5-5-2017

Episodic and Semantic Memory Activation in Sentence Processing

Zachary Ekves
University of Connecticut, zachary.ekves@uconn.edu

Episodic and Semantic Memory Activation in Sentence Processing

Zachary Ekves

B.Phil., University of Pittsburgh, 2014

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
At the
University of Connecticut
2017

APPROVAL PAGE

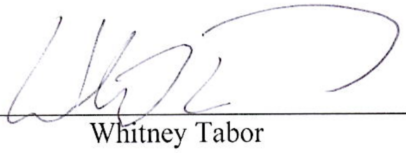
Masters of Science Thesis

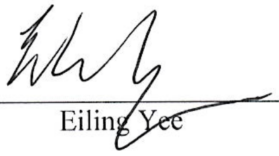
Episodic and Semantic Memory Activation in Sentence Processing

Presented by

Zachary Ekves, B.Phil.

Major Advisor 
Gerry Altmann

Associate Advisor 
Whitney Tabor

Associate Advisor 
Eiling Yee

Associate Advisor 
Jon Sprouse

University of Connecticut

2017

Our ability to successfully comprehend language and interact with the world often relies on the ability to maintain both short- and long-term representations of objects. For example, if you were tasked with retrieving a tomato from a friend's kitchen, you would initially be in search of a red fruit. When you discover that they only have green tomatoes in their kitchen, you are able to maintain this short-term, episodic information about this particular set of tomatoes, while still knowing that tomatoes are red in general. In this study I seek to examine how these memory systems interact during language processing. The review will first focus on findings from the event cognition and sentence processing literature which provide insight into how one updates representations on-line. The focus will then turn to studies of semantic representation that are pivotal in understanding the color priming effects that will be central in this study.

On-line representation of object-states

While there has been a large amount of research conducted on the comprehension of events in language (see Zwaan & Radvansky, 1998 for a review), there has been comparatively little work examining the on-line representations that are constructed and manipulated during language comprehension. Altmann and Kamide (2009) conducted an experiment exploring these changes in on-line representation with a specific focus on location updating. They show that when sentences describe movement of an object in a static visual scene to another location in the scene, subsequent mentions of this object drive looks to where the object was described to have moved to, rather than where it was in the display. Their results not only show that the state of mental object representation during events is continually updated as language unfolds, but they also suggest that the various states of an object (e.g. current and previous) might be simultaneously active

during language comprehension (see also Kukona, Altmann, & Kamide, 2014, for further discussion). Hindy, Altmann, Kalenik, and Thompson-Schill (2012) and Solomon, Hindy, Altmann, and Thompson-Schill (2015) tested this prediction by examining differential brain responses to processing language that described objects undergoing various changes from their canonical state: they suggested that in discourse such as "The man will chop the tomato, and then he will smell the tomato," reference to the sentence final "tomato" requires a need to select a particular state of the tomato (the intact state of the tomato before the chopping, or the cut up state of the tomato after the chopping). When referring to an object that had previously been, for example, weighed, the states of the object are roughly identical, so no competition is observed when referring to it. The results of these studies showed that brain regions involved in resolving competition become more active when language refers to objects that had previously undergone substantial change ("chop"), compared to minimal ("weigh"). And because more than one representation must be active for competition to obtain, these results entail simultaneously active representations of object state. Moreover, these competition effects seem to be token specific. That is, sentences such as "The man will chop the tomato, and then he will smell another tomato," do not activate neural regions associated with competition any more than sentences that refer to a tomato that had previously been weighed.

Taken together, these lines of research show that mental representations of events described by language are updated during processing and that tracking of the states of object tokens seems to be critically important. A more formalized account of these phenomena is described in Altmann (2017). The specifics of the account are omitted for the purposes here, but this account makes one critical prediction: object-state tokens are a

fundamental representational unit of event representation generally and these tokens carry with them information about prior states of the token.

Object and color representation

Although the findings described provide insight into what sort of unit is being tracked as one processes language, it says little about what makes up this unit, i.e. what are the content of these object tokens? Embodied theories of conceptual representation, such as the perceptual symbol systems account of Barsalou (1999), suggest that the long-term representation of any concept contains features that were active during past perceptions of, and interactions with, the object. For example, the semantic representation of "tomato" would contain color information that was perceived through repeated exposure to tomatoes in the world. Therefore the representation of the "tomato" concept would carry with it its color associations, i.e. they are normally red but can plausibly be green, a notion empirically supported through the experiments discussed below (see also Yee, Ahmed, & Thompson-Schill, 2012).

Research using the visual-world paradigm has been used to examine the relationship of these sorts of sensorimotor representations between multiple concepts. Huettig and Altmann (2011) used the paradigm to probe whether the activation of semantic information associated with a concept could drive looks to related objects in a visual scene. Here, critical words were presented in a sentence context ("The boy thought about it carefully and then he spotted the pea and asked whether it was a vegetable too.") while participants viewed a quadrant display with three unrelated distractor objects and one target object that shared the same surface color as a canonical pea (a green jacket). They show that after the onset of "pea" more saccades were directed to the similarly

colored jacket than any of the distractors. This indicates that eye movements are driven by the semantic similarity (color in this case) between the auditory and visual input, and provides validation of the paradigm that will be used in the experiments below.

Given the present study's emphasis on episodic and semantic interplay regarding color associations in unfolding language (i.e. on learning that the tomatoes you are searching for are not red but are green), it becomes important then to have available alternatives to prototypical color associations. This raises the issue of how atypical color features become active during processing. Connell and Lynott (2009) examined the degree to which atypical contexts are able to modulate the activation of typical and atypical colors through language comprehension. Here they presented participants with words with strong color associations in sentence contexts which would prompt one to either imagine the object that the word refers to in a typical or atypical color. For example, "Joe was excited to see a bear in the woods" would prompt one to think of a brown bear, while "Joe was excited to see a bear at the North Pole" would prompt one to think of a white polar bear instead. Using a semantic Stroop task, they examined the degree to which sentences such as these would facilitate successful recognition of the word "bear" presented in a font in the object's typical or atypical color, e.g. in brown or white lettering for the bear example above. Their results showed that when the implied color of the object is typical, priming in the Stroop task was observed for only the typical color font. However, they observed that when the implied color of the object was atypical, there was priming for both the typical and atypical font colors. This suggests that when contexts prompt instantiation of a typical example of an object, only its canonical features

become activated, but when context prompts instantiation of an atypical example of an object, features of both that atypical object and canonical features become active.

The present study

The present study uses the visual world paradigm to probe the relative activation of episodic and semantic memory activation through the course of sentence comprehension. Specifically, it addresses the representational content of newly instantiated object tokens in unfolding language. To this end, two experiments were conducted. The first uses sentence triples that introduce an object with strong color associations in the first sentence, attribute plausible but atypical colors to these objects in the second sentence, and then in the final sentence refer back either to these atypically colored objects (*same-token reference*) or to new tokens of an object of the same type (*different-token reference*). For example, "The man has some tomatoes. They are green. He will take out a scale and weigh some of the tomatoes" (*same-token*) vs. "The man has some tomatoes. They are green. He will take out a scale and weigh some other tomatoes" (*different-token*). The visual scenes in this experiment critically contain an object that shares the same surface color as the typical color of the object mentioned in the sentence (red for "tomato"). The number of looks towards the critical object in the display is taken to indicate activation of the typical color representation from semantic memory.

The previously described findings suggest the following predictions: For the same-token condition, the language explicitly refers to atypically colored objects as the referent. Given this, one straightforward prediction can be made such that upon encountering the noun in the final sentence ("He will… weigh some of the tomatoes"), one should observe no difference in looks between the target object (a red object, for the

tomato example) and distractors – this is because the tomato in question is green, not red, and should therefore engender no more looks to a red object than to any other (unrelated) object. However, given the account of Altmann (2017) and the findings of Connell and Lynott (2009), one could also predict that there will be *more* looks to the target object compared to distractors. The account of Altmann and the account Connell and Lynott predict this pattern for different reasons. The sentences used here are intended to prompt instantiation of a typical object, which is then discovered to be atypically colored. This implies a transformation of the mental representation from typical to atypical, even though these objects are, in fact, different tokens of the same object. For Altmann, this means that the previous typical qualities of the object are carried with it, even though it is atypically colored. That is, by virtue of the atypically colored object being previously thought of as typically colored, canonical color will be activated when the object is referred to. Connell and Lynott, however, suggest that the activation of the atypical state of an object brings with it characteristics of the typical object generally (that is, a property of atypically featured objects is that they activate their typical features also, regardless of any past history or context; the Altmann account can be construed as positing that the typical features are activated only if they had *previously* been activated). A further comparison of these accounts will be discussed later.

For the different-token case, two potential patterns could emerge. If there is an increase in looks to the target object, this suggests that the instantiation of a new object token entails typical characteristics, which would be predicted by Connell and Lynott (2009). If it's the case that there is no difference in looks, this might be indicative of an influence from previous salient episodic information being inherited into new object

tokens. In other words, residual activation from the previous feature set influences the subsequently activated feature set.

A second experiment will be described below, which provides a baseline measure against which to contrast the data from Experiment 1.

Experiment 1

Methods

Participants

Thirty-two participants were recruited from the University of Connecticut Introduction to Psychology subject pool[1]. All participants were native speakers of English, aged 18 or above, and had normal or corrected to normal hearing and vision.

Sentence contexts and visual displays

Two experimental conditions were used for the present study: same-token sentences and different-token sentences.

Same-token reference: "The man has some tomatoes. They are green. He will take out a scale and weigh some of the tomatoes."

Different-token reference: "The man has some tomatoes. They are green. He will take out a scale and weigh some other tomatoes."

A total of 120 trials took place during the experiment with the experimental items, filler items, and comprehension questions interleaved in a fixed (across participants) random order. Thirty-four experimental sentences were presented, counter-balanced across

[1] An additional 10 participants had been run, but were excluded prior to analysis because of computer malfunction which was subsequently corrected or inconsistent calibration.

participants. Sixty-six filler sentences of 3 types were also presented. None of the objects in these sentences necessarily had strong color associations. The first type omitted the second sentence, e.g. “The man has some bread. He will throw away the wrapper and open some other bread.” The second type used a non-color adjective in the second sentence, e.g. “The man has some slippers. They are ragged. He will sit in his chair and put on some other slippers.” The third type used the same structure as the experimental items, though the nouns were singular, e.g. “The child has a sleeping bag. It is blue. He will extinguish the campfire and sleep in the sleeping bag.” Twenty comprehension questions were used during the experiment in order to promote attentive listening. These questions probed for information about the actions and objects mentioned in the sentence, e.g. “Did the man weigh the green tomatoes?”

For each trial a quadrant display was presented to participants. For experimental trials, this display contained one target object (that shared the surface color of the object mentioned in the sentence-final position) and three distractor objects that were not related in any way. For filler trials, no objects in the display were related to the sentential context. The object referred to by the sentence-final noun never appeared in the display. The objects displayed in the visual display were photographic images of various items. See Figure 1 for a sample display.

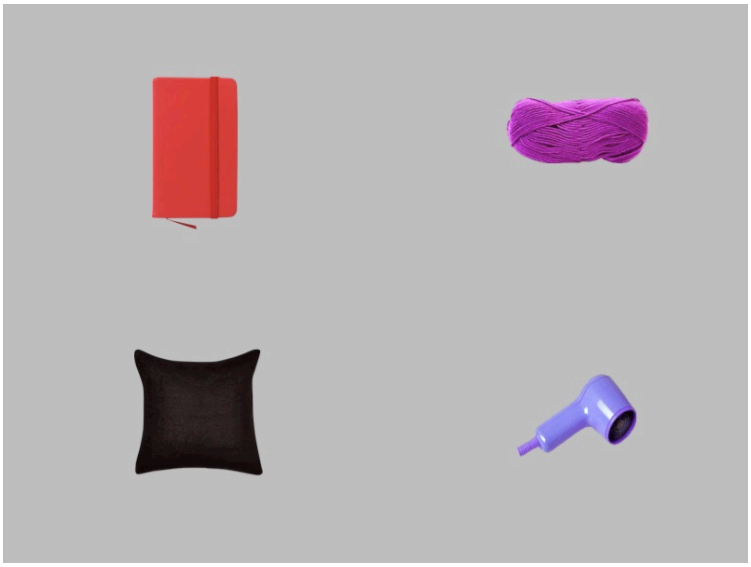


Figure 1. Experimental display corresponding to the sentence contexts “The man has some tomatoes. They are green. He will take out a scale and weigh some of the/some other tomatoes.”

Object color and reference norming

Each of the 34 experimental sentences referenced objects with strong color associations. It was important that all mentioned objects had strong primary color associations and secondary plausible colors, i.e. that tomatoes are primarily red and plausibly green for the sentence “The man has some tomatoes. They are green. He will take out a scale and weigh some of the/some other tomatoes.” In order to ensure that this is the case, the stimuli were normed by 99 native English speakers from the US (recruited using Amazon Mechanical Turk) who gave judgments using Qualtrics presentation software about color likelihood of objects using a fill-in-the-blank form. Participants were asked to provide the most likely, the second through fourth most likely, and the least likely color of 61 different objects. Thirty-four objects were chosen for the experiment from this initial set of 61. For each of these 34 at least 75% of participants agreed that a particular color was the most likely color of an object, and at least 35% of participants agreed that a particular color was the second most likely color of an object. The nature of the items used here

prohibited a more strict secondary color requirement as many items did not have a diagnostic color (e.g. golf balls are normally white, but the secondary color was less agreed upon). Some items did in fact have strong secondary color associations though, e.g. tomatoes are normally red and secondarily green. The paradigm used here only required that the secondary color be plausible, which I take this 35% to be indicative of.

A second norming study was carried out to establish whether, for the example above, participants would assume that weighing "some of the tomatoes" meant weighing green tomatoes, and whether weighing "some other tomatoes" would mean weighing red tomatoes (the canonical color). Seventy-seven native speakers of English from the US (recruited using Amazon Mechanical Turk) provided judgments. Each participant was shown a quadrant visual display with 2 critical items, 2 unrelated distractor items, and the experimental sentence. Each of these objects were cropped images of real objects. The critical items in the display were the critical objects in their primary and secondary colors. For example, a display would be shown with green tomatoes, red tomatoes, two unrelated distractor objects, and the sentence "The man has some tomatoes. They are green. He will take out a scale and weigh some of the tomatoes." Participants were instructed to choose the object that was being referred at the end of the final sentence. Lists were counterbalanced such that each participant was only presented with one of the two experimental sentence contexts for each item. For all experimental sentences, at least 75% of participants selected that "some of the tomatoes" referred to an object in its secondary color (green, in this case), and "some other tomatoes" referred to the object in its primary color (red, in this case).

Procedure

Participants were seated approximately 60 cm from a 24-inch monitor. Eye-movements were recorded using an SR Research Eyelink 1000 Plus eye tracker with remote camera (500 Hz sampling rate). Participant head movements were unrestricted. Participants were told that they should listen to each sentence carefully and that there would be occasional comprehension questions. Eight sets of 5-point calibration were conducted throughout the course of the experiment.

Each trial began after participants fixated a central point on the display. A blank screen was shown as participants were auditorily presented with the first two sentences. At the offset of the second sentence, the blank screen was replaced with the corresponding quadrant display for that experimental item. There was then a 2000 ms delay between the offset of the second sentence and the onset of the third sentence. After the offset of the final sentence, the display remained on the screen for 2000 ms.

The location of the target object was randomized across trials, but was fixed across participants. Each participant heard only one variation of each sentence context, and the trial structure was counter-balanced such that the sentence context for each object was presented to an equal number of participants.

Results

Interest areas for each visual scene were similarly sized rectangular shapes around each of the four objects in the display. Given the predictions described above, the results presented here focus on the final verb phrase in the third sentence. They will be described with reference to the experimental item "The man has some tomatoes. They are green. He will take out a scale and weigh some of the/some other tomatoes". The primary measures

reported here are saccades to, and fixations on, an item in the display. The proportion of saccades is calculated as the proportion of trials in which a saccade was launched to an object in the visual display during the final determiner phrase, while the proportion of fixations is calculated as the proportion of trials in which there was a fixation on an object in the display during this time region. As there is only one critical object in each display (e.g. the red object) and all distractors are theoretically equivalent in this design, looks to distractors have been averaged. 2 (condition: same token, different token) x 2 (object: target, distractor) ANOVAs were conducted for all saccadic and fixation measures. Given the nature of the proportion measures, arcsine transformations were used for all statistical analyses. Descriptive statistics of untransformed data for each segment of the critical region are shown in Tables 1 through 3. Figure 2 shows the untransformed proportion of fixations across for the critical region.

Table 1. Proportion of fixations to target and distractor objects at word onsets and offsets (standard deviation in parentheses).

			Weigh		Some of the/Some other		Tomatoes	
			Onset	Offset	Onset	Offset	Onset	Offset
Subjects	Same-Token	Target	22.44 (11.16)	18.08 (8.42)	16.84 (8.76)	19.60 (11.35)	19.77 (11.37)	22.22 (9.70)
		Distractors	19.98 (5.04)	21.38 (5.46)	21.81 (5.58)	21.02 (5.55)	21.08 (5.50)	19.55 (4.97)
	Different-Token	Target	16.36 (10.32)	16.73 (12.78)	16.36 (13.43)	18.20 (11.99)	18.57 (12.15)	22.24 (10.54)
		Distractors	21.20 (5.99)	21.02 (5.37)	21.02 (5.51)	20.96 (5.23)	20.96 (5.23)	19.85 (5.74)
Items	Same-Token	Target	11.40 (4.47)	9.19 (4.13)	8.55 (3.89)	9.93 (5.21)	10.02 (5.24)	11.31 (5.51)
		Distractors	9.99 (2.10)	10.69 (1.97)	10.91 (1.75)	10.51 (1.74)	10.54 (1.77)	9.77 (1.63)
	Different-Token	Target	8.18 (5.39)	8.36 (5.12)	8.18 (4.98)	9.10 (6.50)	9.28 (6.52)	11.12 (6.20)
		Distractors	10.60 (1.82)	10.51 (1.80)	10.51 (1.76)	10.48 (2.29)	10.48 (2.29)	9.93 (2.58)

Table 2. Proportion of saccades to target and distractor objects from the onset of the listed word to the onset of the next listed word (excluding "tomatoes" which was measured from word onset to word offset; standard deviation in parentheses).

			Weigh	Some of the/Some other	Tomatoes
Subjects	Same-Token	Target	8.64 (6.84)	13.79 (9.64)	15.63 (9.75)
		Distractors	12.25 (6.28)	11.40 (6.23)	12.75 (5.79)
	Different-Token	Target	10.85 (8.98)	14.52 (10.56)	13.97 (8.29)
		Distractors	11.27 (6.59)	12.81 (5.43)	11.58 (6.37)
Items	Same-Token	Target	4.32 (3.60)	6.89 (4.85)	7.81 (5.80)
		Distractors	6.13 (2.73)	5.70 (2.06)	6.37 (2.07)
	Different-Token	Target	5.42 (5.76)	7.26 (4.41)	6.99 (5.81)
		Distractors	5.64 (2.67)	6.40 (2.56)	5.79 (2.80)

Table 3. Proportion of saccades to target and distractor objects from the onset of "some" to the offset of "tomatoes."

			Onset of "Some" to offset of "Tomatoes"
Subjects	Same-Token	Target	28.31 (12.89)
		Distractors	23.28 (10.14)
	Different-Token	Target	27.02 (10.66)
		Distractors	23.47 (9.78)
Items	Same-Token	Target	14.15 (5.04)
		Distractors	11.53 (2.47)
	Different-Token	Target	13.51 (6.28)
		Distractors	11.73 (3.20)

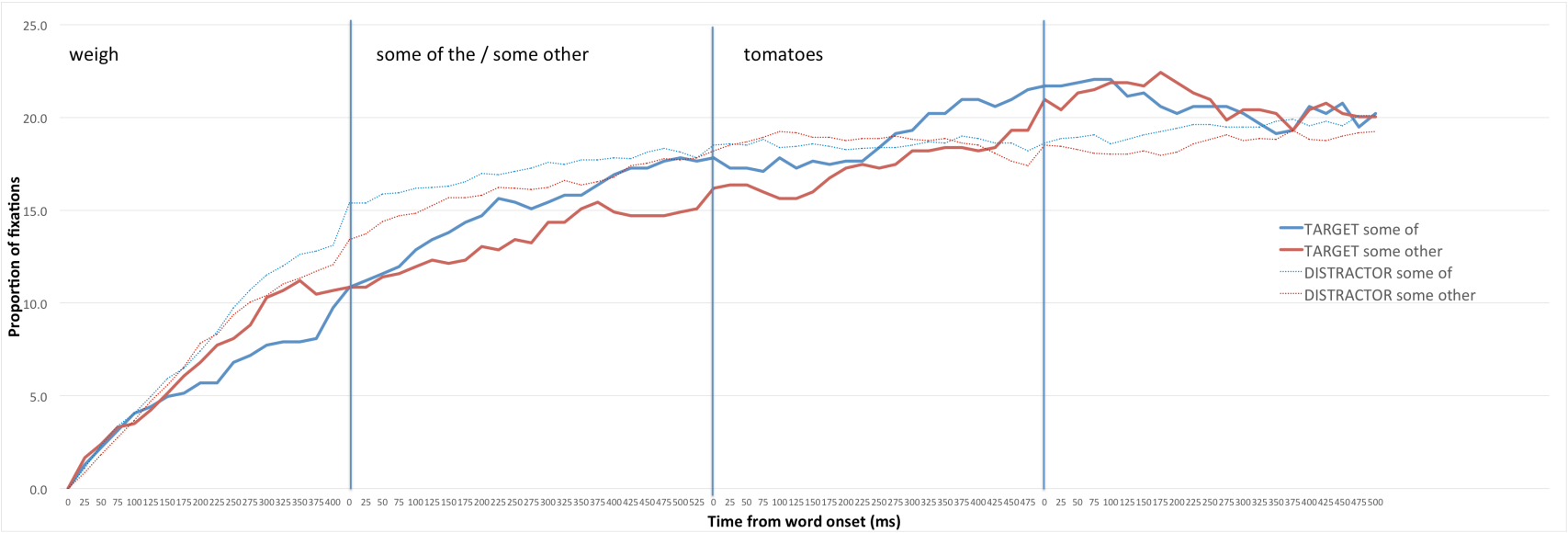


Figure 2. Proportion of fixations to target and distractor objects in the visual scene by condition (binned in 25 ms increments; reset to 0 at the onset of "weigh"). As each sentence differed in length, onset boundaries (marked by vertical lines) were calculated as the average across all sentences.

Examining saccades during the region beginning at the onset of "some" to the offset of "tomato," a marginally significant main effect of object was observed (by-subjects: $F(1,31) = 5.884$, $p = .021$; by-items: $F(1,33) = 1.924$, $p = .175$), with more saccades being launched to the target object than distractor objects. There was no main effect of condition (by-subjects: $F(1,31) = 0.189$, $p = .667$; by-items: $F(1,33) = .321$, $p = .575$) and there was no interaction of object and condition (by-subjects: $F(1,31) = .219$, $p = .643$; by-items: $F(1,33) = .396$, $p = .533$).

In order to ensure that this (marginal) difference in saccades was a result of the language presented during the region described above, fixations at the onset of "some" were also examined (reflecting the eye movement record before the critical difference between the sentence conditions). This analysis shows a main effect of object (by-subjects: $F(1,31) = 12.161$, $p = .001$; by-items: $F(1,33) = 9.406$, $p = .004$), with more fixations being observed on the distractor objects compared to the target object. No main effect of condition (by-subjects: $F(1,31) = 1.416$, $p = .243$; by-items: $F(1,33) = 1.022$, $p = .319$) or interaction (by-subjects: $F(1,31) = .133$, $p = .718$; by-items: $F(1,33) = .250$, $p = .875$) was observed.

Despite more saccades being directed towards the target object compared to distractors from the onset of "some" to the offset of "tomato", analysis of fixations at the offset of "tomato" revealed no difference in fixation probability as a function of object (by-subjects: $F(1,31) = 1.391$, $p = .247$; by-items: $F(1,33) = .060$, $p = .808$) although the same pattern is numerically shown. No main effect of condition (by-subjects: $F(1,31) = .003$, $p = .950$; by-items: $F(1,33) = .005$, $p = .943$) and no interaction between object

and condition (by-subjects: $F(1,31) = .003$, $p = .960$; by-items: $F(1,33) = .015$, $p = .903$) were shown. No other differences in sub-regions or intervening points were significant.

Discussion

This experiment examined the degree to which previous episodic information influenced the representation of subsequent objects that are referred to in language. Here experimental sentences attributed atypical (but plausible) color information to objects with strong canonical color representations (e.g. attributing a green color to a canonically red tomato). These sentences subsequently either referred to objects with this atypical color (same token) or referred to a new set of tokens (different token). The results show that beginning from the onset of the determiner phrase to the offset of the noun ("some of the/some other tomatoes"), more saccades were launched to an object in the display that shared surface color with the canonical color of the object mentioned compared to unrelated distractor objects (although this difference was only significant in a by-subjects analysis) despite a statistically significant bias favoring looks to distractor objects at the onset of the region. No differences were observed in comparing the number of saccades to the target object between the same-token and different-token conditions.

For the same-token case, this pattern (if statistically reliable) would suggest that regardless of the fact that the language is explicitly referring to a set of atypically colored objects, the canonical color becomes active during reference. This pattern of results is predicted based on the accounts of both Altmann (2017) and Connell & Lynott (2009). These accounts differ in their theoretical motivation for making this prediction however.

For Altmann, objects that undergo change carry with them aspects of their histories. To illustrate this, consider the experimental stimulus: "The man has some

tomatoes. They are green. He will take out a scale and weigh some of the tomatoes." This discourse first prompts instantiation of a set of canonical, red tomatoes, which are then modified in the second sentence when they are attributed a new color (green). When the final sentence refers back to these objects, by virtue of the objects having previously been attributed the red feature, this discourse final reference will result in activation of both red and green color features.

According to Connell and Lynott (2009), sentential contexts that prompt instantiation of objects in non-canonical forms also result in the activation of the canonical features as well. Given our experimental stimuli, when sentences prompt the instantiation of atypical objects (e.g. referring to a set of green tomatoes), by-virtue of tomatoes having a canonical color, both red and green color features are activated. Although the current experiment is not able to tease apart these accounts, this account does not seem to be compatible with the findings of Hindy et al. (2012). As previously mentioned, they compared the processing of sentences which described substantially, compared to minimally, changed objects. In addition to showing "competition" effects when referring back to the changed object in sentences such as "The man will chop the tomato, and then he will smell the tomato," they also observed competition in sentences such as "The man will chop the tomato, but first he will smell the tomato." Here the sentence context prompts the instantiation of a tomato representation from before the chopping event, which, according to the account of Connell and Lynott, should only activate the canonical representation. If that's the case, no competition effects should have been observed, but they were. Given this, the account of Altmann seems to have the most explanatory power regarding the results of Hindy et al. (though the account of

Connell and Lynott did not explicitly intend to account for these effects). Regarding the different-token condition, the same pattern was found with more saccades being directed to the target object in the display compared to distractors. This is suggestive that the instantiation of new objects prompts activation of typical features.

However, an unresolved issue in the interpretation of these results (to the extent that we observed a marginally significant effect of object) is that there is no baseline by which we can judge the effects. It could be the case that the attribution of atypical color in this paradigm was ineffective and objects maintain their typical color as a result of this. To this end, a second experiment was conducted where instead of introducing the tomatoes and then changing their color to the atypical green, the tomatoes were introduced but no reference to their color was given, allowing them to maintain their "generic redness". This provides the ability to establish a baseline and judge the sensitivity of the paradigm.

Experiment 2

Methods

Participants

Twenty-two participants were recruited from the University of Connecticut Introduction to Psychology subject pool[2]. No participants who took part in Experiment 1 took part in this experiment. All participants were native speakers of English, aged 18 or above, and had normal or corrected to normal hearing and vision.

[2] An additional 7 participants had been run, but were excluded prior to analysis because of incomplete data or inconsistent calibration.

Sentence contexts and visual displays

The same set of experimental stimuli from Experiment 1 was used here, although the second sentence was omitted:

Same-token reference: "The man has some tomatoes. He will take out a scale and weigh some of the tomatoes."

Difference-token reference: "The man has some tomatoes. He will take out a scale and weigh some other tomatoes.

The filler items and visual scenes were unchanged from Experiment 1, and the same counter-balancing scheme was used here.

Procedure

The procedure was unchanged from Experiment 1.

Results

Interest areas for this experiment were the same as Experiment 1. 2 (condition: same token, different token) x 2 (object: target distractor) ANOVAs were conducted for all saccadic and fixation measures. Arcsine transformations were used for all inferential statistics described below. Descriptive statistics of untransformed data are reported in Tables 4 through 6 and Figure 3 shows the untransformed proportion of fixations across the entire critical sentence region. Again, for simplicity in description the results will be described with reference to the example phrase "The man will weigh some of the/some other tomatoes."

Table 4. Proportion of fixations to target and distractor objects at word onsets and offsets (standard deviation in parentheses).

			Weigh		Some of the/Some other		Tomatoes	
			Onset	Offset	Onset	Offset	Onset	Offset
Subjects	Same-Token	Target	20.32 (11.73)	24.60 (12.66)	24.87 (13.70)	22.73 (15.75)	22.46 (15.58)	22.46 (13.04)
		Distractors	20.23 (4.58)	18.27 (4.62)	18.45 (4.71)	18.45 (7.23)	18.72 (7.00)	19.61 (5.21)
	Different-Token	Target	21.12 (11.43)	16.84 (12.48)	16.58 (12.13)	18.72 (11.14)	18.72 (11.14)	19.52 (15.34)
		Distractors	19.34 (6.69)	20.68 (7.02)	20.68 (7.28)	19.96 (7.02)	19.88 (6.93)	19.88 (7.09)
Items	Same-Token	Target	10.16 (5.39)	12.30 (6.76)	12.43 (6.74)	11.36 (6.55)	11.23 (6.15)	11.23 (7.19)
		Distractors	10.12 (2.68)	9.14 (2.49)	9.22 (2.62)	9.22 (2.31)	9.36 (2.16)	9.80 (2.42)
	Different-Token	Target	10.56 (6.50)	8.42 (5.61)	8.29 (5.30)	9.36 (5.70)	9.36 (5.70)	9.76 (5.83)
		Distractors	9.67 (2.87)	10.34 (2.70)	10.34 (2.68)	9.98 (2.69)	9.94 (2.69)	9.94 (2.67)

Table 5. Proportion of saccades to target and distractor objects from the onset of the listed word to the onset of the next listed word (excluding "tomatoes" which was measured from word onset to word offset; standard deviation in parentheses).

			Weigh	Some of the/Some other	Tomatoes
Subjects	Same-Token	Target	14.97 (9.90)	12.57 (11.94)	12.03 (7.37)
		Distractors	12.21 (5.34)	12.75 (8.27)	12.30 (4.54)
	Different-Token	Target	9.89 (7.12)	13.64 (8.40)	13.64 (10.17)
		Distractors	12.66 (8.04)	13.81 (5.97)	13.99 (6.88)
Items	Same-Token	Target	7.49 (6.89)	6.28 (4.88)	6.02 (5.68)
		Distractors	6.11 (3.11)	6.37 (2.00)	6.15 (2.86)
	Different-Token	Target	4.95 (5.05)	6.82 (4.91)	6.82 (4.91)
		Distractors	6.33 (3.53)	6.91 (2.59)	7.00 (3.34)

Table 6. Proportion of saccades to target and distractor objects from the onset of "some" to the offset of "tomatoes."

			Onset of "Some" to offset of "Tomatoes"
Subjects	Same-Token	Target	24.33 (14.89)
		Distractors	24.51 (10.90)
	Different-Token	Target	26.20 (14.94)
		Distractors	27.09 (10.39)
Items	Same-Token	Target	12.17 (6.96)
		Distractors	12.25 (3.15)
	Different-Token	Target	13.10 (7.65)
		Distractors	13.55 (3.62)

Figure 3. Proportion of fixations to target and distractor objects in the visual scene by condition (binned in 25 ms increments; reset to 0 at the onset of "weigh"). As each sentence differed in length, onset boundaries (marked by vertical lines) were calculated as the average across all sentences.

An examination of saccades from the onset of "some" to the offset of "tomato" shows no main effect of object (by-subjects: $F(1,21) = .311$, $p = .583$; by-items: $F(1,33) = 1.396$, $p = .246$), or condition (by-subjects: $F(1,21) = 1.805$, $p = .583$; by-items: $F(1,33) = .475$, $p = .495$), and no interaction between these variables (by-subjects: $F(1,21) = .037$, $p = .850$; by-items: $F(1,33) = .294$, $p = .591$).

An analysis of fixations at the onset of "some" shows no main effect of object (by-subjects: $F(1,21) = .048$, $p = .829$; by-items: $F(1,33) = .387$, $p = .538$). However, a main effect of condition (by-subjects: $F(1,21) = 5.838$, $p = .025$; by-items: $F(1,33) = 4.222$, $p = .048$) and an interaction between object and condition was observed (by-subjects: $F(1,21) = 5.448$, $p = .030$; by-items: $F(1,33) = 5.447$, $p = .026$). Given that the conditions did not differ by this point in the sentence, and thus reflect an arbitrary grouping of data, no further t-tests were conducted.

An analysis of fixations at the offset of "tomato" shows no main effect of object (by-subjects: $F(1,21) = .013$, $p = .912$; by-items: $F(1,33) = .472$, $p = .497$) or condition (by-subjects: $F(1,21) = 3.054$, $p = .095$; by-items: $F(1,33) = .744$, $p = .395$). No interaction between these variables was shown (by-subjects: $F(1,21) = 1.463$, $p = .240$; by-items: $F(1,33) = .406$, $p = .528$).

Despite no main effects or interaction being shown in saccades from the onset of "some" to the offset of "tomato" an analysis from "tomato" onset to "tomato" offset shows a marginally significant main effect of object (by-subjects: $F(1,21) = .575$, $p = .457$; by-items: $F(1,33) = 4.601$, $p = .039$), with more saccades being directed towards distractor objects compared to the target object. No main effect of condition (by-subjects: $F(1,21) = 1.099$, $p = .306$; by-items: $F(1,33) = 1.197$, $p = .282$) or interaction between

object and condition (by-subjects: $F(1,21) = .018$, $p = .895$; by-items: $F(1,33) = .042$, $p = .838$) was observed. No other sub-region or intervening point showed significant differences.

Discussion

In Experiment 2, the activation of canonical color features during sentence processing in the absence of atypical color attribution was intended to provide a baseline with which to interpret the results of Experiment 1. However, the only significant effect of condition was at the onset of the determiner phrase. Given that by this point in the sentence both conditions are equivalent, this difference must be spurious. Aside from this, there were no significant differences in looks between target object and distractors (not as a function of condition). Further, an examination of saccades during the final noun show that more saccades were launched to distractor objects than the target object, the opposite pattern of what was expected. In fact, a qualitative examination of the data shows that the only point at which targets seem to be fixated more than distractor objects was long after the offset of the critical noun (as seen in Figure 3).

It was predicted that referencing the discourse final noun would prompt activation of canonical features as in Huettig and Altmann (2011). The current results, while not contradictory, lend little support for this, however the paradigm used by Huettig and Altmann differs from the one used here in a few key ways. First is the issue of predictability of stimuli. With reference to "The man has some tomatoes. They are green. He will take out a scale and weigh some of the/some other tomatoes," the stimuli used in this experiment enable one to anticipate the critical noun after "of the" and "other"; all stimuli referred back to the noun mentioned in the first sentence. Consider this sentence

from the Huettig and Altmann stimuli: “The boy thought about it carefully and then he spotted the pea.” Given the type of structures used in their study, all activation associated with color must arise beginning from the mention of the critical noun and no anticipatory activation is possible.

Second, and related to the point above, the sentences used in this study were predictable for the reason that all objects were repeated two times in each sentence context, while they were only presented once in Huettig and Altmann. It’s unclear to what extent this repetition influenced subsequent looks to the target object in the display.

Third, in Huettig and Altmann, their paradigm involved a preview of the quadrant display before any language was presented. The procedure in this study involved the display being shown after some language was presented. This means that there could be pre-existing biases that influenced how the display was viewed even before the critical sentence region.

These differences, along with fact that the color activation effects in Huettig and Altmann were late occurring, could be leading to a further attenuation of activation, resulting in the very late trend that is observed here. The previously mentioned experiment conducted by Connell and Lynott (2009), which showed color priming effects, was conducted by using a semantic Stroop task. The nature of this task means that activation of color features during sentence processing can be late occurring and still produce priming effects in the subsequent Stroop task. Further, Yee et al. (2012) show color priming effects in language processing *only* when participants took part in a Stroop task prior – that is, when they were already focused on color as a relevant feature. Based on these findings, although unexpected, there may be some explanation for why no

difference was observed here. However, the fact that we observed no difference in looks to the target object on hearing the sentence-final reference makes the interpretation of Experiment 1 problematic. Was the marginal effect observed there spurious? Given the previously mentioned research, one would not expect to find an effect in Experiment 1 without also finding an effect in Experiment 2.

General Discussion

Two experiments explored the activation of semantic features of objects in sentential contexts. In Experiment 1, sentences prompted manipulation of the episodic features of these objects and then referred back to these changed objects or to different tokens of the same object type. In Experiment 2, sentence contexts involved instantiating objects in an episodic context (although their canonical properties were unchanged, unlike in Experiment 1), and these objects or different tokens of these objects, were again referenced at the end of the sentence.

The results of Experiment 1 show that, in both conditions, marginally more saccades were launched to objects in a visual display that shared surface color with the canonical form of the object mentioned in the sentence (compared to unrelated distractor objects) from the onset of the determiner phrase preceding the noun to the offset of the noun ("He will weigh some of the/some other tomatoes"). In the discussion of Experiment 1, it was suggested that this supported the theoretical account of event processing from Altmann (2017), which makes the prediction that objects that are altered over time carry with them qualities of their previous states. It also suggests that the instantiation of new object tokens prompts activation of the canonical feature set associated with that object.

In Experiment 2, which was intended to provide a baseline for the interpretation of Experiment 1, no increase in looks to the object that shared surface color with the canonical color of the noun in the discourse, compared to distractor objects, was observed from the onset of the determiner to the offset of the noun as it was in Experiment 1. This may have arisen from the nature of our task, which attempted to manipulate and probe an already subtle activation effect. Given the results of Experiment 2, the interpretation of Experiment 1 is greatly complicated, because they suggest that the paradigm is not sensitive to color overlap in the way that was originally noted by Huettig & Altmann (2011).

Generally, the results presented here are suggestive in that Experiment 1 supports (albeit weakly) the theoretical position that objects carry their histories with them. In order to more strongly support these claims, further work would need to be done, primarily in two areas. First, although previous evidence suggests that sentence contexts can produce color activation effects, the nature of these effects is somewhat small and late occurring. Further work would benefit the current account by using a similar paradigm but probing a more robust effect, such as shape priming (mentioning a "rubber band" in discourse prompting attention to an object with a similar form; Huettig & Altmann, 2004). For example, the sentence "John saw a rubber band on the ground. It was snapped. He will pick up the rubber band" could be used with a visual display in which there was a bracelet. Second, an accurate baseline would need to be attained. A fundamental quality of the stimuli used here is that they rely on referring to objects previously mentioned in the discourse, in contrast to other studies of conceptual priming mentioned here in which objects are only mentioned once. This leads to predictability of

upcoming language and a potential issue with repeated naming effects. In short, it's unclear to what extent these issues influence the priming effects that we are intending to examine here, although this knowledge is crucial in interpreting the results of primary importance.

References

Altmann, G. T. M. (2017). Abstraction and generalization in statistical learning: implications for the relationship between semantic types and episodic tokens. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160060. https://doi.org/10.1098/rstb.2016.0060

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*(1), 55–71. https://doi.org/10.1016/j.cognition.2008.12.005

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioural and Brain Sciences*, *22*, 577–660.

Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review*, *16*(3), 573–577. https://doi.org/10.3758/PBR.16.3.573

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? *Journal of Neuroscience*, *32*(17), 5795–5803. https://doi.org/10.1523/JNEUROSCI.6294-11.2012

Huettig, F., & Altmann, G. T. M. (2004). The online processing of ambiguous and unambiguous words in context: Evidence from head-mounted eye-tracking. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond* (pp. 187–2007). New York: Psychology Press.

Huettig, F., & Altmann, G. T. M. (2011). Looking at anything that is green when hearing "frog": How object surface colour and stored object colour knowledge influence language-mediated overt attention. *The Quarterly Journal of Experimental Psychology*, *64*(1), 122–145. https://doi.org/10.1080/17470218.2010.481474

Kukona, A., Altmann, G. T. M., & Kamide, Y. (2014). Knowing what, where, and when: Event comprehension in language processing. *Cognition*, *133*(1), 25–31. https://doi.org/10.1016/j.cognition.2014.05.011

Solomon, S. H., Hindy, N. C., Altmann, G. T. M., & Thompson-Schill, S. L. (2015). Competition between Mutually Exclusive Object States in Event Comprehension. *Journal of Cognitive Neuroscience*, *27*(12), 2324–2338. https://doi.org/10.1162/jocn_a_00866

Yee, E., Ahmed, S. Z., & Thompson-Schill, S. L. (2012). Colorless Green Ideas (Can) Prime Furiously. *Psychological Science*, *23*(4), 364–369. https://doi.org/10.1177/0956797611430691

Zwaan, R. A., & Radvansky, G. A. (1998). Situation Models in Language Comprehension and Memory. *Psychological Bulletin*, *123*(2), 162–185.