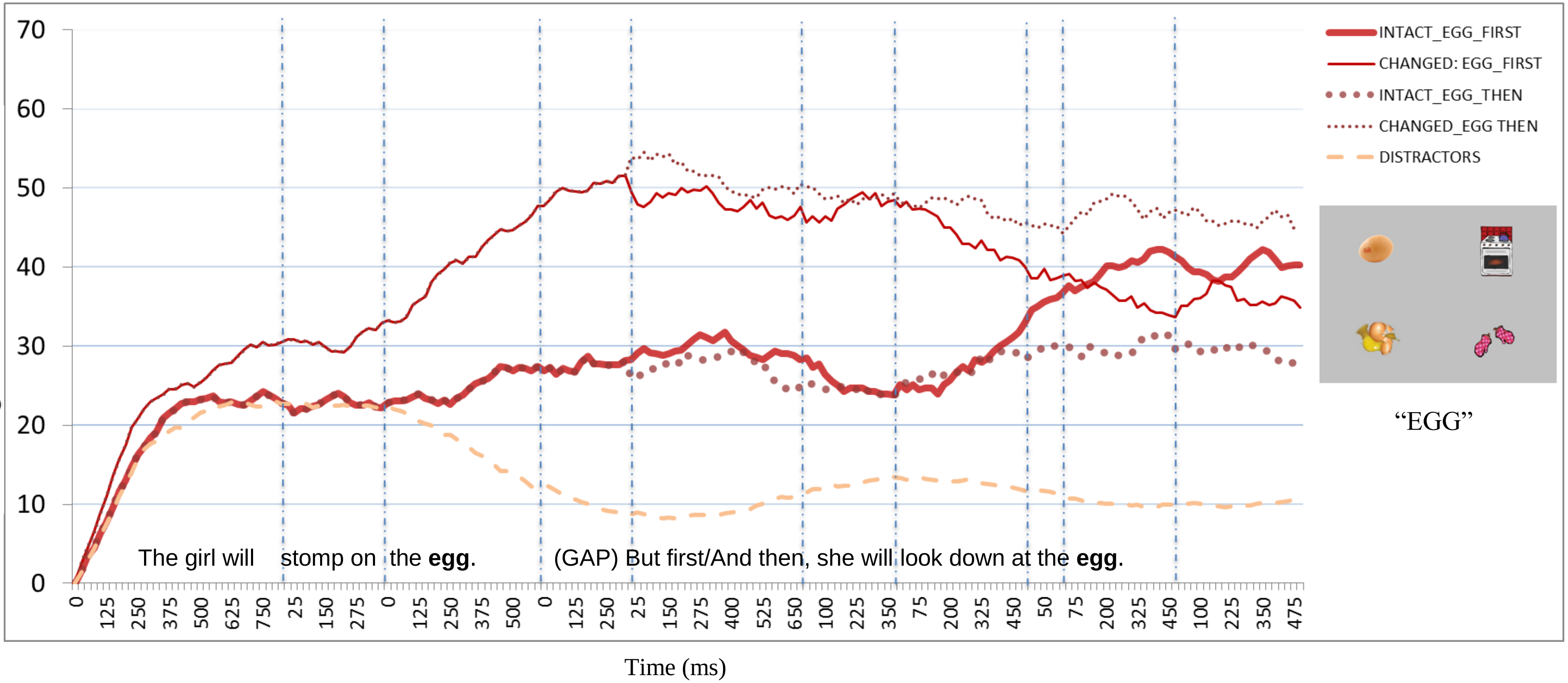


Figure 3.10b Experiment 5: Percentage of trials with fixations on the visual display as the auditory stimuli unfolded. This graph shows conditions in which the target underwent a substantial change in state (e.g., "STOMP ON THE EGG"). Fixations were calculated every 25 ms sequentially from the onset critical words and phrases, as indicated in the x-axis. The y-axis shows mean proportions of fixations that were calculated on a trial-by-trial basis and averaged across subjects. The straight lines are fixations on the two different version of the target object when "BUT FIRST" was mentioned at the beginning of the second sentence, while the dotted lines showing fixations when "AND THEN" was mentioned. The two versions of the target object are different with one being an intact state and the other a changed state. The dashed line shows averaged percentage of fixations on the two distractors.

Data from the first sentence suggested that there were obvious changes of fixations on the two identical versions of the unchanged object (e.g., *penny*) throughout the whole utterance. By contrast, participants predominantly looked at the broken egg compared to the intact egg before the verb "*stomp on*" was mentioned. This pattern remained the same until the final reference to the target object in the second sentence (e.g., *the egg*). Similar to Experiment 4, there was no immediate influence of temporal shifts on eye movements. However, when the noun phrase of mentioned again at the end of the second sentence, there were decreasing fixations on the broken egg but increasing eye movements on the intact egg for "*but first, egg*", but no obvious changes of fixations on the two versions of the egg for "*and then, egg*".

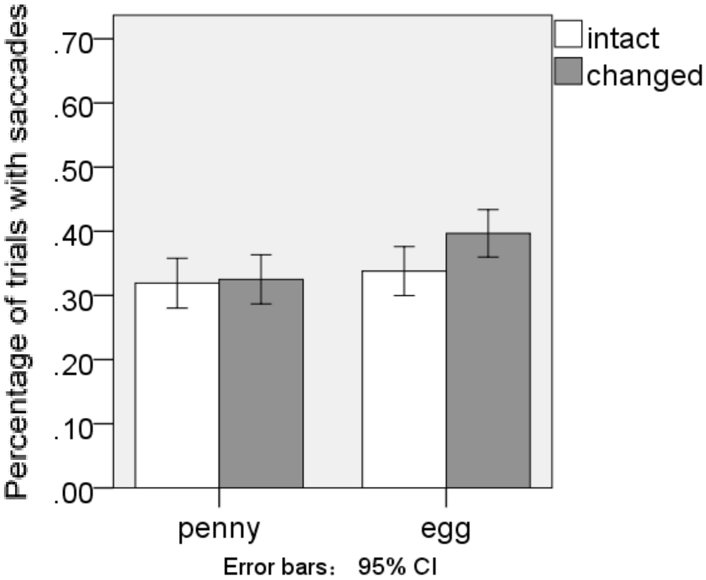
Therefore, the pattern of fixations in Experiment 5 seemed to be largely similar to Experiment 4. The following statistical analysis of saccadic eye movements during the same critical time windows as in Experiment 5 further revealed the changes of eye movements as language unfolded.

3.3.2.2 Saccades

"*The girl will <u>stomp on</u> <u>the penny/egg</u>*"

(i) During "*<u>stomp on</u>*": A two-way repeated measures ANOVA revealed that there was a main effect of object type (penny vs. egg) (just missing significance by items) with more saccades towards the *egg*s than the *pennies* (see Figure 3.11), $F_1(1,63) = 6.54$, $p = .012$, $\eta_p^2 = .05$, $F_2(1,35) = 3.97$, $p = .054$, $\eta_p^2 = .10$, but no main effect of object state despite more initial fixations on the broken egg than the intact egg, $F_1(1,63) = 2.87$, $p = .093$, $\eta_p^2 = .02$, $F_2(1,35) = 2.61$, $p = .115$, $\eta_p^2 = .069$, nor significant interaction, $F_1(1,63) = 1.65$, $p = .202$, $\eta_p^2 = .01$, $F_2(1,35) = 1.75$, $p = .194$,

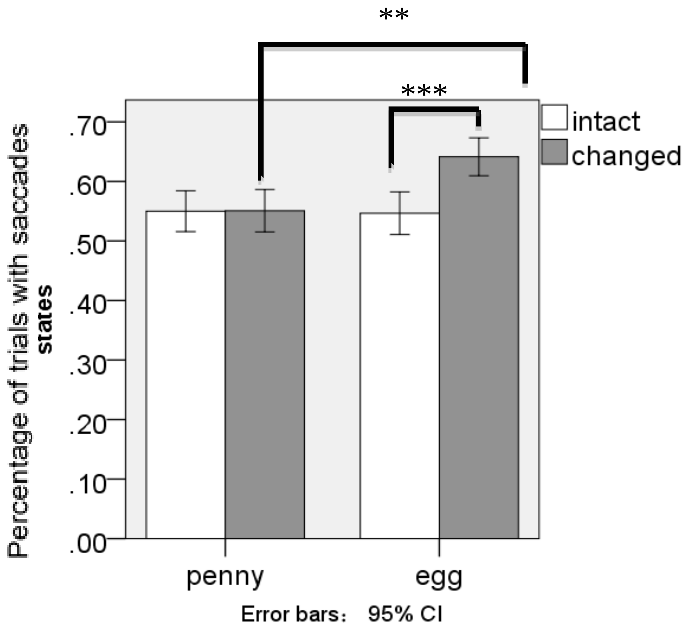
η_p^2 = .05. This pattern suggested that eye movements were largely driven by the visual features of depecited items, and not yet influenced by the linguistic context.



*Note: * p< .05; ** p<.01, *** p<.001*

Figure 3.11 Experiment 5: Percentage of trials with saccades towards the target regions between the onset and offset of the verb (e.g., "stomp on") (*M*=607 ms) in the first sentence. The labels on the x-axis refer to the two target objects with *"penny*" representing the condition of No Change and "egg" the condition of Change of state. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

(ii) *During "the penny/egg"*: There was a main effect of object type on saccades (more looks towards the eggs than the pennies), $F_1(1,63)$ = 8.46, p = .004, η_p^2 = .12, $F_2(1,35)$ = 11.13, p = .002, η_p^2 = .24, a main effect of the object state (more looks towards the changed state than the intact state), $F_1(1,63)$ = 8.53, p = .004, η_p^2 = .12, $F_2(1,35)$ = 4.02, p = .004, η_p^2 = .10, and a significant interaction, $F_1(1,63)$ = 6.09, p = .015, η_p^2 = .10, $F_2(1,35)$ = 13.44, p = .001, η_p^2 = .28, suggesting eye movements were influenced by both object type and the depicted object state. This pattern was similar to that seen in Experiment 4.



Note: * $p< .05$; ** $p<.01$, *** $p<.001$

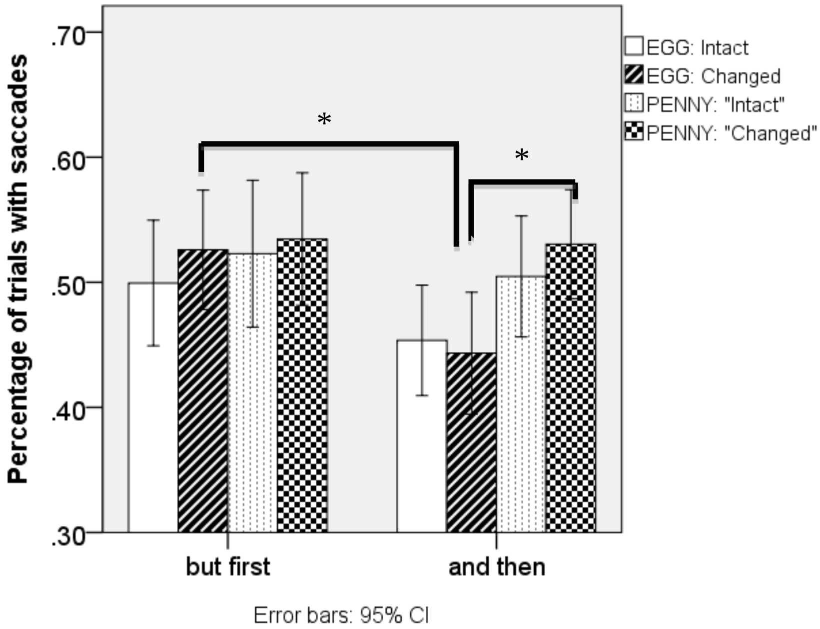
Figure 3.12 Experiment 5: Percentage of trials with saccades towards the target regions between the onset and offset of the reference to the target object (M = 562 ms) in the first sentence. The labels on the x-axis refer to the two described events in the first sentence, with *"penny*" representing the condition of No Change and "egg" the condition of Change of state. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

Pairwise comparisons suggested that despite no difference in looks towards the two versions of "*penny*" (as expected, of course), $t_1(63) = 0.3$, $p = .976$, $t_2(35) = 0.3$, $p = .976$, but a significant difference between the two versions of "egg", with more looks to the changed state ($M = .64$, $SD = .18$) than the intact state ($M = .55$, $SD = .21$, $t_1(63) = 3.87$, $p < .001$, $t_2(35) = 0.3$, $p = .976$), as shown in Figure 3.12. Although this pattern suggested that despite existing fixations on the broken egg, there were more saccades towards it than the intact one.

In sum, data from the first sentence demonstrated that when the target object was not expected to change its state (e.g., "*stomp on the penny*"), there was no difference in looks towards it's the two depicted versions. By contrast, when the object (presumably) was described to change (e.g., "*stomp on the egg*"), the changed state tended to draw more eye movements than the intact state despite existing fixations.

"*But first/And then, she will look down at the penny/egg*"

(iii) *During "but first/and then"*: A three-way repeated measures ANOVA and suggest that there was a significant main effect of the temporal shift (*but first vs. and then*) with more saccades on hearing "*and then*" than "*but first*", $F_1(1, 63) = 4.62$, $p = .035$, $\eta_p^2 = .07$, $F_2(1, 35) = 4.70$, $p = .037$, $\eta_p^2 = .12$, and a main effect of object type (*penny vs. egg*) with more saccades towards the pennies than the eggs (See Figure 3.13) , $F_1(1, 63) = 5.11$, $p = .027$, $\eta_p^2 = .08$, $F_2(1, 35) = 6.05$, $p = .019$, $\eta_p^2 = .15$, but there was no main effect of object state (*intact vs. changed*), $F_1(1, 63) = .57$, $p = .45$, $\eta_p^2 = .01$, $F_2(1, 35) = .54$, $p = .47$, $\eta_p^2 = .02$. No significant interactions were found. No such difference was found for "*penny*". The above data suggested that the increase of eye movements on the visual scene was largely driven by "and then", but not by "but first". Nonetheless, the contrasting versions of the "egg" was not differently influenced, suggesting no immediate impact of the description of temporal shifts on individual representation of the target object.

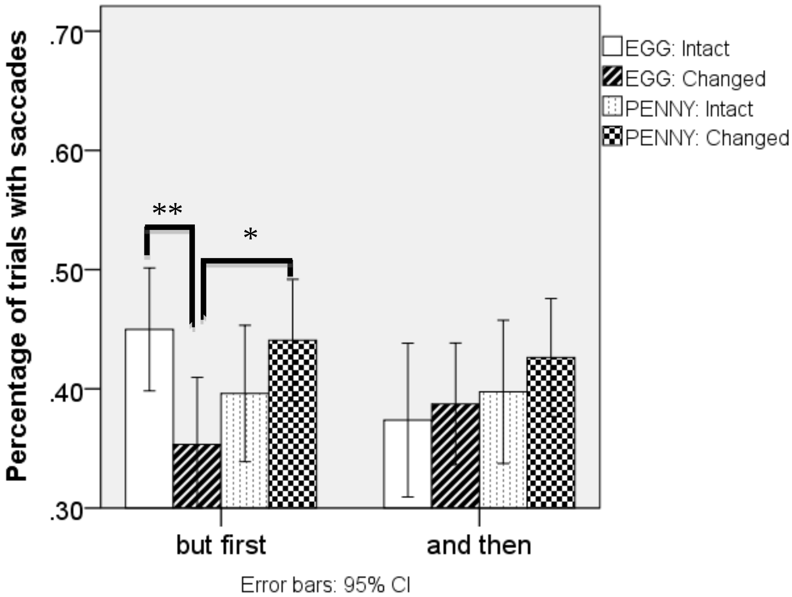


Note: * *p*< .05; ** *p*<.01, *** *p*<.001

Figure 3.13 Experiment 5: Percentage of trials with saccades towards the target regions between the onset and offset of the reference to the target object during the temporal shifts (M = 732 ms). The labels on the x-axis refer to the temporal shifts. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

(iv) *During "look down at"*: To examine whether the influence of linguistic context would be revealed during the verb region, stastical analysis of saccades was conducted in this time window. A three-way repeated measures ANOVA showed that there was no significant main effects of the temporal shift (*but first vs. and then*), object type (*penny vs. egg*), and object state (*intact vs. changed*). However, there was a significant interaction between the depicted state and object type, $F_1(1, 63) = 4.55$, $p = .037$, $\eta_p^2 = .07$, $F_2(1, 35) = = 6.90$, $p = .013$, $\eta_p^2 = .17$. As predicted, pairwise comparisons revealed that the intact version of the "*egg*" was preferred compared to its broken version when a backward temporal shift was mentioned ("*but first*"), $t_1(63) = 2.80$, $p = .007$, $t_2(35) = 2.32$, $p = .027$

(See Figure 3.14). No such difference was found for the "*and then*" condition probably due to exisiting more fixations on the broken one than the intact one. This result suggested that by the time the verb in the second sentence was mentioned, participants were able to anticipate the intact state of the "egg" was the appropriate object state in the context. Thus, compared with Experiment 4, Experiment 5 showed an ealier demonstration of the retrieval of the intact state in the verb region prior to the noun region.

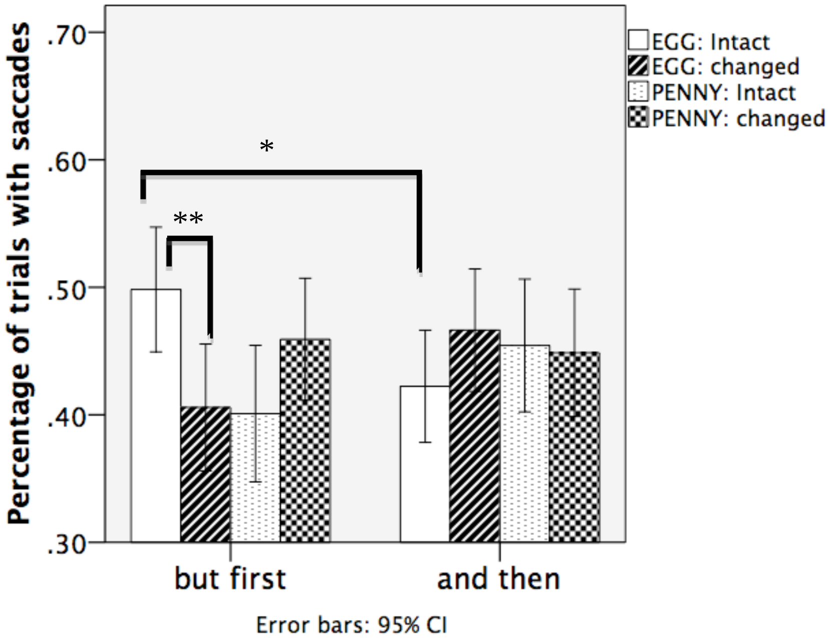


Note: * *p*< .05; ** *p*<.01, *** *p*<.001

Figure 3.14 Experiment 5: Percentage of trials with saccades towards the target regions between the onset and offset of the verb (e.g., *look down at*) (*M* = 586 ms) in the second sentence. The labels on the x-axis refer to the temporal shifts. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

(v) *During "the penny/egg"*: Figure 3.15 shows mean percentage of trials with saccades towards depicted versions of the "egg". A three-way repeated measures ANOVA showed that there was no significant main effect of the temporal shift (*but first vs. and then*), $F_1(1, 63) = .18$, *p*

= .674, η_p^2 = .003, $F_2(1, 35) = .01$, $p = .912$, η_p^2 = .00, object type (*penny vs. egg*), $F_1(1, 63) = .17$, $p = .681$, η_p^2 = .003, $F_2(1, 35) = .02$, p= .880, η_p^2 = .001, nor main effect of object state (*intact vs. changed*), $F_1(1, 63)$ = .01, p= .945, η_p^2= .00, $F_2(1, 35) = .09$, $p = .772$, η_p^2 = .002. However, there was a significant three-way interaction, $F_1(1, 63) = 8.89$, p= .004, η_p^2= .12, $F_2(1, 35) = 7.64$, $p = .009$, η_p^2 = .179.



Note: * p< .05; ** p<.01, *** p<.001

Figure 3.15 Experiment 5: Percentage of saccades towards the depicted versions of the object between the onset and offset of the noun phrase (e.g., "*the penny/egg*") (M = 630 ms) in the second sentence. The labels on the x-axis refer to the two depicted states, while y-axis shows estimated arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

Pairwise t-tests found that there were higher percentages of trials with saccades towards the intact state of the "*egg*" than its changed state when the second sentence began with *"but first*", $t_1(63)$ = 2., p = .042, $t_2(35) = 2.76$, $p = .009$). Besides, there were more looks towards the intact "*egg*" in "*but first*" *than* "and then", $t_1(63) = 2.50$, $p = .015$, $t_2(35)$ =

2.59, *p* = .014). No such difference was found for the "*penny*". This result suggests that despite increased saccades towards the intact egg, participants were able to further increase looks it compared to the broken version. This pattern replicated the findings of Experiment 4.

In sum, data from the second sentence showed that participants were able to retrieve the state of the target object as described in the language context when object name was mentioned at the end of the second sentence and at the verb region prior to it.

3.3.3 Discussion

In Experiment 5, the degree of change was manipulated by using the same verb with two different nouns (e.g., *stomp on the penny/egg*). For each set of auditory stimuli, separate visual scenes were used for the two target objects. Two identical versions were depicted for the object that was not expected to change its state (e.g., *stomp on the penny*), while two different versions (intact vs. changed) were depicted for the object that was expected to change its state with the same action (e.g., *stomp on the egg*).

All eye movement patterns in Experiment 4 were replicated in Experiment 5. No difference was found between eye movements towards the two identical versions of "*the penny*", though they switched from one version of to the other in the "but first" condition. By contrast, despite an initial bias towards the broken *egg* than the intact egg, eye movements were modulated by the linguistic context. After the verb onset in first sentence (e.g., "*stomp* on"), there were increasing inspections to the broken egg than the intact egg. At the end of the second sentence, looks towards the intact version increased but looks towards the changed version decreased in the “but first’ condition. No such changes were found in the “and then” condition. Nonetheless, compared with Experiment 4, the shift of object representation was revealed earlier at the verb region in Experiment 5, which may be resulted from the difference in linguistic structures as the properites of object were important for understanding change in Experiment 5 but not necessarily in Experiment 4.

3.4 General discussion

With the “look and listen” eye-tracking paradigm, Experiments 4 and Experiment 5 aimed to investigate the time course of tracking object-state representation in language comprehension. The target object was depicted in two versions along with two distractors. We have obtained reliable evidence of the influence of linguistic context on eye movements towards the depicted versions of the object. Participants’ eye movements revealed that there was an initial bias towards the changed version of the target object, but this bias was reversed when no change was described to happen on the object. Besides, in the second sentence, the earliest moment eye movements switched from the unwanted changed state to the intact state was at the noun region. In principle, they could favour (and switch to) the appropriate state as soon as they heard the temporal shifts, since they were clearly required to shift back/forward in time when the current object-state representation may not be relevant. The fact that looks towards the appropriate state happened at a later linguistic region was probably due to the processing time that is required to construct situation models, as illustrated in Experiment 3.

What was driving participants' eye movements towards the depicted versions of the object during the utterance? Could eye movements towards a particular object state reflect phonological overlap (Allopenna, Magnuson, & Tanenhaus, 1998) between the described object-state representation and the depicted object state? I might look at an ice cream because I hear the word "ice cream" and this directs my attention to anything to which that label can apply, or I might look at an ice cream because I have activated a conceptual representation and it is this (conceptual) activation that directs my attention to an ice cream (one that most closely matches my conceptual representation). Data from Experiment 4 and Experiment 5 suggested that phonological overlap was not likely to be the case: If it were phonological overlap that drove the eye movements, there should be no differences in looks towards the two depicted states of the ice cream. After all, both of pictures had the same object name.

Alternatively, we might assume that participants were treating the two depicted versions as the same object, the intact version being the initial state and the changed version the end state. The two object states represented how the target object would look like before and after a change of state occurred. Increasing looks at the intact state as required by the backward temporal shift further showed that appropriate object states could be established in language comprehension. On the other hand, participants may have treated these two versions as two separate objects with the same linguistic label. For example, if there are two bottles of juice with exactly the same appearance in front you, one bottle empty and the other one full, you would not assume that the empty bottle of juice is the "future" existence of the full one, but you would assume instead that it is a separate bottle.

Therefore, it might have been the semantic overlap between the depicted object state and the mentally constructed object-state representation that drove eye movements to one object or another. That is, the object-state representation may not necessarily be ‘bound’ to the identity of the depicted version, but it may share core semantic features with the depicted item in the visual scene. For example, they have the same linguistic label and thus both can be referred to with the same object name; second, they are in the same state as indicated in the linguistic context. This semantic feature mapping effect was first reported by Cooper (1974), who found that participants were more likely to fixate items that were semantically related with the spoken word (e.g., more looks towards a snake or a lion for “Africa” than unrelated control words). Thus, aside from the time it took to integrate the linguistic description, there might be other cognitive processes (e.g., semantic feature mapping) involving in this process, which might have delayed early looks towards the appropriate object state.

Taken together, Experiment 4 and Experiment 5 have shown that visual and auditory input can be rapidly integrated with representation of object state, but it is not clear what drove participants’ different eye movements towards the two depicted versions of object – whether they treated the two depicted versions as conflicting states of the same object or mapped semantic features of the described object-state representation onto the depicted object state in the visual scene. In the next Chapter, this issue will be further examined. Answering this question may be useful to provide an account of the level of mental representation activation in online language comprehension.

3.5 Conclusion

To conclude, Experiment 4 and 5 revealed that object-state representation could be established during language comprehension. Eye movement measures revealed rapid mapping between object-state representation and object state in the immediate visual scene. Nevertheless, this effect tended to manifested when object name was explicitly mentioned, but not before. Further studies will be presented in the following chapter to examine what drove participants' eye movements around the visual scene – the binding of the two conflicting states as the before and after versions of the same object, or semantic overlap between object-state representation and the depicted state?

CHAPTER 4

Tracking object-state representation during language comprehension: More evidence from eye movements

4.1 Overview

In Chapter 3, two experiments with the visual world paradigm investigated the linguistically-mediated influence of described events on eye movements towards two conflicting depicted versions (intact vs. changed) of a target object (e.g., ice cream). Initially, participants preferred the changed version (e.g., a dropped ice cream) to the intact version (e.g., an upright ice cream), but they adjusted their eye movements towards an appropriate version of the target object that matched the linguistic context as language unfolded. In the first sentence, they switched from the changed version to the intact version when no change was described as happening to the object (e.g., "*The woman will choose the ice cream"*). In the second sentence, eye movements were again directed to the relevant versions of the ice cream based on the linguistic description. For example, after "*drop, but first*" was mentioned, they started to look away from the dropped ice cream but towards the upright ice cream, suggesting the retrieval of the intact version as the appropriate object-state representation. However, looks towards the intact version did not exceed the changed version until the object name (e.g., "ice cream") was mentioned at the end of the sentence, which may be due to updating of situation models.

What drove participants' eye movements switching from one ice cream to the other in the second sentence? On the one hand, we might assume that the two conflicting versions were bound to the same identity of the target object, with the upright ice cream reflecting the initial state and the dropped ice cream reflecting the end state. Although these two versions could not co-exist in real life, participants were able to track the appropriate one based on the linguistic description. When "*the ice cream*" was mentioned, they just looked at the one that matched the described state.

On the other hand, participants may have treated these two versions as two separate tokens of the target object. Semantic features of spoken words have shown to be able to direct eye movements towards objects in the visual scene (Huettig & Altmann, 2005, 2007; Huettig & McQueen, 2007; Huettig, Quinlan, McDonald, & Altmann, 2006; Yee & Sedivy, 2006). Huettig and Altmann (2005) found that when "piano" was heard, participants were more likely to fixate on a picture of a trumpet than unrelated objects due to overlap of semantic features between "piano" and "trumpet" (i.e. as musical instrument). Yee, Huffstetler, & Thompson-Schill (2011) further demonstrated that objects sharing semantic knowledge, including perceptual (shape) and abstract (function) features can be co-activated. Thus, perhaps participants mapped the semantic features of the described object-state representation in the episodic memory onto the depicted versions of the target object on the visual display. Participants' eye movements were then directed to the one that overlapped most with the object's representation in mind. If this is the case, the depicted versions of the target object were not necessarily bound to the same object identity. Instead, the two versions could be totally different-looking tokens of the same type of object, with looks to each version driven solely by the degree of overlap between that version and whatever features were being 'held in mind'.

The three experiments in this chapter were designed to see whether findings from Experiment 4 and 5 were due to participants' treating the two displayed versions as depicting the same object, or whether eye-movements reflected semantic overlap between the depicted object and the current mental representation of the object – one that was not necessarily 'bound' to the same token of the target object.

Specifically, Experiment 6 intended to investigate whether eye movements could be driven by semantic overlap alone when the two depicted versions of the target object cannot be bound to the same object identity. As shown in Figure 4.1, two tokens of the target object ("the ice cream") were depicted on the visual display. One was in the intact state (e.g., an upright ice cream) and the other one in the changed state (e.g., a dropped ice cream), but the two versions could not be treated as conflicting states of the same object identity due to differences in perceptual features.

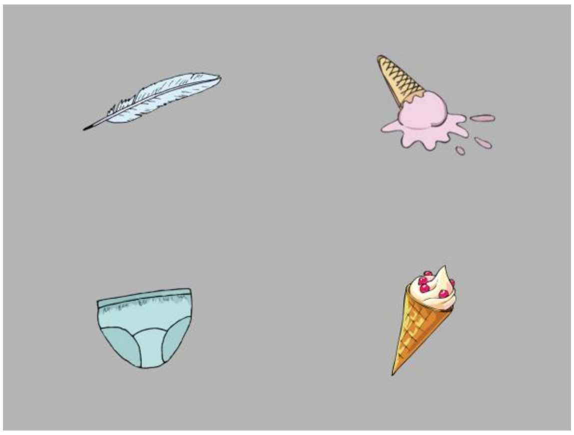


Figure 4.1 Experiment 6: An example visual scene in Experiment 6. On each display, there were four items, including two target two object with the same linguistic label but in conflicting states and two unrelated distractors.

The auditory stimuli in Experiment 6 were the same as in Experiment 4, for example, "*The woman will choose/drop the ice cream. But first/And then, she will look at the ice cream*". We predicted that if eye movements were driven by semantic overlap alone, we would see the same pattern of eye movements in Experiment 6 as in Experiment 4.

In Experiment 7, visual stimuli from Experiments 4 and 6 were combined into a single design (See Figure 4.2).

a. similar b. different

Figure 4.2 Experiment 7: Example visual display in Experiment 7. On (4.2a), two similar-looking versions of the target object were presented together with two distractors. The intact version was intended to resemble the initial state of the object before it undergoes a change of state event, e.g., *drop*, while the changed version as the end state after change. On (4.2b), however, two different-look versions of the target objet were presented. Despite they were in different states, they were not supposed to be treated as conflicting states of the same item as perceptual differences cannot be attributed to a change of state action.

The auditory stimuli in Experiment 7 were modified based on Experiment 6 to introduce either "the ice cream" or "another ice cream", such as "*The woman will drop the ice cream. But first, she will look at* *the/another* *ice cream*". We expect to replicate findings of Experiment 4 and Experiment 6 when "*the ice cream*" was mentioned. When "*another ice cream*" was mentioned, however, participants were expected to establish a new object-state representation that does not necessarily inherit the episodic characteristics of the first ice cream that has been dropped. If this is the case, there would be no difference in looks towards the two upright ice creams between (4.2a) and (4.2b). However, if in Experiment 4, binding of the two representations of the ice cream (upright and dropped) did in fact occur, there would be more looks towards the upright ice cream in (4.2b) than (4.2a).

Consideration of the outcomes of Experiment 7, and their possible explanations, led to a final study in this series: Experiment 8. Three versions of ice creams were presented on the visual display. In this way, this final study distinguishes between alternative interpretations of the prior experiments, and leads to the conclusion that semantic overlap, and not binding, drives eye movements in this paradigm.

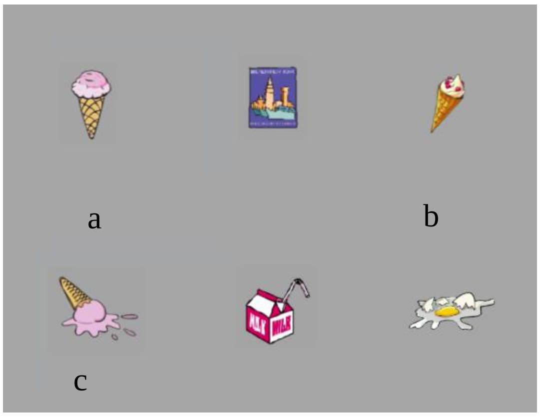


Figure 4.3 Experiment 8: Example visual display in Experiment 8. The objects including a dropped ice cream (c), an upright ice cream that is more similar to the dropped one (a), an upright ice cream that is less similar to the dropped one (b), and three distractors. The letters are shown here for exposition only and were not part of the visual display.

Taken together, with three experiments, this chapter would like to provide a comprehensive examination of the mechanisms that drove eye movements on the visual scene as object-state representation was established during language comprehension.

4.2 Experiment 6

In Experiment 6, we investigated whether eye movements around the visual scene could be driven solely by semantic overlap between the activated concept and the objects depicted on the visual scene. Unlike previous experiments in Chapter 3, the target object (e.g., ice cream) was depicted as two distinct tokens (i.e. two different looking ice creams. See Figure 4.1). Thus, it was not possible for participants to treat them as two conflicting states of the same token of the ice cream and to bind each to the same object identity.

If participants were binding the two depicted versions as the same object identity in Experiment 4, this was not possible in Experiment 6 due to the perceptual differences between them, as they were two different tokens of the ice cream. In this case, we should not see increasing looks (or as many looks) towards the upright ice cream when a backward temporal shift ("*but first*") was mentioned. However, if eye movements were solely driven by semantic overlap between the depicted version and the described object-state representation of the ice cream, we should see the same pattern of eye movements in Experiment 6 as in Experiment 4.

4.2.1 Method

Participants. 64 students (40 female, mean age 21, range 18-24 years old) from the University of York participated in this study. They received either half an hour course credit or three pounds for their contribution. All participants were native speakers of British English and had normal or corrected to normal vision. None of the participants in this study had taken part in previous experiments.

Stimuli. The visual stimuli were modified based on those in Experiment 4. Instead of having two "similar-looking" versions of the target object, two "different-looking" versions were depicted in the visual scene, as shown in Figure 4.1 above. The auditory stimuli were identical to those used in Experiment 4, as shown in the examples below:

a) The woman will choose the ice cream. But first, she will look at the ice cream.

b) The woman will choose the ice cream. And then, she will look at the ice cream.

c) The woman will drop the ice cream. But first, she will look at the ice cream.

d) The woman will drop the ice cream. And then, she will look at the ice cream.

Procedure. Experiment 6 followed the same procedure as Experiment 4 and Experiment 5. The onset of the auditory stimulus occurred 1000 ms after participants previewed the scene. Eye movements were monitored throughout the experiment using an EyeLink II head-mounted eye-tracker, which sampled at 250 Hz. Viewing was binocular but only the right eye was tracked. Each trial lasted 9 seconds and the total length of the experiment was about 30 minutes.

4.2.2 Results

Data were analyzed by the following steps as in Experiment 4 and 5. Both fixations and saccades around the same critical time windows as Experiment 4 were reported.

4.2.2.1 Fixations

Figure 4.4a and Figure 4.4b illustrate percentage of trials with fixations on the depicted ice creams, which were synchronized to the auditory stimuli. Similar to Experiment 4, the two versions of the target object received more fixations than distractors.

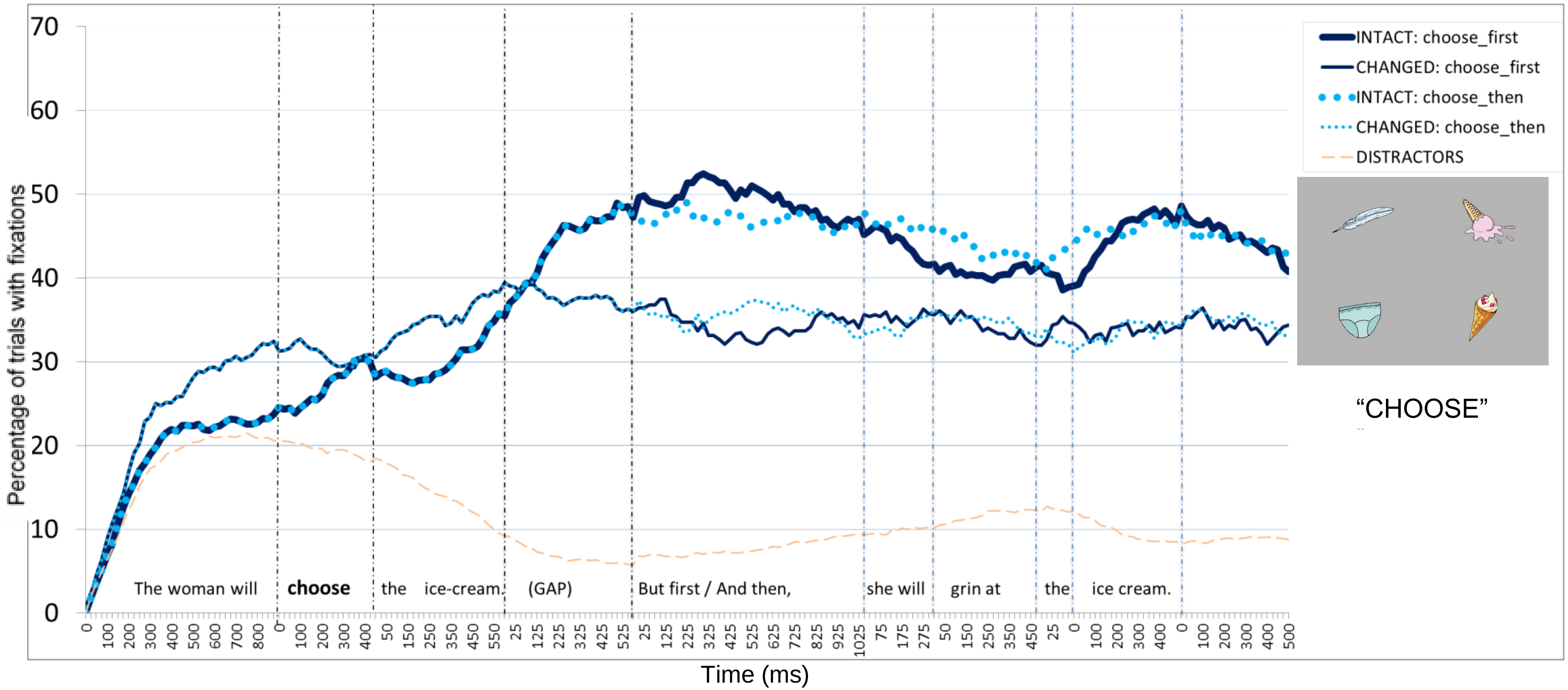


Figure 4.4a Experiment 6*:* Percentage of trials with fixations on the target regions during the presentation of the auditory stimuli. This graph shows conditions in which the target object underwent no change of state (e.g., CHOOSE). Fixations were calculated every 25 ms sequentially from the synchronization point at the onset of critical words and phrases, as shown on the x-axis. Y-axis shows mean proportion of fixations that were calculated on a trial-by-trial basis. The straight lines indicate fixations on the two target items when "BUT FIRST" was mentioned in the second sentence, while the dotted lines show fixations when "AND THEN" was mentioned. Since the temporal shift was mentioned in the second sentence, the fixations of "but first" and "and then" were averaged in the first sentence for the purpose of the graph. Fixations on the two distractors were averaged and indicated by the dashed line.

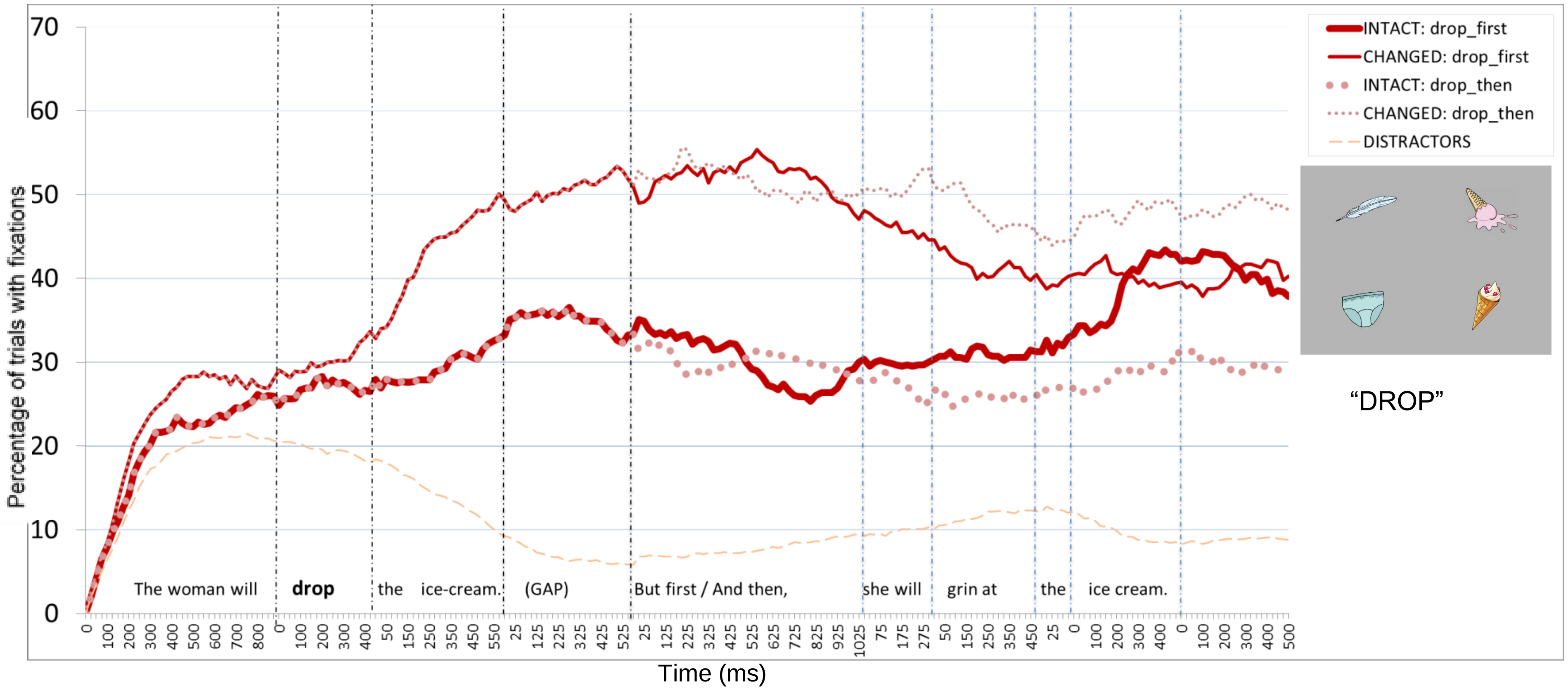


Figure 4.4b Experiment 6*:* Percentage of trials with fixations on the target regions during the presentation of the auditory stimuli. This graph shows conditions in which the target underwent a substantial change in state (e.g., “DROP”). Fixations were calculated every 25 ms sequentially from the synchronization point at the onset of critical words and phrases, as shown on the x-axis. Y-axis shows mean proportion of fixations that were calculated on a trial-by-trial basis. The straight lines indicate fixations on the two target items when “BUT FIRST” was mentioned in the second sentence, while the dotted lines show fixations when no change of “AND THEN” was mentioned. Since the temporal shift was mentioned in the second sentence, the fixations of “but first” and “and then” were averaged in the first sentence for the purpose of the graph. Fixations on the two distractors were averaged and indicated by the dashed line.

At the onset of the verb in the first sentence, there were more fixations on the dropped ice cream than the upright ice cream regardless of the verbs. At the offset of the verb when they had heard "*the woman will choose/drop…"*, there were more fixations on the dropped ice cream for "drop", but no difference of the upright ones for "choose". Nonetless, at the offset of "the ice cream" when they had heard "*the woman will choose/drop the ice cream*", fixations on the two depicted versions of the ice cream were influenced by the described actions in different ways, with more fixations on the dropped ice cream in "*drop*" than "*choose*", but the upright ice cream in "*choose*" than "*drop*".

In the second sentence, shortly after the onset of "*but first/and then",* fixations on both versions of the target object decreased. However, at the onset of "the ice cream", fixations on the upright ice cream rapidly increased when "but first" was mentioned but not when "and then". Nonetheless, at the offset of the noun phrase in the second sentence, there were no difference in fixations on the upright one and the dropped one in "*drop, but first*", which is different from Experiment 4 where there were more fixations on the upright one than the dropped one.

Therefore, in Experiment 6 proportions of fixations were similar to Experiment 4 in the first sentence with more fixations on the dropped ice cream than the upright one, but not in the second sentence, where not necessarily more fixations on the upright ice cream than the dropped one.

4.2.2. 2 Saccades

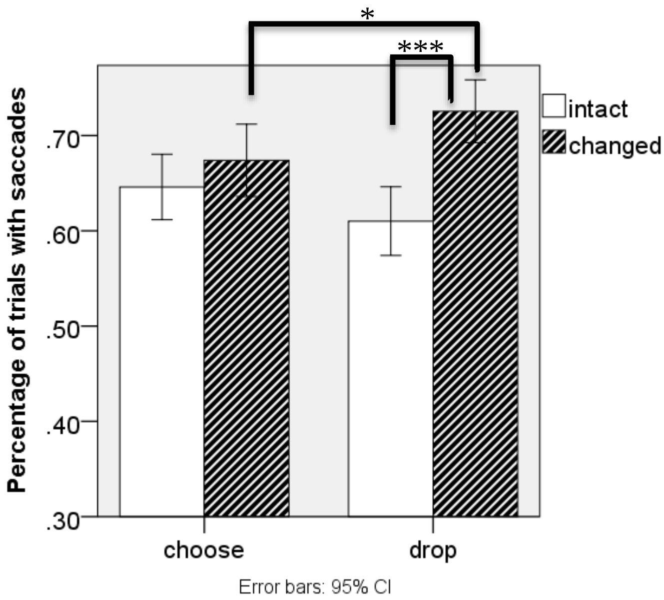
"*The woman will choose/drop the ice cream*"

(i) *During "choose/drop":* A repeated measures two-way ANOVA showed *that* there was no main effect of the verb on percentage of saccades

towards the depicted ice creams, $F_1(1,63) = .34$, $p = .561$, $\eta_p^2 = .01$, $F_2(1,35) = 1.01$, $p = .321$, $\eta_p^2 = .028$, the depicted version, $F_1(1,63) = 1.54$, $p = .217$, $\eta_p^2 = .01$, $F_2(1,35) = .01$, $p = .954$, $\eta_p^2 = .00$, or interaction between them, $F_1(1,63) = .81$, $p = .371$, $\eta_p^2 = .01$, $F_2(1,35) = 1.49$, $p = .231$, $\eta_p^2 = .04$. Hence, saccades seemed to be not influenced by the verbs when they were just mentioned.

(ii) *During "the ice cream"*: There was no main effect of the verb *(choose vs. drop)*, $F_1(1,63) = .30$, $p = .583$, $\eta_p^2 = .01$, $F_2(1,35) = .12$, $p = .729$, $\eta_p^2 = .01$, but there was a significant main effect of the depicted ice creams (more looks overall to the dropped version than the upright version. See Figure 4.5), $F_1(1,63) = 15.44$, $p < .001$, $\eta_p^2 = .11$, $F_2(1,35) = 4.88$, $p = .034$, $\eta_p^2 = .12$, and an interaction between them by subjects, $F_1(1,63) = 5.91$, $p = .016$, $\eta_p^2 = .04$, but not by items, $F_2(1,35) = 3.62$, $p = .065$, $\eta_p^2 = .09$.

To compare the results with Experiment 4, pairwise comparisons revealed that there was no difference in the percentage of saccades towards the upright ice cream depending on the verb *("choose" vs. "drop")*. However, participants tended to look towards the dropped ice cream in "*drop*" ($M = .73$, $SD = .19$) more than in "*choose*" ($M = .67$, $SD = .22$, $t_1(63) = 2.16$, $p = .033$, $t_2(35) = 1.63$, $p = .112$), as shown in Figure 4.5. Hence, saccades were influenced both by the described actions and the depicted object states with more saccades towards the dropped ice cream when "drop" was mentioned, but more towards the upright ice cream when "choose" was mentioned, replicating the findings of Experiment 4.



Note: ** p< .05; ** p<.01, *** p<.001*

Figure 4.5 Experiment 6: Percentage of trials with saccades towards the two depicted versions of the target object ("*the ice cream*") between the onset and offset of noun phrase (*M* = 603 ms) in the first sentence. The labels on the x-axis refer to the two described events in the first sentence, with "*choose*" representing the condition of No Change and "drop" the condition of Change of state. The y-axis shows percentage of trials with saccades that were calculated trial-by-trial and averaged across subjects.

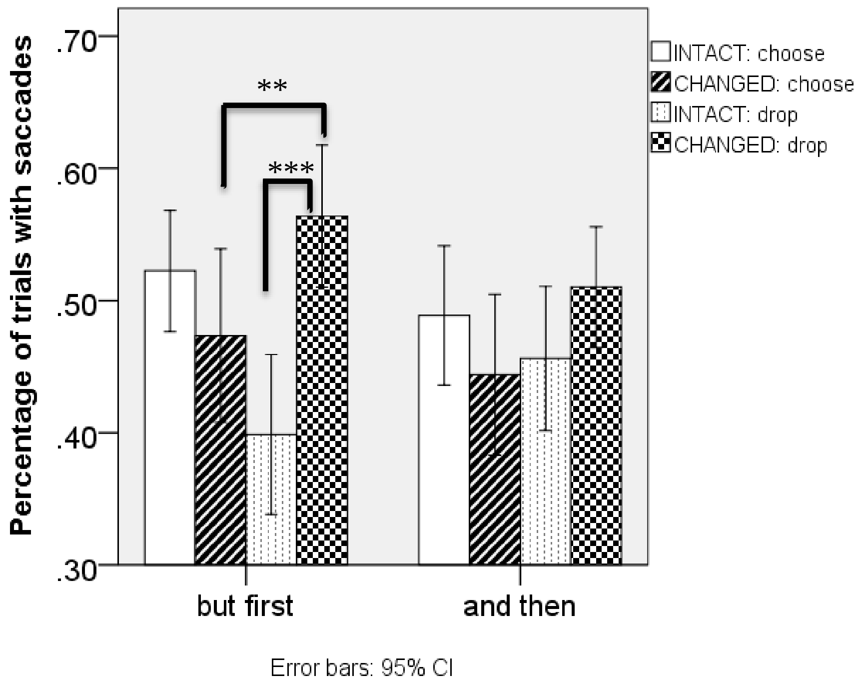
In sum, data from the first sentence showed that eye movements around the visual scene tended to be influenced both by what was described as the appropriate object-state representation in the linguistic context and by the depicted versions of the target object. This effect manifested at the end of the first sentence when the noun phrase (e.g., "the ice cream") was mentioned.

"*<u>But first/And then</u>, she will <u>look at</u> <u>the ice cream</u>*"

(iii) *During "<u>but first/and then</u>":* Figure 4.6 illustrates mean proportions of saccades towards the two depicted versions of the target object that

were averaged across subjects for each condition. A three-way repeated ANOVA showed that there was no significant main effect of the temporal shift (*but first vs. and then*), $F_1(1, 63) = .68$, $p = .412$, $\eta_p^2 = .01$, $F_2(1, 35) = 3.37$, $p = .075$, $\eta_p^2 = .09$, the verbs (*choose vs. drop*), $F_1(1, 63) = .00$, $p = .999$, $\eta_p^2 = .00$, $F_2(1, 35) = .04$, $p =.853$, $\eta_p^2 = .00$, nor the depicted ice creams, $F_1(1, 63) = 2.76$, $p = .102$, $\eta_p^2 = .04$ (not significant by subjects), $F_2(1, 35) = 5.14$, $p = .030$, $\eta_p^2 = .13$ (significant by items), but there was a significant interaction between the verbs and ice creams, $F_1(1, 63) = 15.73$, $p < .001$, $\eta_p^2 = .20$, $F_2(1, 35) = 11.71$, $p = .002$, $\eta_p^2 = .25$.

To confirm the source of the interaction, planned comparison showed that participants preferred to look more at the upright ice cream after "*the woman will choose the ice cream,* <u>*but first*</u> *…*" than after "*the woman will drop the ice cream*, <u>*but first*</u>...". The reverse pattern was found for saccades towards the dropped ice cream with higher probability of eye movements in "*drop, but first" than* "*choose, but first*". This pattern was similar to eye movements in the first sentence. The above results showed that eye movements were not influenced by the temporal shift, even though "*but first/and then*" has just been mentioned.



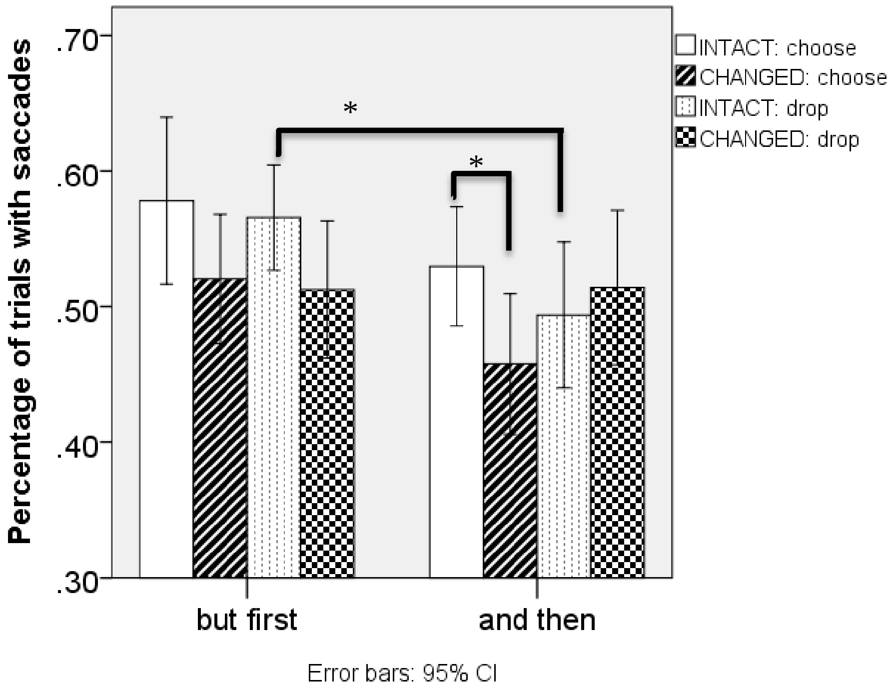
Note: * *p*< .05; ** *p*<.01, *** *p*<.001

Figure 4.6 Experiment 6: Percentage of trials with saccades towards the two versions of the target object on the visual scene between the onset and offset of "*but first/and then*" (*M* = 619 ms) in the second sentence. The labels on the x-axis refer to the two conditions of temporal shifts in the second sentence, with "but first" indicating a backward shift and "and then" a forward shift. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

(iv) *During "look at":* A three-way repeated ANOVA showed that there were no main effects or interactions. Thus, there were no differences of saccades towards the depicted versions of the target object depending on the linguistic context and object type.

(v) *During "the ice cream"*: Figure 4.7 shows mean saccades towards the two versions of the ice creams during this time window. A three-way repeated ANOVA showed that there was a significant main effect of the temporal shift (*but first vs. and then*) (more saccades overall in "*but first*" than "*and then*"), $F_1(1, 63) = 7.52$, $p = .008$, η_p^2

= .12, $F_2(1, 35) = 5.79$, $p = .021$, $\eta_p^2 = .14$, and depicted ice creams by subjects, $F_1(1, 63) = 6.17$, $p = .016$, $\eta_p^2 = .09$, but not by items, $F_2(1, 35) = 2.63$, $p = .114$, $\eta_p^2 = .07$, but there was no main effect of the verbs, $F_1(1, 63) = .01$, $p = .924$, $\eta_p^2 = .00$, $F_2(1, 35) = .01$, $p = .913$, $\eta_p^2 = .00$. No interactions were found.



Note: * *p*< .05; ** *p*<.01, *** *p*<.001

Figure 4.7 Experiment 6: Percentage of trials with saccades towards the target regions between the onset and offset of "the ice cream" (*M* = 565 ms) in the second sentence. The labels on the x-axis refer to the two conditions of temporal shifts in the second sentence, with "but first" indicating a backward shift and "and then" a forward shift. The y-axis shows arcsine transformed proportions of trials with saccades that were calculated trial-by-trial and averaged across subjects.

In sum, data from the second sentence revealed that participants were influenced by the described degree of change and the depicted versions of the target objects when a temporal shift was mentioned, showing a preference to the dropped ice cream for "*drop, but first*", but the upright ice cream for "*choose, but first*". This finding was not consistent with Experiment 4, in which the influence of the temporal shift manifested immediately after the temporal shit was mentioned. However, similar to Experiment 4, at the final reference to the target object, there was an increase of eye movements towards the upright ice cream with more saccades towards it in “*drop, but first*” than “*drop, and then*”. No such difference was found for the dropped ice cream.

4.2.3 Discussion

In Experiment 6, participants listened to a description of a target object and looked at a visual display that illustrated two different versions of the object along with two distractors. One version of the object was in an intact state, while the other version was in a changed state. Unlike Experiment 4, in which the two versions might be associated with two conflicting states of the “same” object identity (see Figure 3.1), in Experiment 6 the two versions were depicted as two “different’ tokens of the target object that cannot be bound into the same object identity (see Figure 4.1). By manipulating the identity of the target object on the visual display, we intended to investigate whether semantic overlap alone can drive eye movements on the visual scene towards the appropriate versions of the object.

We predicted that if participants were binding the two versions into the same object identity, we would see fewer eye movements towards the intact version when participants were, in effect, requested to move backward in time, because it did not correspond to the initial state of, in the example, the dropped ice cream (recall that the two ice creams were manifestly different tokens of an ice cream). Yet, if eye movements were driven by semantic overlap between described object-state representation and depicted object, the identity of the two depicted versions would not be relevant. In this case, the pattern of eye movements in Experiment 6 would be similar to Experiment 4.

Eye movements during the first sentence showed that participants were aware of the consequences of the described events and tended to look at the changed version of the target object after a change of state was described. This effect manifested when the object name was mentioned, but not before. However, unlike in Experiment 4, there was no reliable interaction between verb type (*choose vs. drop)* and object type. In the second sentence, the influence of the temporal shift did not manifest immediately when "*but first/and then*" was heard. Instead, the pattern of eye movements were following the same trend as in the first sentence. Nevertheless, consistent with Experiment 4, at the final reference to the target object there tended to be more looks towards the upright ice cream in "*drop, but first*" than "*drop, and then*". Thus, compared with Experiment 4, eye movements towards the intact version were delayed and reduced in Experiment 6.

These findings were not consistent with the prediction of the identity binding account that participants treated the upright ice cream and the dropped ice cream as two conflicting states of the same object identity in Experiment 4. According to this account, participants were binding the upright ice cream and the dropped one as the same object identity in Experiment 4. If this is the case, we were not expecting to see increasing eye movements towards the upright ice cream in Experiment 6 since the two versions of ice creams were different tokens can could be not bound to the same token of the target object. However, the results were not consistent with the semantic overlap account either, since the patterns were different from Experiment 4, even though there were increasing looks towards the upright ice cream. Perhaps, eye-movements around the visual scene were likely guided both by ascribing the same identity to the two versions of the object and by semantic overlap between the linguistically described object-state representation and a depicted version.

4.3 Experiment 7

Experiment 7 intended to further explore what drove the eye movements towards the depicted versions of the target object, being the binding of two depicted versions as the same object or semantic overlap between the object-state representation in mind and the depicted versions. All linguistic stimuli in Experiment 7 described a change of state in the first sentence and started the second sentences with "*but first*". However, the final noun phrase in the second sentence differed, as either "*the*" object or "*another*" was mentioned (e.g., *the woman will drop the ice cream. But first, she will look at the/another ice cream*). Compared with "*the ice cream*", "*another ice cream*" shares certain characteristics with the ice cream in the first sentence (i.e. it must be of the same type), but it is not expected to inherit its episodic characteristics and consequences (due to it having to be a different token). Instead, on hearing "*another ice cream*", a new token of ice cream that had not been encoded in the current model of language processing was expected to be introduced into the event representation. Based on findings in previous chapters, an intact token would be inferred in the second sentence.

On the visual display, we manipulated the perceptual features of the target object. One visual scene was the same as in Experiment 4 (see Figure 3.1/Figure 4.2a) and the other was the same as in Experiment 6 (see Figure

4.1/Figure 4.2b). For each trial, participants saw either two “similar” versions of the target object that could be associated with its initial state and end state (Figure 4.2a) or two “different” versions that could not be treated as conflicting states of the same object but two separate tokens of the object (Figure 4.2b).

a. similar b. different

Figure 4.2 Experiment 7: Example visual display in Experiment 7. On (4.2a), two similar-looking versions of the target object were presented together with two distractors. The intact version was intended to resemble the initial state of the object before it undergoes a change of state event, e.g., drop, while the changed version the end state after change. On (4.2b), however, two different-look versions of the target objet were presented. Despite they were in different states, they were not supposed to be treated as conflicting states of the same item as perceptual differences cannot be attributed to a change of state action.

If eye movements are driven by participants’ treating the two depicted ice creams in (4.2a) as representing the same actual (i.e. token) ice cream, we would expect to see more looks towards the upright ice cream in (4.2a) than (4.2b) when “the ice cream” is mentioned since the one in (4.2b) could not be treated as the corresponding dropped version. When “*another*” is mentioned, however, there should be fewer eye movements towards the upright ice cream in (4.2a) compared with the one in (4.2b). The upright ice cream in (4.2a) would be interpreted as the same token as the dropped ice cream, which means it would be a poor candidate for “another ice cream”.

By comparison, if eye movements are driven by semantic overlap, when "the" is mentioned there should also be more eye movements towards the upright ice cream in (4.2a) than (4.2b), because the one in (4.2a) was not only upright (i.e. not dropped) but also shared more perceptual features with the dropped ice cream than the one in (4.2b) in episodic memory. When "another" was referred to, however, there would be no differences in eye movements towards the upright versions between (4.2a) and (4.2b), because both versions are intact, and thus fit the newly-established representation of a (presumably intact, but different) ice cream. To summarize, when "the ice cream" was mentioned, we would expect to replicate findings in experiment 4 and Experiment 6, with more looks at the upright ice cream than the dropped one for both visual displays. However, when "another ice cream" was mentioned, there would be differences in eye movements between (4.2a) and (4.2b) according to the binding account, but not according to the semantic overlap account.

4.3.1 Method

Participants. 64 students (41 female, mean age 21, range 18-26 years old) from the University of York participated in this study. They received either half and hour course credit or two pounds for their contribution. All participants were native speakers of British English and had normal or corrected to normal vision. None of the participants in this study had taken part in previous experiments.

Materials. 36 sets of auditory stimuli were included as experimental trials. In each set of stimuli, either "*the" or "another"* was mentioned in the noun phrase at the end of the second sentence. For example,

4a) The woman will drop the ice cream. But first, she will look at <u>the</u> ice cream.

4b) The woman will drop the ice cream. But first, she will look at <u>another</u> ice cream.

There were 36 sets of foil sentences that followed the same sentence structure as the experimental trials. Half of them included "*the*" in the final noun phrase, while the other half included "*another*". All experimental and foil sentences were recorded by a male native speaker of British English at the sampling rate of 44,100 Hz and 16-bit sound resolution in a sound proof booth. Identical words in the recorded stimuli of each set were spliced and re-used to exclude effects of any subtle difference in auditory cues, such as intonation. Each stimulus was presented together with one of the two versions of visual display, as shown in Figure 4.2a and 4.2b.

Procedure. The procedure was the same as previous eye-tracking experiments in this thesis. The total length of the experiment was about 35 minutes.

4.3.2 Results

Following the same data analysis procedure as in previous experiments, fixations and saccades arond the time windows of the verb in the first sentence ("*drop"*), the first reference to the target object ("*the ice cream"*), temporal shifts ("*but first"*), the verb ("*look at*"), and the noun ("*ice cream"*) were reported.

4.3.2.1 Fixations

Figure 4.8a and Figure 4.8b illustrate the percentage of trials with fixations on the depicted versions of the target object, which were synchronized to the auditory stimuli[7]. At the onset of the verb in the first sentence, the dropped ice cream received more fixations than the upright ice cream in both pictures,

[7] The duration of "another" is longer than "the". For the purpose of the fixation graph, we have used the average duration of "another ice cream" and "the ice cream" to synchronize eye movements. However, for statistical analysis, we excluded the determiner and only took the average duration of the noun ("ice cream"), which has been cross-spliced across the linguistic conditions.

replicating previous experiments. This pattern did not change during the remaining section of the first sentence (See Figure 4.8a & 4.8b). At the offset of the “ice cream”, there were more fixations on the dropped ice cream but fewer on the upright ice cream in (4.2a) than (4.2b) (See Figure 4.8a & 4.8b).

In the second sentence, shortly after the onset of “*but first”,* fixations on both of the depicted versions decreased, but the pattern of eye movements remained the same until at the final reference to “the ice cream” when eye movements switched to the upright ice creams. When “*another ice cream*” was mentioned, there were more fixations on the upright version but fewer on the dropped version compared with “*the ice cream*”.

In sum, data of fixations revealed that there were more eye movements on the dropped ice cream than the upright ice cream in the first sentence, replicating previous experiments. In the second sentence, eye movements switched to the upright ice creams when “*the/another ice cream*” was mentioned.

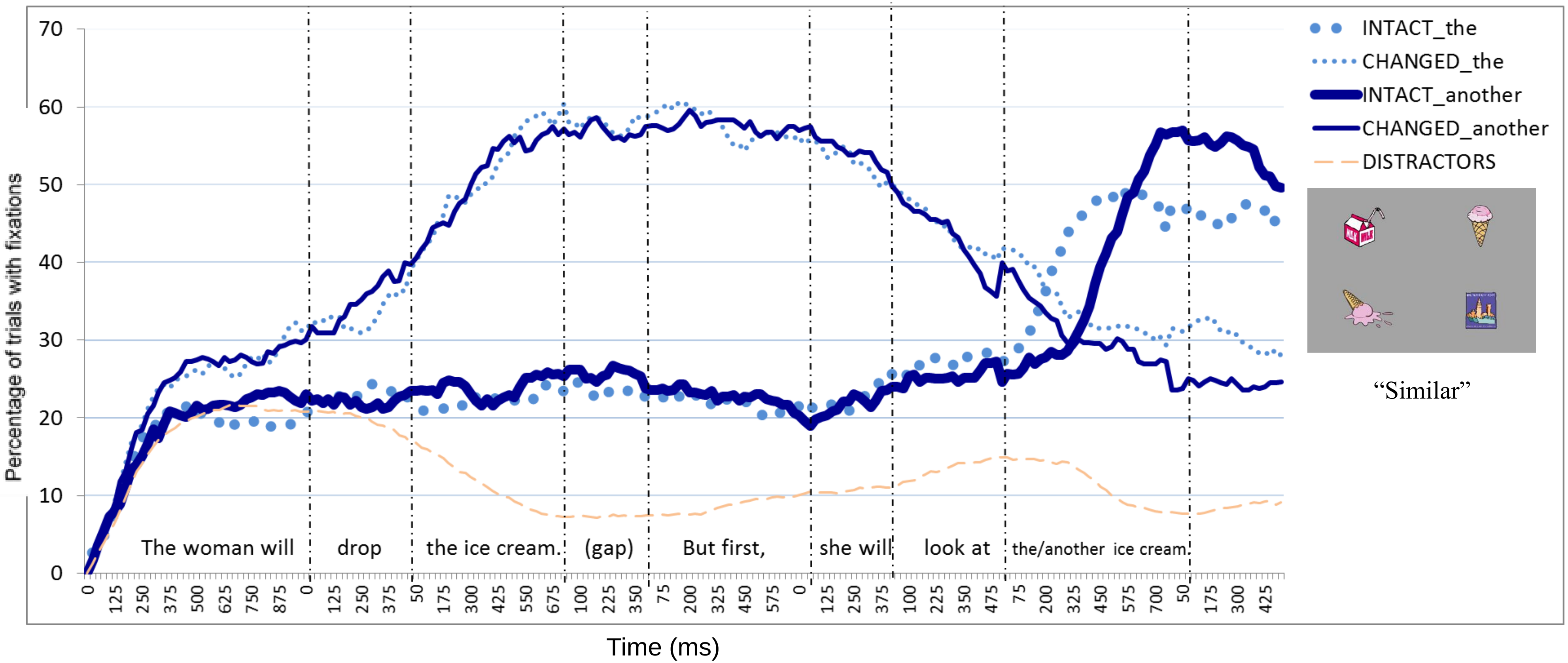


Figure 4.8a Experiment 7: Percentage of trials with fixations on the target regions during the presentation of the auditory stimuli. Two versions of the target object were depicted, which could be associated with the INTACT state and the CHANGED state of the SAME object. Fixations were calculated every 25 ms sequentially from the synchronization point at the onset of critical words and phrases, as shown on the x-axis. Y-axis shows mean proportion of fixations that were calculated on a trial-by-trial basis. The straight lines indicate fixations on the two target items when "ANOTHER" was mentioned in the second sentence, while the dotted lines show fixations when "THE" was mentioned. Fixations on the two distractors were averaged and indicated by the dashed line.

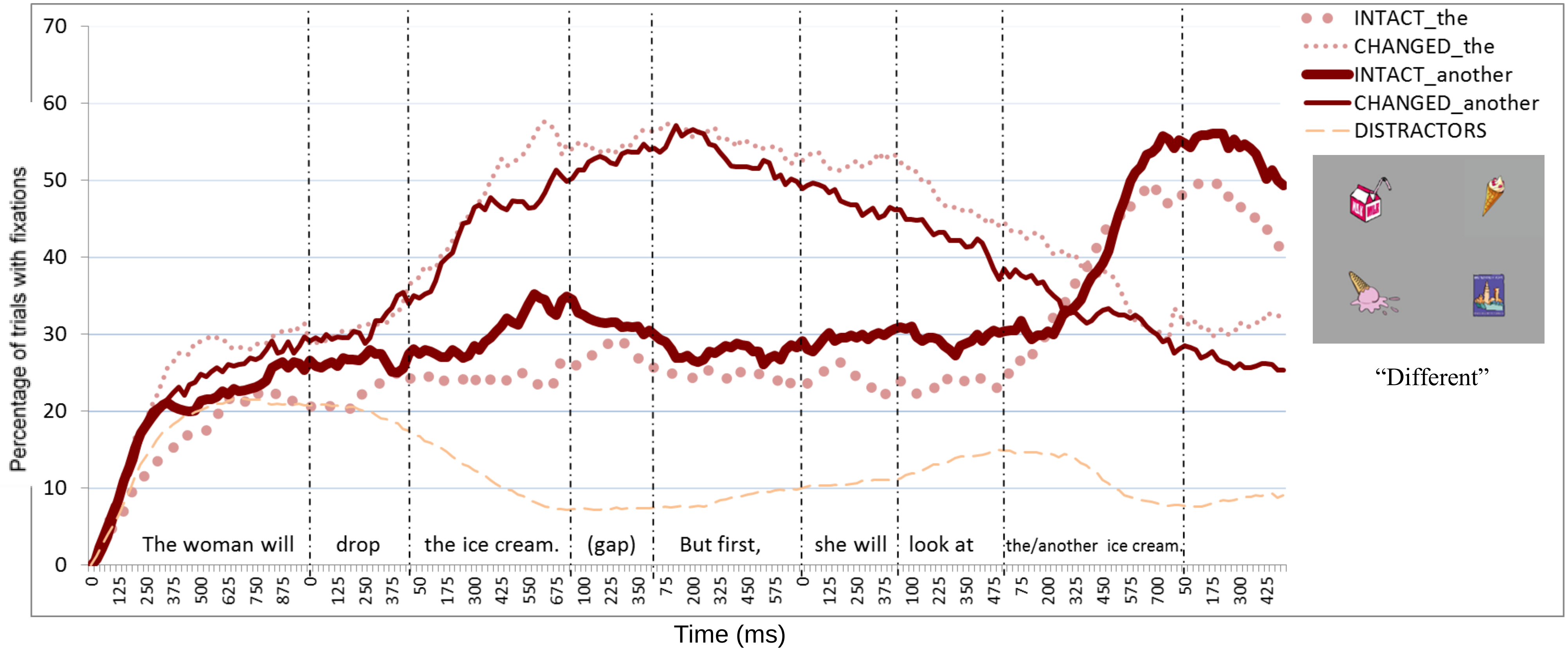
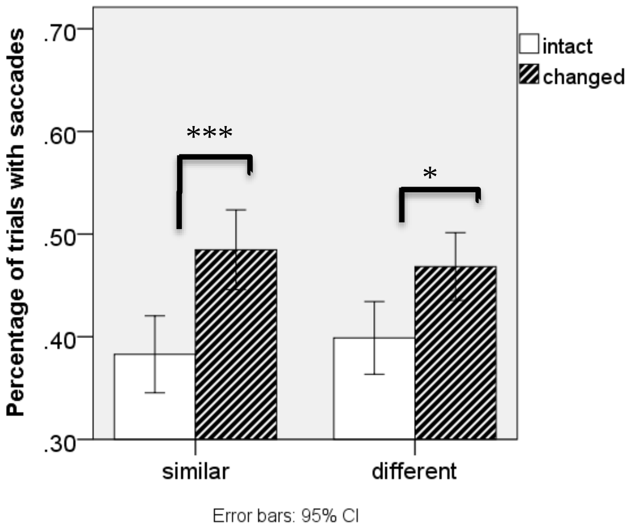


Figure 4.8b Experiment 7: Percentage of trials with fixations on the target regions during the presentation of the auditory stimuli. Two versions of the target object were depicted, which *could not* be possibly associated with the INTACT state and the CHANGED state of the same object, but two DIFFERENT objects sharing the same linguistic label. Fixations were calculated every 25 ms sequentially from the synchronization point at the onset of critical words and phrases, as shown on the x-axis. Y-axis shows mean proportion of fixations that were calculated on a trial-by-trial basis. The straight lines indicate fixations on the two target items when "ANOTHER" was mentioned in the second sentence, while the dotted lines show fixations when "THE" was mentioned. Fixations on the two distractors were averaged and indicated by the dashed line.

4.3.2.2 Saccades

"*The woman will <u>drop the ice cream</u>*"

(i) *During "<u>drop</u>":* There were increasing looks towards the dropped ice cream. A repeated measures two-way ANOVA showed that there was a main effect of the depicted ice creams (more looks at the dropped ice cream than the upright one), $F_1(1,63) = 19.43$, $p < .001$, $\eta_p^2 = .24$, $F_2(1,35) = 10.07$, $p = .003$, $\eta_p^2 = .22$, but there was no main effect of the type of visual displays, $F_1(1,63) = .00$, $p = .989$, $\eta_p^2 = .00$, $F_2(1,35) = .37$, $p = .549$, $\eta_p^2 = .01$, or no interaction, $F_1(1,63) = .84$, $p = .364$, $\eta_p^2 = .01$, $F_2(1,35) = .70$, $p = .409$, $\eta_p^2 = .02$.

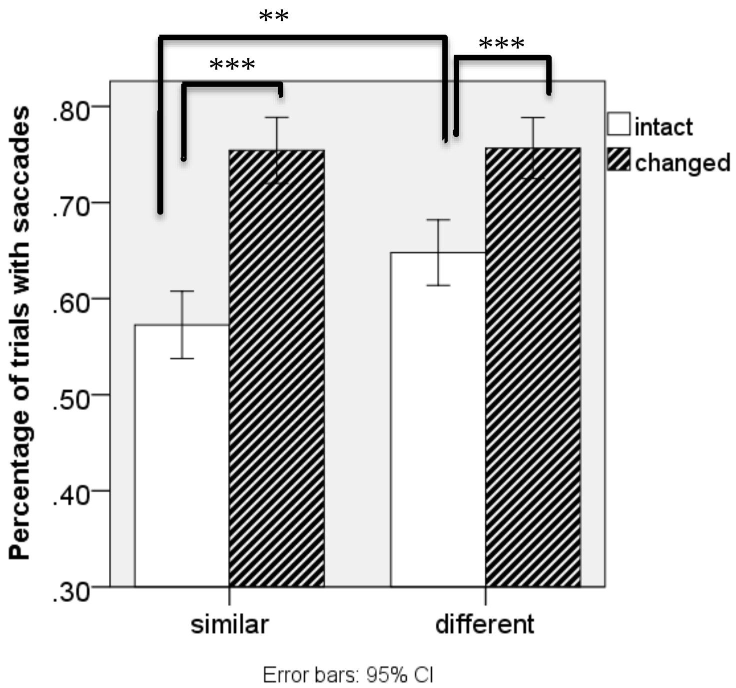


Note: * $p< .05$; ** $p<.01$, *** $p<.001$

Figure 4.9 Experiment 7: Percentage of trials with saccades towards the depicted versions of the target object between the onset and offset of the verb (e.g., "*drop*") (M = 425 ms) in the first sentence. The labels on the x-axis refer to the two types of visual display. "*Similar*" indicates the upright ice cream and the dropped ice cream may be possibility treated as the intact state and the changed state of the ice cream, while "different" means the two ice creams could not be treated as the same object identify but two different tokens. The y-axis shows percentage of trials with saccades that were calculated trial-by-trial and averaged across subjects.

This result is consistent with the fixation data that there were more looks to the dropped ice cream than the upright ice cream regardless the visual displays, as shown in Figure 4.9.

(ii) *During "the ice cream":* Looks towards the dropped ice cream continued to increase. A repeated measures two-way ANOVA showed that there was a significant main effect of the depicted ice creams (more saccades towards the dropped one than the upright one, see Figure 4.10), $F_1(1,63) = 61.61$, $p < .001$, $\eta_p^2 = .49$, $F_2(1,35) = 23.31$, $p < .001$, $\eta_p^2 = .40$, a main effect of the type of visual display by subjects (more looks at the display with two different-looking ice creams than two similar ones), $F_1(1,63) = 6.05$, $p = .017$, $\eta_p^2 = .09$, but not by items, $F_2(1,35) = 3.26$, $p = .080$, $\eta_p^2 = .09$. A significant interaction was found, $F_1(1,63) = 4.34$, $p = .041$, $\eta_p^2 = .06$, $F_2(1,35) = 4.78$, $p = .036$, $\eta_p^2 = .12$. As shown in Figure 4.10, there was no difference in saccades towards the dropped ice cream between the two visual displays, but there were more saccades towards the upright ice cream when the two versions of the target object were less similar, $t_1(63) = 3.56$, $p =.001$, $t_2(35) = 2.99$, $p = .005$. Thus, these findings were consistent with Experiment 4 and Experiment 6.



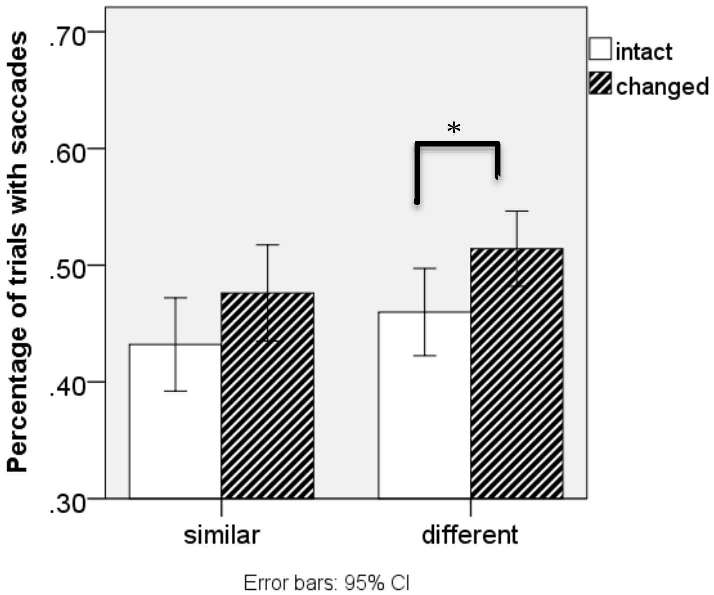
Note: * *p< .05; *** *p<.01,* *** *p<.001*

Figure 4.10 Experiment 7: Percentage of trials with saccades towards the depicted versions of the target object between the onset and offset of the noun phrase (e.g., "*the ice cream*") (*M* = 667 ms) in the first sentence.

In sum, in the first sentence, there were more saccades towards the dropped ice cream than the upright ice cream in both visual displays, replicating the findings in Experiment 4 and Experiment 6. Nevertheless, when comparing across the visual displays, there were more saccades towards the upright ice cream when the two depicted versions of the ice creams had less identical perceptual features.

"<u>*But first*</u>*, she will look at* <u>*the/another ice cream*</u>"

(iii) *During "<u>but first</u>":* A repeated measures two-way ANOVA showed that there was a significant main effect of the depicted ice creams (more saccades towards the dropped one than the upright one, see Figure 4.11), $F_1(1,63) = 5.53$, $p = .022$, $\eta_p^2 = .08$, $F_2(1,35) = 7.91$, $p = .008$, $\eta_p^2 = .18$, no main effect of the type of visual displays, $F_1(1,63) = 2.86$, $p = .096$, $\eta_p^2 = .04$, $F_2(1,35) = 1.89$, $p = .178$, $\eta_p^2 = .05$, or interaction, $F_1(1,63) = .08$, $p = .782$, $\eta_p^2 = .00$, $F_2(1,35) = .01$, $p = .973$, $\eta_p^2 = .00$.

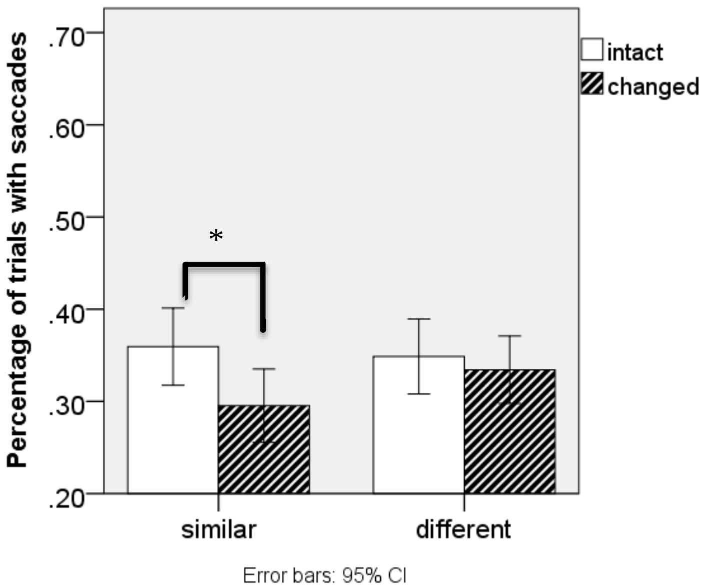


Note: ** p< .05; ** p<.01, *** p<.001*

Figure 4.11 Experiment 7: Percentage of trials with saccades towards the depicted versions of the target object between the onset and offset of "*but first*" (*M* = 685 ms) in the second sentence.

(iv) *During "look at":* Figure 4.12 showed the mean percentage of trials with saccades towards the two versions of the ice creams on each display during this time window. A repeated measures two-way ANOVA showed that there was a marginal main effect of the depicted ice creams (more saccades towards the upright one than the dropped one, see Figure

4.11), $F_1(1,63) = 3.33$, $p = .073$, $\eta_p^2 = .05$, $F_2(1,35) = 6.39$, $p = .016$, $\eta_p^2 = .15$, but no main effect of the type of visual display, $F_1(1,63) = .34$, $p = .563$, $\eta_p^2 = .00$, $F_2(1,35) = .27$, $p = .609$, $\eta_p^2 = .01$, nor interaction, $F_1(1,63) = 2.58$, $p = .113$, $\eta_p^2 = .04$, $F_2(1,35) = 1.48$, $p = .232$, $\eta_p^2 = .04$.

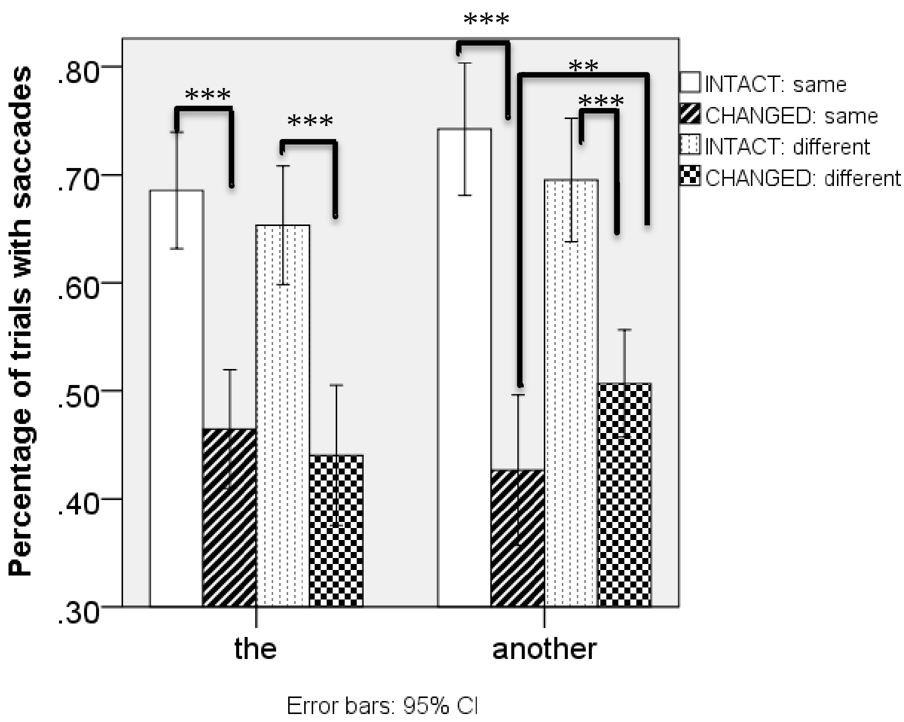


Note: * *p< .05; ** p<.01, *** p<.001*

Figure 4.12 Experiment 7: Percentage of trials with saccades towards the depicted versions of the target object between the onset and offset of "look at" (*M* = 491 ms) in the second sentence.

(v) *During "ice cream":* Due to the difference of duration between "*the*" and "*another*", only the time window for "ice cream" was included for statistical analysis of saccades. A three-way repeated measures ANOVA showed that there was a significant main effect of the depicted ice creams (more saccades towards the upright one than the dropped one, see Figure 4.13), $F_1(1,63) = 67.79$, $p < .001$, $\eta_p^2 = .52$, $F_2(1,35) = 62.72$, $p < .001$, $\eta_p^2 = .64$, but no reliable main effect of the determiner (*the vs. another*), $F_1(1,63) = 3.17$, $p = .080$, $\eta_p^2 = .05$, $F_2(1,35) = 2.81$, $p = .103$, $\eta_p^2 = .07$, nor main effect of the type of the visual displays, $F_1(1,63) = .12$,

p = .735, η_p^2 = .00, $F_2(1,35)$ = .15, p = .697, η_p^2= .01. No interactions were found.



Note: * $p< .05$; ** $p<.01$, *** $p<.001$

Figure 4.13 Experiment 7: Percentage of trials with saccades towards the depicted versions of the target object between the onset and offset of the final noun ("ice cream") (M = 512 ms) in the second sentence.

In sum, in the second sentence saccades switched from the dropped ice cream to the upright ice cream on the visual displays for both "the ice cream" and "another ice cream".

4.3.3 Discussion

The results from Experiment 7 showed that participants' eye movements towards the visual displays were influenced by the linguistic context. Consistent with previous eye-tracking experiments in this thesis, there were more fixations on the dropped ice cream than the upright ice cream at the beginning of the first sentence. Looks towards the dropped ice cream continued to increase when a change of state action ("*drop*") was described to happen on the ice cream. In the second sentence, after "*but first*" was mentioned, participants started to look away from the dropped ice cream and towards the upright ice cream. When "*the ice cream*" was referred to at the end of the second sentence, there were more saccades towards the upright ice cream than the dropped ice cream on both visual displays.

Similarly, when "*another ice cream*" was mentioned, there were also more saccades towards the dropped ice cream than the upright one regardless of the visual displays. Thus, it seems that participants were not treating the upright and dropped ice creams as conflicting states of the same ice cream and did not bind them to the same object identity (4.2a). Otherwise, it would be difficult for them to switch to the upright ice cream when "*another*" was mentioned. Instead, eye movements seemed to be predominantly guided by semantic overlap as there were increasing saccades towards the upright ice cream on both visual displays regardless of the determiners ("the"/"another"). Nonetheless, a difference in saccades towards the upright ice cream and the dropped ice cream emerged when the verb ("look at") was mentioned in the second sentence for (4.2a) but not in (4.2b). Thus, although the results were not consistent with the prediction of the binding account, participants did seem able to switch to the upright ice cream earlier when the depicted versions were more similar to each other.

These findings suggest that semantic mapping is also important for discourse comprehension, which is consistent with studies on individual words (Huettig & Altmann, 2005; Yee & Sedivy, 2001, 2006). To conclude, Experiment

7 suggested that eye movements on the visual scene were driven by semantic overlap between the depicted version of the object and the described object-state representation.

One limitation of Experiment 7, however, is that only a single intact version of the target object was depicted in the visual display. Thus, participants had no alternative but to opt for that particular upright ice cream when they heard "another". Experiment 8 addresses this.

4.4 Experiment 8

In Experiment 6 and Experiment 7, we have demonstrated that sematic overlap between linguistic and visual information is likely to drive eye movements towards a specific version of the target object on the visual display. For example, when a change of state was described (e.g., "the woman will drop the *ice cream*"), there were more eye movements towards a dropped ice cream than an upright one. However, this pattern of eye movements reversed in the subsequent discourse when "*the/another ice cream*" was mentioned, as participants were directed to travel back in time ("*but first*") and retrieve the intact version, i.e. an upright ice cream. However, when they heard "*another ice cream*", they most likely had an upright ice cream in mind, which may or may not have the same perceptual features as the one in the visual display. As there was only one upright ice cream on the visual display, participants had no alternative but to look at that ice cream when looking for "another ice cream". Hence, in Experiment 8, we depicted two intact ice creams along with one dropped ice cream (matching one of the intact ones). Distractors were also included (See Figure 4.3). The same auditory stimuli in Experiment 7 were used. The “and then” condition was added so as to compare how the influence of temporal shifts on eye movements(e.g., *The woman will drop the ice cream. But first/And then, she will look at the/another ice cream*).

Therefore, with this manipulation in Experiment 8, we were able to directly compare eye movements towards three different versions on the visual display as the language unfolded. This was particularly useful for identifying which version of the two upright ice creams could be considered as “*another ice cream*”. On the one hand, we may assume no difference between (a) and (b), since both versions were intact and may match the default object-state representation of “another ice cream”. On the other hand, we may predict that (a) would be preferred to (b) for “another ice cream”, since new information is integrated into the current situation model in which (c) is the current object-state representation. Also, (a) shared more “episodic characteristics” of (c) than (b) did. Thus, perhaps it is easier to modify (c) into (a) than (b). As for “the ice cream”, we may expect more eye movements towards (a) than (b) regardless of whether there was an influence of the current situation model since (b) had fewer overlapping semantic features with (c) compared with (a).

4.4.1 Method

Participants. Thirty-eight students[8] (22 female, mean age 20, range 18-23 years old) from the University of York participated in this study, receiving 1-hour course credit or a payment of three pounds for their contribution. All participants were native speakers of British English and had normal or corrected to normal vision. None of the participants in this study had taken part in previous

[8] There were fewer participants in Experiment 8 (N=38) than other experiments (N=64). Due to time constraints it was not possible to run the intended number of participants. The conclusions, given the data, should accordingly be treated with a degree of caution.

experiments.

Stimuli. Six items were depicted on each visual display, as shown in Figure 4.3. The auditory stimuli were shown in following examples:

(a) The woman will drop the ice cream. But first, she will look at the ice cream.

(b) The woman will drop the ice cream. But first, she will look at another ice cream.

(c) The woman will drop the ice cream. And then, she will look at the ice cream.

(d) The woman will drop the ice cream. And then, she will look at another ice cream.

Procedure. The procedure was the same as in Experiment 7, except that the calibration was based on 13 points rather than 9 points. The experiment was run on an Eyelink II remote eye-tracker, which sampled at 250 Hz from the right eye. The total length of the experiment was about 45 minutes.

4.4.2 Results

Data were analyzed by the following steps as in previous experiments. Both fixations and saccades on the visual display were reported by using the same critical time windows as in previous experiments.

4.4.2.1 Fixations

Figure 4.14a and Figure 4.14b illustrate proportion of fixations on items on the visual display. At the onset of the verb in the first sentence, there were more fixations on the dropped ice cream than the upright ice creams, replicating previous experiments. This difference became greater as the first sentence unfolded: At the offset of the verb when participants had just heard "*The woman will drop* … ".

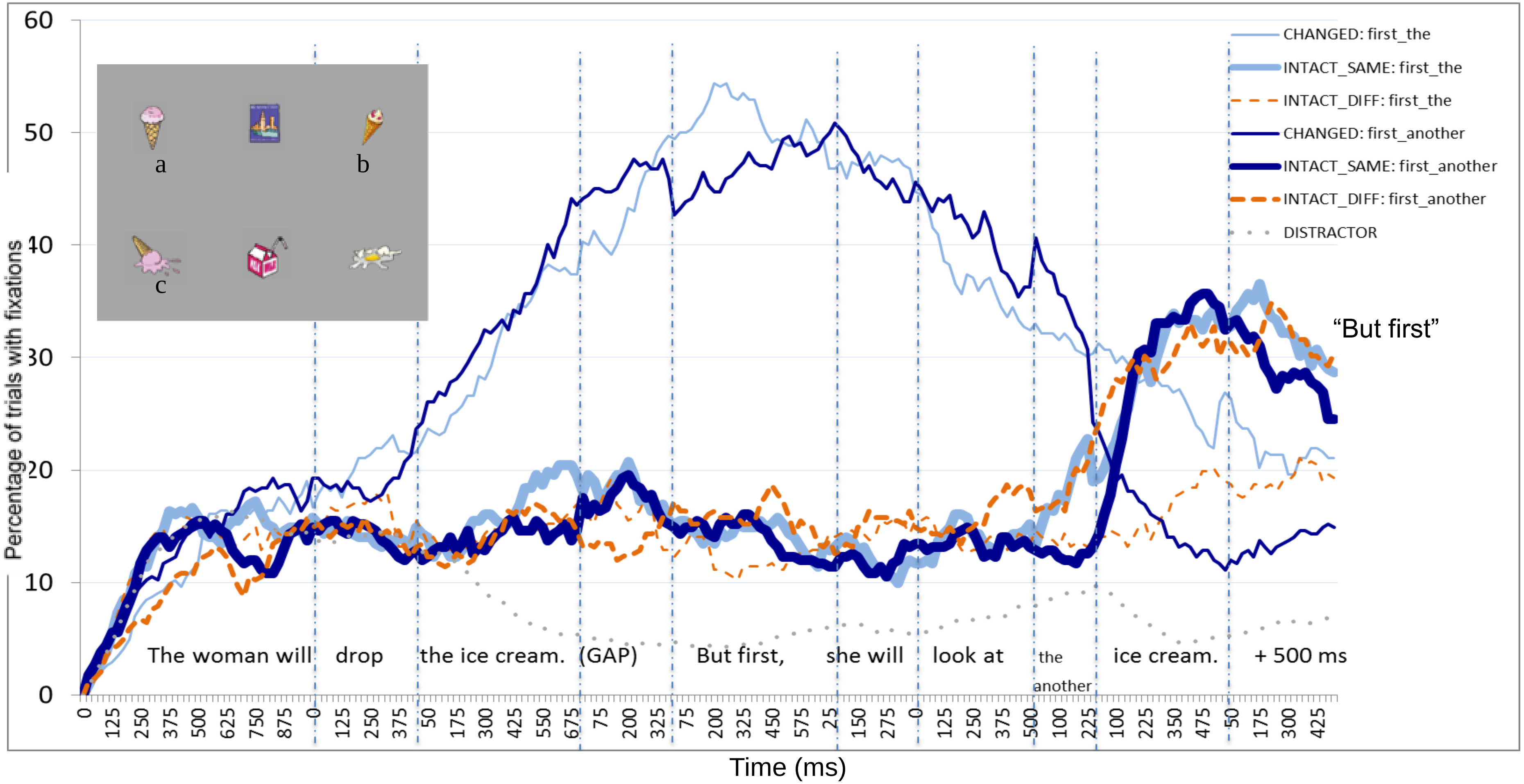


Figure 4.14a Experiment 8: Percentage of trials with fixations on items on the visual display when "BUT FIRST" was mentioned in the second sentence. Fixations were calculated every 25 ms sequentially from the synchronization point at the onset of critical words and phrases, as shown on the x-axis. Y-axis shows mean proportion of fixations that were calculated on a trial-by-trial basis. The straight blue lines indicate fixations on INTACT_SAME (a) and CHANGED (c), while the dashed orange lines reveal fixations on INTACT_DIFF(ERENT) (b). The dotted grey lines the averaged fixations on unrelated distractors.

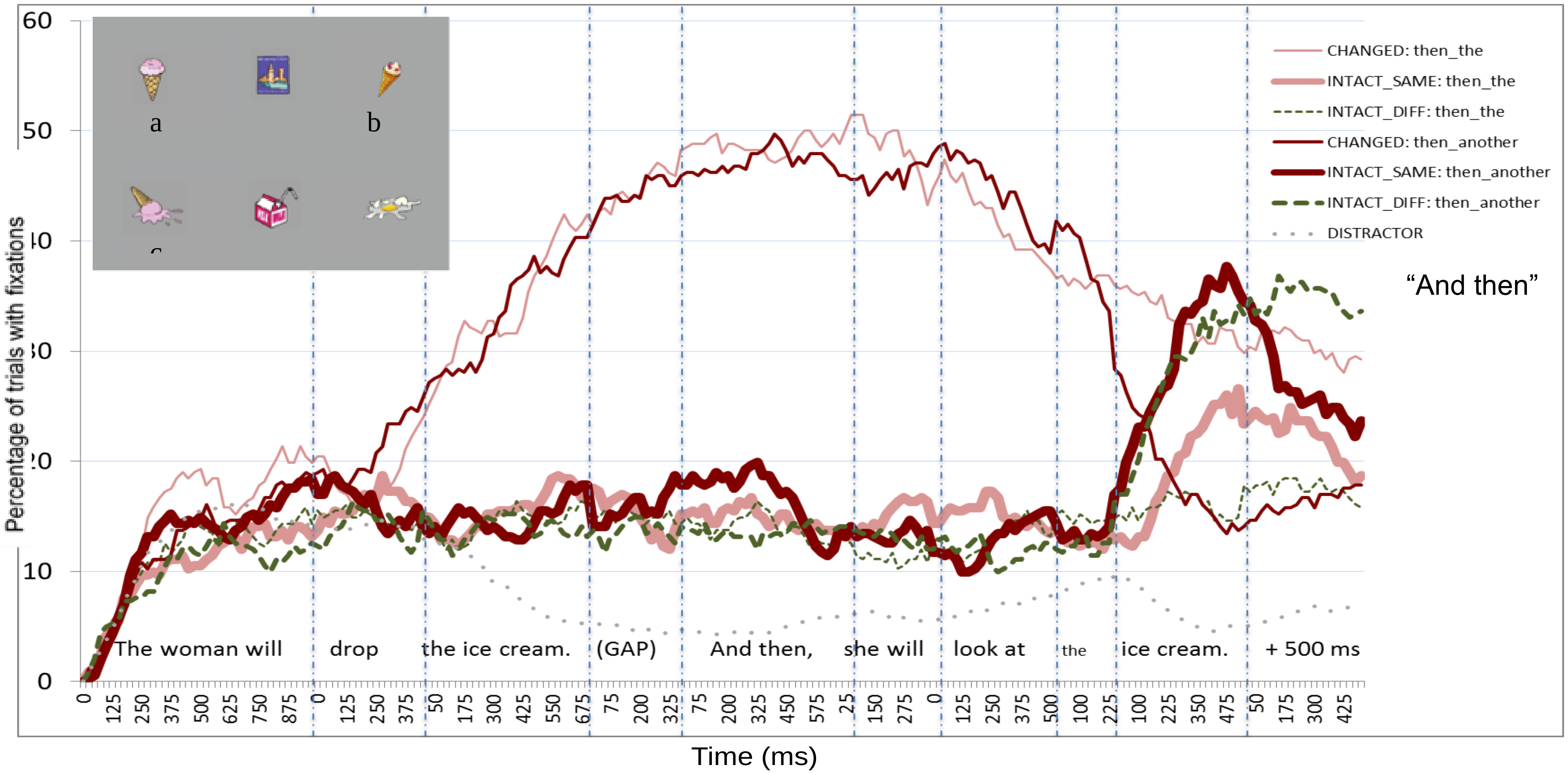


Figure 4.14b Experiment 8: Percentage of trials with fixations on items on the visual display when "BUT FIRST" was mentioned in the second sentence. Fixations were calculated every 25 ms sequentially from the synchronization point at the onset of critical words and phrases, as shown on the x-axis. Y-axis shows mean proportion of fixations that were calculated on a trial-by-trial basis. The straight blue lines indicate fixations on INTACT_SAME (a) and CHANGED (c), while the dashed orange lines reveal fixations on INTACT_DIFF(ERENT) (b). The dotted grey lines the averaged fixations on unrelated distractors.

In the second sentence, despite decreasing fixations on the dropped ice cream after “but first” was mentioned, fixations on the upright ice creams did not exceed the dropped one until “the/another ice cream” was mentioned, suggesting a late influence of the linguistic context. Specically, when “*but first*” was mentioned (see Figure 4.14a), there was no difference in fixations on (a) INTACT_SAME between “the ice cream” and “another ice cream” at the offset of “ice cream”, but more fixations on (b) INTACT_DIFFERENT for “another ice cream” than “the ice cream”. This pattern indicated that (a) was considered as an appropriate candidate of object representation for both “the” and “another” ice creams for “but first” condition. However, (b) INTACT_DIFFERENT was treated as a more appropriate representation for “another” than “the” ice cream. By contrast, when “*and then*” was mentioned (see Figure 4.14b), there was no difference in fixations on (a) INTACT_SAME and (b) INTACT_DIFFERENT. Thus, this result indicated that both (a) INTACT_SAME and (b) INTACT_DIFFERENT appropriate object-state representation for “another ice cream” and “the ice cream” for “and then” condition. Interestialy, overall there were more fixations on both (a) INTACT_SAME and (b) INTACT_DIFFERENT in “another ice cream” than “the ice cream”.

In sum, fixations on the depicted ice creams suggested that participants predominantly looked at the dropped ice cream rather than at the upright ones in the first sentence. This finding is consistent with previous eye-tracking experiments reported in this thesis. In the second sentence, when “but first, the” was mentioned, there were increasing looks towards both upright ice creams, but the upright one that was more similar to the dropped one attracted more fixations than the less similar one for “the ice cream”. The opposite pattern was found for “but first, another”, with more fixations on the different-looking ice cream than the same-looking one. As for “and then, the/another”, both upright ice creams were considered as more appropriate for “another ice cream” than for “the ice cream”, because the former requires any intact ice cream, while the latter requires continued attention to the dropped ice cream.

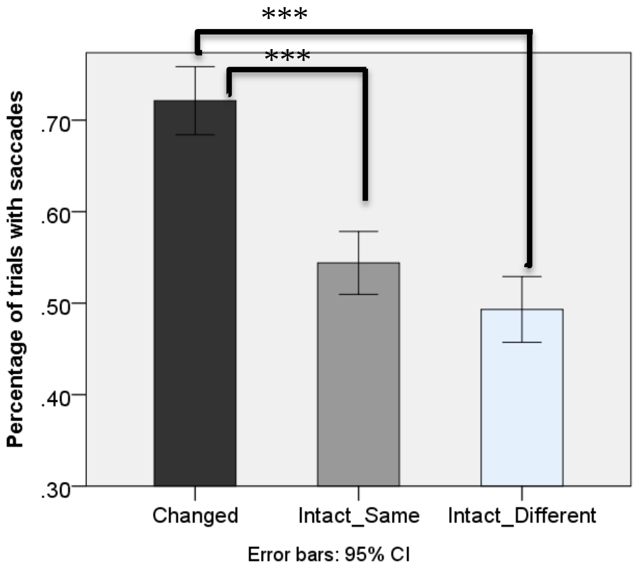
4.4.2.2 Saccades

“*The woman will* <u>*drop*</u> *the ice cream*”

(i) *During “*<u>*drop*</u>*”:* A repeated measures one-way ANOVA showed that there was no main effect of the depicted ice creams, $F_1(2,36) = 1.25$, $p = .290$, $\eta_p^2 = .02$, $F_2(2,34) = 2.23$, $p = .124$, $\eta_p^2 = .12$, suggesting no difference of saccades towards the ice creams.

(ii) *During ‘*<u>*the ice cream*</u>*”:* A repeated measures one-way ANOVA showed *that* there was a significant main effect of the depicted ice creams on saccades, $F_1(2,36) = 45.33$, $p < .001$, $\eta_p^2 = .72$, $F_2(2,34) = 15.36$, $p < .001$, $\eta_p^2 = .48$. Pairwise comparisons suggested that there were more saccades towards the dropped ice cream than the upright ice creams, with $t_1(37) = 7.19$, $p < .001$, t_2 $(35) = 4.25$, $p <.001$, $t_1(37) = 9.30$, $p < .001$, t_2 $(35) = 5.54$, $p <.001$, respectively (see Figure 4.15).

In sum, in the first sentence, there were more saccades towards the dropped ice cream than the upright ice cream. This pattern replicated preivous findings in Experiment 4-7.

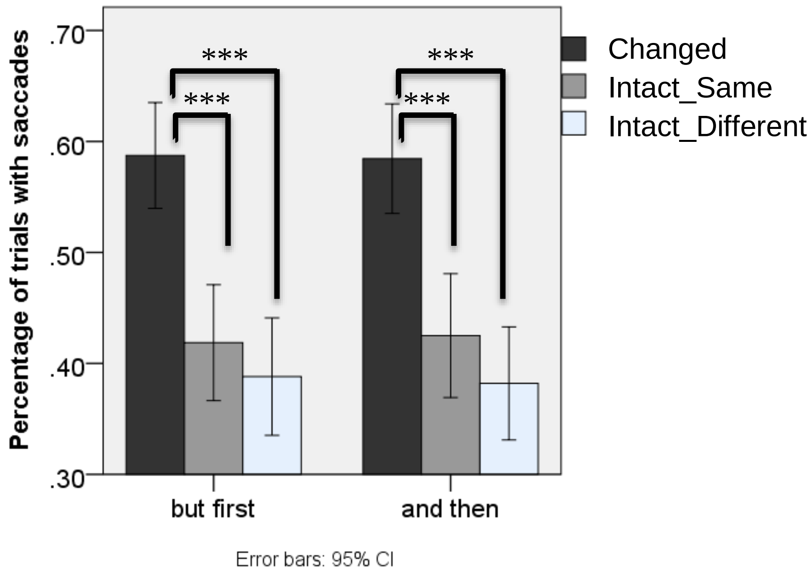


Note: * *p< .05; *** *p<.01, **** *p<.001*

Figure 4.15 Experiment 8: Percentage of trials with saccades towards the items on the visual display when the target object was mentioned (e.g., "the ice cream") (*M* = 667 ms) in the first sentence. "Changed" refers to the dropped ice cream; "Intact_Same" refers to the upright ice cream that is more similar to the dropped one, while "Intact_Different" refers the other upright ice cream that is less similar to the dropped one.

"*<u>But first/And then</u>, she will look at the/another <u>ice cream</u>*"

(iii) *During "<u>but first'/and then</u>":* A repeated measures two-way ANOVA showed that there was a significant main effect of the depicted ice creams (more saccades towards the dropped one than the upright ones. See Figure 4.16), $F_1(2,36) = 33.56$, $p < .001$, $\eta_p^2 = .65$, $F_2(2,34) = 23.25$, $p < .001$, $\eta_p^2 = .58$, but there was no significant main effect of the temporal shifts, $F_1(1,37) = .002$, $p = .961$, $\eta_p^2 = .00$, $F_2(1,35) = .05$, $p = .823$, $\eta_p^2 = .001$, nor interaction, $F_1(2,36) = .04$, $p = .961$, $\eta_p^2 = .002$, $F_2(2,34) = .56$, $p = .576$, $\eta_p^2 = .03$. Thus, the pattern of eye movements was similar to the first sentence, suggesting no immediate influence of the temporal shifts on eye movements. This result was also consistent with previous findings in Experiment 4-7.



Note: * *p< .05; ** p<.01, *** p<.001*

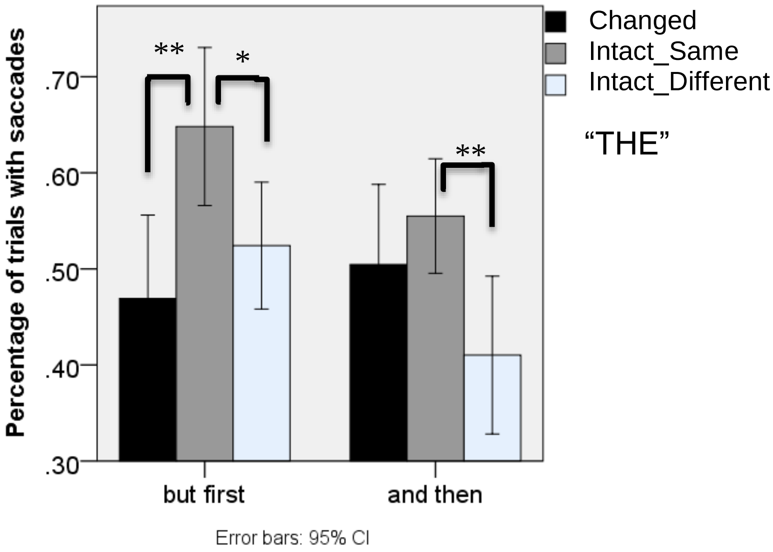
Figure 4.16 Experiment 8: Percentage of trials with saccades towards the items on the visual display during temporal shifts ("*but first/and then*") (*M* = 487 ms) in the second sentence. "Changed" refers to the dropped ice cream; "Intact_Same" refers to the upright ice cream that is more similar to the dropped one, while "Intact_Different" refers the other upright ice cream that is less similar to the dropped one.

(iv) *During "look at"*: A repeated measures two-way ANOVA showed that there was a significant main effect of the depicted ice creams (more saccades towards the dropped one than the upright ones), $F_1(2,36) = 5.87$, $p = .004$, $\eta_p^2 = .14$, $F_2(2,34) = 19.52$, $p < .001$, $\eta_p^2 = .48$, but there was no significant main effect of the temporal shifts, $F_1(1,37) = .28$, $p = .598$, $\eta_p^2 = .004$, $F_2(1,35) = .14$, $p = .756$, $\eta_p^2 = .01$, nor interaction, $F_1(2,36) = .47$, $p = .626$, $\eta_p^2 = .01$, $F_2(2,34) = .50$, $p = .720$, $\eta_p^2 = .04$, suggesting the pattern did not change from the last time window (*"but first/and then"*), thus no anticipatory eye movements towards the appropriate version of the ice cream.

(v) *During "ice cream":* This is the most important time window in

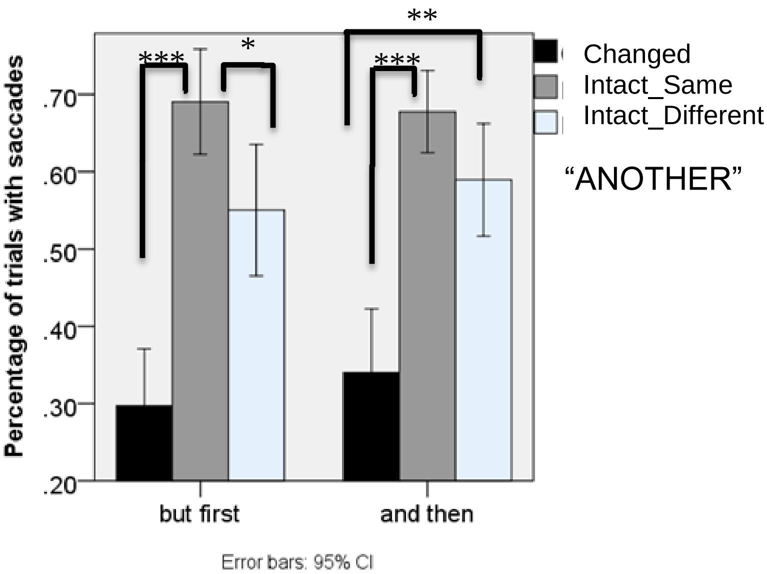
Experiment 8. Figure 4.17a and Figure 4.17b illustrate the proportion of trials with saccades towards the ice creams. Since we have cross-spliced the same words in each set of the stimuli, the duration of the final noun was identical across conditions. A repeated measures three-way ANOVA showed that there was a significant main effect of the depicted ice creams (more saccades overall towards the same-looking upright ice cream than the different-looking one), $F_1(2,36) = 55.36$, $p < .001$, $\eta_p^2 = .83$, $F_2(2,34) = 50.29$, $p < .001$, $\eta_p^2 = .82$, but there was no main effect of the determiners, $F_1(1,37) = .58$, $p = .452$, $\eta_p^2 = .02$, $F_2(1,35) = .84$, $p = .367$, $\eta_p^2 = .02$, or temporal shifts, $F_1(1,37) = .12$, $p = .740$, $\eta_p^2 = .00$, $F_2(1,35) = .00$, $p = .985$, $\eta_p^2 = .00$. A significant interaction between the depicted ice creams and the determiners was found, $F_1(2,36) = 8.21$, $p < .001$, $\eta_p^2 = .41$, $F_2(2,34) = 10.14$, $p < .001$, $\eta_p^2 = .48$.

Planned comparisions were conducted by separting the conditions based on the determiners ("the" vs. "another"). As shown in Figure 4.17a below, when the linguistic context was "*but first, the*", there were most saccades towards the same-looking upright ice cream than the different-looking one. Nonetheless, as shown in Figure 4.17b, when "*but first, another*" was mentioned, there were more saccades towards the same-looking upright ice cream than the different-looking one. These results suggest that when participants needed to retrieve an intact object-state representation as indicated by "*but first*", they tended to associate it with the same-looking upright ice cream during both "the/another ice cream".



Note: * *p< .05; ** p<.01, *** p<.001*

Figure 4.17a Experiment 8: Percentage of trials with saccades towards the items on the visual display during the noun ("ice cream") (*M* = 513 ms) after "the" was mentioned in the second sentence.



Note: * *p< .05; ** p<.01, *** p<.001*

Figure 4.17b Experiment 8: Percentage of trials with saccades towards the items on the visual display during the noun ("ice cream") (*M* = 513 ms) after "another" was mentioned in the second sentence. "Changed" refers to the dropped ice cream; "Intact_Same" refers to the upright ice cream that is more similar to the dropped one, while "Intact_Different" refers the other upright ice cream that is less similar to the dropped one.

By comparison, when "*and then*" was mentioned, there were also more saccades towards the same-looking upright ice cream than the different-looking for one during "the ice cream". However, there was no difference in saccades between the two upright ice creams during "another ice cream", suggesting both of them were likely to be treated as "another" one.

In sum, data from the second sentence showed that immediatey after "*but first*" or "*and then*" was mentioned, saccades were not influenced by the described temporal shifts. Instead, the pattern of saccades was similar to the first sentence with more saccades towards the dropped ice creams than other the other two. During "ice cream", when "*but first*" was mentioned at the beginning of the second sentence, there were more saccades towards both upright ice creams than the dropped one regardless of the determiners. There were also more saccades towards the same-looking one than the different-looking one for both "the/another ice cream". However, when "*and then*" was mentioned at the beginning of the second sentence, during "ice cream", there were more saccades overall towards the same-looking upright ice cream than the different-looking one for "the", but not for "another"

4.4.3 Discussion

Experiment 8 intended to further explore whether semantic overlap between described object-state representation and the depicted versions drove eye movements on the visual display. Similar to previous experiments, participants listened to auditory stimuli while their eye movements on the visual stimuli were monitored. There were two sentences in the auditory stimuli. The first sentence described a change of state event, and the second sentence began with a backward temporal shift ("*but first*") or a forward temporal shift ("*and then*") and referred to either "*the*" or "*another*" target object. On the visual display, three versions of the target object were depicted, including a changed one (e.g., the dropped ice cream), an intact one (e.g., the upright ice cream that is similar to the dropped one), and an intact competitor (e.g., the upright ice cream that is different from the dropped one).

Eye movements in the first sentence replicated previous experiments with more looks towards the dropped ice cream than towards the upright ones. This pattern suggests that participants were sensitive to the consequences of the change of state event. In the second sentence, when "*but first/and then*" was mentioned the eye movement pattern was similar to the first sentence, suggesting no immediate influence of temporal shifts. However, when "ice cream" was subsequently referred to, there were more eye movements towards the upright one that was more similar to the dropped one than towards the different-looking upright one after both "but first, the" and "but first, another". By contrast, if "*and then*" was mentioned at the beginning of the second sentence, there were no difference in eye movements between the two upright ice creams after "and then, another", but more saccades towards the upright one that was more similar to the dropped one than towards the less similar one after "and then, the"

Therefore, it seems that participants tended to look at both upright ice creams for "another ice cream". This finding suggests that when "*another ice cream*" was mentioned, participants had an intact version in mind; the canonical object-state representation (See Experiment 2). Rather than looking at the different-looking upright ice cream, they looked at the same-looking one. This result is consistent with Experiment 7 that participants were not binding the dropped and upright ice creams as belonging to the same object identity. Instead, “another ice cream” may have inherited the episodic characteristics of the dropped ice cream in the previous sentence. Perhaps, participants established an object-state representation for “another ice cream” based on existing situation models. Similarly, participants preferred the same-looking ice cream to the different-looking for both “*but first, the*” and “*and then, the*”, probably due to the inheritance of semantic features in the episodic memory rather than binding the same-looking ice cream and the dropped one as the same object identity.

4.5 General Discussion

In the three studies reported in this chapter, visual and linguistic contexts were manipulated as a target object was mentioned in two sentences. We examined whether semantic overlap between the representation in mind and the depicted versions of the target object on the visual scene drove the switch in looks from one version to the other, or whether it was due to identity binding of them.

First, the experiments replicated results of Experiment 4 & 5. Initially, participants were more interested in the dropped ice cream than the upright one at the onset of the verb in the first sentence. This difference was increased by the description of a change of state event (e.g., *drop the ice cream*). Subsequently, eye movements switched to the upright ice cream from the dropped one due to the introduction of a backward temporal shift (“*but first*”). However, this effect was not manifest immediately, but later when “ice cream” was mentioned again. This is consistent with Altmann & Kamide (2009) that eye movements can be directed to a particularly depicted item on the visual display, at least in part, by a dynamically modifiable representation of the object.

Second, we found that semantic overlap alone could not account for the effect. In Experiment 6, two tokens of the ice cream were depicted on the visual display. One was in the intact state (e.g., an upright ice cream) and the other one in the changed state (e.g., a dropped ice cream), but the two versions could not be treated as conflicting states of the same object identity due to differences in perceptual features. As predicted, there were increasing looks towards the upright ice cream at the end of the second sentence. However, the upright ice cream did not receive more saccades than the dropped one, suggesting reduced preference in the different-looking upright ice cream.

This lead to Experiment 7, where a new token of the ice cream was introduced at the end of the second sentence by saying “another ice cream”. In this way, we were able to find out how this new token of ice cream drove eye movements on both type of visual displays in Experiment 4 (two ice creams that might be bound to a single object identity) and Experiment 6 (two ice creams that could not be bound to the same object). We found that eye movements were not only driven by the overlap of semantic features between the noun phrase and the current visual display, but also inherited features in the previous sentence. For example, after “*But first/And then, she will look at another…*” was mentioned, the same eye movement patterns were found for the visual scenes, suggesting participants were treating both the upright ice creams as “another ice cream”. Nonetheless, an alternative account may be the visual constraint. When there was only one intact version on the visual scene (Experiment 6 &

Experiment 7), eye movements could only be directed to the upright ice cream that was available on the visual scene.

This was further explored in Experiment 8, which demonstrated that eye movements were directed to the object that matched the new information and that was similar to the previous object. When there were two upright ice creams on the visual display, eye movements were directed to the one that resembled previously constructed object-state representation (i.e. the dropped one) (Experiment 8). This finding was consistent with the idea that object-state representation was modified on the basis of its previous mental representation, as proposed by Zwaan and Radvansky (1998).

4.6 Conclusion

To conclude, Experiment 6, Experiment 7, and Experiment 8 revealed eye movements on the visual scene were driven by semantic overlap between dynamically modifiable object-state representation and the depicted versions of the object in the visual scene, which were also influenced by the previous object-state representation in mind.

Chapter 5

Event-related brain potentials reflect processing of object-state representations in language comprehension

5.1 Overview

We have so far demonstrated that a single object can be associated with different mental representations depending on the linguistic context. With the sentence-picture verification paradigm, we found that participants reacted faster to a depicted version of the target object when it matched what was described in the language than mismatched. For instance, when an ice cream was described to experience a substantial change (e.g., "*The woman will drop the ice cream"*), response latencies to a dropped ice cream was shorter than when it was described to experience a minimal change (e.g., "*The woman will choose the ice cream"*). Subsequently, with the visual world paradigm, we further demonstrated that eye movements towards conflicting versions of a target object could be modulated as language unfolded. Participants tended to look the version that matched the consequence of the described action. When a later sentence refer back to the object introduced in an earlier sentence, participants tended to switch look at an item that matched both prior and current contexts, suggesting object representations are tracked in real-time language processing.

However, it is not clear what cognitive processes are involved to keep track of an object's representation. In this chapter, we report an experiment using the

Event Related Potentials (ERP) to examine the electrophysiological evidence of tracking an object's representations in language comprehension. Specifically, we explored whether a change-of-state event (e.g., *"The woman will drop the ice cream"*) would elicit a different ERP relative to a no-change-of-state event (e.g., "*The woman will choose the ice cream*") in the subsequent sentence (e.g., "*And then, she will look at the ice cream*"), and if so, to what extent this effect resembles ERP effects observed in other studies. Extant models of language comprehension do not lead to strong predictions about what ERP components to expect.

To further explore whether keeping track of object representation and conflict resolution is correlated in neural activity, we additionally administered a modified Stroop colour identification task based on Hindy et al. (2012). In this task, the response is limited to three colours (yellow, green, and blue). Participants were instructed to press a key corresponding to the typeface color of the word. There were three types of trials: congruent trials, response-eligible incongruent trials, and response-ineligible incongruent trials. For response-eligible incongruent trials, the word (yellow, green, or blue) matched one of the response keys, but mismatched with the typeface colour. For example, the word was "YELLOW" in the font colour of blue. For response-ineligible incongruent trials, the word (orange, lime, or red) mismatched the response keys and the typeface colour. For example, the word was "RED" in the font colour of yellow. This manipulation allows for the separation of the source of conflict. In the response-eligible condition, participants may push a wrong button because of the meaning of the printed word matches one of the response keys. Thus, the conflict may come from both response and representational levels. In the response-ineligible condition, the meaning of the target word is not shown in the response keys, but it may still competes with the font colour at a representational level. Stroop conflict at the representational level has been shown to predict conflict during language processing, e.g., syntactic ambiguity resolution (January, Trueswell, & Thompson-Schill, 2009) and object representational conflict (Hindy et al., 2012). Correlation analyses between the two kinds of Stroop effect and the ERPs in separate time windows were thus conducted to explore whether resolution of Stroop conflict was related to ERPs in current experiment.

The goal of the present experiment is to examine the ERPs that may associate with keeping track of object-states in language comprehension. In order to understand a sentence such as “She will look at the ice cream”, we must engage processing not only at semantic and syntactic levels but also at discourse level (e.g., what is talked about). Recent fMRI studies (Hindy et al., 2012) and previous experiments in this thesis further demonstrated that “what is talked about” is not limited to different objects, but also related to the state of the object itself.

Figuring out object representation may be an issue of conflict resolution. As shown by Hindy et al. (2012) understanding sentences describing a substantial change of state activate the same brain region (pVLPFC) involved in Stroop conflict. The N400 effect, a negative-going component with a broad centro-parietal scalp distribution, is the ERP component that was discovered to associate with conflict resolution at the semantic level (Brown and Hagoort, 1993; Kutas & Federmeier, 2000; Kutas, Van Petten & Kluender, 2006). For example, the semantically anomalous condition (e.g., *He spread his warm* *bread with socks)* elicited a more negative going waveform compared to semantically coherent condition (e.g., *He spread his warm bread with butter)* (Kutas and Hillyard, 1980). Moreover, N400 is associated with conflict resolution at the discourse level (Nigam, Hoffman, & Simons, 1992). Compared to related/matched word or picture, an unrelated/mismatched one elicited the N400 (Nigam et al., 1992), regardless whether the picture was presented prior to sentence reading (Coppens, Gootjes, & Zwaan, 2012) or afterwards (Knoeferle, Urbach, & Kutas, 2011).

Alternatively, referential ambiguity resolution may be involved in keeping track of object representation. As shown in previous eye tracking experiments, at least two object representations (the changed state vs. the intact state) can be activated when the target object is described to experience a substantial change. When the object is encountered in a subsequent sentence, perhaps we have to choose beween these two representations. If this is the case, we may find the Nref effect of referential ambiguity resolution (Nieuwland & Van Berkum, 2008; Van Berkum, et al., 1999; Van Berkum, Brown, Hagoort, and Zwitserlood, 2003; Yu, Zhang, & Boland, 2015).

Van Berkum, et al. (1999) asked participants to read short stories, such as the following texts (English translation of Dutch stimuli):

(1a) *David had asked* ***the boy and the girl*** *to clean up their room before lunchtime. But the boy had stayed in bed all morning, and the girl had been on the phone all the time. David told the* ***girl*** *that had been on the phone to hang up.*

(1b) David had asked the ***two girls*** *to clean up their room before lunchtime. But one of the girls had stayed in bed all morning, and the other had been*

on the phone all the time. David told the ***girl*** *that had been on the phone to hang up.*

In (1a), the critical noun "girl" had only one eligible antecedent. In (1b) there were two equally possible referents, but only one eligible antecedent is possible for the referent under the language context. This manipulation leads to a clear and reliable effect of negative ERP waveforms emerging about 300 ms after onset of the referentially ambiguous nouns (as shown in 1b) largely at the anterior sites. This effect has been replicated when using pronouns and when stimuli were presented in auditory format (Van Berkum, et. al, 2003).

Nonetheless, unlike individual entities, multiple representations of the same object may not co-exist. Perhaps, selective retrieval of an object representation is required when the object name is encountered, which may evoke the old/new parietal effect. Studies of recognition memory have revealed left posterior distributions of negative-going ERPs about 400 – 800 ms after critical word onset when participants correctly retrieved details or other contextual information. In these studies, participants were required to identify whether the word/object was presented before (e.g., Curran, 2000; Cycowicz, Friedman, Snodgrass, 2001; Johansson, Stenberg, Lindgren, & Rosén, 2002; Johansson & Mecklinger, 2003; Senkfor and van Petten, 1998) or they had to complete a recognition test of actions (e.g., Leynes, Crawford, & Bink, 2005; Leynes, Grey, & Crawform, 2006). As Johansson & Mecklinger (2003) pointed out, this left late posterior negativity might reflect "search for and/or retrieval/evaluation of the attributes in modality-specific cortical regions". Importantly, recent studies suggested that the old/new effect can also be evoked by old items (Bridger, Sprondel, & Mecklinger, 2015; Herron & Rugg, 2003). Bridger et al. (2015)

presented items in two different context. Participants were required to accept items from a specific context as old (targets) but to reject items from the other class (nontargets), the found a larger old/new effect for targets than nontargets.

The above-mentioned ERP effects are different in time windows and scalp distributions. The N400 often begins about 250 ms and peaks about 400 ms at the centroparietal site after poststimulus onset. By contrast, Nref is a late effect that peaks at anterior sites between 400 -1000 ms. The left parietal old/new effect usually appears about 500-800 ms post-stimulus onset, but it peaks at the left parietal sites. To summarize the predictions: first, the N400 effect may be expected after the target object is described to experience a substantial change compared with no change due conflict resolution; second, we shall see an Nref effect for the substantial change due to the activation of two competing states of the same object, but not for no change; third, perhaps the left parietal old/new effect will be found due to selective retrieval of object representations. N400 effect is expected to appear 300-500 ms post stimulus onset over centralparietal sites, but Nref and the old/new effects are about 400-1000 ms over anterior and left posterior, respectively.

5.2 Method

Participants

After giving informed consent, 29 students from the University of York (19 females, 18-24 years old) participated in this study for 10 pounds of payment or 2 hours' course credit. All participants were native speakers of British English. None of them had any neurological impairment or any neurological trauma. Two participants were excluded for data analysis due to sleepiness and excessive blinks during the experiments.

Materials

Sentence comprehension task

120 sets of stimuli based on Hindy et al (2012) were used in this experiment. They were divided into three conditions. Each stimulus is consisted of three sentences. The first sentence describes an event that either involved a minimal change (labeled as “no change”) or a substantial change (labeled as “change”) that will happen on the target object. Verbs in the first sentence were matched in ratings of lexical ambiguity and frequency of use. The second sentence refers to the object again. The third sentence describes the object, which is not relevant to previous events. For each set of stimuli, an implausible stimulus that followed the same structure was added. The first sentence of the Implausible condition indicated no change by taking the verb from the second sentence of the experimental conditions. The verb in the second sentence was not expected to appear with the target noun, creating anaolmous meanings (e.g., iron the ice cream). The third sentence is the same as the one in experimental conditions. Three counterbalanced lists were used. Participants were exposed to all conditions, but never saw more than one version of each stimulus. Table 5.1 shows two sets of example stimuli.

Table 5.1 Example stimuli in Experiment 9.

Condition	Examples
No Change	The woman will **choose** the ice cream. And then, she will look at the **ice cream**. It is a Wall's.
Change	The woman will **drop** the ice cream. And then, she will look at the **ice cream**. It is a Wall's.
Implausible	The woman will look at the ice cream. And then, she will **iron** the **ice cream**. It is transparent.
No Change	The chef will **select** the onion. And then, he will smell the **onion**. It is fresh.
Change	The chef will **chop** the onion. And then, he will smell the **onion**. It is fresh.
Implausible	The chef will smell the onion. And then, he will **telephone** the **onion**. It is odourless.

Ratings for the degree of change of the first sentence were collected through online surveys. As mentioned in Hindy et al. (2012), 85 participants from the University of Pennsylvania rated "*the degree to which the depicted object will be at all different after the action occurs that it had been before the action occurred*" on a 7-point scale ranging from "just the same" to "completely changed". The Change condition (e.g., *The woman will drop the ice cream*) (M=4.64, SD=0.84) significantly had higher ratings than the No Change condition (e.g., *The woman will choose the ice cream*) (M=1.97, SD=0.57, t (119) = 27.63, p < .001, by items). Ratings for imageability and ambiguity of the events that were described in the first sentence were also collected. The Change condition (M=4.89, SD=0.64) was rated as more imageable than the No Change condition (M=5.46, SD=0.41, t (119) = 8.24, p = .001), but no difference was found for ambiguity (M_{change}=3.20, SD= .55, M_{No_change}=3.62, SD=.60, t (119) = 5.66, p = .179).

As for the second sentence, a separate pool of 95 participants rated the degree of change (Hindy et al., 2012). The results confirmed that the second sentence conveyed a minimal change (*M*=1.90, *SD*=4.7). Furthermore, the likelihood of the second sentence following the first sentence of each stimulus were also collected, on a 7-point scale from a further 93 participants. No difference was found between the No Change condition (*M*=4.06, *SD*=0.77) and the Change condition in the likelihood ratings (*M*=4.34, *SD*=0.79, t (119) = 2.82, p = .920), suggesting there was no bias on the occurrence of the second sentence despite the differences of verbs in the first sentence.

Stroop task

There were 216 trials in total, which were composed of three types: congruent trials, response-eligible trials, and response-ineligible trials. Participants were presented with a single word for each trial and instructed to press the key corresponding to the typeface color of the word. Three keys on the keyboard attached with yellow, green, and blue stickers were used as response keys. In response-eligible trials, the word denoted a colour that was included in the response keys (i.e., yellow, green, or blue). In the response-ineligible trials, the word denoted a colour that was not a potential answer (i.e., orange, lime, or red). Stroop effect was calculated by taking the difference of accuracy rates and reaction times between incongruent trials and congruent trials for both response-eligible and response- ineligible trials, respectively.

Procedure

EEG data of language comprehension and behavioral data of Stroop conflict were collected in a single session (the language task before the Stroop task) with rest periods. Participants were tested individually. In the language comprehension task, participants were informed that they would read short stories and judge whether the situation described in the sentences is plausible or not. Sentences were presented word-by-word using a serial visual presentation procedure, wherein each word was presented in white font on a black background for 300 ms (except the critical words), followed by a 200 ms blank screen before the next word appeared. The critical word at the end of the second sentence (e.g., ice cream) was presented for 800 ms and there was a 500 ms gap before it. After all words have been presented, the plausibility question “Plausible or Implausible? Y for “yes”, N for “no”” appeared directly on the screen. Participants were expected to answer yes” to experimental trials, but “no” to implausible trials. The experiment was divided into 3 blocks, separated by rest reminders. Each trial was separated from the next by a blank screen for 1000 ms followed by a fixation cross for another 1000 ms. Participants only saw one version of each stimulus, but were exposed to all stimuli and all conditions. Participants were instructed to minimize body movement during EEG recording. No additional task demands were imposed. After completing the language comprehension task, participants were instructed to take a break before performomg the Stroop colour identification task. No EEG signals were recorded for this task. We followed the same procedure as described in Hindy et al. (2012). Total time-on-task was about 80 minutes in total for both tasks.

EEG recording and ERP data analysis

The EEG was recorded with 30 electrodes in an EEG cap from 30 scalp locations (Fpz, Fp1/2, Fz, F3/4, F7/8, FC1/2, FC5/6, Cz, C3/4, T7/8, CP1/2, CP5/6, Pz, P3/4, P7/8, POz, Oz, O1/2). The EEG signals were amplified by using an ANT amplifier, digitized at 500 Hz and re-referenced to the algebraic average of the two mastoids. Electrooculographic activity (EOG) was assessed using four additional electrodes above and below the right eye (VEOG) and over the outer canthi (HEOG). Impedances were kept below 5 kΩ. A digital band-pass filter (0.1-30 Hz) was applied.

Epochs were extracted beginning 200 ms prior to and ending 1000 ms after the onset of stimulus presentation and re-referenced to a baseline period of 200 ms pre-stimulus baseline period. Ocular and muscular artifacts were corrected using Independent Component Analysis. Any epochs containing artifacts after the ICA were removed based on visual inspection. Data from two participants were excluded from data analysis due to excessive movement artifacts. Statistical analyses on average amplitudes were conducted separately for two critical time windows chosen based on the literature (e.g., Kutas & Hillyard, 1980; van Berkum et al., 1999) and on visual inspection of the data: 300 ms - 500 ms for the N400, and 400 ms -1000 ms for Nref and the old/new effect. Separate analyses were conducted for mean amplitudes in each time window. Data were averaged across four regions of interest, including left anterior (F3, F7, FC5, C3), left posterior (CP5, P3, P7, O1), right anterior (F4, F8, FC4, C4) and right posterior (CP6, P4, P8, O2). Central/midline electrodes were not included. Repeated measures ANOVA was conducted using anteriority (anterior vs. posterior) and laterality (left vs. right) as topographic factors.

5.3 Results

5.3.1 Comprehension questions

Participants' plausibility judgments were 93.56% accurate on average. A repeated measures ANOVA revealed that there was a main effect of condition, F (2, 25) = 1446.22, p < .001. Participants reliably rejected the semantically anomalous sentences and accepted the experimental sentences ("*yes*" responses were given on only 4.83% of trials in the implausible condition). Sentences in the No Change condition (e.g., *The woman will choose the ice cream. And then, she will look at the ice cream. It is a Wall's*) were judged as more plausible than the Change conditions, (e.g., *The woman will drop the ice cream. And then, she will look at the ice cream. It is a Wall's*), t (26) = 4.30, p < .001.

5.3.2 Stroop colour-word interference

The Stroop effect was measured by the differences of responses between congruent trials versus incongruent trials. The Stroop inhibitory ability was calculated by taking the differences of response accuracy between congruent trials and incongruent trials, while the Stroop inhibitory efficiency was indexed by the differences of reaction times. The outcome indicated participants' ability and efficiency of making the right choice by inhibiting inappropriate choice. Overall, participants correctly answered 96% of all trials. The average response times (excluding errors) were 562 ms for congruent trials and 628 ms for incongruent trials. Repeated-measures ANOVA revealed a main effect of trial type in both accuracy (higher for congruent trials than incongruent ones), F (2, 26) = 11.32, p < .001 and reaction times (faster for congruent trials than incongruent ones), *F* (2, 26) = 54.31, *p* < .001. Moreover, within-subjects t-tests showed response-eligible trials elicited greater Stroop effect than response-ineligible trials, *t* (28) = 3.91, *p* = .001, for Stroop inhibitory ability, *t* (28) = 4.94, *p* < .001, for Stroop inhibitory efficiency.

5.3.3 Event-related Potentials (ERPs).

Figure 5.1 displays the grand average ERPs of all subjects (N=27) at all 30 electrode sites and corresponding scalp distributions time-locked to the onset of the noun phrase in the second sentence (e.g., *"And then, she will look at <u>the ice cream</u>"*). Visual inspection of the figure suggested that the Change condition elicited a more negative ERP waveform than the No Change condition during 400 ms after post stimulu onset. On the basis of earlier studies (e.g., van Berkum et al., 1999; Kutas & Hillyard, 1980) two time windows were chosen for statistical analysis, including: (a) 300-500 ms; (b) 400 – 1000 ms post onset of "*the ice cream*".

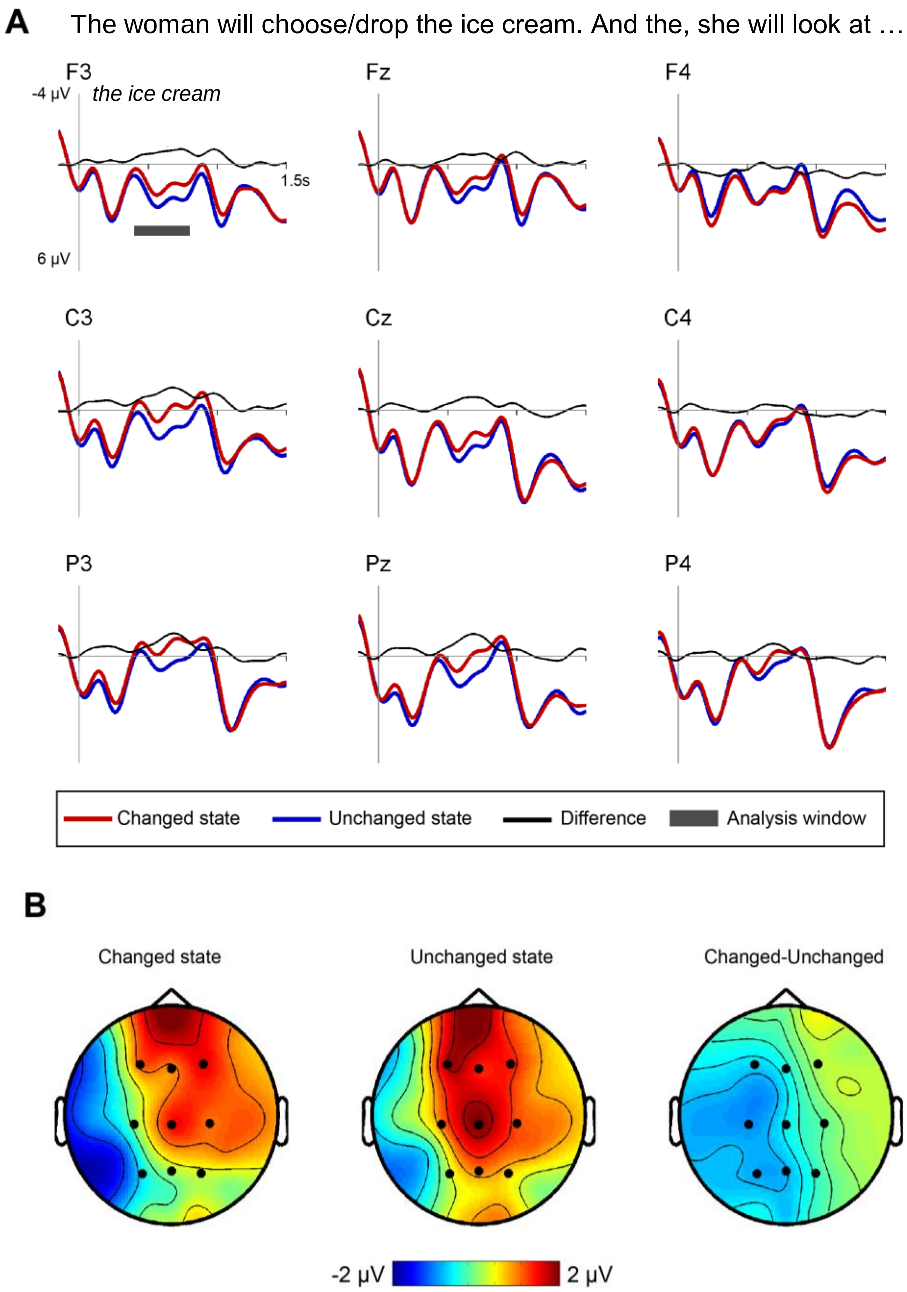


Figure 5.1 Experiment 9: Grand average ERPs waves (A) and scalp distribution (B) after onset of the critical noun phrase in the second sentence involving No Change (blue line) and Change (red line) in the first sentence, and corresponding scalp distributions during the time windows (300 ms – 1000 ms).

Table 5.2 presents pairwise comparisons of the Change and Implausible condition compared to No Change condition, averaged across four electrodes sites. Results suggested that the Implausible condition elicited more negative ERPs relative to No Change condition at posterior sites in time window 300 – 500 ms. By contrast, in time window 400 – 1000 ms, the Change condition elicited more negative ERP relative to No Change condition at left hemisphere.

Table 5.2 Experiment 9: *F, MSE,* and *p* values for main effect of degree of change, and implausibility in two time windows (computed across all electrodes, and for four electrode ROIs, separately). All statistical tests in this table involve pairwise comparisons of each condition compared to the No Change condition.**p* <= .05, ** *p* <=.01, *** *p* <=.001.

	Electrodes	300-500 ms			400-1000 ms		
		F	*p*	*MSE*	*F*	*p*	*MSE*
Change	All	.47	.50	.87	2.25	*.146*	.48
	Left	1.60	.217	.96	6.31	*.019**	.52
	Right	.00	.981	.82	.14	*.713*	.54
	Anterior	.19	.665	1.19	1.49	*.234*	.60
	Posterior	.83	.370	.76	2.63	*.117*	.48
	All	.97	.335	1.27	.02	*.889*	.44
Implausible	Left	.20	.659	1.48	.15	*.703*	.49
	Right	2.38	.135	1.19	.01	*.911*	.46
	Anterior	.21	.652	1.57	.39	*.54*	.66
	Posterior	5.80	.023*	1.35	.20	*.66*	.52

Time window 300 ms – 500 ms

Results of repeated measures ANOVAs with conditions (No Change, Change, Implausible), Hemisphere (left, right), Anteriority (anterior, posterior) suggested that there was no main effect of conditions, $F(2, 25) = 1.22$, $p = .311$, $\eta_p^2 = .09$, and anteriority, $F(1,26) = 0.69$, p=.794, η_p^2 = .003, but there was a main effect of hemisphere (more negative over right than left), F (1,26) = 9.70, p =.004, η_p^2 = .27, and significant interactions between Condition and Hemisphere, $F(1,26)$ = 5.22, p = .013, η_p^2 = .30, and between Condition and Anteriority, $F(1,26)$ = 10.43, p = .001, η_p^2 = .46. Planned comparison suggested that implausibility trials elicited more negative ERPs over the right posterior sites relative to No Change condition, $t(26)$ = 3.15, p =.004, and Change condition, t (26) =2.95, p =.007, but no such difference was found between experimental conditions, t (26) =.37, p =.717. Thus, the Change condition did not elicit different ERPs relative to No Change condition, but the Implausible condition evoked more negative N400 relative to experimental conditions, but no difference in N400 was found between experimental conditions.

Moreover, correlation analyses using the difference of ERP amplitudes (Implausible minus No Change, Implausible minus Change, Change minus No Change) over right posterior sites and Stroop effects (inhibitory ability and inhibitory efficiency) revealed a marginally significant correlation between the Stroop effect (inhibitory ability of response-ineligible trials) the Implausible (minus No Change) condition, r = -.38, p=.058. No significant relationship between Stroop effects and other conditions. The pattern suggests that for participants who were more accurate in the response-ineligible trials tended to show a more robust N400 effect over right posterior sites, but this effect was revealed only when N400 effect presented.

Time window 400 ms – 1000 ms

Repeated measures ANOVAs yielded no significant main effect of the conditions, *F*(2,25) = 1.48, *p* = .247, η_p^2 = .11, but there were main effects of anteriority (more negative over posterior than anterior), *F*(1,26) = 6.84, *p* = .015, η_p^2 = .21, and hemisphere (more negative over left than right), *F*(1,26) = 4.39, *p* = .046, η_p^2 = .14. Besides, there was a significant interaction between conditions and hemisphere, *F*(1,26) = 7.70, *p* = .002, η_p^2 = .38. Planned comparisons revealed but the Change condition elicited more negative ERPs over the left hemisphere than No Change, *t* (26) =2.51, *p* =.019, and Implausible conditions, *t* (26) =2.44, *p* =.022, but the scalp distribution is more negative over posterior sites than anterior sites for both experimental conditions, *t* (26) =2.76, *p* =.010 (No Change), *t* (26) =3.12, *p* =.004 (Change). Correlation analyses using the difference of ERP amplitudes revealed no significant correlations.

5.4 Discussion

In this study, we explored the ERP evidence on updating object representation in language comprehension. The temporal resolution of the ERP method allowed us to characterize the time course of neural signals. Specifically, we recorded neural activity while participants read short stories about a target object that was described to experience either a change or no change of its original state in the first sentence. In experimental conditions, the second sentence was the same regardless of the first sentence, in which he object was described to experience no further changes. In the Implausible condition, however, the object was described to experience an anomalous action. ERPs at the time windows of 300 ms – 500 ms and 400 ms – 1000 ms after onset of the object name in the second sentence were analyzed. Our results revealed that there was no difference in ERP amplitudes between experimental condition in the time window of 300 – 500 ms, but differences were found in the time window of 400 – 1000 ms. After onset of the object name, the Change condition induced more negative ERPs than the No Change condition at both the anterior and posterior areas of the left hemisphere.

The ERPs in the time window of 400 – 1000 ms post stimulus onset were different in scalp distributions from studies on referential ambiguity resolution by Van Berkum and colleagues (e.g., Nieuwland & Van Berkum, 2006, 2008; Nieuwland et al., 2007a; Van Berkum et al., 1999; Van Berkum eta al., 2003), but similar to the old/new effect in episodic memory retrieval (e.g., Johansson & Mecklinger, 2003; Leynes, Grey, & Crawform, 2006; Senkfor and van Petten, 1998;). Typically, the old items elicit more positive ERPs than new items, which is maximal over the left parietal sites. Anterior old/new effect that onsets at approximately 400 ms at right frontal sites has also been observed. These two kinds of old/new ERP effects are known to be associated with familiarity and recollection, respectively (Curran, 2000). Thus, the similarity of the current ERP effect and the left parietal old/new effect suggested that the episodic retrieval (recollection) of the target object tended to be different depending on the prior context. Compared to no change, the object that is described to experience a substantial change seem to resemble a ‘new’ object, suggesting difficult in episodic retrieval. This result is consistent with Kalenik (2012) that longer reading time is needed to read the repeated target noun phrase in the Change condition than No Change.

The lack of correlation between the Stroop effect and ERPs amplitudes in our experiment may be due to the difference in the time course of activation of the neural resources. Studies on the EEG activity of the Stroop task suggested that increased fronto-central negativity appeared around 400 ms for incongruent trials compared to congruent and neutral items (Hanslmayr, et al., 2008; Liotti, Woldorff, Prez III, & Mayberg, 2000). Using amplitudes in the time window of 300-500 ms may not map to neural activity of the Stroop conflict. Also, as only behavioral outcomes of the Stroop effect was collected, this may lead to difficult in map Behaviour responses to ERP amplitudes. Studies that revealed significant correlation between behavioral and ERP measures often involved the same task rather than two separate tasks (e.g., Knoeferle et a., 2011).

In sum, our findings provided supportive evidence for the existence of difference in neural activity after an object was described to experience a change of state compared with no change. It shows that the processing system can very rapidly determine whether an object is representationally different in one situation from the other. About 400 ms after onset of the object name, the readers have determined whether the object had been described differently in in earlier linguistic context. This effect may not be attributed to the same neural resources as the Nref. Instead, it is more similar to the posterior distribution of the old/new effect.

5.5 Conclusion

We have shown that when an object was described as undergoing substantial change, keeping track of its representations led to differences in electrophysiological responses over the scalp in contrast to a minimal/no change when the object name was mentioned in subsequent context. We compared the effect that we observed with other effects with similar latencies and topography in existing studies. The effect resembled the old/new effect in recognition memory, suggesting recognition memory plays an important role in updating object-state representations.

CHAPTER 6

Conclusion: Aligning hands, eyes and the brain

6.1 Overview

This thesis set out to explore the keeping track of object-state representations in language comprehension. The construction of object-state representations that encode different states at different event times is key to serve as an index of events in situation model. However, there has been limited evidence on whether specific object states, and their change across time, are encoded during language comprehension. Drawing on evidence from three experimental techniques with different measures – button-press reaction times, eye movements, and event-related potentials (ERPs), the data unequivocally tells us that a single object can have multiple instantiations in the mental representation of the event, and that this representational complexity must be resolved during the interpretation of unfolding sentences. Among these techniques, the picture verification paradigm has been used for studying the relationship between internal representations and external referents (e.g., Zwaan & Pecher, 2012), while eye movements and ERPs are excellent for capturing temporal dynamics and real-time responses from participants in language comprehension (e.g., Altmann & Kamide, 2009; Van Berkum, et al., 1999). By combining the techniques, we intended to provide a more comprehensive understanding of the construction of object-state representations. In this chapter, I synthesize the empirical findings reported in previous chapters and discuss theoretical and methodological implications of these findings.

6.2 Synthesis of empirical findings

6.2.1 Do we establish multiple object-state representations in language comprehension?

The first three experiments in Chapter 2 with the picture verification task provided evidence for the activation of object-state representations in language comprehension. We found that the intact state was responded to faster than the changed state when no contextual information was provided (Experiment 1). Nevertheless, when linguistic context was provided, the representation that matched the consequences of the described event was activated (Experiment 2). The time available for comprehension was also shown to be important for activating object-state representations (Experiment 3). When there was a short delay (250 ms) between sentence reading and picture verification, verification response times were not influenced by contextual information but the typicality of the intact state. By contrast, when a long delay (1250 ms) was provided, picture verification responses to pictures that matched the context were then shorter than that mismatches the context after a forward temporal shift. Thus, the results suggest that different object-state representations can be constructed in language comprehension but sufficient processing time is required.

6.2.2 What is the time course of establishing an appropriate object-state representation?

Chapter 3 and 4 reported evidence from a series of visual world eye-tracking experiments. Participants' eye movements were recorded while they viewed a clipart visual display and listened to description of a target object. Eye movements on the display were thus time-locked to the unfolding utterances. Experiments 4 and 5 explored when during the time course of comprehension listeners establish an object-state representation. The results showed that participants switched their eye movements to an intended object-state upon hearing the object name but not before. These findings were replicated by Experiments 6, 7, and 8, which also revealed that eye movements were primarily driven by semantic overlap between the depicted object state and the described object-state representation.

6.2.3 Is there ERP evidence for keeping track of an object's representations when that object is described as changing of state?

In Chapter 5, Experiment 9 examined the ERP "marker" that may be relevant when object-state representation was established. No N400 was revealed between the Change condition than the No Change condition at the time window of 300 ms – 500 ms after onset of the critical noun phrase in the second sentence ("the ice cream"), but greater negativity was found at the time window of 400 ms – 1000 ms. This effect was distributed in a similar time window as the Nref of referiential ambiguity resolution and old/new effect. However, the Nref is typically distributed at the anterior sites, but the negativity in our experiment was largely at the posterior sites. In this case, the effect was more similar to the late posterior negativity that was found in studies of recognition memory (e.g., Ecker, et al., 2007). Therefore, differences in ERPs between the Change and No

Change condition suggest that object-state representations may be processed differently due to episodic memory retrieval (cf. recognition memory).

Taken together, the findings suggest that different instantiations of the same entity can be constructed, maintained and updated in language comprehension. The linguistic context provides criteria for the kind of object-state representation being established. When a visual display was provided, the established instantiation can be mapped onto the appropriate depicted state of the object.

6.3 Methodological Implications

Each of the techniques has their strengths and weaknesses. Offline picture verification measures can capture the established object-state representation in an efficient and relatively low-cost manner. However, pictures were presented after the sentence was read and thus cannot capture moment-by-moment sentence processing.

In the thesis, this issue has been addressed by using eye tracking that provide an excellent way of investigating online language comprehension. As pointed out by Tanenhaus & Trueswell (2006), eye-tracking measures reflect when and where people fixate on the visual display as the utterance unfolds, and then use that information to draw inferences about the underlying processes and representations.

Besides, to explore the construction of object-state representations in reading comprehension, ERPs data were also collected by using similar stimuli in picture verification and eye-tracking.

So far, there has been limited interaction between these paradigms. This is probably due to differences in stimuli presentation (e.g., a serial order of words in ERPs but paragraph/visual stimuli in eye tracking) and data analysis, it is difficult to capitalize on the strengths of both techniques. Nonetheless, data from this thesis provide an opportunity to explore the same issue by using different paradigms. The use of multiple techniques provides the opportunity to obtain a comprehensive view of this issue compared with using a single method alone. As each paradigm was intended to address different aspects of the same research questions, data in this thesis were comprehensive, covering behavioral consequences to neural responses.

6.4 Theoretical implications

The experimental data presented in this thesis provided empirical evidence for the construction and updating of multiple object-state representations in language comprehension, reflecting the changes of state that an object may experience as a consequence of, or during, the described event. In respect of prior work on event representation, most current models focus on *event segmentation* than on the actual content of the representations. The work reported here focuses less on how we segment events and more on what the content is of event representation and how that content is dynamically updated as sentences unfold. Four main theoretical implications are discussed.

6.4.1 *Object-state representations are included in situation models of language comprehension*

Data in this thesis suggest that contextual information (*no change vs. change*) activates a particular object-state representation, indicating contextual information activates candidate object-state representations in a top-down manner. Our data are consistent with episodic object recognition (Ecker, et al., 2007) that reaction times increase when the size, orientation, or colour of a target object is changed from study to test. These findings also support the situated/grounded view of language comprehension that language understanding is grounded in motor and perceptual systems (Barsalou, 1999) that reflect real-world experience with the described actions and objects.

6.4.2 *Semantic overlap plays a role in online language comprehension, particularly the updating of object-state representation.*

Semantic overlap of perceptual (shape), motoric (action), and abstract (function) features between spoken words and objects direct eye movements towards the named objects (e.g., hear "*piano*" and look at a piano) and related ones (e.g., hear "*piano*" but look at a trumpet) in the visual scene However, until the studies in Chapter 4, it was not clear whether semantic overlap also influences eye movements towards objects overlapping in respect of object *state*. We demonstrated that when an object name (e.g., "*ice cream*") was repeated in subsequent discourse ("*the woman dropped the ice cream. Then she looked at another ice cream*"), eye movements were directed to the object that matched the linguistic context and shared perceptual features with the ice cream held in episodic memory (and matching the depicted dropped ice cream in the stimuli we used).

6.4.3 An object's changed state activated the action that caused its current state.

In all eye-tracking experiments, we have shown that participants preferred to look at the dropped ice cream than the upright ice cream even before the verbs were mentioned. Perhaps, perceptual properties of the dropped ice cream indicated a "*drop*" action has happened to the object upon seeing it. This finding is consistent with previous studies suggest that actions can be automatically activated during object processing (e.g., Campanella & Shallice, 2011; Davey et al., 2015; Martin et al., 1995). For example, Martin *et* al. (1995) asked participants to generate an action word related to an object. They found that participants often activated the conceptual representation of the object that is corresponding to an action verb (e.g., a picture of a pair of *scissors* and the verb *cut* in response). Campanella & Shallice (2011) have shown that if the target object and another object share similar manipulation features, matching an object picture to a word is less accurate than when they do not share these features.

6.4.4 Change-of-state may be used to explore cognition and language.

Understanding change-of-state events also speaks to the theories of cognitive representation and conceptual development (Galazka & Ganea, 2014; Ganea, et al., 2007; Osina, Saylor, & Ganea, 2013; Ganea & Harris, 2013). Recent studies revealed that toddlers as young as 22 months can use verbal information to update the representation of an absent toy when the object was described as undergoing a change in state, but 19-months-olds failed to perform the task （Ganea, et al., 2007). Data in this thesis suggest that representations of the same object token in the discourse may be processed differently from multiple tokens. This finding opens the possibility of further studies on what cognitive mechanism may be involved in updating object-state representations. Such research, coupled with studies on child cognitive development (e.g., Ganea, et al., 2007), will provide important implications for understanding the interaction between language and cognition.

6.5 Conclusion

The studies presented in this thesis explored when and what type of object-state representations are constructed and updated as the target object is described to experience a change of state event. Our findings suggest that linguistic context modulated the activation of object-state representation in language processing. Responses to the representation when it matched the linguistic description were significantly shorter than when they mismatched. In the eye movement studies, the corresponding effect manifested when the object name was explicitly mentioned, but not before. These findings correspond well theories of situation models of language comprehension and go further by outlining the time course of the updating process. Besides, as shown by the ERP study, keeping track of object-state representation is similar to episodic memory retrieval. Thus, the studies presented in this thesis provided detailed information in terms of establishing and updating object-state representations in online language comprehension.

The work reported here has three important contributions: first, it highlights the need for accounts of event representation to incorporate *object state change*; second, and related, it requires theories to take account of the dynamics with which such object representations change as events (and sentences describing events) unfold; finally, it highlights the need for models of event representation to take account of the interplay between general semantic knowledge about objects, and the episodic knowledge introduced by the sentential context and on which basis the event representation must be updated. The ERP data suggest that research on episodic memory may be particularly relevant in respect of understanding the neural signatures associated with event processing.

APPENDICES

Appendix I Experimental Stimuli
(Words and Sentences)

A1.1 Experimental words – Experiment 1

1	gate
2	dustbin
3	gift
4	acorn
5	banana
6	jump rope
7	candle
8	plant
9	blackboard
10	sunglasses
11	ice cream
12	glasses
13	gum
14	onion
15	vase
16	rug
17	pizza
18	blowtorch
19	drawing
20	blueberries
21	cucumber
22	bagel
23	piglet
24	pumpkin
25	ball
26	egg
27	bed
28	sleeping bag
29	tree
30	umbrella
31	apple
32	wall

A1.2 Experimental sentences – Experiment 2

No.	Condition A: No Change	Condition B: Change
1	The man has jumped over the gate.	The man has shut the gate.
2	The teenager has looked into the dustbin.	The teenager has tipped over the dustbin.
3	The bride has accepted the gift.	The bride has unwrapped the gift.
4	The squirrel has sniffed the acorn.	The squirrel has cracked the acorn.
5	The gorilla has inspected the banana.	The gorilla has peeled the banana.
6	The girl has examined the jump rope.	The girl has cut the jump rope.
7	The young man has smiled at the candle.	The young man has blown out the candle.
8	The child has crouched behind the plant.	The child has knocked down the plant.
9	The teacher has inspected the blackboard.	The teacher has wiped clear the blackboard.
10	The customer has picked out the sunglasses.	The customer has unfolded the sunglasses.
11	The woman has chosen the ice cream.	The woman has drop the ice cream.
12	The customer has picked out the glasses.	The customer has broken the glasses.
13	The girl has inspected the stick of gum.	The girl has chewed the stick of gum.
14	The chef has weighed the onion.	The chef has chopped the onion.
15	The boy has photographed the vase.	The boy has broken the vase.
16	The waiter has praised the rug.	The waiter has unrolled the rug.
17	The student has sniffed the pizza.	The student has cut up the pizza.
18	The welder has clutched the blowtorch.	The welder has lit the blowtorch.
19	The child has admired the drawing.	The child has coloured in the drawing.
20	The grocer has weighed the blueberries.	The grocer has puréed the blueberries.
21	The woman has squeezed the cucumber.	The woman has squeezed the cucumber.
22	The man has chosen the bagel.	The man has sliced the bagel.
23	The man has fed the piglet.	The man has roasted the piglet.
24	The farmer has weighed the pumpkin.	The farmer has carved the pumpkin.
25	The boy has patted the ball.	The boy has deflated the ball.
26	The woman has selected the egg.	The woman has broken the egg.
27	The maid has sat upon the messy	The maid has made up the

	bed.	messy bed.
28	The traveler has laid on the sleeping bag.	The traveler has rolled up the sleeping bag.
29	The lumberjack has measured the tree.	The lumberjack has chopped down the tree.
30	The woman has inspected the umbrella.	The woman has opened the umbrella.
31	The woman has picked out the apple.	The woman has halved the apple.
32	The construction worker has leaned on the wall.	The construction worker has knocked down the wall.

A1.3 Experimental sentences –Experiment 3, 4, & 6

No.	Stimuli a = no change, but first; b = no change, and then; c = change, but first; d = change, and then
01a	The man will jump over the gate. But first, he will lean on the gate.
01b	The man will jump over the gate. And then, he will lean on the gate.
01c	The man will shut the gate. But first, he will lean on the gate.
01d	The man will shut the gate. And then, he will lean on the gate.
02a	The teenager will look into the dustbin. But first, he will walk around the dustbin.
02b	The teenager will look into the dustbin. And then, he will walk around the dustbin.
02c	The teenager will tip over the dustbin. But first, he will walk around the dustbin.
02d	The teenager will tip over the dustbin. And then, he will walk around the dustbin.
03a	The bride will accept the gift. But first, she will give thanks for the gift.
03b	The bride will accept the gift. And then, she will give thanks for the gift.
03c	The bride will unwrap the gift. But first, she will give thanks for the gift.
03d	The bride will unwrap the gift. And then, she will give thanks for the gift.
04a	The squirrel will sniff the acorn. But first, it will lick the acorn.
04b	The squirrel will sniff the acorn. And then, it will lick the acorn.
04c	The squirrel will crack the acorn. But first, it will lick the acorn.
04d	The squirrel will crack the acorn. And then, it will lick the acorn.
05a	The gorilla will inspect the banana. But first, he will grunt at the banana.
05b	The gorilla will inspect the banana. And then, he will grunt at the banana.
05c	The gorilla will peel the banana. But first, he will grunt at the banana.
05d	The gorilla will peel the banana. And then, he will grunt at the banana.
06a	The girl will examine the jump rope. But first, she will complain about the jump rope.
06b	The girl will examine the jump rope. And then, she will complain about the jump rope.
06c	The girl will cut the jump rope. But first, she will complain about the jump rope.
06d	The girl will cut the jump rope. And then, she will complain about the jump rope.
07a	The young man will smile at the candle. But first, he will reach for the candle.
07b	The young man will smile at the candle. And then, he will reach for the candle.
07c	The young man will blow out the candle. But first, he will reach for the candle.
07d	The young man will blow out the candle. And then, he will reach for the candle.
08a	The child will crouch behind the plant. But first, he will jump over the plant.

08b The child will crouch behind the plant. And then, he will jump over the plant.
08c The child will knock down the plant. But first, he will jump over the plant.
08d The child will knock down the plant. And then, he will jump over the plant.
09a The teacher will inspect the blackboard. But first, she will point at the blackboard.
09b The teacher will inspect the blackboard. And then, she will point at the blackboard.
09c The teacher will wipe clear the blackboard. But first, she will point at the blackboard.
09d The teacher will wipe clear the blackboard. And then, she will point at the blackboard.
10a The customer will pick out the sunglasses. But first, she will ask about the sunglasses.
10b The customer will pick out the sunglasses. And then, she will ask about the sunglasses.
10c The customer will unfold the sunglasses. But first, she will ask about the sunglasses.
10d The customer will unfold the sunglasses. And then, she will ask about the sunglasses.
11a The woman will choose the ice cream. But first, she will grin at the ice cream.
11b The woman will choose the ice cream. And then, she will grin at the ice cream.
11c The woman will drop the ice cream. But first, she will grin at the ice cream.
11d The woman will drop the ice cream. And then, she will grin at the ice cream.
12a The customer will pick out the glasses. But first, she will ask about the glasses.
12b The customer will pick out the glasses. And then, she will ask about the sunglasses.
12c The customer will break the glasses. But first, she will ask about the sunglasses.
12d The customer will break the glasses. And then, she will ask about the sunglasses.
13a The girl will inspect the stick of gum. But first, she will complain about the gum.
13b The girl will inspect the stick of gum. And then, she will complain about the gum.
13c The girl will chew the stick of gum. But first, she will complain about the gum.
13d The girl will chew the stick of gum. And then, she will complain about the gum.
14a The chef will weigh the onion. But first, she will smell the onion.
14b The chef will weigh the onion. And then, she will smell the onion.
14c The chef will chop the onion. But first, she will smell the onion.
14d The chef will chop the onion. And then, she will smell the onion.

15a	The boy will photograph the vase. But first, he will examine the vase.
15b	The boy will photograph the vase. And then, he will examine the vase.
15c	The boy will break the vase. But first, he will examine the vase.
15d	The boy will break the vase. And then, he will examine the vase.
16a	The waiter will praise the rug. But first, he will step over the rug.
16b	The waiter will praise the rug. And then, he will step over the rug.
16c	The waiter will unroll the rug. But first, he will step over the rug.
16d	The waiter will unroll the rug. And then, he will step over the rug.
17a	The student will sniff the pizza. But first, he will blow on the pizza.
17b	The student will sniff the pizza. And then, he will blow on the pizza.
17c	The student will cut up the pizza. But first, he will blow on the pizza.
17d	The student will cut up the pizza. And then, he will blow on the pizza.
18a	The welder will clutch the blowtorch. But first, he will fiddle with the blowtorch.
18b	The welder will clutch the blowtorch. And then, he will fiddle with the blowtorch.
18c	The welder will light the blowtorch. But first, he will fiddle with the blowtorch.
18d	The welder will light the blowtorch. And then, he will fiddle with the blowtorch.
19a	The child will admire the drawing. But first, she will display the drawing.
19b	The child will admire the drawing. And then, she will display the drawing.
19c	The child will colour in the drawing. But first, she will display the drawing.
19d	The child will colour in the drawing. And then, she will display the drawing.
20a	The grocer will weigh the blueberries. But first, he will smell the blueberries.
20b	The grocer will weigh the blueberries. And then, he will smell the blueberries.
20c	The grocer will purée the blueberries. But first, he will smell the blueberries.
20d	The grocer will purée the blueberries. And then, he will smell the blueberries.
21a	The woman will squeeze the cucumber. But first, she will talk about the cucumber.
21b	The woman will squeeze the cucumber. And then, she will talk about the cucumber.
21c	The woman will squeeze the cucumber. But first, she will talk about the cucumber.
21d	The woman will squeeze the cucumber. And then, she will talk about the cucumber.
22a	The man will choose the bagel. But first, he will smell the bagel.
22b	The man will choose the bagel. And then, he will smell the bagel.
22c	The man will slice the bagel. But first, he will smell the bagel.
22d	The man will slice the bagel. And then, he will smell the bagel.
23a	The man will feed the piglet. But first, he will examine the piglet.

23b The man will feed the piglet. And then, he will examine the piglet.
23c The man will roast the piglet. But first, he will examine the piglet.
23d The man will roast the piglet. And then, he will examine the piglet.
24a The farmer will weigh the pumpkin. But first, he will photograph the pumpkin.
24b The farmer will weigh the pumpkin. And then, he will photograph the pumpkin.
24c The farmer will carve the pumpkin. But first, he will photograph the pumpkin.
24d The farmer will carve the pumpkin. And then, he will photograph the pumpkin.
25a The boy will pat the ball. But first, he will throw the ball.
25b The boy will pat the ball. And then, he will throw the ball.
25c The boy will deflate the ball. But first, he will throw the ball.
25d The boy will deflate the ball. And then, he will throw the ball.
26a The woman will select the egg. But first, she will examine the egg.
26b The woman will select the egg. And then, she will examine the egg.
26c The woman will break the egg. But first, she will examine the egg.
26d The woman will break the egg. And then, she will examine the egg.
27a The maid will sit upon the messy bed. But first, she will complain about the bed.
27b The maid will sit upon the messy bed. And then, she will complain about the bed.
27c The maid will make up the messy bed. But first, she will complain about the bed.
27d The maid will make up the messy bed. And then, she will complain about the bed.
28a The traveller will lay on the sleeping bag. But first, he will moan about the sleeping bag.
28b The traveller will lay on the sleeping bag. And then, he will moan about the sleeping bag.
28c The traveller will roll up the sleeping bag. But first, he will moan about the sleeping bag.
28d The traveller will roll up the sleeping bag. And then, he will moan about the sleeping bag.
29a The lumberjack will measure the tree. But first, he will point at the tree.
29b The lumberjack will measure the tree. And then, he will point at the tree.
29c The lumberjack will chop down the tree. But first, he will point at the tree.
29d The lumberjack will chop down the tree. And then, he will point at the tree.
30a The woman will inspect the umbrella. But first, she will compliment at the umbrella.
30b The woman will inspect the umbrella. And then, she will compliment at the umbrella.
30c The woman will open the umbrella. But first, she will compliment at the umbrella.
30d The woman will open the umbrella. And then, she will compliment at

the umbrella.

31a The woman will pick out the apple. But first, she will talk about the apple.

31b The woman will pick out the apple. And then, she will talk about the apple.

31c The woman will halve the apple. But first, she will talk about the apple.

31d The woman will halve the apple. And then, she will talk about the apple.

32a The construction worker will lean on the wall. But first, she will frown at the wall.

32b The construction worker will lean on the wall. And then, she will frown at the wall.

32c The construction worker will knock down the wall. But first, she will frown at the wall.

32d The construction worker will knock down the wall. And then, she will frown at the wall.

33a The man will install the mirror. But first, he will examine the mirror.

33b The man will install the mirror. And then, he will examine the mirror.

33c The man will smash the mirror. But first, he will examine the mirror.

33d The man will smash the mirror. And then, he will examine the mirror.

34a The man will lick the chocolate. But first, he will smell the chocolate.

34b The man will lick the chocolate. And then, he will smell the chocolate.

34c The man will bite the chocolate. But first, he will smell the chocolate.

34d The man will bite the chocolate. And then, he will smell the chocolate.

35a The child will ask for the tomato. But first, he will wash the tomato.

35b The child will ask for the tomato. And then, he will wash the tomato.

35c The child will squeeze the tomato. But first, he will wash the tomato.

35d The child will squeeze the tomato. And then, he will wash the tomato.

36a The secretary will praise the coffee. But first, she will make the coffee.

36b The secretary will praise the coffee. And then, she will make the coffee.

36c The secretary will spill the coffee. But first, she will make the coffee.

36d The secretary will spill the coffee. And then, she will make the coffee.

A1.4 Experimental sentences – Experiment 5

No.	Stimuli
	a = penny, but first; b = penny, and then; c = egg, but first; d = egg, and then
01a	The girl will stomp on the penny. But first, she will look down at the penny.
01b	The girl will stomp on the penny. And then, she will look down at the penny.
01c	The girl will stomp on the egg. But first, she will look down at the egg.
01d	The girl will stomp on the egg. And then, she will look down at the egg.
02a	The boy will jump on the hay. But first, he will look at the hay.
02b	The boy will jump on the hay. And then, he will look at the hay.
02c	The boy will jump on the box. But first, he will look at the box.
02d	The boy will jump on the box. And then, he will look at the box.
03a	The boy will pound the beach ball. But first, he will glare at the beach ball.
03b	The boy will pound the beach ball. And then, he will glare at the beach ball.
03c	The boy will pound the sand castle. But first, he will glare at the sand castle.
03d	The boy will pound the sand castle. And then, he will glare at the sand castle.
04a	The teenager will drop the football. But first, he will run away from the football.
04b	The teenager will drop the football. And then, he will run away from the football.
04c	The teenager will drop the piggy bank. But first, he will run away from the piggy bank.
04d	The teenager will drop the piggy bank. And then, he will run away from the piggy bank.
05a	The man will squeeze the tennis ball. But first, he will frown at the tennis ball.
05b	The man will squeeze the tennis ball. And then, he will frown at the tennis ball.
05c	The man will squeeze the biscuit. But first, he will frown at the biscuit.
05d	The man will squeeze the biscuit. And then, he will frown at the biscuit.
06a	The teenager will tip over the empty mug. But first, she will reach for the mug.
06b	The teenager will tip over the empty mug. And then, she will reach for the mug.
06c	The teenager will tip over the coffee. But first, she will reach for the coffee.
06d	The teenager will tip over the coffee. And then, she will reach for the coffee.
07a	The girl will play with the rattle. But first, she will inspect the rattle.
07b	The girl will play with rattle. And then, she will inspect the rattle.

07c	The girl will play with the play dough. But first, she will inspect the play dough.
07d	The girl will play with the play dough. And then, she will inspect the play dough.
08a	The toddler will squeeze the doll. But first, he will glare at the doll.
08b	The toddler will squeeze the doll. And then, he will glare at the doll.
08c	The toddler will squeeze the toothpaste. But first, he will glare at the toothpaste.
08d	The toddler will squeeze the toothpaste. And then, he will glare at the toothpaste.
09a	The girl will pinch the doll. But first, she will scowl at the doll.
09b	The girl will pinch the doll. And then, she will scowl at the doll.
09c	The girl will pinch the baby. But first, she will scowl at the baby.
09d	The girl will pinch the baby. And then, she will scowl at the baby.
10a	The boy will stand on the card. But first, he will look at the card.
10b	The boy will stand on the card. And then, he will look at the card.
10c	The boy will stand on the glasses. But first, he will look at the glasses.
10d	The boy will stand on the glasses. And then, he will look at the glasses.
11a	The girl will blow on the dice. But first, she will look at the dice.
11b	The girl will blow on the dice. And then, she will look at the dice.
11c	The girl will blow on the dandelion. But first, she will look at the dandelion.
11d	The girl will blow on the dandelion. And then, she will look at the dandelion.
12a	The girl will step on the brick. But first, she will stare at the brick.
12b	The girl will step on the brick. And then, she will stare at the brick.
12c	The girl will step on the fly. But first, she will stare at the fly.
12d	The girl will step on the fly. And then, she will stare at the fly.
13a	The girl will drop the fork. But first, she will laugh about the fork.
13b	The girl will drop the fork. And then, she will laugh about the fork.
13c	The girl will drop the ice cream. But first, she will laugh about the ice cream.
13d	The girl will drop the ice cream. And then, she will laugh about the ice cream.
14a	The girl will poke the teddy bear. But first, she will laugh about the teddy bear.
14b	The girl will poke the teddy bear. And then, she will laugh about the teddy bear.
14c	The girl will poke the card tower. But first, she will laugh about the card tower.
14d	The girl will poke the card tower. And then, she will laugh about the card tower.
15a	The racing driver will shake the empty wine glass. But first, he will look at the wine glass.
15b	The racing driver will shake the empty wine glass. And then, he will look at the wine glass.
15c	The racing driver will shake the champagne. But first, he will look at champagne.

15d	The racing driver will shake the champagne. And then, he will look at the champagne.
16a	The child will push the doll. But first, he will run away from the doll.
16b	The child will push the doll. And then, he will run away from the doll.
16c	The child will push the dominoes. But first, he will run away from the dominoes.
16d	The child will push the dominoes. And then, he will run away from the dominoes.
17a	The man will sit on the log. But first, he will frown at the log.
17b	The man will sit on the log. And then, he will frown at the log.
17c	The man will sit on the guitar. But first, he will frown at the guitar.
17d	The man will sit on the guitar. And then, he will frown at the guitar.
18a	The child will kick the tree stump. But first, he will stand next to the tree stump .
18b	The child will kick the tree stump. And then, he will stand next to the tree stump.
18c	The child will kick the bin. But first, he will stand next to the bin.
18d	The child will kick the bin. And then, he will stand next to the bin.
19a	The teenager will drive over the road. But first, she will look at the road.
19b	The teenager will drive over the road. And then, she will look at the road.
19c	The teenager will drive over the skateboard. But first, she will look at the skateboard.
19d	The teenager will drive over the skateboard. And then, she will look at the skateboard.
20a	The toddler will bite the Lego. But first, he will sniff the Lego.
20b	The toddler will bite the Lego. And then, he will sniff the Lego.
20c	The toddler will bite the ice-cream. But first, he will sniff the ice-cream.
20d	The toddler will bite the ice-cream. And then, he will sniff the ice-cream.
21a	The girl will blow on the mirror. But first, she will stand by the mirror.
21b	The girl will blow on the mirror. And then, she will stand by the mirror.
21c	The girl will blow on the candle. But first, she will stand by the candle.
21d	The girl will blow on the candle. And then, she will stand by the candle.
22a	The man will shoot at the tank. But first, he will run away from the tank.
22b	The man will shoot at the tank. And then, he will run away from the tank.
22c	The man will shoot at the window. But first, he will run away from the window.
22d	The man will shoot at the window. And then, he will run away from the window.
23a	The woman will knock over the book. But first, she will look for the book.
23b	The woman will knock over the book. And then, she will look for the book.
23c	The woman will knock over the wine glass. But first, she will look for

	the wine glass.
23d	The woman will knock over the wine glass. And then, she will look for the wine glass.
24a	The maid will drop the rugby ball. But first, she will stand over the rugby ball.
24b	The maid will drop the rugby ball. And then, she will stand over the rugby ball.
24c	The maid will drop the vase. But first, she will stand over the vase.
24d	The maid will drop the vase. And then, she will stand over the vase.
25a	The boy will punch the tree. But first, he will stare at the tree.
25b	The boy will punch the tree. And then, he will stare at the tree.
25c	The boy will punch the mirror. But first, he will stare at the mirror.
25d	The boy will punch the mirror. And then, he will stare at the mirror.
26a	The woman will heat up the frying pan. But first, she will stand beside the frying pan.
26b	The woman will heat up the frying pan. And then, she will stand beside the frying pan.
26c	The woman will heat up the chocolate. But first, she will stand beside the chocolate.
26d	The woman will heat up the chocolate. And then, she will stand beside the chocolate.
27a	The woman will chew the spoon. But first, he will look at the spoon.
27b	The woman will chew the spoon. And then, he will look at the spoon.
27c	The woman will chew the chewing gum. But first, he will look at the chewing gum.
27d	The woman will chew the chewing gum. And then, he will look at the chewing gum.
28a	The man will step on the doormat. But first, he will look at the doormat.
28b	The man will step on the doormat. And then, he will look at the doormat.
28c	The man will step on the beer can. But first, he will look at the beer can.
28d	The man will step on the beer can. And then, he will look at the beer can.
29a	The boy will bite the chopsticks. But first, he will smell the chopstick.
29b	The boy will bite the chopsticks. And then, he will smell the chopsticks.
29c	The boy will bite the chocolate. But first, he will smell the chocolate.
29d	The boy will bite the chocolate. And then, he will smell the chocolate.
30a	The boy will push the statue. But first, he will look at the statue.
30b	The boy will push the statue. And then, he will look at the statue.
30c	The boy will push the flower pot. But first, he will look at the flower pot.
30d	The boy will push the flower pot. And then, he will look at the flower pot.
31a	The girl will throw the balloon. But first, she will laugh about the balloon.
31b	The girl will throw the balloon. And then, she will laugh about the

balloon.

31c The girl will throw the puzzle. But first, she will laugh about the puzzle.

31d The girl will throw the puzzle. And then, she will laugh about the puzzle.

32a The cowboy will snap the leather whip. But first, he will examine the whip.

32b The cowboy will snap the leather whip. And then, he will examine the whip.

32c The cowboy will snap the ruler. But first, he will examine the ruler.

32d The cowboy will snap the ruler. And then, he will examine the ruler.

33a The worker will lean on the wall. But first, he will laugh about the wall.

33b The worker will lean on the wall. And then, he will laugh about the wall.

33c The worker will lean on the pile of papers. But first, he will laugh about the pile of papers.

33d The worker will lean on the pile of papers. And then, he will laugh about the pile of papers.

34a The toddler will sit on the bed. But first, she will look at the bed.

34b The toddler will sit on the bed. And then, she will look at the bed.

34c The toddler will sit on the tomato. But first, she will look at the tomato.

34d The toddler will sit on the tomato. And then, she will look at the tomato.

35a The woman will throw the Frisbee. But first, she will pick up the Frisbee.

35b The woman will throw the Frisbee. And then, she will pick up the Frisbee.

35c The woman will throw the plate. But first, she will pick up the plate.

35d The woman will throw the plate. And then, she will pick up the plate.

36a The man will fan the tent. But first, he will inspect the tent.

36b The man will fan the tent. And then, he will inspect the tent.

36c The man will fan the campfire. But first, he will inspect the campfire.

36d The man will fan the campfire. And then, he will inspect the campfire.

A1.5 Experimental sentences – Experiment 7

a = but first, the; b = but first, another

01a	The man will shut the gate. But first, he will lean on the gate.
01b	The man will shut the gate. But first, he will lean on another gate.
02a	The teenager will tip over the dustbin. But first, he will walk around the dustbin.
02b	The teenager will tip over the dustbin. But first, he will walk around another dustbin.
03a	The bride will unwrap the gift. But first, she will give thanks for the gift.
03b	The bride will unwrap the gift. But first, she will give thanks for another gift.
04a	The squirrel will crack the acorn. But first, it will lick the acorn.
04b	The squirrel will crack the acorn. But first, it will lick another acorn.
05a	The gorilla will peel the banana. But first, he will grunt at the banana.
05b	The gorilla will peel the banana. But first, he will grunt at another banana.
06a	The girl will cut the jump rope. But first, she will complain about the jump rope.
06b	The girl will cut the jump rope. But first, she will complain about another jump rope.
07a	The young man will blow out the candle. But first, he will reach for the candle.
07b	The young man will blow out the candle. But first, he will reach for another candle.
08a	The child will knock down the plant. But first, he will jump over the plant.
08b	The child will knock down the plant. But first, he will jump over another plant.
09a	The teacher will wipe clear the blackboard. But first, she will point at the blackboard.
09b	The teacher will wipe clear the blackboard. But first, she will point at another blackboard.
10c	The customer will unfold the sunglasses. But first, she will ask about the sunglasses.
10d	The customer will unfold the sunglasses. But first, she will ask about another sunglasses.
11a	The woman will drop the ice cream. But first, she will grin at the ice cream.
11b	The woman will drop the ice cream. But first, she will grin at another ice cream.
12a	The customer will break the glasses. But first, she will ask about the sunglasses.
12b	The customer will break the glasses. But first, she will ask about another sunglasses.
13a	The girl will chew the stick of gum. But first, she will complain about the gum.
13b	The girl will chew the stick of gum. But first, she will complain about another gum.

14a The chef will chop the onion. But first, she will smell the onion.
14b The chef will chop the onion. But first, she will smell another onion.
15a The boy will break the vase. But first, he will examine the vase.
15b The boy will break the vase. But first, he will examine another vase.
16a The waiter will unroll the rug. But first, he will step over the rug.
16b The waiter will unroll the rug. But first, he will step over another rug.
17a The student will cut up the pizza. But first, he will blow on the pizza.
17b The student will cut up the pizza. But first, he will blow on another pizza.
18a The welder will light the blowtorch. But first, he will fiddle with the blowtorch.
18b The welder will light the blowtorch. But first, he will fiddle with another blowtorch.
19a The child will colour in the drawing. But first, she will display the drawing.
19b The child will colour in the drawing. But first, she will display another drawing.
20c The grocer will purée the blueberries. But first, he will smell the blueberries.
20d The grocer will purée the blueberries. But first, he will smell another blueberries.
21a The woman will squeeze the cucumber. But first, she will talk about the cucumber.
21b The woman will squeeze the cucumber. But first, she will talk about another cucumber.
22a The man will slice the bagel. But first, he will smell the bagel.
22b The man will slice the bagel. But first, he will smell another bagel.
23a The man will roast the piglet. But first, he will examine the piglet.
23b The man will roast the piglet. But first, he will examine another piglet.
24a The farmer will carve the pumpkin. But first, he will photograph the pumpkin.
24b The farmer will carve the pumpkin. But first, he will photograph another pumpkin.
25a The boy will deflate the ball. But first, he will throw the ball.
25b The boy will deflate the ball. But first, he will throw another ball.
26a The woman will break the egg. But first, she will examine the egg.
26b The woman will break the egg. But first, she will examine another egg.
27a The maid will make up the messy bed. But first, she will complain about the bed.
27b The maid will make up the messy bed. But first, she will complain about another bed.
28a The traveller will roll up the sleeping bag. But first, he will moan about the sleeping bag.
28b The traveller will roll up the sleeping bag. But first, he will moan about another sleeping bag.
29a The lumberjack will chop down the tree. But first, he will point at the tree.
29b The lumberjack will chop down the tree. But first, he will point at

another tree.

30c The woman will open the umbrella. But first, she will compliment at the umbrella.

30d The woman will open the umbrella. But first, she will compliment at another umbrella.

31a The woman will halve the apple. But first, she will talk about the apple.

31b The woman will halve the apple. But first, she will talk about another apple.

32a The construction worker will knock down the wall. But first, she will frown at the wall.

32b The construction worker will knock down the wall. But first, she will frown at another wall.

A1.6 Experimental sentences – Experiment 8

No.	Stimuli
	a = but first, the; b = but first, another; c = and then, the; d = and then, another
01a	The man will shut the gate. But first, he will lean on the gate.
01b	The man will shut the gate. But first, he will lean on another gate.
01c	The man will shut the gate. And then, he will lean on the gate.
01d	The man will shut the gate. And then, he will lean on another gate.
02a	The teenager will tip over the dustbin. But first, he will walk around the dustbin.
02b	The teenager will tip over the dustbin. But first, he will walk around another dustbin.
02c	The teenager will tip over the dustbin. And then, he will walk around the dustbin.
02d	The teenager will tip over the dustbin. And then, he will walk around another dustbin.
03a	The bride will unwrap the gift. But first, she will give thanks for the gift.
03b	The bride will unwrap the gift. But first, she will give thanks for another gift.
03c	The bride will unwrap the gift. And then, she will give thanks for the gift.
03d	The bride will unwrap the gift. And then, she will give thanks for another gift.
04a	The squirrel will crack the acorn. But first, it will lick the acorn.
04b	The squirrel will crack the acorn. But first, it will lick another acorn.
04c	The squirrel will crack the acorn. And then, it will lick the acorn.
04d	The squirrel will crack the acorn. And then, it will lick another acorn.
05a	The gorilla will peel the banana. But first, he will grunt at the banana.
05b	The gorilla will peel the banana. But first, he will grunt at another banana.
05c	The gorilla will peel the banana. And then, he will grunt at the banana.
05d	The gorilla will peel the banana. And then, he will grunt at another banana.
06a	The girl will cut the jump rope. But first, she will complain about the jump rope.
06b	The girl will cut the jump rope. But first, she will complain about another jump rope.
06c	The girl will cut the jump rope. And then, she will complain about the jump rope.
06d	The girl will cut the jump rope. And then, she will complain about another jump rope.
07a	The young man will blow out the candle. But first, he will reach for the candle.
07b	The young man will blow out the candle. But first, he will reach for another candle.
07c	The young man will blow out the candle. And then, he will reach for the candle.
07d	The young man will blow out the candle. And then, he will reach for another candle.
08a	The child will knock down the plant. But first, he will jump over the

plant.

08b The child will knock down the plant. But first, he will jump over another plant.

08c The child will knock down the plant. And then, he will jump over the plant.

08d The child will knock down the plant. And then, he will jump over another plant.

09a The teacher will wipe clear the blackboard. But first, she will point at the blackboard.

09b The teacher will wipe clear the blackboard. But first, she will point at another blackboard.

09c The teacher will wipe clear the blackboard. And then, she will point at the blackboard.

09d The teacher will wipe clear the blackboard. And then, she will point at another blackboard.

10a The customer will unfold the paper airplane. But first, she will ask about the paper airplane.

10b The customer will unfold the paper airplane. But first, she will ask about another paper airplane.

10c The customer will unfold the paper airplane. And then, she will ask about the paper airplane.

10d The customer will unfold the paper airplane. And then, she will ask about another paper airplane.

11a The woman will drop the ice cream. But first, she will grin at the ice cream.

11b The woman will drop the ice cream. But first, she will grin at another ice cream.

11c The woman will drop the ice cream. And then, she will grin at the ice cream.

11d The woman will drop the ice cream. And then, she will grin at another ice cream.

12a The customer will break the glasses. But first, she will ask about the glasses.

12b The customer will break the glasses. But first, she will ask about another pair of glasses.

12c The customer will break the glasses. And then, she will ask about the glasses.

12d The customer will break the glasses. And then, she will ask about another pair of glasses.

13a The girl will snap the ruler. But first, she will pick up the ruler.

13b The girl will snap the ruler. But first, she will pick up another ruler.

13c The girl will snap the ruler. And then, she will pick up the ruler.

13d The girl will snap the ruler. And then, she will pick up another ruler.

14a The chef will chop the onion. But first, she will smell the onion.

14b The chef will chop the onion. But first, she will smell another onion.

14c The chef will chop the onion. And then, she will smell the onion.

14d The chef will chop the onion. And then, she will smell another onion.

15a The boy will break the vase. But first, he will examine the vase.

15b The boy will break the vase. But first, he will examine another vase.

15c The boy will break the vase. And then, he will examine the vase.
15d The boy will break the vase. And then, he will examine another vase.
16a The waiter will unroll the rug. But first, he will step over the rug.
16b The waiter will unroll the rug. But first, he will step over another rug.
16c The waiter will unroll the rug. And then, he will step over the rug.
16d The waiter will unroll the rug. And then, he will step over another rug.
17a The student will cut up the pizza. But first, he will blow on the pizza.
17b The student will cut up the pizza. But first, he will blow on another pizza.
17c The student will cut up the pizza. And then, he will blow on the pizza.
17d The student will cut up the pizza. And then, he will blow on another pizza.
18a The welder will light the blowtorch. But first, he will fiddle with the blowtorch.
18b The welder will light the blowtorch. But first, he will fiddle with another blowtorch.
18c The welder will light the blowtorch. And then, he will fiddle with the blowtorch.
18d The welder will light the blowtorch. And then, he will fiddle with another blowtorch.
19a The child will colour in the drawing. But first, she will display the drawing.
19b The child will colour in the drawing. But first, she will display another drawing.
19c The child will colour in the drawing. And then, she will display the drawing.
19d The child will colour in the drawing. And then, she will display another drawing.
20a The customer will unfold the sunglasses. But first, she will ask about the sunglasses.
20b The customer will unfold the sunglasses. But first, she will ask about another pair of sunglasses.
20c The customer will unfold the sunglasses. And then, she will ask about the sunglasses.
20d The customer will unfold the sunglasses. And then, she will ask about another pair of sunglasses.
21a The secretary will spill the coffee. But first, she will make the coffee.
21b The secretary will spill the coffee. But first, she will make another coffee.
21c The secretary will spill the coffee. And then, she will make the coffee.
21d The secretary will spill the coffee. And then, she will make another coffee.
22a The man will slice the bagel. But first, he will smell the bagel.
22b The man will slice the bagel. But first, he will smell another bagel.
22c The man will slice the bagel. And then, he will smell the bagel.
22d The man will slice the bagel. And then, he will smell another bagel.
23a The man will roast the piglet. But first, he will examine the piglet.
23b The man will roast the piglet. But first, he will examine another piglet.
23c The man will roast the piglet. And then, he will examine the piglet.

23d The man will roast the piglet. And then, he will examine another piglet.

24a The farmer will carve the pumpkin. But first, he will photograph the pumpkin.

24b The farmer will carve the pumpkin. But first, he will photograph another pumpkin.

24c The farmer will carve the pumpkin. And then, he will photograph the pumpkin.

24d The farmer will carve the pumpkin. And then, he will photograph another pumpkin.

25a The boy will deflate the ball. But first, he will throw the ball.

25b The boy will deflate the ball. But first, he will throw another ball.

25c The boy will deflate the ball. And then, he will throw the ball.

25d The boy will deflate the ball. And then, he will throw another ball.

26a The woman will break the egg. But first, she will examine the egg.

26b The woman will break the egg. But first, she will examine another egg.

26c The woman will break the egg. And then, she will examine the egg.

26d The woman will break the egg. And then, she will examine another egg.

27a The cleaner will tidy the messy desk. But first, she will complain about the desk.

27b The cleaner will tidy the messy desk. But first, she will complain about another desk.

27c The cleaner will tidy the messy desk. And then, she will complain about the desk.

27d The cleaner will tidy the messy desk. And then, she will complain about another desk.

28a The traveller will roll up the sleeping bag. But first, he will moan about the sleeping bag.

28b The traveller will roll up the sleeping bag. But first, he will moan about another sleeping bag.

28c The traveller will roll up the sleeping bag. And then, he will moan about the sleeping bag.

28d The traveller will roll up the sleeping bag. And then, he will moan about another sleeping bag.

29a The lumberjack will chop down the tree. But first, he will point at the tree.

29b The lumberjack will chop down the tree. But first, he will point at another tree.

29c The lumberjack will chop down the tree. And then, he will point at the tree.

29d The lumberjack will chop down the tree. And then, he will point at another tree.

30a The woman will open the umbrella. But first, she will compliment at the umbrella.

30b The woman will open the umbrella. But first, she will compliment at another umbrella.

30c The woman will open the umbrella. And then, she will compliment at the umbrella.

30d The woman will open the umbrella. And then, she will compliment at another umbrella.

31a The woman will halve the apple. But first, she will talk about the apple.

31b The woman will halve the apple. But first, she will talk about another apple.

31c The woman will halve the apple. And then, she will talk about the apple.

31d The woman will halve the apple. And then, she will talk about another apple.

32a The construction worker will knock down the wall. But first, she will frown at the wall.

32b The construction worker will knock down the wall. And then, she will frown at another wall.

32c The construction worker will knock down the wall. But first, she will frown at the wall.

32d The construction worker will knock down the wall. And then, she will frown at another wall.

33a The student will snap the ruler. But first, she will pick up the ruler.

33b The student will snap the ruler. But first, she will pick up another ruler.

33c The student will snap the ruler. And then, she will pick up the ruler.

33d The student will snap the ruler. And then, she will pick up another ruler.

34a The dog will empty his food bowl. But first, he will smell the food bowl.

34b The dog will empty his food bowl. But first, he will smell another food bowl.

34c The dog will empty his food bowl. And then, he will smell the food bowl.

34d The dog will empty his food bowl. And then, he will smell another food bowl.

35a The man will smash the mirror. But first, she will examine the mirror.

35b The man will smash the mirror. But first, she will examine another mirror.

35c The man will smash the mirror. And then, she will examine the mirror.

35d The man will smash the mirror. And then, she will examine another mirror.

36a The man will bite the chocolate. But first, he will smell the chocolate.

36b The man will bite the chocolate. But first, he will smell another chocolate.

36c The man will bite the chocolate. And then, he will smell the chocolate.

36d The man will bite the chocolate. And then, he will smell another chocolate.

A1.7 Experimental sentences – Experiment 9

No.	A. No Change, And then	B. Change, And then	C. Implausible
1	The man will buy the hat. And then, he will ask about the hat. It is made of wool.	The man will ruin the hat. And then, he will ask about the hat. It is made of wool.	The man will buy the hat. And then, he will punch the hat. It is smiling.
2	The airline pilot will land the airplane. And then, he will steer the airplane. It is a Boeing 747.	The airline pilot will crash the airplane. And then, he will steer the airplane. It is a Boeing 747.	The pilot will slow down the airplane. And then, he will bloom the airplane. It is rotten.
3	The boy will chase the mouse. And then, he will follow the mouse. It is very scared.	The boy will injure the mouse. And then, he will follow the mouse. It is very scared.	The boy will tease the mouse. And then, he will justify the mouse. It is blue.
4	The music conductor will wave the baton. And then, he will complain about the baton. It is made of plastic.	The music conductor will snap the baton. And then, he will complain about the baton. It is made of plastic.	The music conductor will shake the baton. And then, he will jail the baton. It is happy.
5	The student will purchase the blank canvas. And then, he will ask for the canvas. It is good quality.	The student will paint the blank canvas. And then, he will ask for the canvas. It is good quality.	The student will stare at the blank canvas. And then, he will facility the blank canvas. It is stormy.
6	The woman will study the sculpture. And then, she will describe the sculpture. It was created by Rodin.	The woman will alter the sculpture. And then, she will describe the sculpture. It was created by Rodin.	The woman will appreciate the sculpture. And then, she will upset the sculpture. It is embarrassed.
7	The man will sit beside the fan. And then, he will comment on the fan. It was bought 10 years ago.	The man will repair the fan. And then, he will comment on the fan. It was bought 10 years ago.	The man will put away the fan. And then, he will heal the fan. It is silky.
8	The woman will seek the chocolate. And then, she will grimace at the chocolate. It has an orange flavour.	The woman will melt the chocolate And then, she will grimace at the chocolate. It has an orange flavour.	The woman will taste the chocolate. And then, she will knit the chocolate. It is wet.
9	The babysitter will put away the toy truck. And then, she will grumble about the truck. It is	The babysitter will disintegrate the toy truck. And then, she will grumble about the	The babysitter will polish the toy truck. And then, she will smoke the truck. It

	made of rubber.	truck. It is made of rubber.	is warm.
10	The barista will point to the latte. And then, she will make the latte. It is the most popular coffee.	The barista will spill the latte. And then, she will make the latte. It is the most popular coffee.	The barista will make the latte. And then, she will weigh the latte. It is solid.
11	The bartender will hold the tequila bottle. And then, he will talk about the bottle. It has an interesting design.	The bartender will break the tequila bottle. And then, he will talk about the bottle. It has an interesting design.	The bartender will empty the tequila bottle. And then, he will welcome the bottle. It is terrified.
12	The lumberjack will circle around the tree. And then, he will push against the tree. It is an pine tree.	The lumberjack will chop down the tree. And then, he will push against the tree. It is an pine tree.	The lumberjack will choose the tree. And then, he will unpack the tree. It is sad.
13	The bicyclist will inspect the tire. And then, she will complain about the tire. It is dirty.	The bicyclist will puncture the tire. And then, she will complain about the tire. It is dirty.	The bicyclist will moan about the tire. And then, she will warm up the tire. It is wooden.
14	The botanist will collect the plant. And then, he will document the plant. It is rare.	The botanist will dissect the plant. And then, he will document the plant. It is rare.	The botanist will discover the plant. And then, he will refuse the plant. It has wings.
15	The boxer will step towards the training bag. And then, he will look at the bag. It is made in Italy.	The boxer will cut apart the training bag. And then, he will look at the bag. It is made in Italy.	The boxer will punch the training bag. And then, he will injure the bag. It is sharp.
16	The boy will try on the shoes. And then, he will brag about the shoes. They are made in China.	The boy will wear out the shoes. And then, he will brag about the shoes. They are made in China.	The boy will share the shoes. And then, he will upset the shoes. They are made of paper.
17	The boy will smile at the candle. And then, he will reach for the candle. It is red.	The boy will blow out the candle. And then, he will reach for the candle. It is red.	The boy will touch the candle. And then, he will motivate the candle. It is frantic.
18	The boy will play with the broken zipper. And then, he will complain about the zipper. It is his brothers'	The boy will fix the broken zipper. And then, he will complain about the zipper It is his brothers'	The boy will purchase the broken zipper. And then, he will spell the zipper. It is slippery.

19	The boy will ask for the kite. And then, he will complement the kite. It is colourful.	The boy will crease the kite. And then, he will complement the kite. It is colourful.	The boy will present the kite. And then, he will punish the kite. It is juicy.
20	The boy will admire the coin. And then, he will grin at the coin. It is a penny.	The boy will spin the coin. And then, he will grin at the coin. It is a penny.	The boy will collect the coin. And then, he will rehabilitate the coin. It is watery.
21	The toddler will shake the empty mug. And then, he will giggle about the mug. It is green.	The toddler will smash the empty mug. And then, he will giggle about the mug. It is green.	The toddler will possess the empty mug. And then, he will fax the mug. It is sleepy.
22	The boy will show off the Easter egg. And then, he will guard the egg. It has a moat.	The boy will hammer the Easter egg. And then, he will guard the egg. It has a moat.	The boy will restore the Easter egg. And then, he will undress the egg. It is tired.
23	The boy will swing at the snowman. And then, he will laugh at the snowman. It has a carrot for a nose.	The boy will destroy the snowman. And then, he will laugh at the snowman. It has a carrot for a nose.	The boy will sketch the snowman. And then, he will slay the snowman. It is bewildered.
24	The boy will ask about the Halloween costume. And then, he will try on the costume. It is cheap.	The boy will tear the Halloween costume. And then, he will try on the costume. It is cheap.	The boy will ask for the Halloween costume. And then, he will smoke the costume. It is slow.
25	The boy will study the advent calendar. And then, he will frown at the calendar. It is creepy.	The boy will open the advent calendar. And then, he will frown at the calendar. It is creepy.	The boy will mend the advent calendar. And then, he will rely on the calendar. It is multiplied.
26	The bride will accept the gift. And then, she will give thanks for the gift. It looks expensive.	The bride will unwrap the gift. And then, she will give thanks for the gift. It looks expensive.	The bride will moan about the gift. And then, she will mentor the gift. It is shaking.
27	The man will present the fragile ornament. And then, he will uncover the ornament. It is an antique.	The man will chip the fragile ornament. And then, he will uncover the ornament. It is an antique.	The man will protect the fragile ornament. And then, he will promise the ornament. It is melted.
28	The man will praise the rug. And then, he will	The man will unroll the rug. And then, he will	The man will lie on the rug. And then,

	step over the rug. It is from Egypt.	step over the rug. It is from Egypt.	he will overtake the rug. It is delicious.
29	The camper will go into the tent. And then, he will chat about the tent. It is waterproof.	The camper will put up the tent. And then, he will chat about the tent. It is waterproof.	The camper will put together the tent. And then, he will interrupt the tent. It is delirious.
30	The carpenter will wear the jacket. And then, he will frown at the jacket It is dusty.	The carpenter will rumple the jacket. And then, he will frown at the jacket. It is dusty.	The carpenter will put on the jacket. And then, he will mentor the jacket. It is noisy.
31	The carpenter will stand on the floor. And then, he will frown at the floor. It is muddy.	The carpenter will tile the floor. And then, he will step on the floor. It is muddy.	The carpenter will tile the floor. And then, he will sew the floor. It is envious.
32	The chef will select the onion. And then, she will smell the onion. It is fresh.	The chef will chop the onion. And then, she will smell the onion. It is fresh.	The chef will smell the onion. And then, she will telephone the onion. It is odourless.
33	The child will hide behind the plant. And then, he will jump over the plant. It is leafy.	The child will pluck the plant. And then, he will jump over the plant. It is leafy.	The child will water the plant. And then, he will transcribe the plant. It is tired.
34	The child will lick the ice lolly. And then, she will grin at the lolly. It is delicious..	The child will bite the ice lolly. And then, she will grin at the lolly It is delicious..	The child will grin at the ice lolly. And then, she will revise the lolly. It is boiling.
35	The teenager will take the flyer. And then, he will grin at the flyer. It is about guitar lessons.	The teenager will rip the flyer. And then, he will grin at the flyer. It is about guitar lessons.	The teenager will grin at the flyer. And then, he will rescue the flyer. It is amused.
36	The child will poke the gingerbread. And then, she will laugh at the gingerbread. It is fragile.	The child will decorate the gingerbread. And then, she will laugh at the gingerbread. It is fragile.	The child will laugh at the gingerbread. And then, she will unplug the gingerbread. It is shining.
37	The child will find the ketchup bottle. And then, he will comment on the bottle. It is Heinz.	The child will open the ketchup bottle. And then, he will comment on the bottle. It is Heinz.	The child will comment on the ketchup bottle. And then, he will forgive the bottle. It is cautious.

38	The clown will rub the balloon. And then, he will laugh at the balloon. It is pink.	The clown will inflate the balloon. And then, he will laugh at the balloon. It is pink.	The clown will laugh at the balloon. And then, he will write the balloon. It is courageous.
39	The construction worker will stand by the wall. And then, he will frown at the wall. It is old.	The construction worker will knock down the wall. And then, he will frown at the wall. It is old.	The construction worker will frown at the wall. And then, he will repeat the wall. It is cloudy.
40	The contractor will sit by the swimming pool. And then, he will walk around the pool. It is clean.	The contractor will drain the swimming pool. And then, he will walk around the pool. It is clean.	The contractor will walk around the swimming pool. And then, he will write the pool. It is jealous.
41	The customer will try on the sunglasses. And then, she will ask about the sunglasses. They are expensive.	The customer will unfold the sunglasses. And then, she will ask about the sunglasses. They are expensive.	The customer will ask about the sunglasses. And then, she will spray the sunglasses. They are hungry.
42	The dentist will brush the tooth. And then, he will talk about the tooth. It is a wisdom tooth.	The dentist will drill the tooth. And then, he will talk about the tooth. It is a wisdom tooth.	The dentist will talk about the tooth. And then, he will tame the tooth. It is dangerous.
43	The boy will pick up the homework. And then, he will frown at the homework. It is difficult.	The boy will tear up the homework. And then, he will frown at the homework. It is difficult.	The boy will frown at the homework. And then, he will recruit the homework. It is enthusiastic.
44	The girl will grab the glass. And then, she will smell the glass. It is stinky.	The girl will drop the glass. And then, she will smell the glass. It is stinky.	The girl will smell the glass. And then, she will comb the glass. It is faithful.
45	The employee will stand beside the mannequin. And then, she will laugh at the mannequin. It has red hair.	The employee will undress the mannequin. And then, she will laugh at the mannequin. It has red hair.	The employee will laugh at the mannequin. And then, she will invent the mannequin. It is helpless.
46	The woman will touch the blanket. And then, she will look at the blanket. It is white.	The woman will bleach the blanket. And then, she will look at the blanket. It is white.	The woman will look at the blanket. And then, she will disarm the blanket.

			It is asleep.
47	The expert will study the oil. And then, he will recommend the oil. It is from Texas.	The expert will burn the oil. And then, he will recommend the oil. It is from Texas.	The expert will recommend the oil. And then, he will wrestle the oil. It is selfish.
48	The exterminator will spot the cockroach. And then, he will talk about the cockroach. It is brown.	The exterminator will poison the cockroach. And then, he will talk about the cockroach. It is brown.	The exterminator will talk about the cockroach. And then, he will interview the cockroach. It is sparkling.
49	The farmer will photograph the pumpkin. And then, he will chat about the pumpkin. It is heavy.	The farmer will carve the pumpkin. And then, he will chat about the pumpkin. It is heavy.	The farmer will chat about the pumpkin. And then, he will command the pumpkin. It is thoughtful.
50	The fireman will search for the fire hose. And then, he will ask about the hose. It is not working.	The fireman will unravel the fire hose. And then, he will ask about the hose. It is not working.	The fireman will ask about the fire hose. And then, he will interpret the hose. It is busy.
51	The chef will gaze at the duck. And then, he will ask for the duck. It is big.	The chef will defrost the duck. And then, he will ask for the duck. It is big.	The chef will ask for the duck. And then, he will forecast the duck. It is cautious.
52	The food critic will taste the coffee. And then, he will ask about the coffee. It smells nice.	The food critic will finish the coffee. And then, he will ask about the coffee. It smells nice.	The food critic will ask about the coffee. And then, he will excite the coffee. It is adventurous.
53	The girl will admire the flower. And then, she will sniff the flower. It is pretty.	The girl will pluck the flower. And then, she will sniff the flower. It is pretty.	The girl will sniff the flower. And then, she will compete with the flower. It is aggressive.
54	The girl will get the stick of gum. And then, she will complain about the gum. It is cinnamon flavour.	The girl will chew the stick of gum. And then, she will complain about the gum. It is cinnamon flavour.	The girl will complain about the stick of gum. And then, she will marry the gum. It is alert.

55	The man will carry the bomb. And then, he will find the bomb. It is well hidden.	The man will detonate the bomb. And then, he will find the bomb. It is well hidden.	The man will find the bomb. And then, he will boil the bomb. It is alive.
56	The girl will lie beside the diary. And then, she will read the diary. It is written by her mother.	The girl will rip apart the diary. And then, she will read the diary. It is written by her mother.	The girl will read the diary. And then, she will catch the diary. It is amused.
57	The girl will pose behind the shirt. And then, she will ask about the shirt. It is for tomorrow's party.	The girl will crumple the shirt. And then, she will ask about the shirt. It is for tomorrow's party.	The girl will ask about the shirt. And then, she will begin the shirt. It is arrogant.
58	The gorilla will grab the banana. And then, he will grunt at the banana. It smells nice.	The gorilla will peel the banana. And then, he will grunt at the banana. It smells nice.	The gorilla will grunt at the banana. And then, he will justify the banana. It is ashamed.
59	The grandfather will photograph the turkey. And then, he will pick up the turkey. It is huge.	The grandfather will cut up the turkey. And then, he will pick up the turkey. It is huge.	The grandfather will pick up the turkey. And then, he will install the turkey. It is bewildered.
60	The grandmother will look toward the lamp. And then, she will ask about the lamp. It is cute.	The grandmother will turn on the lamp. And then, she will ask about the lamp. It is cute.	The grandmother will ask about the lamp. And then, she will drink the lamp. It is fluid.
61	The grocer will weigh the blueberries. And then, he will smell the blueberries. They are from New Zealand.	The grocer will blend the blueberries. And then, he will smell the blueberries. They are from New Zealand.	The grocer will smell the blueberries. And then, he will melt the blueberries. They are transparent.
62	The groomer will pet the dog. And then, he will feed the dog. It is a Labrador.	The groomer will shave the dog. And then he will feed the dog. It is a Labrador.	The groomer will feed the dog. And then he will convert the dog. It is talkative.
63	The grouch will curse at the alarm clock. And then, he will hit the clock. It is loud.	The grouch will turn off the alarm clock. And then, he will hit the clock. It is loud.	The grouch will hit the alarm clock. And then, he will mentor the clock. It is bulbous.

64	The gym teacher will show the basketball. And then, he will comment on the basketball. It is still new.	The gym teacher will deflate the basketball. And then, he will comment on the basketball. It is still new.	The gym teacher will comment on the basketball. And then, he will invent the basketball. It is brave.
65	The gymnast will adjust the jump rope. And then, she will complain about the rope. It made of plastic.	The gymnast will cut the jump rope. And then, she will complain about the rope. It made of plastic.	The gymnast will complain about the jump rope. And then, she will forgive the rope. It is cautious.
66	The hairdresser will caress the wig. And then, she will compliment the wig. It is made of real hair.	The hairdresser will loosen the wig. And then, she will compliment the wig. It is made of real hair.	The hairdresser will compliment the wig. And then, she will freeze the wig. It is cloudy.
67	The housepainter will climb the ladder. And then, he will walk around the ladder. It is tall.	The housepainter will extend the ladder. And then, he will walk around the ladder. It is tall.	The housepainter will walk around the ladder. And then, he will dramatize the ladder. It is clumsy.
68	The hunter will examine the rifle. And then, he will frown at the rifle. It is old.	The hunter will dismantle the rifle. And then, he will frown at the rifle. It is old.	The hunter will frown at the rifle. And then, he will rescue the rifle. It is equal.
69	The hunter will stalk the wildebeest. And then, he will follow the wildebeest. It is a male.	The hunter will kill the wildebeest. And then, he will follow the wildebeest. It is a male.	The hunter will follow the wildebeest. And then, he will hypothesize the wildebeest. It is fragile.
70	The illustrator will copy the cartoon. And then, he will display the cartoon. It is interesting.	The illustrator will colour in the cartoon. And then, he will display the cartoon. It is interesting.	The illustrator will display the cartoon. And then, he will ride the cartoon. It is frosty.
71	The librarian will stand behind the bookshelf. And then, she will check the bookshelf. It is made of oak.	The librarian will tidy up the bookshelf. And then, she will check the bookshelf. It is made of oak.	The librarian will check the bookshelf. And then, she will interrupt the bookshelf. It is dizzy.
72	The little girl will kiss the teddy bear. And then, she will talk to the teddy	The little girl will dress the teddy bear. And then, she will talk to	The little girl will talk to the teddy bear. And then,

	bear. It is her favourite.	the teddy bear. It is her favourite.	she will spill the teddy bear. It is hungry.
73	The lumberjack will point at the tree branch. And then, he will reach for the branch. It is long.	The lumberjack will saw off the tree branch. And then, he will reach for the branch. It is long.	The lumberjack will reach for the tree branch. And then, he will stretch the branch. It is faithful.
74	The maid will sit on the messy bed. And then, she will complain about the bed. It has a velvet sheet.	The maid will make up the messy bed. And then, she will complain about the bed. It has a velvet sheet.	The maid will complain about the messy bed. And then, she will excuse the bed. It is healthy.
75	The man will look at the wristwatch. And then, he will brag about the wristwatch. It is a counterfeit.	The man will reset the wristwatch. And then, he will brag about the wristwatch. It is a counterfeit.	The man will brag about the wristwatch. And then, he will proofread the wristwatch. It is homeless.
76	The man will criticize the brochure. And then, he will read the brochure. It is informative.	The man will fold the brochure. And then, he will read the brochure. It is informative.	The man will publicize the brochure. And then, he will sow the brochure. It is made of metal.
77	The man will choose the bagel. And then, he will smell the bagel. It is freshly baked.	The man will eat the bagel. And then, he will smell the bagel. It is freshly baked.	The man will smell the bagel. And then, he will punch the bagel. It is red.
78	The man will sketch the label. And then, he will read the label. It is in the wrong font.	The man will remove the label. And then, he will read the label. It is in the wrong font.	The man will read the label. And then, he will squeeze the label. It is joyous.
79	The man will kick the gate. And then, he will lean on the gate. It is made of iron.	The man will shut the gate. And then, he will lean on the gate. It is made of iron.	The man will lean on the gate. And then, he will summarize the gate. It is lazy.
80	The man will lean on the dishwasher. And then, he will comment on the dishwasher. It is a Bosch.	The man will take apart the dishwasher. And then, he will comment on the dishwasher. It is a Bosch.	The man will comment on the dishwasher. And then, he will obey the dishwasher. It is windproof.
81	The man will look for the mobile phone. And then, he will gripe about the phone. It is a Nokia.	The man will smash the mobile phone. And then, he will gripe about the phone. It is a	The man will gripe about the mobile phone. And then, he will hunt the

		Nokia.	phone. It is mushy.
82	The mechanic will examine the car engine. And then, he will ask about the engine. It is a diesel.	The mechanic will disassemble the car engine. And then, he will ask about the engine. It is a diesel.	The mechanic will ask about the car engine. And then, he will boil the engine. It is obedient.
83	The mother will select the egg. And then, she will talk about the egg. It is medium sized.	The mother will poach the egg. And then, she will talk about the egg. It is medium sized.	The mother will talk about the egg. And then, she will confront the egg. It is pleasant.
84	The child will try the saxophone. And then, she will ask about the saxophone. It is a birthday present.	The child will strap on the saxophone. And then, she will ask about the saxophone. It is a birthday present.	The child will ask about the saxophone. And then, she will drink the saxophone. It is prickly.
85	The musician will play the piano. And then, he will rave about the piano. It is black.	The musician will tune the piano. And then, he will rave about the piano. It is black.	The musician will rave about the piano. And then, he will drive the piano. It is puzzled.
86	The new employee will stare at the safe. And then, he will ask about the safe. It looks complicated.	The new employee will open up the safe. And then, he will ask about the safe. It looks complicated.	The new employee will ask about the safe. And then, he will convert the safe. It is shy.
87	The ninja will grip the sword. And then, he will grin at the sword. It is a gift from his father.	The ninja will unsheathe the sword. And then, he will grin at the sword. It is a gift from his father.	The ninja will grin at the sword. And then, he will load the sword. It is repulsive.
88	The office worker will type on the keyboard. And then, he will complain about the keyboard. It is dirty.	The office worker will plug in the keyboard. And then, he will complain about the keyboard. It is dirty.	The office worker will complain about the keyboard. And then, he will irritate the keyboard. It is talented.
89	The man will aim at the field mouse. And then, he will follow the mouse. It is small.	The man will shoot the field mouse. And then, he will follow the mouse. It is small.	The man will follow the field mouse. And then, he will exhibit the mouse. It is chatty.
90	The palaeontologist will dig around the fossil. And then, he will talk about the fossil. It looks	The palaeontologist will dig up the fossil. And then, he will talk about the fossil. It	The palaeontologist will talk about the fossil. And then, he

	like a skull.	looks like a skull.	will motivate the fossil. It is smoggy.
91	The parking attendant will check the parking meter. And then, she will rest against the meter. It is broken.	The parking attendant will empty the parking meter. And then, she will rest against the meter. It is broken.	The parking attendant will rest against the parking meter. And then, she will determine the meter. It is witty.
92	The police woman will investigate the violent crime. And then, she will receive reports about the crime. It is a serious case.	The police woman will solve the violent crime. And then, she will receive reports about the crime. It is a serious case.	The police woman will receive reports about the violent crime. And then, she will print the crime. It is splendid.
93	The passer-by will watch the bird. And then, she will call to the bird. It has red feathers.	The passer-by will frighten the bird. And then, she will call to the bird. It has red feathers.	The passer-by will call to the bird. And then, she will bleach the bird. It is vast.
94	The plumber will leave behind the pipe. And then, he will frown at the pipe. It is leaking.	The plumber will bend the pipe. And then, he will frown at the pipe. It is leaking.	The plumber will frown at the pipe. And then, he will fry the pipe. It is wicked.
95	The sailor will hold onto the sail. And then, he will scowl at the sail. It is ragged.	The sailor will take down the sail. And then, he will scowl at the sail. It is ragged.	The sailor will scowl at the sail. And then, he will salt the sail. It is woolly.
96	The scuba diver will peer through the diving mask. And then, he will tug at the mask. It is his brother's.	The scuba diver will wipe clear the diving mask. And then, he will tug at the mask. It is his brother's.	The scuba diver will tug at the diving mask. And then, he will compute the mask. It is stormy.
97	The secretary will search for the file. And then, she will gripe about the file. It is top secret.	The secretary will shred the file. And then, she will gripe about the file. It is top secret.	The secretary will gripe about the file. And then, she will mourn the file. it is thoughtful.
98	The squirrel will sniff the acorn. And then, it will lick the acorn. It is tiny.	The squirrel will crack the acorn. And then, it will lick the acorn. It is tiny.	The squirrel will lick the acorn. And then, it will fold the acorn. It is worried.
99	The student will smell the pizza. And then, he will blow on the pizza. It	The student will slice the pizza. And then, he will blow on the	The student will blow on the pizza. And then, he will

	is pepperoni.	pizza. It is pepperoni.	forgive the pizza. It is nervous.
100	The student will measure the iron beam. And then, he will stare at the beam. It is for his project.	The student will engrave the iron beam. And then, he will stare at the beam. It is for his project.	The student will stare at the iron beam. And then, he will frighten the beam. It is tired.
101	The tailor will measure the dress. And then, he will ask about the dress. It is glamorous.	The tailor will shorten the dress. And then, he will ask about the dress. It is glamorous.	The tailor will ask about the dress. And then, he will rhyme the dress. It is weary.
102	The politician will support the welfare program. And then, he will praise the program. It is for working mothers.	The politician will expand the welfare program. And then, he will praise the program. It is for working mothers.	The politician will praise the welfare program. And then, he will itch the program. It is quaint.
103	The teacher will point at the blackboard. And then, she will read from the blackboard. It is huge.	The teacher will wipe clear the blackboard. And then, she will read from the black board. It is huge.	The teacher will read from the blackboard. And then, she will divert the black board. It is yellow.
104	The teenager will bang on the dustbin. And then, he will walk around the dustbin. It smells bad.	The teenager will tip over the dustbin. And then, he will walk around the dustbin. It smells bad.	The teenager will walk around the dustbin. And then, he will offend the dustbin. It is misty.
105	The teenager will interpret the tarot cards. And then, she will chat about the cards. They are mysterious.	The teenager will shuffle the tarot cards. And then, she will chat about the cards. They are mysterious.	The teenager will chat about the tarot cards. And then, she will chop the cards. They are round.
106	The toddler will look over the puzzle. And then, he will marvel at the puzzle. It is about jungle animals.	The toddler will break apart the puzzle. And then, he will marvel at the puzzle. It is about jungle animals.	The toddler will marvel at the puzzle. And then, he will float the puzzle. It is obedient.
107	The man will join the political commune. And then, he will study the commune. It is not well known.	The man will split the political commune. And then, he will find the commune. It is not well known.	The man will join the political commune. And then, he will warm up the commune. It is tasty.

108	The traffic cop will observe the CCTV camera. And then, he will photograph the camera. It is facing the wrong direction.	The traffic cop will stop the CCTV camera. And then, he will photograph the camera. It is facing the wrong direction.	The traffic cop will recognize the CCTV camera. And then, he will sew the camera. It is lonely.
109	The trainer will pat the horse. And then, he will feed the horse. It is a black stallion.	The trainer will saddle the horse. And then, he will feed the horse. It is a black stallion.	The trainer will feed the horse. And then, he will extract the horse. It is barking.
110	The traveller will lay on the blanket. And then, he will complain about the blanket. It is a bit wet.	The traveller will roll up the blanket. And then, he will complain about the blanket. It is a bit wet.	The traveller will complain about the blanket. And then, he will frame the blanket. It is proud.
111	The ventriloquist will address the dummy. And then, he will make fun of the dummy. It looks freaky.	The ventriloquist will seat the dummy. And then, he will make fun of the dummy. It looks freaky.	The ventriloquist will make fun of the dummy. And then, he will persuade the dummy. It is fluorescent.
112	The welder will clutch the blowtorch. And then, he will fiddle with the blowtorch. It is his friend's.	The welder will light the blowtorch. And then, he will fiddle with the blowtorch. It is his friend's.	The welder will fiddle with the blowtorch. And then, he will misspell the blowtorch. It is depressed.
113	The woman will pick up the potato. And then, she will talk about the potato. It is a Maris Piper.	The woman will mash the potato. And then, she will talk about the potato. It is a Maris Piper.	The woman will talk about the potato. And then, she will evacuate the potato. It is modern.
114	The woman will pick out the apple. And then, she will talk about the apple. It is sweet.	The woman will halve the apple. And then, she will talk about the apple. It is sweet.	The woman will talk about the apple. And then, she will interfere the apple. It is flying.
115	The woman will look into the pram. And then, she will kneel beside the pram. It is from Mothercare.	The woman will fold up the pram. And then, she will kneel beside the pram. It is from Mothercare.	The woman will kneel beside the pram. And then, she will flash the pram. It is leafy.
116	The woman will peer behind the curtain. And then, she will frown at the curtain. It has a	The woman will wash the curtain. And then, she will frown at the curtain. It has a floral	The woman will frown at the curtain. And then, she will land the curtain. It is stringy.

	floral pattern.	pattern.	
117	The woman will inspect the cello. And then, she will kneel beside the cello. It is a gift.	The woman will restring the cello. And then, she will kneel beside the cello. It is a gift.	The woman will kneel beside the cello. And then, she will pour the cello. It is smiling.
118	The woman will tap the table. And then, she will walk around the table. It is from IKEA.	The woman will set the table. And then, she will walk around the table. It is from IKEA.	The woman will walk around the table. And then, she will fail the table. It is hairy.
119	The architect will rate the room. And then, she will comment on the room. It is for a client's daughter.	The architect will redecorate the room. And then, she will comment on the room.. It is for a client's daughter.	The architect will comment on the room. And then, she will diagnose the room. It is dripping.
120	The zookeeper will pat the animal. And then, he will chat about the animal. It is furry.	The zookeeper will rescue the animal. And then, he will chat about the animal. It is furry.	The zookeeper will chat about the animal. And then, he will hover the animal. It is sugary.

Appendix II
Supplement Tables

Table a2.1

Mean fixations and saccades towards the depicted versions of the target objects on the visual display in Experiment 4

DEPICTED STATE (PICTURE)	CHANGE OF STATE (VERB)	Mean (SD) (Percentage of trials %)	The woman will	choose /drop	the ice cream.	But first/ But first,		she will		look at		the ice cream.	
						FIRST	THEN	FIRST	THEN	FIRST	THEN	FIRST	THEN
INTACT STATE	CHOOSE	Saccades	40 (14)	21 (9)	38 (11)	26 (12)	22 (18)	9 (11)	9 (13)	18 (13)	16 (12)	32 (18)	29 (17)
		Fixations at the onset	0 (0)	22 (10)	26 (11)	50 (19)	53 (20)	43 (23)	47 (22)	44 (23)	45 (20)	39 (21)	36 (18)
		Fixations at the offset	N/A	26 (10)	42 (12)	47 (20)	48 (20)	44 (23)	45 (20)	39 (21)	36 (18)	45 (20)	39 (21)
	DROP	Saccades	43 (15)	21 (9)	28 (12)	22 (16)	19 (14)	9 (12)	7 (9)	17 (14)	16 (15)	39 (18)	30 (17)
		Fixations at the onset	0 (0)	24 (10)	28 (10)	31 (15)	32 (18)	26 (16)	26 (17)	26 (17)	24 (17)	28 (17)	23 (16)
		Fixations at the offset	N/A	26 (10)	30 (13)	26 (15)	27 (17)	26 (17)	24 (16)	28 (17)	23 (16)	46 (19)	34 (16)
CHANGED STATE	CHOOSE	Saccades	43 (17)	21 (9)	34 (11)	27 (16)	24 (17)	11 (13)	10 (11)	19 (14)	22 (16)	31 (18)	26 (18)
		Fixations at the onset	0 (0)	29 (11)	30 (9)	33 (16)	32 (16)	31 (18)	33 (17)	32 (20)	34 (16)	34 (16)	37 (19)
		Fixations at the offset	N/A	29 (9)	35 (10)	33 (16)	33 (18)	32 (20)	34 (16)	34 (16)	37 (19)	35 (14)	37 (17)
	DROP	Saccades	45 (15)	24 (12)	36 (11)	28 (16)	22 (15)	9 (9)	9 (11)	18 (13)	21 (15)	24 (18)	31 (18)
		Fixations at the onset	0 (0)	30 (11)	41 (14)	54 (17)	55 (23)	49 (24)	52 (22)	46 (21)	50 (23)	41 (18)	45 (20)
		Fixations at the offset	N/A	35 (12)	53 (15)	53 (23)	52 (22)	47 (22)	50 (23)	41 (18)	45 (20)	35 (18)	47 (19)

Table a2.2

Mean fixations and saccades towards the depicted versions of the target objects on the visual display in Experiment 5

DEPICTED STATE	OBJECT TYPE	Mean (SD) (Percentage of trials %)	The girl will	stomp on	the penny/egg.	But first/ And then,		she will		look at		the penny/egg.	
						FIRST	THEN	FIRST	THEN	FIRST	THEN	FIRST	THEN
INTACT VERSION	PENNY	Saccades	26 (12)	14 (8)	29 (11)	28 (15)	26 (15)	15(12)	14(11)	19(14)	19(14)	18(13)	22(14)
		Fixations at the onset	0 (0)	16 (10)	18 (12)	31 (14)	31 (14)	28(16)	29(14)	30 (17)	27 (12)	28(16)	29(14)
		Fixations at the offset	N/A	18 (12)	27 (10)	28 (16)	29 (14)	30(17)	27(12)	22(20)	22(16)	27(16)	32(15)
	EGG	Saccades	29 (11)	15 (10)	29 (12)	24 (15)	24 (14)	10(10)	12(10)	22(13)	18(15)	25(13)	19(11)
		Fixations at the onset	0 (0)	21 (11)	20 (12)	26 (16)	24 (15)	25(14)	22(14)	23 (14)	22 (13)	29(14)	26(15)
		Fixations at the offset	N/A	20 (12)	24 (11)	25 (14)	22(14)	23(14)	22(13)	22(16)	19(16)	35(17)	26(12)
CHANGED VERSION	PENNY	Saccades	29 (11)	14 (9)	29 (12)	27 (13)	26 (15)	12(12)	15(12)	21(13)	20(12)	22(13)	21(14)
		Fixations at the onset	0 (0)	19 (9)	18 (9)	32 (15	32 (16)	30 (16)	28(17)	24 (15)	29(16)	30(18)	28(15)
		Fixations at the offset	N/A	18 (9)	27 (9)	30 (16)	28 (17)	24(15)	29(16)	21 (18)	22(15)	34(15)	30(15)
	EGG	Saccades	31(10)	18 (8)	37 (12)	41 (14)	40 (16)	15(12)	13(10)	16(12)	17(11)	18(12)	22(14)
		Fixations at the onset	0 (0)	28 (9)	30 (11)	42 (15	43 (17)	40 (17)	42 (16)	41(17)	40(16)	35(17)	36(17)
		Fixations at the offset	N/A	30 (11)	41 (11)	40 (17)	42 (16)	41(17)	40(16)	28(20)	30(18)	29(16)	39(16)

***Note*:** The two versions of objects that were not expected to experience any change, such as the penny, were the identical on the visual display. Thus, "intact/changed version" of the "*penny*" refer to the corresponding locations of the "*egg*" rather than change of state.

Table a2.3

Mean fixations and saccades towards the depicted versions of the target object on the visual display in Experiment 6.

DEPICTED STATE (PICTURE)	*DESCRIBED CHANGE (VERB)*	*Mean (SD) (Percentage of trials %)*	*The woman will*	*choose /drop*	*the ice cream.*	*But first/ And then,*		*she will*		*look at*		*the ice cream.*	
						FIRST	*THEN*	*FIRST*	*THEN*	*FIRST*	*THEN*	*FIRST*	*THEN*
INTACT VERSION	*CHOOSE*	*Saccades*	*40(15)*	*23(13)*	*37(14)*	*27(14)*	*25(15)*	*8(10)*	*11(10)*	*17(12)*	*19(18)*	*33(19)*	*28(14)*
		Fixations at the onset	*0 (0)*	*25(12)*	*28(11)*	*50(17)*	*45(17)*	*45(19)*	*48(18)*	*42(20)*	*46(19)*	*41(18)*	*41(17)*
		Fixations at the offset	*N/A*	*28(10)*	*37(14)*	*49(21)*	*47(20)*	*42(20)*	*46(19)*	*41(18)*	*41(17)*	*49(17)*	*48(17)*
	DROP	*Saccades*	*40(14)*	*21(9)*	*34(14)*	*19(14)*	*22(15)*	*9(10)*	*8(11)*	*18(14)*	*17(12)*	*31(13)*	*26(15)*
		Fixations at the onset	*0 (0)*	*25(11)*	*28(10)*	*35(17)*	*31(14)*	*30(17)*	*28(15)*	*30(15)*	*27(13)*	*33(19)*	*27(16)*
		Fixations at the offset	*N/A*	*28 (9)*	*35(12)*	*27(15)*	*31(16)*	*30(15)*	*26(13)*	*33(19)*	*27(16)*	*42(19)*	*31(16)*
CHANGED VERSION	*CHOOSE*	*Saccades*	*41(14)*	*20(11)*	*40(15)*	*25(18)*	*22(16)*	*9(11)*	*11(10)*	*18(14)*	*19(14)*	*28(15)*	*23(15)*
		Fixations at the onset	*0 (0)*	*31(12)*	*31(12)*	*36(15)*	*37(15)*	*36(16)*	*33(16)*	*36(19)*	*36(16)*	*33(16)*	*34(15)*
		Fixations at the offset	*N/A*	*31 (13)*	*39(12)*	*33(19)*	*36(19)*	*36(19)*	*35(16)*	*33(16)*	*34(15)*	*34(15)*	*34(17)*
	DROP	*Saccades*	*43(13)*	*22(12)*	*44(13)*	*31(17)*	*26(14)*	*9(10)*	*11(12)*	*18(13)*	*19(16)*	*27(15)*	*28(16)*
		Fixations at the onset	*0 (0)*	*29(10)*	*33(11)*	*49(19)*	*53(17)*	*48(18)*	*51(20)*	*45(18)*	*51(21)*	*39(19)*	*45(20)*
		Fixations at the offset	*N/A*	*32 (12)*	*48(12)*	*53(17)*	*51(20)*	*45(18)*	*51(21)*	*39(19)*	*45(20)*	*40(19)*	*47(17)*

Note: The two versions are not expected to be associated with the same object - being the intact state and the changed state respectively.

Table a2.4

Mean fixations and saccades towards the depicted versions of the target objects on the visual displays in Experiment 7

VISUAL DISPLAY (PICTURE)	DESCRIBED STATE	Mean (SD) (Percentage of trials %)	The woman will	drop	the ice cream.	But first,	she will	look at	the/another		ice cream	
									THE	ANOTHER	THE	ANOTHER
SIMILAR VERSIONS	INTACT STATE	Saccades	42(15)	39(13)	31(11)	21(12)	13(10)	17(11)	6(8)	17(15)	41(19)	46(20)
		Fixations at the onset	0 (0)	22(11)	23(10)	23(14)	21(12)	24(15)	27(16)	25(19)	29(18)	32(20)
		Fixations at the offset	N/A	23(13)	25(11)	21(12)	24(15)	26(14)	29(18)	32(20)	47(20)	57(22)
	CHANGED STATE	Saccades	44(14)	54(14)	47(13)	24(12)	11(9)	13(9)	4(8)	9(8)	23(16)	22(19)
		Fixations at the onset	0 (0)	31(9)	38(11)	58(18)	56(22)	51(22)	42(24)	40(25)	41(24)	30(21)
		Fixations at the offset	N/A	38(17)	58(14)	56(22)	51(22)	41(22)	41(24)	30(21)	32(17)	24(17)
DIFFERENT VERSIONS	INTACT VERSION	Saccades	39(16)	41(11)	37(13)	23(11)	12(9)	16(10)	7(9)	18(13)	38(18)	42(19)
		Fixations at the onset	0 (0)	24(12)	26(9)	28(11)	27(10)	27(11)	24(16)	30(19)	28(17)	36(18)
		Fixations at the offset	N/A	26(14)	31(12)	27(17)	27(11)	27(14)	27(17)	35(18)	47(19)	54(16)
	CHANGED STATE	Saccades	42(14)	51(12)	47(14)	26(10)	13(8)	14(9)	4(7)	12(10)	23(17)	26(15)
		Fixations at the onset	0 (0)	29(11)	35(11)	55(13)	51(19)	50(18)	45(23)	39(21)	44(211)	32(17)
		Fixations at the offset	N/A	35(14)	52(13)	51(19)	50(18)	42(19)	44(21)	32(17)	34(17)	28(14)

Table a2.5a

Mean fixations and saccades towards the depicted versions of the target objects on the visual displays in Experiment 8 when a backward temporal shift "but first" was indicated at the beginning of the second sentence.

VISUAL DISPLAY (PICTURE)	Mean (SD) (Percentage of trials %)	The woman will	drop	the ice cream.	But first,	she will	look at	the/another		ice cream	
								THE	ANOTHER	THE	ANOTHER
INTACT STATE	Saccades	34(8)	15(7)	29(9)	20(11)	12(8)	13(8)	7(9)	11(9)	38(22)	41(19)
	Fixations at the onset	0 (0)	16(6)	14(5)	15(10)	12(8)	13(8)	13(10)	13(13)	19(14)	13(9)
	Fixations at the offset	N/A	14(6)	17(6)	12(8)	13(8)	13(8)	18(14)	13(10)	33(19)	35(18)
INTACT COMPETITOR	Saccades	32(9)	15(5)	25(8)	18(10)	12(9)	14(7)	1(4)	16(11)	27(15)	30(20)
	Fixations at the onset	0 (0)	15(6)	15(7)	14(10)	13(9)	15(10)	15(11)	16(11)	14(11)	23(14)
	Fixations at the offset	N/A	15(7)	15(6)	13(9)	15(10)	15(7)	14(11)	22(14)	19(14)	30(18)
CHANGED STATE	Saccades	34(9)	17(8)	44(11)	33(13)	17(9)	19(8)	2(5)	10(11)	25(18)	13(11)
	Fixations at the onset	0 (0)	19(6)	24(8)	48(19)	48(18)	45(20)	33(19)	41(24)	30(17)	24(19)
	Fixations at the offset	N/A	24(8)	42(11)	48(18)	45(20)	37(18)	30(18)	25(20)	26(18)	12(10)
CHANGED COMPETITOR	Saccades	25(12)	12(8)	19(8)	12(8)	8(6)	13(8)	1(3)	11(13)	13(12)	10(12)
	Fixations at the onset	0 (0)	11(6)	13(8)	7(5)	9(8)	9(8)	13(11)	10(12)	11(10)	10(12)
	Fixations at the offset	N/A	13(8)	9(4)	9(8)	9(8)	11(8)	11(11)	10(12)	8(10)	7(9)

Table a2.5b

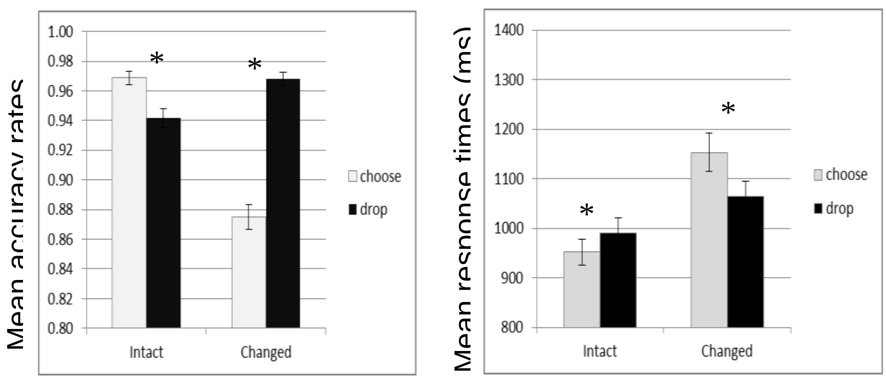
Mean fixations and saccades towards the depicted versions of the target objects on the visual displays in Experiment 8 when a forward temporal shift "and then" was indicated at the beginning of the second sentence.

VISUAL DISPLAY (PICTURE)	Mean (SD) (Percentage of trials %)	The woman will	drop	the ice cream.	And then	she will	look at	the/another		ice cream	
								THE	ANOTHER	THE	ANOTHER
INTACT STATE	Saccades	34(8)	15(7)	29(9)	21(12)	11(9)	15(9)	3(5)	15(15)	29(14)	40(15)
	Fixations at the onset	0 (0)	16(6)	14(5)	17(10)	13(9)	13(10)	14(12)	14(11)	13(11)	17(12)
	Fixations at the offset	N/A	14(6)	17(6)	13(9)	13(10)	14(8)	13(11)	17(12)	23(16)	35(14)
INTACT COMPETITOR	Saccades	32(9)	15(5)	25(8)	18(10)	9(7)	14(11)	3(6)	13(11)	20(15)	33(17)
	Fixations at the onset	0 (0)	15(6)	15(7)	14(7)	13(9)	12(9)	15(14)	12(12)	15(13)	16(13)
	Fixations at the offset	N/A	15(7)	15(6)	13(9)	12(9)	13(10)	15(14)	16(13)	18(14)	33(18)
CHANGED STATE	Saccades	34(9)	17(8)	44(11)	312(15)	18(10)	19(12)	5(8)	10(12)	27(18)	16(14)
	Fixations at the onset	0 (0)	19(6)	24(8)	47(18)	48(18)	46(20)	37(25)	42(25)	36(24)	28(20)
	Fixations at the offset	N/A	24(8)	42(11)	48(18)	46(20)	39(23)	36(24)	28(20)	30(18)	14(14)
CHANGED COMPETITOR	Saccades	25(12)	12(8)	19(8)	12(9)	8(7)	13(10)	4(6)	9(9)	13(10)	10(9)
	Fixations at the onset	0 (0)	11(6)	13(8)	8(6)	9(6)	10(8)	13(11)	12(11)	11(10)	10(10)
	Fixations at the offset	N/A	13(8)	9(4)	9(6)	10(8)	12(8)	11(10)	11(10)	8(10)	6(7)

Appendix III
Supplement Experiments

Experiment A1

In this experiment, 211 participants were recruited via Amazon's Mechanical Turk. They read a sentence such as "*The woman will choose the ice cream*" or "*The woman will drop the ice cream*" and judged whether an object in the probe picture was mentioned in the sentence. The probe picture was presented 500 ms after they finished reading the sentence. For experimental items, the probe picture showed a matched object in either an intact state or a changed state, thus required "yes" answers. There were also foil items with mismatched sentences and objects, thus "no" answers were required. A within-subjects design was used and 4 counterbalanced lists were included. In each list, there were 32 experimental items and 32 foil items. After excluding non-native speakers and participants with low-accuracy rates (less than 3 standard deviation from the mean), 176 participants (44 per list) were included in data analysis. Figure a1.1 illustrates mean verification accuracy rates and latencies.



Note: * p< .05; ** p<.01, *** p<.001

Figure a1.1 Mean accuracy rates (left) and response times (right) of correct responses in Experiment A1.

A two-way repeated measures ANOVA revealed a main effect of event, $F_1(1, 175) = 37.67$, $p < .001$, $\eta_p^2 = .18$ (by subjects), $F_2(1, 31) = 2.62$, $p = .116$, $\eta_p^2 = .08$ (by items), and a main effect of picture probe, $F_1(1, 175) = 34.31$, $p < .001$, $\eta_p^2 = .16$ (by subjects), $F_2(1, 31) = 1.82$, $p = .188$, $\eta_p^2 = .06$ (by items). However, there was a significant interaction, $F_1(1, 175) = 102.75$, $p < .001$, $\eta_p^2 = .37$ (by subjects), $F_2(1, 31) = 8.91$, $p = .006$, $\eta_p^2 = .22$ (by items), replicating the finding in Experiment 2. As shown in Figure a1.1, the intact state was verified more accurately and faster when no change was described. By contrast, the changed state was verified more accurately and faster when a change of state was described. Therefore, the results suggest that responses to the probe pictures were influenced by the linguistic context. Participants responded more accurately and faster to probe pictures that matched the linguistic context compared to ones that mismatched the linguistic context.

Experiment A2

In this experiment, we have investigated whether an event in between would influence the keeping track of of object representations. Participants were 64 native speakers of British English. According to event boundary theory (e.g., Radvansky & Copeland, 2006), once a model has been updated due to changes (e.g., location shifts), what was in the previous model would not be as accessible compared to the case when no updating has happened. Here, we aimed to test whether similar effect would be shown when another object was mentioned. Thus, we predicted that the difference of eye movements between (a) and (b) would not be found between (c) and (d). The results confirmed our predictions. When an additional event was mentioned participants' overall looks towards target regions were reduced, and difference of looks between "penny" and "egg" disappeared. The data thus suggest that adding a new event not only reduces the accessibility of target objects, but also interferes the memory retrieval process of the intended object representation.

a) *The girl will stomp on the **penny**. But first, she will look down at the **penny**.*

b) *The girl will stomp on the **egg**. But first, she will look down at the **egg**.*

c) *The girl will stomp on the **penny**. But first, she will put on her gloves and look down at the **penny**;*

d) *The girl will stomp on the **egg**. But first, she will put on her gloves and look down at the **egg**.*

Figure A2 Example visual stimuli

Experiment A3

To control for the visual difference between the two types of target objects (e.g., penny vs. egg), they were not shown in the visual scene (see Figure A3). Instead, the same visual scene was used for both conditions. Half of the linguistic stimuli were the same as in Experiment A2, and the other half were in “and then” condition, as shown in examples below. We expected to see more eye movements towards the plant when the linguistic stimuli were in “egg” condition than “penny” condition,

a) The girl discovered a ***penny*** *behind the plant. She will stomp on the penny.* ***But first****, she will look down at the penny.*

b) The girl discovered a ***penny*** *behind the plant. She will stomp on the penny.* ***And then****, she will look down at the penny.*

c) The girl discovered an ***egg*** *behind the plant. She will stomp on the egg.* ***But first****, she will look down at the egg.*

d) The girl discovered an ***egg*** *behind the plant. She will stomp on the egg.* ***And then****, she will look down at the egg.*

Figure A3 Example of a visual scene for Experiment A3

96 students from the University of York participated in this study. During the final reference to the target objects (e.g., penny/egg). A two-way repeated measures ANOVA was conducted on saccades towards the “plant” when the “egg” or “penny” was mentioned at the end of the third sentence. The results showed that there was neither significant main effect of the type of objects, F (1, 95) = .14, p=.71, temporal shifts, F (1, 95) = .94, p=.34. Nevertheless, there was a marginal interaction between the change of state and the temporal shifts by subjects, F (1, 95) = 3.61, p=.06. The results indicate that it’s easier to retrieve an object’s initial state when change was described to happen on it, but it’s easier to maintain an object’s current state when no change was associated with it. Thus, the data suggest that keeping track of an intact state (regardless due to the properties of the object itself or as the initial state before changes happen) was not as difficult as keeping tracking of a changed state.

REFERENCES

Aginsky, V., & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition, 7*, 147–162.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6,* 451-474.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419−439.

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition, 93,* 79–87.

Altmann, G.T.M., Garnham, A., Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*, 685-712.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition, 73,* 247–264.

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57,* 502–518.

Altmann, G.T.M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition*, 111, 55-71.

Arnold, J., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J.C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eye-tracking. *Cognition, 76*-B13-B26.

Barsalou, L. W. (1999). Perceptual symbol systems. T*he Behavioral and Brain Sciences, 22(4),* 577–609; discussion 610–60.

Barsalou, L. W., & Wilson, C. D., & Hasenkamp. (2010). On the vices of nominalization and the virtues of contextualizing. In B. Mesquita, L. Feldman Barrett & E. Smith (Eds.), *The Mind in Context* (pp. 334–360)

Barwise, J. & Perry, J. (1981). Situations and Attitudes. *Journal of Philosophy*

78 (11): 668-691. New York: Guilford Press.

Bekkering, H., Adam, J.J., Kingma, H., Huson, A., & Whiting, H.T.A. (1994). Reaction time latencies of eye and hand movements in single- and dual-task conditions. *Experimental Brain Research, 97,* 471-476.

Bentin, S., McCarthy G., Wood, C.C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60* (4): 1985, 343-55.

Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition, 28(1),* 108–124.

Bohnemeyer, J. & Pederson (Eds.)((2011). *Event Representation in Language and Cognition.* Cambridge: Cambridge University Press.

Boland, J. E. 2005. Visual arguments. *Cognition, 95*.237–74.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition, 75*, 1-28.

Bransford, J.D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3,* 193-209.

Bridger, E. K., Sprondel, V., & Mecklinger, A. (2015). Control over recollection varies with context-type: ERP evidence from the exclusion task. *Journal of Cognitive Neuroscience, 6 (1),* 31-38.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience. 5*, 34–44.

Carroll, Mary & von Stutterheim, Christiane (2011). Event representation, time event relations, and clause structure. In Bohnemeyer, J. & Pederson (Eds). *Event Representation in Language and Cognition.* Cambridge, UK: Cambridge University Press.

Camblin, C.C., Gordon, P.C., Swaab, T.Y., (2006). The interplay of discourse congruence and lexical association during sentence processing: evidence from ERPs and eye tracking. *Journal of Memory and Language*.

Campanella, F., & Shallice, T. (2011). Refractoriness and the healthy brain: a behavioural study on semantic access. *Cognition, 118(3)*, 417–431.

Chao, L.L., Haxby, J.V., Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*: 913-919.

Connell L. (2007). Representing object colour in language comprehension. *Cognition*, 102, 476-485.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: a new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84–107.

Coppens, L.C., Gootjes, L., & Zwaan, R.A. (2012). Incidental picture exposure affects later reading: Evidence from the N400. *Brain and Language*, 122(1): 64-69.

Coulson, S. & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters 316:* 71-74.

Coulson, S. & Lovett, C. (2004). Handedness, hemispheric asymmetry, and joke comprehension. *Cognitive Brain Research 19:* 275-288

Curran, T. (2000). Brain potentials of recollection and familiarity*. Memory and Cognition, 28,* 923–938

Curran, T., Tucker, D.M., Kutas, M., & Posner, M.I. (1992). Topography of the N400: brain electrical activity reflecting semantic expectancy. *lectroencephalography and Clinical Neurophysiology, 88*: 188 – 209.

Cycowicz, Y. M., Friedman, D., Snodgrass, J. G., & Duff., M. (2001). Recognition and source memory for pictures in children and adults. *Neuropsychologia, 39,* 55-267.

Dahan, D., & Tanenhaus, M.K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 30(2): 498-513.

Dahan, D., & Tanenhaus, M.K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, 12(3): 453-459.

Damasio, A. R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences, 90,* 4957-4960.

Davey, J., Rueschemeyer, S., Costigan, A., Murphy, N., Krieger-Redwood, K., Hallam, G., & Jefferies, E. (2015). Shared neural processes support semantic control and action understanding. *Brain and Language, 142,* 24-35.

Davidson, D. (1980). *Essays on Actions and Events.* Oxford University Press. Oxford, UK.

Delong, K.A., Urbach, T.P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117-1121.

Debruille, J. B. (2007). The N400 potential could index a semantic inhibition. *Brain Research Reviews, 56(2)*: 472-477.

Debruille, J. B., Ramirez, D., Wolf, Y., Schaefer, A., Nguyen, T.V., Bacon, B.A., Renoult, L., Brodeur, M. (2007). Knowledge inhibition and N400: a within- and a between subjects study with distractor words. *Brain Research, 1187*: 167-183.

Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition, 73(1):* B1-B15.

Dowty, D.R. (1991). Thematic proto-roles and argument selection. *Language 67,* 547-619.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses & Review. *Journal of Experimental Psychology: General, Vol. 140* (4), 552-572.

Ecker, U.K.H., Zimmer, H., & Groth-Bordin, C. (2007). Color and context: An ERP sudy on instrinsic and extrinsic feature binding in episodic memory. *Memory and Cognition, 35 (6)*: 1483 – 1501.

Federmeier, K. D., Segal, J. B., Lombrozo, T., and Kutas, M. (2000) Brain responses to nouns, verbs and class- ambiguous words in context. *Brain, 123:* 2552–2566.

Ferguson, H. J., Tresh, M., & Leblond, J. (2013). Examining mental simulations of uncertain events. *Psychonomic Bulletin & Review, 20*(2), 391–399.

Friedman, D. & Johnson, R. Jr. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. Microscopy Research Technique, 51, 6 – 28.

Galazka, M.A., & Ganea, P. A. (2014). The role of representational strength in verbal updating: Evidence from 19- and 24-month-olds. *Journal of*

Experimental Child Psychology, 121, 156-168.

Ganea, P.A. & Harris, P. H. (2013). Early limits on the verbal updating of an object's location. *Journal of Experimental Child Psychology, 114,* 89-101.

Ganea, P.A., Shutts, K., Spelke, E.S., & DeLoache, J.S. (2007). Thinking of things unseen: Infants' use of language to update mental representations. *Psychological Science. 18*: 734-739.

Ganis, G, Kutas,M., & Sereno, M.I. (1996). The search for the common sense--an electrophysiological study of the comprehension of words. *Journal of Cognitive Neuroscience, 8,* 89-106.

Gerrig, R.J., & McKoon, G. (1998). The readiness is all: the functionality of memory-based text processing. *Discourse Processes, 26,* 67-86.

Gentner, D. (1981). Verb semantic structures in memory for sentences: Evidence for componential representation. *Cognitive Psychology, 13,* 56-83.

Glenberg, A.M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. J*ournal of Memory and Language, 26,* 69–83.

Goldberg, A. E. (1995) *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press

Gordon, P.C., Hendrick, R., 1998. The representation and processing of coreference in discourse. *Cognitive Science*. 22, 389–424.

Gordon, P.C., Hendrick, R., Ledoux, K., Yang, C.L., 1999. Processing of reference and the structure of language: an analysis of complex noun phrases. *Language Cognitive Processes*. 14, 353–379.

Gordon, P.C., Camblin, C.C., Swaab, T.Y., 2004. On-linemeasures of coreferential processing. In: Carreiras, M., Clifton, C. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP,and Beyond.* Psychology Press, New York, NY, pp. 139–150.

Grodner, D., Gibson, E., Watson, D. (2005). The influence of contextual contrast on syntactic processing: evidence for strong-interaction in sentence comprehension. *Cognition, 95:* 275-296.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15(6),* 883-899.

Hagoort, P. & Brown, C. (1994). Brain responses to lexical ambiguity resolution and parsing, in: Clifton Jr.C., Frazier, L., Rayner, K. (Eds.), *Perspectives on sentence processing.* Erlbaum, Hillsdale, NJ, pp. 45–81.

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. Language and *Cognitive Processes, 8,* 439-483.

Hamm, J. P., Johnson, B. W., Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology, 113:* 1339-1350.

Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., Klimesch, W. (2008) The Electrophysiological Dynamics of Interference during the Stroop Task. *Journal of Cognitive Neuroscience, 20*, 215-225.

Henderson, J. & Ferreira, F. (Eds.) (2004). *The Interface of Language, Vision, and Action: Eye Movements and the Visual World.* New York: Psychology Press.

Henderson, J., Luke, S., Schmidt, J., & Richards, J.E. (2013). Co-registration of eye movements and event-related potentials in connected-text paragraph reading. *Frontiers in Systems Neuroscience, 7, 28.* doi:10.3389/fnsys.2013.00028

Herron, J.E., & Rugg, M.D. (2003). Strategic influences on recollection in the exclusion task: Electrophysiological evidence. *Psychonomic Bulletin & Review, 10,* 703-710.

Hindy, N. C., Altmann, G.T.M., Kalenik, E., & Thompson-Schill, S.L. (2012). The effect of object-state changes on event processing: Do objects compete with themselves?. *The Journal of Neuroscience, 32(17)*, 5795–5803.

Hindy, N.C., Solomon, S.H., Altmann, G.T.M., Thompson-Schill, S.L. (2015). A cortical network for the encoding of object change. *Cerebral Cortex. 5 (25):*884–894.

Horton J. J., Chilton L. B. (2010). The labor economics of paid crowdsourcing. *In Proceedings from EC '10: The 11th ACM Conference on Electronic Commerce (pp. 209–218).* New York, NY: ACM.

Hoover, M. A. & Richardson, D. C. (2008). When Facts Go Down the Rabbit Hole: Contrasting Features and Objecthood as Indexes to Memory. *Cognition 108(2)*, 533-42

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96(1),* B23-B32.

Huettig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15(8)*, 985-1018.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57(4),* 460-482.

Huettig, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica, 121(1),* 65-80.

Huettig, F., Rommers, J., Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137*: 151-171.

Hurley, R.S., Paller, K.A., Rogalksi, E.J., & Mesulam, M.M. (2012). Neural mechanisms of object naming and word comprehension in primary progressive aphasia. *Journal of Neuroscience, 32* (14): 4848-4855.

Hutzler, F., Braun, M., Vo, M.L.-H., Engl, V., Hofmann, M., Dambacher, M., Leder, H., Jacobs, A.M. (2007). Welcome to the real world: Validating fixation-related brain potentials for ecologically valid settings. *Brain Research, 1172,* 124-129.

Jaeger, A. & Parente, M. (2008). Event-related potentials and the study of memory retrieval: A critical review, *2(4):*248-255

Jackendoff, R. (1990). *Semantic Structures.* Cambridge, MA: MIT Press.

January D., Trueswell J.C., Thompson-Schill S.L. (2009) Co-localization of Stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience, 21*:2434–2444.

Johansson, M., & Mecklinger, A. (2003). Action monitoring and episodic memory retrieval: An ERP evaluation. *Biological Psychology, 64* (1-2), 99-125.

Johansson, M., Stenberg, G., Lindgren, M., & Rosen, I. (2002). Memory for perceived and imagined pictures: An event-related potential study.

Neuropsychologia, 40: 986–1002

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.

Just, M. A., & Carpenter, P.A. (1980). A theory of reading: From eye-fixations to comprehension. *Psychological Review, 87,* 329–354.

Just, M. A., Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149

Kable, J.W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience, 14,* 795-805.

Kalenik, E. (2012). *Competing Object Instantiations in a Self-Paced Reading Task*. Unpublished Master's Thesis. University of Pennsenviania, USA.

Kamide, Y., Altmann, G. T. M., & Haywood, S. (2003). The time course of prediction in incremental sentence processing: evidence from anticipatory eye-movements. *Journal of Memory and Language, 49,* 133–156.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32,* 37–55.

Kant, I. (1787/1996). *Critique of Pure Reason*. Hackett Publishing Company, Indiana, US.

Kaup, B., & Zwaan, R. A. (2003). Effects of negation and situational presence on the accessibility of text information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29(3)*, 439-446.

King, J. W., and Kutas, M. (1995) Who did what and when? Using word- and clause-related ERPs to monitor working memory usage in reading.*Journal of Cognitive Neuroscience 7* (3), 376-395.

Klepousniotou, E., & Baum, S.R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics, 20* (1): 1-24.

Kliegl, R., Dambacher, M., Dimigen, O., Jacobs, A.M., & Sommer, W. (2012). Eye movements and brain electric potentials during reading.

Psychological Research, 76 (2): 145-158.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. J*ournal of Cognitive Neuroscience, 5*, 196-214.

Knoeferle, P., & Crocker, M. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science, 30,* 481–529.

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. J*ournal of Memory and Language, 57,* 519–543.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95,* 95–127.

Knoeferle, P., Urbach, T.P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: insights from ERPs and picture-sentence verification. *Psychophysiology, 48,* 495-506.

Kukona, A., Altmann, G. T. M., Kamide, Y. (2014). Knowing what, where, and when: Event comprehension in language processing. *Cognition 133* (1): 25-31.

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 4*, 463-470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203-208.

Kutas, M., & King, J.W. (1996). The potentials for basic sentence processing: Differenting integrative processes. In I. Ikeda & J.L. McClelland (Eds.), *Attention and Performance* (Vol. 16, pp.501-546). Cambridge, MA: MIT Press.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1995-2005. Traxler, M., & Gernsbacher, M.A. (eds.). *Handbook of Psycholinguistics.* 2nd ed. New York: Elsevier, pp. 659-724.

Lee, D., & Chun, M. M. (2001). What are the units of visual short-term memory: Objects or spatial locations? *Perception & Psychophysics, 63,* 253-257.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation.* University of Chicago Press, Chicago, IL.

Leynes, P.A., Crawford, J.T., & Bink, M.L. (2005). Interrupted actions affect output monitoring and event-related potentials (ERPs). *Memory, 13,* 759-772.

Leynes, P.A., Grey, J.A., & Crawford, J.T. (2006). Event-related potential (ERP) evidence for sensory-based action memories. International *Journal of Psychophysiology, 62,* 193-202.

Liotti, M., Woldorff, M.G., Perez, R., & Mayberg, H.S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia, 38 (5):* 701-711.

Lüdtke, J., Friedrich, C., De Filippis, M., & Kaup, B. (2008). ERP correlates of negation. *Journal of Cognitive Neuroscience, 20,* 1355-1370

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology – Paris, 102*, 59–70.

Mandler, J. M. (1986). The development of event memory. In F. Klix & H. Hagendorf (Eds.), *Human Memory and Cognitive Capabilities.* Amsterdam: Elsevier.

Martin A, Haxby JV, Lalonde M, Wiggs CL, Ungerleider LG (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science 270*:102–105.

Makeig S., Jung T-P, Bell A.J., and Sejnowski, T.J. (1997). Blind separation of auditory event-related brain responses into independent components. Proceedings of National Academy of Sciences. USA, *94:*10979-10984.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25-45.

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of colour and knowledge of action. *Science, 270,* 102–105.

Meyer, A.M., & Federmeier, K. D. (2008). The divided visual world paradigm: Eye tracking reveals hemispheric asymmetries in lexical ambiguity resolution. *Brain Research, 1222:* 166-183.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227–234.

Milham M. P., Banich M.T., Webb A., Barad V., Cohen N.J., & Wszalek, T., & Kramer A. F. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research, 12*, 467– 473.

Miller, G. A. & Johnson-Laird. P. N. (1979). *Language and Perception.* Cambridge, Mass: Harvard University.

Mǜeller, H.M., King, J.W., & Kutas, M. (1997). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research, 5*: 193-203.

Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature, 395*(6697), 71–73.

Nebes, R. D., & Halligan, E. M. (1996). Sentence context influences the interpretation of word meaning by Alzheimer patients. *Brain and Language, 54,* 233-245.

Nieuwland, M. S., & Kuperberg, G. R. (2008). An Event-Related Potential Study on the Pragmatics of Negation. *Psychological Science,12,*1213 -1218.

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain & Language, 106,* 119-131.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language, 63*(3), 324–346.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 and semantic anomaly with pictures and words. *Journal of Cognitive Neuroscience, 4,* 15-22.

Odekar, A., Hallowell, B., Kruse, H., Moates, D., & Lee, C.Y. (2009). Validity of eye movements methods and indices for capturing semantic (associative) priming effects. *Journal of Speech, Language, and Hearing Research, 52*, 31-48.

Osina, M.A., Saylor, M.M., & Ganea, P. A. (2013). When Familiar Is Not Better: 12-Month-Old Infants Respond to Talk About Absent Objects. *Developmental Psychology, 49,* 138-145

Otten, M., Nieuwland, M.S., & Van Berkum, J.J.A. (2007). Great expectations:

specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience, 8,* 89.

Otten, M., Van Berkum, J.J.A. (2007). What makes a discourse contraining? Comparing the effects of discourse message and scenario fit on the discourse dependent N400 effect. *Brain Research, 1153,* 116-177.

Otten, M., & Van Berkum, J.J.A. (2008). Discourse-based word anticipation during language processing: prediction or priming? *Discourse Processes, 45,* 464-496.

Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience, 4,* 375-391.

Palmeri, T.J., & Tarr (2008). Object recognition and long-term visual memory for objects. In S. Luck & A. Hollingsworth (Eds.), *Visual Memory.* Oxford University Press.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition 108:* 155-184.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge, MA: MIT Press.

Proverbio, A. M., Burco, F., del Zotto, M., & Zani, A. (2004). Blue piglets? Electrophysiological evidence for the primacy of shape over color in object recognition. *Cognitive Brain Research, 18,* 288–300.

Pulvermüller, F. (1999). Words in the brain’s language. *Behavioral and Brain Sciences, 22,* 253-336.

Radvansky, G. A. (1999a) Memory retrieval and suppression: The inhibition of situation models. *Journal of Experimental Psychology: General, 128,* 563-579.

Radvansky, G. A. (1999b). The fan effect: A tale of two theories. *Journal of Experimental Psychology: General, 128,* 198-206.

Radvansky, G. A. (2005). Situation models, propositions, and the fan effect. *Psychonomic Bulletin & Review, 12(3)*: 478-483.

Radvansky, G. A. (2009). Spatial directions and situation model organization. *Memory & Cognition, 37,*796-806..

Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science, 21*, 269-272.

Radvansky, G. A. & Copeland, D. E. (2000). Functionality and spatial relations in situation models. *Memory & Cognition, 28,* 987-992.

Radvansky, G. A., & Copeland, D. E. (2001). Working memory and situation model updating. *Memory & Cognition, 29,* 1073-1080

Radvansky, G. A., & Copeland, D. E. (2006).Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition, 34(5),* 1150-1156.

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,* 210-216.

Radvansky, G. A., Spieler, D. H., & Zacks, R. T. (1993). Mental model organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 95-114.

Radvansky, G. A., Wyer, R. S., Curiel, J. C., & Lutz, M. F. (1997). Situation models and abstract ownership relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1233-1246.

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

Radvansky, G. A., Zwaan, R. A., Federico, T., & Franklin, N. (1998). Retrieval from temporally organized situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1224-1237.

Rappaport Hovav, M. and B. Levin (1998). *Morphology and lexical semantics*, in A. Zwicky and A. Spencer, eds., *Handbook of Morphology*, Blackwell, Oxford, UK, 248-271.

Rebai M, Bernard C, Lannou J. (1997). The Stroop's test evokes a negative brain potential, the N400. *The International Journal of Neuroscience, 91*: 85–94

Reinitz, M.T., Wright, E., Loftus, G.R. (1989). Effects of semantic priming on visual encoding of pictures. *Journal of Experimental Psychology: General, 118*(3): 280-297.

Richardson, D., Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, 102, 129-138.

Richardson, D. C. & Spivey, M. J. (2000). Representation, space and Hollywood

Squares: Looking at things that aren't there anymore. *Cognition, 76,* 269-295.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46,* 245-266.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science, 28,* 89-104.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology-General,* 104(3), 192-233.

Rösler, F., Heil, M., & Glowalla, U. (1993). Memory retrieval from long-term memory by slow event related brain potentials. *Psychophysiol 30*:170–182.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11,* 251-257.

Senkfor, A. J., & Van Petten, C. (1998). Who said what: An event-related potential investigation of source and item memory. *Journal of Experimental psychology: Learning, Memory & Cognition, 24,* 1005-1025.

Scialfa, C.T., Joffe, K.M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perceptual Psychophysics, 60,* 1067-82.

Shinkareva, S.V., Malave, V.L., Mason, R. A., Mitchell, T.M., & Just, M. A. (2011). Commonality of neural representations of words and pictures. *NeuroImage, 53,* 2418-2425.

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia, 45 (12):* 2802-2810.

Slobin, D. I. (1991). Learning to think for speaking: native language, cognition, and rhetorical style. *Pragmatics*. 1. (1), 7-25.

Slobin, D. I. (1996). From "thought and language" to "thinking to speaking". In J. J. Gumperz & S. C. Levinson (Eds.). *Rethinking Linguistic Relativity,* 70-96. Cambridge: Cambridge University Press.

Solomon, S.H., Hindy, N.C., Altmann, G.T.M., & Thompson-Schill, S.L. (in press). Competition between mutually exclusive object states in event comprehension. *Journal of Cognitive Neuroscience.*

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,* 125-140.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. Psychological Science, 18(5), 449–455.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65,* 235–241.

Spivey, M., Tanenhaus, M.K., Eberhard, K.M., & Sedivy, J. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45,* 447-481.

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

Sussman, S. R., & Sedivy, J. C. (2003). The time-course of processing syntactic dependencies: evidence from eye-movements. *Language and Cognitive Processes, 18,* 143–163.

Swaab, T., Brown, C. M., & Hagoort, P. (2003). Understanding words in sentence contexts: The time course of ambiguity resolution. *Brain and Language, 86(2),* 326-343.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General,*138, 236–257.

Swallow, K.M., Barch, D.M., Head, D., Maley, C.J., Holder, D., Zacks, J.M. (2011). Changes in events alter how people remember recent information. *Journal of Cognitive Neuroscience*, 23:5, 1052-1064.

Talmy, L. (1975). *Semantics and syntax of motion*, in J.P. Kimball, ed., Syntax and Semantics 4, Academic Press, New York, NY, 181-238.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 632–634.

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & Cognition, 34,* 78–89.

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: a dynamic pragmatic account. *Quarterly Journal of Experimental Psychology , 63(12)*, 2305–2312.

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

Van Berkum, J.J., Hagoort, P., Brown, C.M. (1999b). Semantic integration in sentences and discourse: evidence from the N400. *Journal of Cognitive Neuroscience. 11,* 657–671.

Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken-language comprehension. *Psychophysiology, 40,* 235–248.

Van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

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

Wyer, R. S., Jr., & Radvansky, G. A. (1999). The comprehension and validation of social information. *Psychological Review, 106,* 89-118.

Yang, C.L., Perfetti, C.A., & Schmalhofer, F. (2005). Event-related potential indicators of text integration across sentences boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 55-89.

Yee, E., & Sedivy, J. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*(1), 1-14.

Yee, E., Huffstetler, S., & Thompson-Schill, S.L. (2011). Function follows form: Activation of shape and function features during object identification. *Journal of Experimental Psychology: General, 140(3)*, 348-363.

Yonelinas, A.P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *Journal of Memory and Language. 33,* 441–517.

Yu, J., Zhang, Y., Boland, J.E., & Cai, L. (2015). The interplay between referential processing and local syntactic/semantic processing: ERPs to written Chinese discourse. *Brain Research, 1597,* 139-158.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127,* 3-21.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin, 133*, 273–293.

Zemplen, M., Renken, R., Hoeks, J.C.J., Hoodguin, J.M., & Stowe, L.A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage. 34 (3).* 1270-1279.

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

Zwaan, R.A., Stanfield, R.A., & Yaxley, R.H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13,* 168–171.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science, 28:* 611-619.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition,21*(2), 386-397.

Zwaan. R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS ONE.7(12)*: e51382. doi:10.1371/journal.pone.0051382.

Zwaan, R. A. & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123 (2)*: 162-185.

Zwaan, R. A., Radvansky, G. A., & Whitten, S. N. (2002). Situation models and themes. M. M. Louwerse & W. van Peer & (Eds.) *Thematics: Interdisciplinary Studies, pp. 35-53*. Amsterdam/Philadelphia, John Benjamins.

Zwaan, R. A., Stanfield, R.A., Yaxley, R.H. (2002). Do language comprehenders routinely represent the shapes of objects? *Psychological Science, 13,* 168-171.

Zwaan, R.A., & Taylor, L. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135* (1): 1-11.