12-13-2017

# Investigating How Anticipation of Object States Drives Event Comprehension

Kyra Krass
kyra.krass@uconn.edu

# Investigating How Anticipation of Object States Drives Event Comprehension

Kyra Leigh Krass

B.S., The Pennsylvania State University, 2015

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2017

# APPROVAL PAGE

Masters of Science Thesis

## Investigating How Anticipation of Object States Drives Event Comprehension

Presented by

Kyra Leigh Krass, B.S.

Major Advisor _____
Gerry Altmann

Associate Advisor _____
James Magnuson

Associate Advisor _____
Letitia Naigles

Associate Advisor _____
Eiling Yee

University of Connecticut

2017

# Acknowledgements

I would like to express my gratitude to those who made this research possible. Firstly, I would like to thank my advisor Dr. Gerry Altmann for imparting his wisdom and encouraging me to think outside of the box. I appreciate his guidance and patience through this process. I would also like to thank my committee members Drs. James Magnuson, Letitia Naigles, and Eiling Yee for their help and feedback throughout the development of this project. Additionally, a heartfelt thanks to Dr. Gitte Joergensen for her constant guidance and encouragement. To the Altmann and Yee labs, a sincere thank you for their welcomed feedback and assistance, especially with running participants. Finally, I thank my family for their love and support and my friends both near and far who have always kept me grounded and been gracious editors at various stages over the past several years.

The research presented in this thesis was funded by National Science Foundation IGERT grant DGE-1144399 to the University of Connecticut.

Table of Contents

## Introduction

Several theories have been proposed to help accurately capture our mental representation of language. Though each theory differs slightly, mental models and situation models, theories of how the mind represents what happens in the world, both suggest that we comprehend language by incorporating information about the world into our representation of an event (e.g., Glenberg & Langston, 1992; Kintsch, 1988; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). According to these views, any visual information, context, and semantic knowledge that an individual has contributes to their mental representation of language.

James Gibson also researched the idea that individuals have a relationship with the world around them. He coined the term *affordances* to better explain this relationship. Affordances are the possibilities that an object or some other part of the environment gives to an individual (Gibson, 1979). In other words, affordances reflect how individuals can interact with their environment. For example, water affords drinking and quenching thirst. Affordances are now widely used to describe objects in a variety of contexts and can be thought of as a type of predict. Other researchers have studied the role that affordances play in sentence comprehension as well as the understanding of actions and events. Similarly, work has been done to argue that affordances are crucial for understanding the meaning of language and our environment (Glenberg & Robertson, 1999; Glenberg & Robertson, 2000; Kaschak & Glenberg, 2000).

To understand affordances, one must delve deeper into the role of prediction and anticipation with respect to language. For our purposes, prediction and anticipation are effectively the same. Affordances are a type of prediction; it is the very definition of affordances that allows us to make predictions about how we can interaction with an object. We anticipate that we can eat a cake because it *affords* that action. In our lives, we anticipate many things. If

you see a grey cloud, you anticipate rain and prepare by bringing an umbrella. We carry our past experiences with us, and it drives us forward. This can be said of language as well. Anticipation is utilized as a sentence unfolds. A simple recurrent network by Elman (1990) suggests that language can be modeled by creating something that uses previous linguistic context to predict which words would appropriately follow in a sentence. The model learns information about frequency and word type (e.g., verbs and nouns) to predict upcoming linguistic stimuli. Similarly, event segmentation theory, put forth by Zacks, Speer, Swallow, Braver, and Reynolds (2007), suggests that prediction is vital in our ability to understand and "segment" events. The theory proposes that when we are unable to make a prediction about what comes next in a series, this signals that an event has occurred.

Eye-tracking and the visual world paradigm were initially used to track eye-movements as various objects were mentioned during the process of language comprehension (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). This was later adapted to investigate the role of affordances and anticipation in linguistically-primed event representation and processing. One study by Chambers, Tanenhaus, Eberhard, Filip, and Carlson (2002) found that individuals' eye-movements are driven by affordances when processing actions. For instance, when participants heard "Put the cube inside the can", eye-movements were motivated by which objects presented would afford holding the cube (Chambers et al., 2002). Additionally, if two cans were present, individuals looked to the one that was big enough to contain the cube. This demonstrates that not only are eye-movements linked to language, but their control is also sensitive to affordances. Similarly, work by Chambers, Tanenhaus, and Magnuson (2004) showed that individuals' eye movements are driven by the affordances of the objects in a scene

when hearing temporarily ambiguous sentences such as “Pour the egg in the bowl over the flour”.

Our study aims to better understand the links between anticipation, language, and event processing. We will be conducting several eye-tracking experiments based on previous literature which diverges in its results. In a pivotal study by Altmann and Kamide (1999), participants heard sentences like “The boy will eat the cake”. While hearing the stimuli, a display depicted, for this example, a boy, a cake, and several unrelated objects. When participants heard *eat*, they looked to the cake. This was interpreted to mean that participants anticipate the object in the scene that affords the action (e.g., in this case, the cake affords eating). In a later study by Altmann and Kamide (2007), participants heard sentences like “The woman will eat the cake/has eaten the scones” and saw a scene with objects like a cake, a plate with crumbs, and two distractor objects. The previous Altmann and Kamide (1999) study showed that, in the future tense, individuals looked to the initial state of the cake (i.e., the state that affords the action) arguably because the action had not yet happened. However, the later study by Altmann and Kamide (2007) investigated how individuals interpret an action that has already occurred. It was predicted that individuals would look to the end state in the past tense case, knowing that the action had already been performed. They suggested that the end state could attract looks given that an empty plate, for example, affords that an eating action had taken place. Their results showed that participants showed a preference for the empty plate as a function of tense. There were more looks to the empty plate in the past tense than the future tense, but looks to the full cake were not significantly different across the two tenses (Altmann & Kamide, 2007). These results show that individuals’ eye-movements differ when presented with the past and future tense.

In a more recent eye-tracking study by Kang (2015), participants were given sentences like “The woman will drop the ice cream. But first, she will look at the ice cream.” while simultaneously seeing an upright ice cream, a dropped ice cream, and two distractors. This study differs from the Altmann and Kamide (1999) study because this study depicts both the initial and end states of the action, whereas the former study showed only the initial state. When participants heard *drop*, they looked to the dropped ice cream or the end state of the action. This is unexpected given the results of the initial two studies, as the participants anticipated the end state or goal state of the action and not, as in Altmann and Kamide (1999), the object that afforded the action; in the example, the *upright* ice cream afforded dropping *not* the dropped one. Another crucial difference between the Kang (2015) study and those by Altmann and Kamide (1999; 2007) is the verb type. Kang used change of state verbs where there is a clear initial and end state. However, Altmann and Kamide used destruction verbs where the initial state is clear, but the object itself then disappears (e.g., after eating a cake it is now gone, at least visually). These differences in the visual and linguistic stimuli lead to the question of whether specific verb types are processed differently from one another and if this difference in processing would account for the different patterns of eye movements.

Though differing stimuli could be the driving force between the seemingly conflicting results of Kang (2015) and Altmann and Kamide (1999; 2007), one possibility is that goals drove the results of the Kang (2015) study. One can ask whether the action itself drives understanding of actions or if the goal is what drives comprehension. Several lines of research have suggested that goals drive event comprehension. This has been shown in adults as well as children. Research with children shows that a preference for goals starts at a young age. One study by Loucks and Meltzoff (2013) looked at how children at three years old were able to remember

actions presented to them. Children were shown the actions "grouped" together according to goal (e.g., multiple actions related to being fed followed by multiple actions related to the bath) or shown the actions "interleaved" (e.g., several actions related to being fed intermixed with actions related to the bath). They were able to recall the same number of actions and were able to group the actions together according to goal regardless of which condition they were presented (i.e., "grouped" or "interleaved"). In other words, the children prioritized the goal of the action over the chronological order in which it was presented to them (Loucks & Meltzoff, 2013). Rather than simply imitating the actions in the exact order that they were presented to them, children favored putting the actions together according to a higher goal. Another study involved an experimenter showing 18-month-old children an action. Either the action was fully completed, or the action was attempted and failed (Meltzoff, 1995). The children then imitated the actions of the experimenter. The children demonstrated the same number of actions regardless of whether the action was completed or if it was merely attempted (Meltzoff, 1995). The children who saw the uncompleted actions did not imitate what they saw, but they instead completed the intended goal. This shows that children were inferring the goal, and this seemed to drive their own actions even when the goal was not explicitly presented to them.

Some additional research shows that children between the age of four and seven show a preference for expressing the goal of an action rather than the origin of the action when describing a video that they have seen (Lakusta & Landau, 2005). Additionally, children as young as 12 months old show a preference for the goals of an action by showing longer looks at actions where there were multiple goals as opposed to those with multiple origins (Lakusta, Wagner, O'Hearn, & Landau, 2007). All of these studies suggest that young children are able to understand and bring about the goal of an action. They show a preference for the goal rather than

the initial state of the object and seem to motivate their actions with respect to the goal of the action.

One would expect that if goals are shown to be the primary driver of behavior in children, the same would be found in adults. Several studies address this. One idea, the *common-coding principle*, suggests that the representation of actions is tightly connected to the effects that come about from those actions (Prinz, 1997). Research by Hommel (2009) suggests that anticipation plays a crucial role in understanding and representing actions. His theory of event coding (TEC) suggests that we anticipate the goal in order to complete an action (Hommel, 2009). TEC emphasizes the nature of humans to anticipate the effect of their actions as well. However, Hommel also suggests that perception of an action and the action itself are so interwoven that both *completing* an action and *perceiving* one would yield similar results (i.e., anticipating the goal of the action). Perception can be thought of as being related to the experience of having completed actions. Although this particular suggestion was not addressed or tested in the Hommel (2009) study, it could account for the results found in Kang's (2015) study. This potential connection between action and action perception would lead to predictions that when participants hear that an action has taken place, they focus on the goal of that action. However, further studies need to be assembled to specifically test this theory. The following studies were conducted to address whether it is affordances or goal states that drive anticipation. If individuals anticipate what affords an action, this would manifest itself as looks toward the initial state (e.g., looks toward a closed umbrella because it affords opening). If individuals anticipate the goal state, this would be demonstrated with looks toward the end state (e.g., looks toward an open umbrella, the goal of opening).

## Current study

The current study explores sentence processing and event comprehension using the visual world paradigm. Three experiments were conducted, each using single sentence auditory stimuli paired with related objects. The purpose was to track the object representations that participants maintain when processing these sentences using a passive listening task. There are several differences between the Kang (2015) study and the Altmann and Kamide (2007) study which were addressed briefly above. These studies showed different results (i.e., anticipating the goal state vs. anticipating the state that affords the action, respectively), and a notable difference in the stimuli used was the verb type. Kang (2015) used verbs of change which could be easily depicted in each state, whereas Altmann and Kamide used verbs of destruction, such as *eat* or *drink*, where the end state is inferred and not an altered version of the original object. In the Kang (2015) study, both the initial and end states were depicted, with the end state being a changed version of the initial state. For example, she showed an upright ice cream cone as the initial state and, as the end state, a dropped ice cream. However, for Altmann and Kamide (2007), the end state of *eat the cake* was a plate with crumbs. A plate with crumbs on it is not a cake anymore. It must be inferred that the cake was eaten to understand that the plate once contained a cake. The end state is not easily depicted for these verbs of destruction. Because of this key difference, we chose to investigate verb type to determine how this would affect eye-movements. Three different types of verbs were used across the three experiments: reversible verbs for which the initial and end states can be directly depicted (cf. Kang, 2015; e.g., *open/close*), destruction verbs for which only the initial state can be directly depicted (cf. Altmann & Kamide, 1999, 2007; e.g., *eat/drink*), and creation verbs for which only the end state can be directly depicted (e.g., *knit/bake*).

Furthermore, Altmann and Kamide (2007) found opposite patterns of eye-movements when presenting participants with the past and future tenses (i.e., more looks to the initial state in the future tense and more looks to the inferable end state in the past tense), so this suggests that verb tense influences eye-movements. However, this particular pattern was only found to be statistically reliable in the original study's second experiment. Therefore, we decided to create stimuli that would more easily facilitate anticipating the end state to determine how strong the original effect was. For instance, one item from the Altmann and Kamide (2007) study was "The man will drink the beer/has drunk the wine". For this particular item, and most of the others, the initial and end states depicted were not of the same item. The initial state was a full glass of beer, and the end state was an empty glass of wine. In our stimuli, we chose items where the same item was referred to across conditions (e.g., *The woman will eat/has eaten the cake*).

The current study was run in two parts. The first study included reversible action verbs in the future tense and destruction verbs in the past and future tense (e.g., *will open/close* and *has/will eat*). The second study included reversible action verbs in the past tense and creation verbs in the past and future tenses (e.g., *has open/close* and *has/will knit*). However, for ease of conceptualization, the following study has been broken up into three separate experiments with the motivation of each detailed below. See Table 1 for clarification of the chronological order of experiments.

Table 1

*Chronological Order of Experiments*

| **Group 1: Tested Fall 2016** | **Group 2: Tested Spring 2017** |
|---|---|
| 20 future reversible verbs (Exp 1) | 20 past tense reversible verbs (Exp 1) |
| 12 destruction verbs (Exp 2) | 12 creation verbs (Exp 3) |
| 64 fillers | 61 fillers |
| 96 total items (+ 12 comprehension questions) | 93 total items (+ 11 comprehension questions) |

Research to date does not address whether individuals anticipate what affords an action rather than the goal of an action or whether they might anticipate both when reading or hearing action verbs. If they anticipate the goal state more as suggested by Kang (2015), we would predict that individuals look more to the end state of an action when given action verbs that entail a change in state from an initial state to an end state. Experiment 1 tests this. We used sentences such as "The pedestrian will open/close the umbrella" and "The pedestrian has opened/closed the umbrella" where both the initial state of the verb and the end state of the verb were depicted for each trial. For example, a closed umbrella and an open umbrella were seen when participants heard *open the umbrella*. Individuals who heard "open" did not also hear "close". If participants anticipate the goal, this would manifest as anticipating an open umbrella when hearing *open* and a closed umbrella when hearing *closed*. It was important for us to test both the future tense and the past tense to determine if these types of verbs would be affected by tense in the same way as the destruction verbs of Altmann and Kamide (2007).

Additionally, we used reversible action verbs to ensure that each object could serve as its own control, ruling out differences in looks to the end vs. initial state that would could be due to variables such as differences in object type, size, or salience. Therefore, looks to an open umbrella can be compared when that umbrella is the end state (following the verb *open*) and when that same umbrella is the initial state (following the verb *close*).

Establishing a pattern for reversible action verbs does not address how individuals anticipate other type of verbs, such as destruction verbs like those used in Altmann and Kamide (1999; 2007). Given that we have presented both the past tense and the future tense for Experiment 1 and would predict no difference across tense based on the findings of Kang (2015), it is especially important to determine if the Altmann and Kamide (2007) results can be

replicated. If we are able to replicate those results, we predict that individuals anticipate the end state for the past tense but the initial state for the future tense. In the past tense, the action has already occurred, so individuals' eye-movements should be driven towards the end state because it affords action. If individuals are anticipating the next action to be performed, this action would be completed on the end state. In the future tense, the state which affords an action is the initial state. Experiment 2 uses sentences like "The woman will eat/has eaten the cake". It is worth noting that this type of verb has a less easily depicted end state than the reversible verbs. Both object states are not the same object. For example, a cake is the initial state of eating, whereas the end state is an empty plate (i.e., not a cake). However, the results of Altmann & Kamide (1999; 2007) suggest that participants will anticipate the end state (e.g., a plate with crumbs) for the past tense and the initial state (e.g., a cake) for the future tense regardless of this difference.

One other type of verb that is not addressed by the above experiments is creation verbs. Creation verbs have a clear *end* state but a less clear *initial* state. The object state that depicts the result of the action is not the same as the object that is acted upon. For example, in "The grandmother will knit/has knit the sweater", a ball of yarn is acted upon, yielding an end state of a sweater. The initial state, though related, is not the same object as the end state. This leads to two possible predictions for this class of verbs: 1) creation verbs will pattern like the destruction verbs with more looks to the end state in the past tense and more looks to the initial state in the future tense or 2) creation verbs will pattern like the reversible action verbs with more looks to the goal state for both tenses. Given that we have two predictions for our results, it is crucial that we run an experiment using creation verbs. Experiment 3 tests such creation cases, asking where the eyes move during verbs such as "knit" – to the ball of yarn (the object that affords the action of knitting) or to the sweater (the goal of the knitting). If the creation verbs pattern like the

reversible ones, we predict individuals will look to the goal of knitting for both tenses. However, if they pattern similarly to the destruction verbs, we predict individuals will look to the object that affords action: the end state for the past tense and the initial state for the future tense.

## Experiment 1

### Methods

**Participants.** Thirty-nine participants took part in the future tense conditions and a second group of forty participants took part in the past tense conditions. Participants were recruited from the University of Connecticut subject pool of undergraduate psychology students. An additional five participants were run for the past tense condition but were excluded due to poor calibrations. All participants were over the age of 18, native speakers of English, and had normal or corrected-to-normal vision.

**Stimuli.** This experiment had two groups. One group of participants was given 20 experimental items with action verbs that were reversible and presented in the future tense (e.g., will open/close or will tie/untie). The main manipulation was the verb: The pedestrian will *open* the umbrella vs. The pedestrian will *close* the umbrella. If an individual heard *will open*, they never heard *will close*. The experimental items were counter-balanced using a within-subjects design so that participants heard one version of each sentence. Participants heard a total of 20 reversible items (see Appendix A for a full list of sentences). Each sentence was accompanied by a visual scene depicting the initial state of the action, the end state, and two unrelated distractors.

The rest of the items seen by this group were the 12 destruction verbs (e.g., eat or drink) detailed in Experiment 2. Each participant also heard 64 filler sentences. These sentences used change of state verbs but were divided into three types based on the visual stimuli depicted: (i)

two pairs of objects (one target pair and one distractor pair similar to the visual items for the experimental items); (ii) one pair of objects and two individual objects (one object in the pair was the target object or one of the individual objects was the target); and (iii) four individual objects on the screen (one target and three distractors). See Appendix B for a list of filler stimuli. Forty-three out of the sixty-four fillers were presented in the past tense. Therefore, of the items shown, 49 were in the past tense, and 47 were in the future tense. Participants were also given 12 yes/no comprehension questions after filler items to help maintain focus on the experiment.

A second group of participants was run using the past tense versions of the action verbs given to the first group (e.g., The pedestrian has *opened*/*closed* the umbrella). If an individual heard *has opened*, they never heard *has closed*. The rest of the items seen by this group were the 20 creation verb items detailed in in Experiment 3 as well as the same filler items and comprehension questions as the other group was given but with fewer items (fillers n=61, questions n=11 with 22 presented in the past tense). [1] Each participant saw a total of 93 items (past tense n= 48, future tense n= 45). See Figure 1 for an example of the visual stimuli.

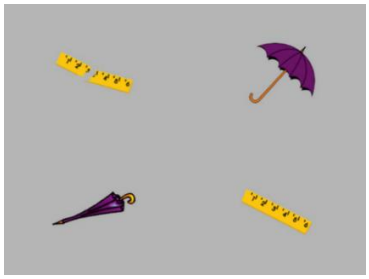

*Figure 1*. This display would be presented with the sentences "The pedestrian will open/close the umbrella" and "The pedestrian has opened/closed the umbrella".

[1] Individuals saw fewer fillers in the past tense of Experiment 1 due an overlap in items between the fillers and the experiment. For example, one filler that was rejected said "The grandmother has knit the sweater" which was too similar to an experimental item for Experiment 3.

**Procedure.** Stimuli were presented on a 24-inch wide computer screen with participants seated approximately 60 cm from the screen. Participants' eye-movements were tracked via an Eyelink 1000Plus eye tracker at a sampling rate of 500 Hz. Participants were told that the sentences presented to them would describe something depicted by the pictures. They were instructed to listen and try to relate the sentences to the pictures. Both groups of participants were told that the purpose of the eye-tracking camera was to track pupil size during the experiment. This was to avoid participants being overly conscious of their eye movements during the study. They then were given three practice trials. For all practice and experimental trials, participants were given a 1000 ms preview of the four objects before hearing the sentences. After the practice session, participants moved to the experimental portion. After each item, there was a drift correction which adjusts for any head movements which change the angle between the eye and the camera. After every twelve items, and at the start of the experiment, there was a five-point calibration to ensure accurate tracking of the eye-movements. The entire experiment lasted between 25-30 minutes, after which participants were fully debriefed.

## Results

Our primary interest in analyzing the data was to determine if participants were already showing a preference to one object state before the noun (e.g., looks by "The pedestrian will open the"). We used multiple measures to test the strength of our results. Two of these measures were saccades and fixations. Saccades were analyzed over a specific time window. This window started at the onset of the verb and ended at the offset of the determiner (e.g., open the). [2] This

[2] Although this window contains co-articulation, this is constant across all conditions. The noun is the same in both conditions (e.g., open the *umbrella*, close the *umbrella*).

analysis determines where the eyes were moving towards before the final noun. Fixations were analyzed at one instant in time. We analyzed the onset of the determiner. This measure tells us where the eye already was fixated before the final noun. Both windows/points were chosen to capture anticipatory looks.

We used 2 x 2 (state x tense) mixed ANOVAs as all data satisfied the assumptions of normal distribution and homogeneity of variance. Two groups were run for Experiment 1, so tense was manipulated between-subjects. Arcsine transformations were used for all data analyses involving proportions/percentages. Eye-movements were analyzed by-participant and by-item ($F_1$ and $F_2$, respectively).

Fixations were analyzed to determine which of the two object states was fixated on more. If there are more fixations to one object state, one would expect more eye-movements that give rise to those fixations (i.e., saccades). A greater number of saccades toward an object state would predict that there should be fewer saccades away from this object state. Therefore, we analyzed saccades to and away from each object state. If one object state is more salient based on the analysis of fixations and saccades, one would also predict longer fixations to this more salient object state. Longer duration of fixations on an object state would also predict that if an individual's eyes land on the less salient state, they would switch to the more salient state. If individuals have already landed on the more salient state, they would stay fixated on this state and not switch. Because of the relationships between the measures, all of the aforementioned analyses have been reported below. See Figures 2 and 3 for proportion of fixations over the course of the sentence.

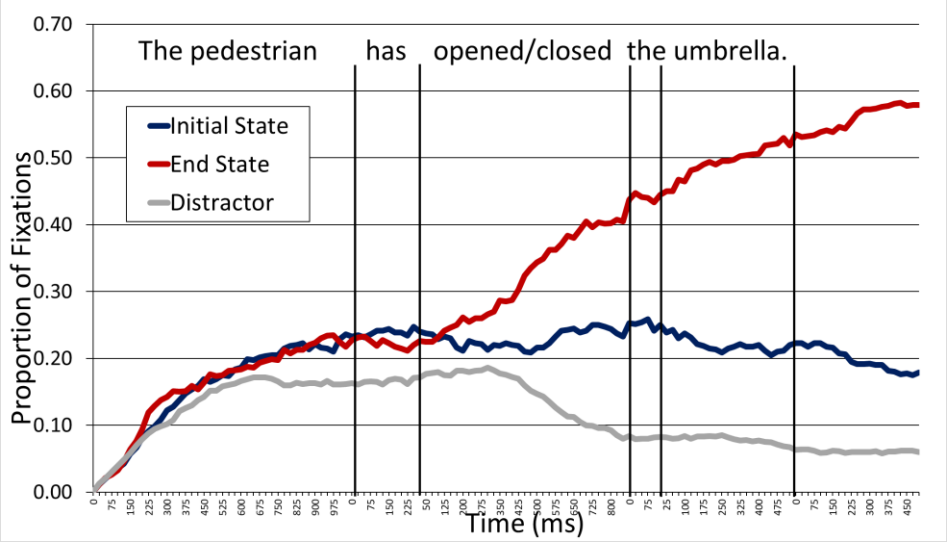


*Figure 2*. Proportion of fixations for past tense reversible verbs such as "The pedestrian has opened/closed the umbrella". Vertical lines indicate the onset of each word shown above. Eye movements are synchronized to each such onset (hence zero on the x-axis at each vertical). For example, the first vertical line shows the onset of *has*, with fixations synchronized to this onset.

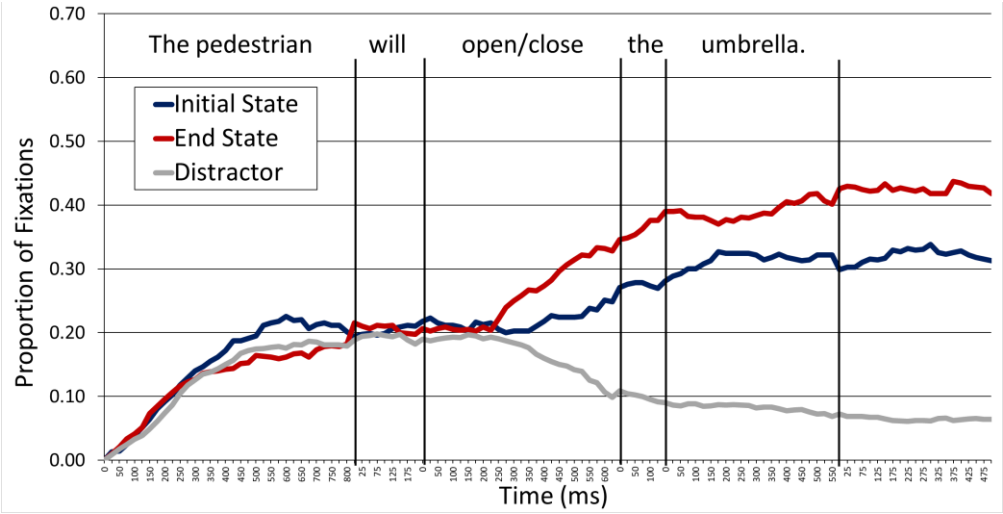


*Figure 3*. Proportion of fixation for future tense reversible verbs such as "The pedestrian will open/close the umbrella". [3]

[3] Fixations were plotted comparing the initial and end states with each tense shown in separate graphs because tense was between-subjects.

**Fixations**. At the onset of the determiner, there was a main effect of state ($F_1(1,77)=29.46$, $p<0.001$; $F_2(1,38)=47.74$, $p<.001$) with a larger proportion of trials with fixations on the end state. There was also a marginal main effect of tense with significant differences by-participants but not by-items ($F_1(1,77)=4.68$, $p=.034$; $F_2(1,38)=2.87$, $p=.098$) with a larger proportion of trials with fixations in the past tense. Additionally, there was an interaction of state by tense at this time point ($F_1(1,77)=5.70$, $p=.019$; $F_2(1,38)=8.36$, $p<.01$). Planned comparisons showed no difference in looks to the initial state in the future tense compared to the past tense ($F_1<1$; $F_2<1$) but significantly more looks to the end state in the past tense than the future tense ($F_1(1,77)=8.06$, $p<.01$; $F_2(1,38)=7.36$, $p=.01$).

Additionally, analyses showed a main effect of state ($F_1(1,77)=157.06$, $p<.001$; $F_2(1,38)=100.69$, $p<.001$) with more looks to the initial state than the distractors. [4] This significant difference between the experimental objects and the distractors was found across all fixation and saccade measures for Experiment 1. See Table 2 for the proportion of trials with fixations to the initial or end state.

Table 2
*Mean proportion of trials with at least one fixation to the initial state or end state at the onset of the determiner (e.g., the). Standard deviations are in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Determiner (e.g., the) | 0.25 (0.11) | 0.27 (0.10) | 0.44 (0.15) | 0.35 (0.13) |

[4] An average value between the two distractors was used to accurately compare two individual objects.

**Saccades**. Proportion of saccades were analyzed over from the onset of the verb to the offset of the determiner (e.g., open the) and calculated by participant and by item. This measure is the proportion of trials with at least one saccade. There was a main effect of object state ($F_1(1,77)=41.37$, $p<.001$; $F_2(1,38)=42.71$, $p<.001$) with a larger proportion of looks to the end state. Additionally, there was a main effect of tense ($F_1(1,77)=14.00$, $p<.001$; $F_2(1,38)=11.71$, $p<.01$) with a larger proportion of looks in the past tense during the same window. There was also an interaction of state by tense for the same window ($F_1(1,77)=5.51$, $p=.021$; $F_2(1,38)=5.33$, $p=.026$) with a significant difference in saccades to the end state comparing the past and future tense ($F_1(1,77)=17.22$, $p<.001$; $F_2(1,38)=7.36$, $p=.01$) but no difference in saccades to the initial state in the past vs. future tense ($F_1(1,77)=3.16$, $p=.08$; $F_2<1$).

Saccades away from each object were also analyzed in the same time window of the onset of the verb to the offset of the determiner (e.g., open the). Participants showed a main effect of state ($F_1(1,77)=13.67$, $p<.001$; $F_2(1,38)=12.00$, $p<0.01$) with a significantly larger proportion of trials with at least one saccade away from the initial state than the end state. Participants also showed a main effect of tense ($F_1(1,77)=8.99$, $p<.01$; $F_2(1,38)=9.92$, $p<0.01$) with a significantly larger proportion of trials with saccades away for the past tense during the same window. There was no significant interaction of state by tense ($F_1 < 1$; $F_2 < 1$). See Table 3 for proportion of saccades to and away from each object state.

Table 3

*Mean proportion of trials with at least one saccade to the initial or end state, from the onset of the verb to the offset of the determiner (e.g., open the). Mean proportion of trials with at least one saccade away from the initial or end state during the same window. Standard deviations are in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Saccades to the object | 0.43 (0.13) | 0.38 (0.11) | 0.58 (0.16) | 0.45 (0.10) |
| Saccades away from the object | 0.36 (0.16) | 0.28 (0.11) | 0.31 (0.14) | 0.22 (0.10) |

Table 4

*Mean duration of fixations (in ms) to the initial state or end state from the onset of the verb to the offset of the determiner (e.g., open the). Standard deviations are listed in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Verb to Offset of Determiner (e.g., open the) | 435 (110) | 505 (253) | 844 (364) | 592 (282) |

**Durations**. Duration of fixations were analyzed looking at the time window from the onset of the verb to the offset of the determiner (e.g., "open the"). If individuals are looking more to the end state before the final noun, one would predict longer looks to the end state as well. A main effect of state was found ($F_1(1,77)=34.19$, $p<.001$; $F_2(1,38)=51.42$, $p<.001$) with longer looks to the end state. A main effect of tense was also found ($F_1(1,77)=4.47$, $p=.038$; $F_2(1,38)=5.43$, $p=.025$) with longer looks in the past tense. Additionally, there was a significant interaction of state by tense ($F_1(1,77)=14.48$, $p<.001$; $F_2(1,38)=14.81$, $p<.001$). Planned comparisons showed no difference in looks to the initial state in the future tense compared to the past tense ($F_1(1,77)=2.57$, $p=.11$; $F_2(1,38)=1.59$, $p=.17$) but longer looks to the end state in the past tense compared to the future tense ($F_1(1,77)=11.80$, $p<.01$; $F_2(1,38)=12.14$, $p<.01$). See Table 4 for duration of saccades.

**Switches**. We analyzed the number of trials that had at least one switch from one state directly to the other from the onset of the verb to the offset of the determiner (e.g., open the). This analysis showed a main effect of state ($F_1(1,77)=17.15$, $p<.001$; $F_2(1,38)=14.17$, $p<.01$) with a significantly larger proportion of trials that had at least one switch from the initial state to the end state than trials with at least one switch directly from the end state to the initial state. There was also a main effect of tense ($F_1(1,77)=6.39$, $p=.014$; $F_2(1,38)=4.17$, $p=.048$) with a larger proportion of trials in the past tense with at least one switch overall as compared to the proportion of trials with at least one switch in the future tense. There was no significant

Table 5
*Mean proportion of trials with at least one switch from one state to another from the onset of the verb to the offset of the determiner (e.g., open the). Standard deviations are listed in parentheses.*

| Object | Initial to End State | | End to Initial State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Verb to Offset of Determiner (e.g., open the) | 0.28 (0.13) | 0.21 (0.08) | 0.21 (0.12) | 0.15 (0.08) |

interaction between direction of switches and the tense of the sentence ($F_1$<1; $F_2$<1). See Table 5 for mean proportion of trials with at least one switch.

**Discussion**

Overall, participants looked more to the target items than the distractors. Our results show a strong preference for the end state for all items in both the future and past tenses. Both fixations and saccades showed more looks to the end state. Participants looked away from the initial state more than the end state, and when looking to the end state, they switched to the initial state far less often than they switched from the initial state to the end state. Finally, there were longer looks to the end state. All of these measures suggest that, for what we are calling "reversible action verbs" (e.g., *open/close* or *tie/untie*), participants anticipate the end state when hearing that an action *has* occurred as well as when hearing that an action *will* occur.

These results are consistent with the results of Kang (2015). Participants look more to the end state than the initial state. Kang used different stimuli to the ones we used here. In that design, the preference for the end state manifested as a preference for one object in the scene over another. This required comparison of eye movements towards different objects. While we replicate her finding in Experiment 1, we can also rule out the possibility that participants in her study preferred one object because it was more interesting (e.g., a dropped ice cream). We show that participants look to the dropped ice cream more when they hear *drop*, but they will look to

the upright ice cream when hearing *pick up*. The data from Experiment 1 also show that tense does not change this bias for the end state; participants look to the end state regardless of tense.

However, the results of Experiment 1 showed a stronger effect in the past tense with a larger difference between the initial and end states in the past tense than in the future tense. In the past tense, the initial state cannot be acted on. The end state, however, is the state that can be acted on as well as the goal of the action. For the future tense, the initial state is the object that is acted on, but the end state is the goal of the action. This suggests that eye movements are driven by the goal of the action, but the object that affords the action manipulates looks as well.

One important consideration for this experiment is that, for these reversible action verbs, participants looked to the end state in both the past *and* the future tense. This was our prediction. However, this does not fit with the results of Altmann and Kamide (2007), who showed an effect of tense. That study showed more looks to the end state in the past tense and more looks to the initial state in the future tense. Given that our results are not consistent with those in Altmann and Kamide (2007), it was necessary for us to run an additional experiment to see if their results could be replicated. Experiment 2 uses destruction verbs similar to those used in the Altmann and Kamide (2007) study, such as *eat* and *drink*. However, differently from the original Altmann and Kamide (2007) study, our items will be the same for both tenses. For instance, participants will hear "The woman will eat/has eaten the cake" maintaining the same final noun. This allows us to confirm that individuals are able to anticipate the initial and end states of each action.

## Experiment 2

### Methods

**Participants.** For Experiment 2, 39 participants were run from the University of Connecticut subject pool of undergraduate psychology students. These participants were those

that were used in the Experiment 1 future tense group. All participants were over the age of 18 and native speakers of English. They were screened for normal or corrected-to-normal vision.

**Stimuli.** Participants were given 12 experimental items each containing destruction verbs (e.g., eat/drink). Each item appeared either in the future tense or in the past tense: The woman will eat the cake vs. The woman has eaten the cake. The experimental items were counter-balanced so that all participants heard just one version of each item. For example, if a participant heard "The woman will eat the cake", they did not hear "The woman has eaten the cake". Each participant therefore had six items from each condition (see Appendix C for a full list of sentences used). This experiment was folded in to the Experiment 1 future tense experiment. There were few items used for this manipulation, but we would consider this a first pass to see if the results from Altmann and Kamide (2007) were able to be replicated.

The rest of the items seen in Experiment 2 were the 20, future tense experimental items from Experiment 1, and the 64 filler sentences and 12 comprehension questions from Experiment 1. Each participant saw 69 items in total, with 49 in the past tense and 47 in the future tense. See Figure 4 for an example of the visual stimuli used in Experiment 2.

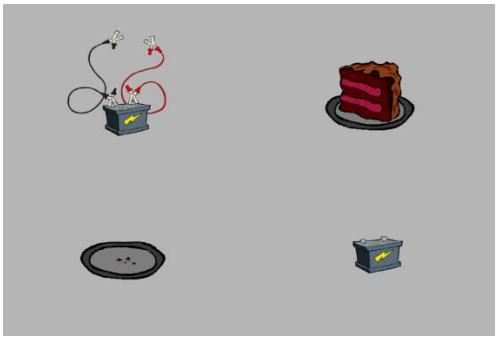

*Figure 4*. Display presented with the sentences "The woman has eaten/will eat the cake."

**Norming**. Stimuli for this experiment were normed via Mechanical Turk for typicality. Participants were given one object state and asked how typical the image was on a scale of one to seven (one being not typical and seven being very typical). For example, one participant saw a full cake and rated how typical a cake this image was. Another participant saw an empty plate with crumbs and rated how typical a cake this image was. A total of sixty participants were run. The participants used for this norming were the same as those used for the typicality norming of Experiment 1. Two lists were made so that participants did not see both object states. Four items were omitted from analysis due to error. The initial state mean rating was 5.04 (SD=0.89), and the end state mean was 3.38 (SD=1.25. These mean ratings were significantly different ($t(12)=-3.06$, $p=.009$).

It was briefly mentioned above that the destruction verbs are particularly interesting because the initial and end states of these verbs are not the same object. For instance, the initial state of eating is a full cake, but the end state is a plate with crumbs (i.e., not a cake). We wanted to be sure that looks to an object were not due to the fact that one object was more associated with the action than the other. In other words, we wanted to determine if, when hearing "eat", looks were driven to one state because it has a significantly higher association with the action of eating. Therefore, the stimuli for this experiment were also normed for associations via a survey given in the University of Connecticut subject pool of undergraduate psychology students. Participants were given one object state and asked how much they associate the image with a verb on a scale of one to seven (one being not associated and seven being completely associated). For example, one participant saw a full cake and rated how much they associate *cake* with the verb *eat*. Another participant saw an empty plate with crumbs and rated how much they associate the *plate* with the verb *eat.* A total of 60 participants were run. The participants used

for this norming were the same as those used for the typicality norming of Experiment 2. Two lists were made so that participants did not see both object states. The mean rating for the initial state was 4.84 (SD= 1.60), and the mean for the end state was 4.51 (SD= 1.11). There was not a significant difference between the ratings for the initial state and end state ($t(19)=-0.60$, $p=.55$).

**Procedure**. The procedure for this experiment matched the procedure presented in Experiment 1.

## Results

For analysis of this data, we analyzed all saccades and fixations. All data in Experiment 2 satisfied the assumptions of normal distribution and homogeneity of variance. Arcsine transformations were performed on all data except duration of fixations. Just as in Experiment 1, we analyzed fixations, saccades to each object state, saccades away from each object state, duration of fixations, and switches from one object state to another. However, to replicate the analysis in Altmann & Kamide (2007) we completed pairwise comparisons to examine looks to the initial state (e.g., cake) in the past tense vs. looks to the initial state in the future tense and looks to the end state (e.g., plate with crumbs) in the past tense vs. future tense. See Figures 5 and 6 for proportion of fixations for each object state across the time course of the entire sentence.

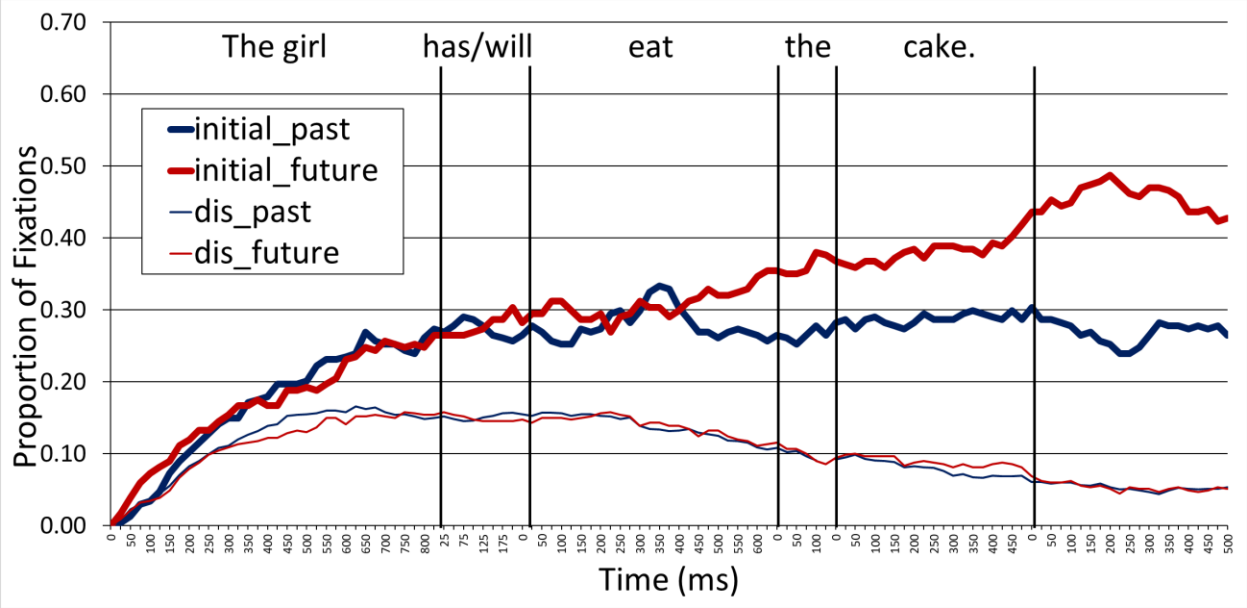


*Figure 5*. Proportion of fixations to the initial state and to the distractors for the destruction verbs, e.g., "The woman has eaten/will eat the cake". An average value was taken across the two distractors shown in the visual stimuli.

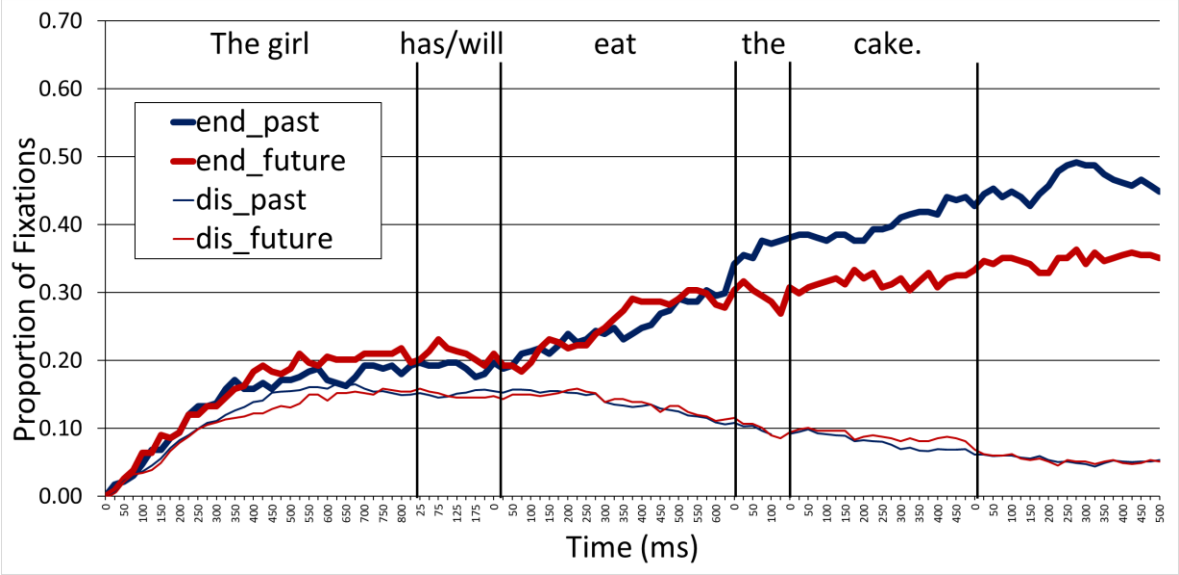


*Figure 6*. Proportion of fixations to the end state and the distractors for destruction verbs, e.g., "The woman has eaten/will eat the cake". An average value was taken for the two distractors presented in the visual stimuli.

Table 6
*Mean proportion of trials with at least one fixation to the initial state or end state at the onset of the determiner (e.g., the) and the offset of the noun (e.g., cake). Standard deviations are in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Determiner (e.g., the) | 0.26 (0.20) | 0.33 (0.19) | 0.38 (0.19) | 0.27 (0.18) |

Table 7
*Mean proportion of trials with at least one saccade to the initial state or end state from the onset of the verb to offset of the determiner (e.g., eat the). Mean proportion of trials with at least one saccade away from the initial or end state in the same window. Standard deviations are in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Saccades to the object | 0.31 (0.21) | 0.42 (0.24) | 0.41 (0.17) | 0.34 (0.20) |
| Saccades away from the object | 0.26 (0.22) | 0.33 (0.19) | 0.22 (0.16) | 0.24 (0.19) |

**Fixations.** At the onset of the determiner (e.g., the), there was no significant difference between the initial state in the future tense compared to the past tense ($F_1$<1; $F_2$(1,22)=3.38, $p$=.08). There was no significant difference between the end state in the future tense compared to the past tense ($F_1$<1; $F_2$<1). However, analyses showed a main effect of state ($F_1$(1,76)=44.12, $p$<.001; $F_2$(1,22)=22.18, $p$<.001) with more looks to the initial state than the distractors. This significant difference between the experimental objects and the distractors was found across all fixation and saccade measures for Experiment 2. See Table 6 for fixation data.

**Saccades**: From the onset of the verb to the offset of the determiner (e.g., eat the), there was no difference in saccades to the initial state in the past tense compared to the future tense ($F_1$(1,76)=1.50, $p$=.22; $F_2$(1,22)=2.40, $p$=.14). There was no difference in saccades to the end state in the past tense compared to the future tense ($F_1$(1,76)=2.54, $p$=.12; $F_2$<1). See Table 7 for saccade data.

Table 8
*Mean duration of fixations (in ms) to the initial state or end state from the onset of the verb to the offset of the determiner (e.g., eat the). Standard deviations are listed in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Verb to Offset of Determiner (e.g., eat the) | 508 (208) | 483 (240) | 584 (413) | 447 (366) |

Table 9
*Mean proportion of trials with at least one switch from one state to another from the onset of the verb to the offset of the determiner (e.g., eat the). Standard deviations are listed in parentheses.*

| Object | Initial to End State | | End to Initial State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Verb to Offset of Determiner (e.g., eat the) | 0.18 (0.16) | 0.21 (0.17) | 0.14 (0.15) | 0.17 (0.18) |

In the same window, there was no difference in saccades away from the initial state in the past tense compared to the future tense ($F_1$<1; $F_2$<1). Additionally, there was no difference in saccades away from the end state in the past tense compared to the future tense ($F_1$<1; $F_2$<1).

**Durations**. From the onset of the verb to the offset of the determiner (e.g., the), there was no significant difference in duration of fixations between the initial state in the future tense compared to the initial state in the past tense ($F_1$<1; $F_2(1,22)=1.02$, $p$=.32). Additionally, there was no significant difference in duration of fixations between the end state in the future tense compared to the end state in the past tense ($F_1$<1; $F_2$<1). See Table 8 for duration data.

**Switches**. From the onset of the verb to the offset of the determiner (e.g., eat the), there was no difference in switches from the initial to the end state in the past tense compared to the future tense ($F_1$<1; $F_2$<1). There was also no difference in switches from the end state to the initial in the past tense compared to the future tense ($F_1$<1; $F_2$<1). See Table 9 for duration data.

## Discussion

Our results showed more looks to the experimental objects than the distractors. However, there were no significant effects when comparing the initial state to the end state. Though our

results were not significant, the patterns shown in Experiment 2 appear in the same direction that was found by Altmann and Kamide (2007). There were more looks to the end state in the past tense compared to the future tense but more looks to the initial state in the future tense compared to the past tense. Despite being able to replicate the pattern of the original study, this does not explain why destruction verbs pattern differently from the reversible action verbs from Experiment 1. Those verbs showed more looks to the end state regardless of tense.

One key difference between the reversible action verbs and the destruction verbs is that the depicted end state for the destruction verbs is *not* the goal state of the action. For example, the goal of eating is not to have an empty plate. In other words, the process of eating is not conducted solely with the goal of achieving the end state of an empty plate. The goal would be to satisfy your hunger. We hypothesize that, for the destruction verbs, eye-movements are not driven to the goal state because it is not depicted. Therefore, anticipation of each object state seems to follow the following pattern: goal state > initial state > "correlated state". The term "correlated state" is used here to describe the end state for these destruction verbs. Being that the end state is not the goal, it can be described as a "correlated state". For example, the empty plate is the result of the action of eating, but is not the goal of eating. It is merely correlated with the action. Participants in Experiment 2 anticipated objects according to this pattern (i.e., looks to initial state > correlated state) for the future tense. Additionally, Experiment 1 can also be accounted for by this proposal because there were more looks to the goal state than the initial state in both tenses.

However, the past tense findings of Experiment 2 are not accounted for by this proposal. The past tense elicited more looks to the end state (or correlated state) than the initial state. The results suggest that the future tense and past tense are processed differently. For the past tense,

participants look to the object that affords an action. In the past tense, the action has already occurred, so the end state is the state which affords action. Therefore, participants' eye movements seem to be drawn to the end state. This preference for the end state in the past tense is also found in Experiment 1.

Our norming data supports our proposal that eye movements are drawn to the object that affords the action in the past tense. Our typicality ratings showed that there was a significant difference between the initial state and the end state typicality with higher typicality for the initial state. However, our results showed more looks to the end state in the past tense compared to more looks to the initial state in the future tense. If typicality were the sole driver of our results, we would not have found this effect of tense. Additionally, our association ratings did not show a significant difference between the initial and end state. Therefore, association can be ruled out as the sole driver of our results.

Experiment 1 showed a preference for the end state in both tenses, but Experiment 2 showed a preference for the end state in the past tense and the initial state in future tense. This shows a difference of results when looking at tense. We chose to address this difference by running a third class of verbs: creation verbs. Creation verbs have a less easily depicted initial state (e.g., a ball of yarn is the initial state of knitting a sweater). This makes creation verbs similar to destruction verbs which have a less easily depicted end state. However, it is unclear how this property affects eye-movements. Experiment 3 will answer the question of whether individuals are able to anticipate the goal when the initial state is less clear. We chose to run a third experiment to determine what motivates looks for creation verbs. If creation verbs pattern similarly to the reversible action verbs used in Experiment 1, we would predict more looks to the end state in both the past tense and future tense. However, if creation verbs pattern similarly to

the destruction verbs used in Experiment 2, we would predict more looks to the end state in the past tense and more to the initial state in the future tense.

## Experiment 3

### Methods

**Participants.** For this experiment, 40 participants were run from the University of Connecticut subject pool of undergraduate psychology students. An additional five participants were run but were excluded due to poor calibrations. These participants are the same participants that were used for the past tense verbs in Experiment 1. All participants were over the age of 18 and native speakers of English. They were screened for normal or corrected-to-normal vision.

**Stimuli.** Participants were given 12 experimental items of creation verbs (e.g., *bake/knit*). The condition for this experiment was tense: The grandmother will knit the sweater vs. The grandmother has knit the sweater. The experimental items were balanced so that all participants heard one version of each item. For example, if a participant heard "The grandmother will knit the sweater", they did not hear "The grandmother has knit the sweater." Each participant had six items from each condition (see Appendix D for a full list of sentences used). See Figure 7 for an example of visual stimuli.

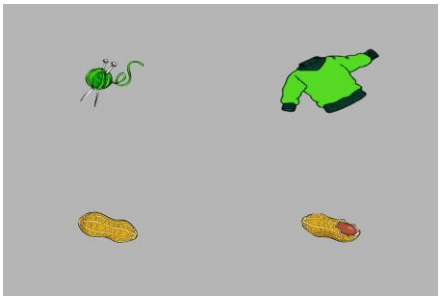

*Figure 7.* This display would be presented with the sentences "The grandmother has knit/will knit the sweater.".

Each participant also heard the 20 past tense items used in Experiment 1. Additionally, the 61 filler sentences and 11 comprehension questions from Experiment 1 were shown. Participants were given 93 total items: 48 in the past tense and 45 in the future tense.

**Norming**. The stimuli for this experiment were normed via a survey given in the University of Connecticut subject pool of undergraduate psychology students. Participants were given one object state and asked how typical the image was on a scale of one to seven (one being not typical and seven being very typical). For example, one participant saw a sweater and rated how typical a sweater this image was. Another participant saw a ball of yarn and completed the same task by rating how typical a sweater the image was. Sixty participants were run in total. Two lists were made so that participants did not see both object states. The mean rating for the initial state was 3.73 (SD=1.18), and the mean for the end state was 5.75 (SD=0.59). There was an overall significant difference between the mean ratings ($t(16)=5.29$, $p<.001$).

Additionally, the stimuli for this experiment were normed for associations via a survey given in the University of Connecticut subject pool of undergraduate psychology students. Participants were given one object state and asked how much they associate the image with a verb on a scale of one to seven (one being not associated and seven being completely associated). For example, one participant saw a sweater and rated how much they associate a *sweater* with the verb *knit*. Another participant saw a ball of yarn and rated how much they associate the *yarn* with the verb *knit*. A total of 60 participants were run. Two lists were made so that participants did not see both object states. The mean rating for the initial state was 5.08 (SD=1.36), and the mean for the end state was 5.38 (SD=1.23). The overall mean ratings for each item of the initial and end states were not significantly different ($t(22)=0.58$, $p=.57$).

**Procedure**. The procedure for this experiment matched that of Experiments 1 and 2.

**Results**

For analysis of the data, 2 x 2 (tense x state) repeated measures ANOVAs were performed on all saccades and fixations. All data satisfied the assumptions of normal distribution and homogeneity of variance. Arcsine transformations were performed on all stimuli except duration. Eye-movements were analyzed by-participant and by-item ($F_1$ and $F_2$, respectively). Additionally, to replicate the analysis in Experiment 2, we completed pairwise comparisons to examine looks to the initial state in the past tense vs. looks to the initial state in the future tense and looks to the end state in the past tense vs. future tense. See Figures 8 and 9 for graphs of the proportion of looks over the course of the sentence.

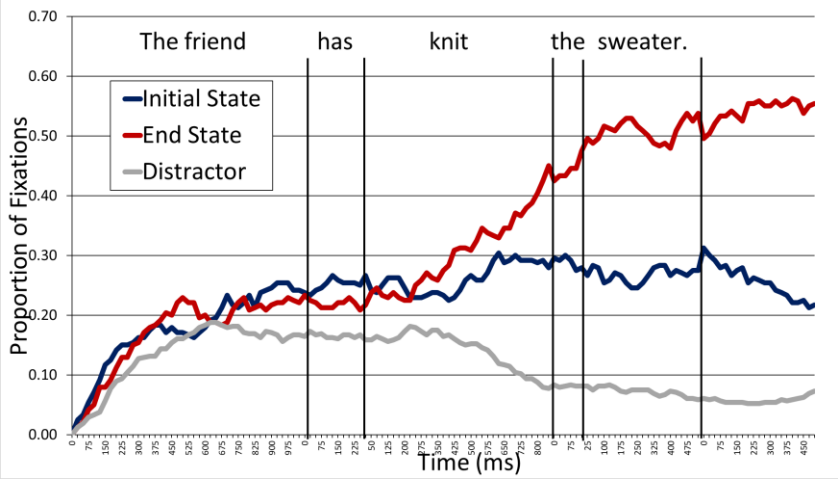


*Figure 8*. Proportion of fixations for past tense creation verbs: "The friend has knit the sweater".

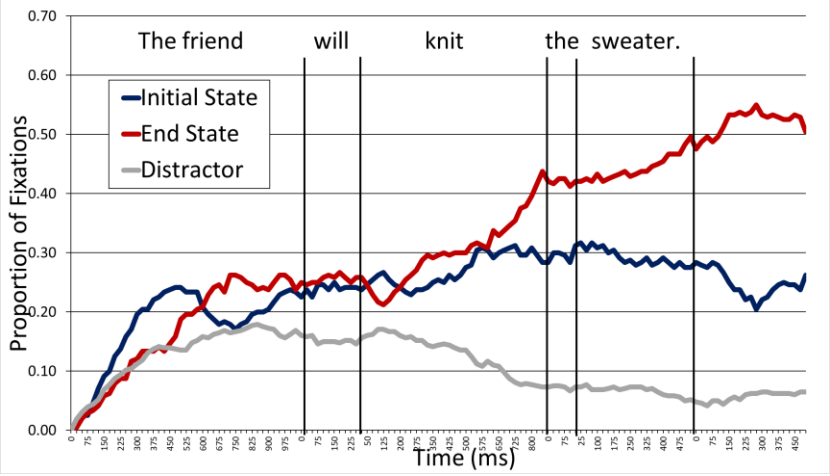


*Figure 9*. Proportion of fixations for the future tense: "The friend will knit the sweater".

Table 10
*Mean proportion of trials with at least one fixation to the initial state or end state at the onset of the determiner (e.g., the) and the onset of the noun (e.g., sweater).*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Determiner (e.g., the) | 0.30 (0.19) | 0.28 (0.18) | 0.43 (0.20) | 0.42 (0.22) |

Table 11
*Mean proportion of trials with at least one saccade to the initial state or end state from onset of the verb to the offset of the determiner (e.g., knit the). Mean proportion of trials with at least one saccade away from the initial or end state during the same window. Standard deviations are listed in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Saccades to the object | 0.42 (0.21) | 0.42 (0.24) | 0.52 (0.19) | 0.49 (0.20) |
| Saccades away from the object | 0.31 (0.19) | 0.29 (0.17) | 0.27 (0.20) | 0.28 (0.19) |

**Fixations.** At the onset of the determiner (e.g., the) there was a main effect of state by-participants but not by-items ($F_1(1,39)=11.85$, $p<.01$; $F_2(1,11)=2.37$, $p=.15$) with more fixations on the end state than the initial state. The pairwise comparisons showed no significant differences between fixations on the initial state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$) or the end state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$). Additionally, analyses showed a main effect of state ($F_1(1,786)=8.75$, $p<.01$; $F_2(1,22)=22.42$, $p<.001$) with more looks to the initial state than the distractors[4]. This significant difference between the experimental objects and the distractors was found across all fixation and saccade measures for Experiment 3. See Table 10 for fixation data.

**Saccades.** See Table 11 for saccade data. From the onset of the verb to the offset of the determiner (e.g., knit the) there was a main effect of state by-participants but not by-items ($F_1(1,39)=8.53$, $p<.01$; $F_2(1,11)= 1.44$, $p=.26$) with more saccades to the end state than the initial state. The pairwise comparisons showed no significant differences between saccades to the initial

Table 12
*Mean duration of fixations (in ms) to the initial state or end state from the onset of the verb to the offset of the object (e.g., knit the sweater). Standard deviations are listed in parentheses.*

| Object | Initial State | | End State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Verb to Offset of Determiner (e.g., knit the) | 454 (204) | 513 (316) | 777 (427) | 653 (409) |

state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$) or the end state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$). Additionally, there was no main effect for saccades away from the objects ($F_1(1,39)=2.73$, $p=.11$; $F_2<1$). The pairwise comparisons showed no significant differences between saccades away from the initial state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$) or the end state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$).

**Durations**. Duration of fixations showed a marginal main effect of state with a significant difference by-participants but not by-items ($F_1(1,39)=13.81$, $p<.01$; $F_2(1,11)=3.18$, $p=.10$) for the time window from the onset of the verb to the offset of the determiner (e.g., knit the). There were longer looks to the end state than the initial. The pairwise comparisons showed no significant differences between durations of fixations on the initial state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$) or the end state in the past tense compared to the future ($F_1$ (1,78)=1.70, $p=.20$; $F_2 < 1$). See Table 12 for mean durations.

**Switches**. When analyzing the number of trials that had at least one switch from one state to the other, there was a main effect of state by-participants but not by-items ($F_1(1,39)=7.73$, $p<.01$; $F_2(1,11)=4.02$, $p=.07$) with a significantly larger percentage of trials with a switch from the initial state to the end state than from the end state to the initial state. This was significant over the time window from the onset of the verb to the offset of the determiner (e.g., knit the). The pairwise comparisons showed no significant differences between switches from the initial to the end state in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$) or from the end state to the

Table 13
*Mean proportion of trials with at least one switch from one state to another from the onset of the verb to the offset of the determiner (e.g., knit the). Standard deviations are listed in parentheses.*

| Object | Initial to End State | | End to Initial State | |
|---|---|---|---|---|
| Condition | Past Tense | Future Tense | Past Tense | Future Tense |
| Onset of Verb to Offset of Determiner (e.g., knit the) | 0.23 (0.15) | 0.24 (0.19) | 0.24 (0.18) | 0.17 (0.16) |

initial in the past tense compared to the future ($F_1 < 1$; $F_2 < 1$). See Table 13 for mean proportion of trials with at least one switch from one state to the other.

**Discussion**

Overall, participants looked more to the experimental objects than the distractors. The results of Experiment 3 replicated the results shown in Experiment 1. These results show a strong preference for the end state for all items in both the future and past tenses. The analyses showed more looks to the end state. There were more fixations on and saccades to the end state than the initial state. Participants looked away from the initial state more, and when looking to the end state, they switched to the initial far less often than they switched from the initial state to the end state. These measures demonstrate that not only are participants able to anticipate the end state of an action, but they are able to do so when presented with different types of verbs.

Not only do reversible action verbs such as *open* and *close* show a preference for the end state, but Experiment 3 shows that creation verbs like *knit* also show this preference. Although creation verbs lack an easily depicted initial state, the participants were still able to infer the initial state. This is demonstrated by the fact that participants looked significantly more toward the initial state than the distractors. Participants in Experiment 3 follow the proposals mentioned in the discussion of Experiment 2. For the past tense, participants anticipate the end state which is the object that affords action in the current state of the world. Additionally, this experiment follows the pattern suggested for the future tense in the discussion of Experiment 2: goal state >

initial state > correlated state. The future tense of Experiment 3 shows more looks to the end state than the initial state. For the creation verbs, the end state corresponds with the goal state.

It is worth noting that the end state was rated as significantly more typical than the initial state in our typicality norming study. Therefore, we cannot rule out typicality as a contributing factor. However, we can rule out associations as the sole driver of our results because there were no significant differences between the ratings of the initial and end states.

## General Discussion

Previous literature shows that goals are important for action comprehension and performance of actions (e.g., Hommel, 2009; Loucks & Meltzoff, 2013). Our results support those found in these developmental and action studies and show that goals are important for anticipation. Three experiments were conducted to determine how verb type and tense influence how participants process actions during a passive listening task. Experiment 1 showed how reversible change of state verbs are processed. Experiment 2 investigated destruction verbs, and Experiment 3 examined creation verbs. All three experiments used the past tense (e.g., has opened) and the future tense (e.g., will open).

Experiment 1 showed that for reversible action verbs, participants anticipate the end state of an action when processing sentences in both the future tense and the past tense. For these verbs, the end state coincides with the goal of the action. For example, the action of *opening the umbrella* has an end state of an open umbrella. The goal of opening an umbrella is to get to this end state of the umbrella being open. It was proposed in the discussion for Experiment 1 that participants' eye movements are driven by the goal state (e.g., an open umbrella being the goal of opening) rather than just what affords the action (e.g., a closed umbrella that affords opening),

as was previously suggested by Altmann and Kamide (1999; 2007). Although it seems that participants are able to anticipate what affords the action, the goal state appeared to be more salient.

The results from Experiment 2 seem to deviate from those found in Experiment 1. The destruction verbs showed different patterns of results by tense which was not found with the reversible action verbs. The past tense of Experiment 2 showed a similar pattern to Experiment 1, where participants tended to look more to the end state than the initial state. However, in the future tense, there was a trend toward more looks to the initial state. It was proposed in the discussion for Experiment 2 that there can be two explanations for what drives eye movements. For the past tense, it was proposed that participants are driven to look at the state that affords action. The current state of the world for the individual coincides with the end state. The action has occurred, so the end state is the current state of the object and the object that affords action. However, for the future tense, participants follow this pattern of eye-movements: goal state > initial state > correlated state. Participants are motivated by the goal of the action. However, this is not depicted in the scene for destruction verbs. Therefore, the next object which draws attention is the object which affords action. In the future tense, the initial state is the object that affords action. Therefore, eye-movements are drawn to the initial state followed by the correlated state.

Experiment 3 showed that creation verbs pattern similarly to the reversible action verbs in Experiment 1 rather than the destruction verbs presented in Experiment 2. Participants anticipate the end state of the action in both the future tense and the past tense. For these creation verbs, the end state coincides with the goal of the action. For example, the goal of *knitting the sweater* is a sweater. A sweater is also the end state of this action.

The two suggested proposals predict the results found in Experiments 1, 2, and 3. For the past tense, all three experiments showed more looks to the end state of the object which corresponds with the object that affords action. For the future tense, Experiments 1, 2, and 3 follow the pattern of goal state > initial state > correlated state. One would predict this pattern with event comprehension in mind. In order to understand that an event has taken place we keep the goal in mind, but we need the initial state to understand what change has occurred. Any other states correlated with the action would then follow. Experiments 1 and 3 both have end states that correspond with the goal state, so more eye-movements toward the end state than the initial state can also be explained as more looks to the goal state than the initial state. Experiment 2, however, does not have a depicted goal state, so there is a preference for the initial state over the depicted end state, which is the correlated state.

Although the explanations for the future and past tense differ slightly, participants overall seem to be motivated by what will happen next in time. Whether this manifests as looking to the object that affords action in the past tense or the goal state in the future tense, we are still constantly looking forward and anticipating the next action and the consequences of the next action. This seems to be a function of action as a whole but can be studied further using language that describes action. This can be seen in all three experimental results. These experiments provide good evidence in favor of the presented theories. However, more work is needed to determine how strong these theories are. For instance, although the "correlated state" theory accounts for our data, this is only one set of studies. Additionally, this proposal was made post-hoc, so the studies themselves were not designed to test this new correlated state theory. Further research would need to be done to test the boundaries of the relationship between the end state,

initial state, and correlated state. One way to continue testing this theory would be to present all three states at once to show us more directly which states are anticipated.

One issue not addressed by this study is that for Experiment 1, participants anticipated the end state (e.g., an open umbrella when they heard *open*). However, there is no way to distinguish if these looks were due solely to the possibility that when participants heard *open* they simply look toward an open object. This would be a concern when looking at Experiments 1 and 3 individually. However, in Experiment 2, participants anticipated the end state for the past tense but the initial state for the future tense. The initial state for the destruction verbs was not a description of the verb. For instance, when hearing *will eat*, participants looked at the full cake which is not the *eaten* object. This differs from Experiment 1 where participants looked to the *open* umbrella which is an *opened* object. Unfortunately, because the results from Experiment 2 were not significant, further studies would need to be conducted to rule out this verb/adjective issue. Similarly, additional research would be needed to test the creation verbs. Given that our norming data showed that the end state was more typical for the creation verbs, we could not rule out typicality as a driver of our final results. An additional study would be necessary to rule out the possibility that typicality was in fact the reason participants look to the end state when they hear creation verbs. Finally, Experiments 2 and 3 had low power; there were only 12 total items for the destruction and creation verbs with only six items per condition for each individual. Additional research should be conducted with more items to confirm the patterns found in this study.

# Appendix A

## Experiment 1 Auditory Stimuli

1. The construction worker has knocked down/built the wall.
   The construction worker will knock down/build the wall.
2. The customer will has broken/mended the glasses.
   The customer will break/mend the glasses.
3. The server has lit/blown out the candle.
   The server will light/blow out the candle.
4. The pedestrian has closed/opened the umbrella.
   The pedestrian will close/open the umbrella.
5. The chef has covered/uncovered the pot.
   The chef will cover/uncover the pot.
6. The hair stylist has curled/straightened the hair.
   The hair stylist will curl/straighten the hair.
7. The assistant has inflated/deflated the balloon.
   The assistant will inflate/deflate the balloon.
8. The waiter has dropped/picked up the ice cream.
   The waiter will drop/pick up the ice cream.
9. The presenter has written/erased the name.
   The presenter will write/erase the name.
10. The farmer has filled/emptied the glass.
    The farmer will fill/empty the glass.
11. The bride has unwrapped/wrapped the gift.
    The bride will unwrap/wrap the gift.
12. The man has wrinkled/ironed the pants.
    The man will wrinkle/iron the pants.
13. The actress has turned off/turned on the lamp.
    The actress will turn off/turn on the lamp.
14. The biker has locked/unlocked the padlock.
    The biker will lock/unlock the padlock.
15. The boy scout has tightened/loosened the knot.
    The boy scout will tighten/loosen the knot.
16. The student has unfolded/folded the paper airplane.
    The student will unfold/fold the paper airplane.
17. The shop owner has unrolled/rolled up the rug.
    The shop owner will unroll/roll up the rug.
18. The cowboy has zipped/unzipped the boot.
    The cowboy will zip/unzip the boot.
19. The barber has wet/dried the towel.
    The barber will wet/dry the towel.
20. The teacher has divided/combined the cookies.
    The teacher will divide/combine the cookies.

## Appendix B

### Filler Auditory Stimuli

1. The florist has watered the sunflowers.
2. The maid has fetched the broom.
3. The employee will pick the mushroom.
4. The jeweler has sold the watch.
5. The cousin will apply the lipstick.
6. The athlete has tied the shoe.
7. The pedestrian has watched the helicopter.
8. The camper has gotten out the flashlight.
9. The shop owner will polish the horn.
10. The girl scout will utilize the compass.
11. The son-in-law has pitched the tent.
12. The nephew has lathered the soap.
13. The woman has bought the hairbrush.
14. The worker has retrieved the stamp.
15. The uncle will set up the telescope.
16. The niece has avoided the butterfly.
17. The daughter has spilled the nail polish.
18. The child will lick the popsicle.
19. The girl will dress the doll.
20. The lawyer has shredded the files.
21. The father has struck the match.
22. The explorer has climbed the mountain.
23. The stepmother has matched the socks.
24. The driver will grab the gas can.
25. The swimmer has examined the seashell.
26. The stepfather has brought the golf club.
27. The grandfather has baked the pie.
28. The aunt has threaded the needle.
29. The toddler has played with the train.
30. The chef will steam the broccoli.
31. The son has cooked the toast.
32. The stepson has unplugged the fan.
33. The lifeguard will apply the sun tan lotion.
34. The homeowner has lit the fire.
35. The lady has packaged the jam.
36. The teenager will plug in the electric guitar.
37. The boy has chosen the candy.
38. The nanny has warmed the bottle.
39. The equestrian will fix the horseshoe.
40. The mother-in-law will replenish the tissues.
41. The assistant has lost the paperclips.
42. The mother has cut the orange
43. The employer has signed the contract.
44. The father-in-law has driven the car.
45. The customer will return the refrigerator.
46. The volunteer has assembled the clipboard.
47. The farmer has built the barn.
48. The janitor will squirt the bottle.
49. The pilot will fly the plane.
50. The expert has moved the chess piece.
51. The gentleman has decorated the door.
52. The handyman has replaced the door hinge.
53. The scientist has corked the test tube.
54. The heiress has tried on the dress.
55. The audience will look through the binoculars.
56. The nurse will fill the syringe.
57. The hostess has roasted the pig.
58. The godson will spin the globe.
59. The cook has pureed the blueberries.
60. The student will fasten the backpack.
61. The man has deposited the cash.
62. The guest will make the bed.
63. The grandmother has knit the sweater.
64. The musician has played the piano.

## Appendix C

### Experiment 2 Auditory Stimuli

1. The general has blown up/will blow up the bomb.
2. The teenager has burned/will burn the photograph.
3. The boy has melted/will melt the ice.
4. The wizard has vanished/will vanish the rabbit.
5. The girl has drunk/will drink the lemonade.
6. The worker has smoked/will smoke the cigarette.
7. The father has deleted/will delete the file.
8. The woman has eaten/will eat the cake.
9. The host has run out of/will run out of toilet paper.
10. The daughter has lost/will lose a tooth.
11. The nurse has dissolved/will dissolve the tablet.
12. The gardener has weeded/will weed the garden.

## Appendix D

### Experiment 3 Auditory Stimuli

1. The housekeeper has filled/will fill the pool.
2. The gardener has planted/will plant the flower.
3. The chef has baked/will bake the bread.
4. The aunt has started/will start the fire.
5. The employee has printed/will print the document.
6. The sister has brewed/will brew the coffee.
7. The grandmother has weaved/will weave the basket.
8. The seamstress has sewn/will sew the skirt.
9. The friend has knit/will knit the sweater.
10. The uncle has sketched/will sketch the drawing.
11. The clown has made/will make the balloon animal.
12. The nanny has frozen/will freeze the popsicle.

References

Altmann, G.T.M. and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Altmann, G.T.M. and 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(4), 502-518.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing Referential Domains during Real-Time Language Comprehension. *Journal of Memory and Language*, *47*(1), 30–49.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of experimental psychology: Learning, memory, and cognition*, 30(3), 687.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton-Mifflin.

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of memory and language*, 31(2), 129-151.

Glenberg, A.M., & Robertson, D.A. (1999). Indexical understanding of instructions. *Discourse Processes*, 28(1), 1-26.

Glenberg, A.M., & Robertson, D.A. (2000). Symbol Grounding and Meaning: A Comparison of High-Dimensional and Embodied Theories of Meaning. *Journal of Memory and Language*, 43(3), 379-401.

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research PRPF*, *73*(4), 512-526.

Kang, X. (2015). Establishing object-state representation in language comprehension: Evidence from picture verification, eye-tracking and ERPs. (Unpublished doctoral dissertation). University of York, UK.

Kaschak, M.P., & Glenberg, A.M. (2000). Constructing Meaning: The Role of Affordances and Grammatical Constructions in Sentence Comprehension. *Journal of Memory and Language*, 43(3), 508-529.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review*, 95(2), 163.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, *96*(1), 1-33.

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language learning and development*, *3*(3), 179-197.

Loucks, J., & Meltzoff, A. N. (2013). Goals influence memory and imitation for dynamic human action in 36-month-old children. *Scandinavian journal of psychology*, *54*(1), 41-50.

Meltzoff, A. (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology, 31*(5), 838-850.

Prinz, W. (1997). Perception and action planning. *European journal of cognitive psychology*, 9(2), 129-154.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science, 268*(5217), 1632-1634.

Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

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

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.