



Published in final edited form as:

Lang Cogn Neurosci. 2017 ; 32(2): 190–204. doi:10.1080/23273798.2016.1238494.

The amount and structure of prior event experience affects anticipatory sentence interpretation

Arielle Borovsky, Ph.D.¹

¹Department of Psychology, Florida State University, Tallahassee, FL

Abstract

Listeners easily interpret speech about novel events in everyday conversation; however, much of research on mechanisms of spoken language comprehension, by design, capitalises on event knowledge that is familiar to most listeners. This paper explores how listeners generalise from previous experience during incremental processing of novel spoken sentences. In two studies, participants initially heard stories that conveyed novel event mappings between agents, actions and objects, and their ability to interpret a novel, related event in real-time was measured via eye-tracking. A single exposure to a novel event was not sufficient to support generalisation in real-time sentence processing. When each story event was repeated with either the same agent or a different, related agent, listeners generalised in the repetition condition, but not in the multiple agent condition. These findings shed light on the conditions under which listeners leverage prior event experience while interpreting novel linguistic signals in everyday speech.

Keywords

prediction; sentence processing; event-learning; linguistic generalisation; eye-tracking

Everyday language is peppered with utterances that we have never heard before and may never encounter again (Chomsky, 1965; Hockett, 1960). At face value, the limitless creative potential of language coupled with the rapid and fleeting nature of the speech signal should pose significant barriers for comprehension. Despite these challenges, language users interpret speech in novel and familiar circumstances with incredible speed and fluency. An abundance of research now suggests that listeners comprehend spoken language by incrementally decoding a wide array of linguistic and non-linguistic cues in real-time, as soon as they become available in the unfurling speech signal, and by predictively generating expectations about upcoming referents. (Altmann & Mirkovi , 2009; DeLong, Troyer, & Kutas, 2014). The majority of prior work on predictive language processing, however, has explored conditions where the relevant referents and events are relatively familiar to the listener. These highly familiar circumstances may not in fact be highly representative of the majority of language comprehension contexts. Indeed, pragmatic conventions of language suggest that much of spoken language does not simply relay information that is mutually understood and highly familiar to all conversational participants (Grice, 1975). Rather, it

may be more pragmatically appropriate to convey novel information about situations where the relationships among event participants have not been previously observed. Little is known about whether and how anticipatory language processes are involved in these novel circumstances.

While the current language processing literature often poses that prediction may be a central and fundamental mechanism in language processing (DeLong, et al., 2014; Pickering & Garrod, 2013), recent proposals have also called into question whether prediction is necessary for language comprehension (Huettig & Mani, 2015). The use of prediction in language processing is especially relevant when considered with respect to language learning processes (Chang, Kidd, & Rowland, 2013; Rabagliati, Gambi, & Pickering, 2015). Consider a learner who has no prior experience with a novel event. What role, if any, does prediction play in language that surrounds a novel context? At some level, it is expected that a learner must have some basic knowledge to anticipate appropriate outcomes. However, the current empirical and theoretical accounts of prediction in language have little to say concerning what the learning conditions that support subsequent predictive processes may be.

Some prior work has established that learners can anticipate and activate recently trained combinatorial relationships during incremental interpretation of written (Amato & MacDonald, 2010) and spoken language (Borovsky, Sweeney, Elman, & Fernald, 2014). For example, Borovsky and colleagues (2014) presented adults and children with stories depicting novel (cartoon) relationships between agents, actions and objects (e.g. a monkey that rides in a car and a dog that rides in a bus). After hearing these stories, participants completed an eye-tracked sentence comprehension task, where the objects from the stories reappeared on a computer screen (e.g. car, bus and other story objects). They found that when adults and school-aged children (aged 5 to 10) heard sentences depicting combinatorial relations from the unique events from the previous story (*The monkey rides in the...*), participants quickly generated predictive fixations towards the appropriate thematic object (CAR). This result indicated that listeners could use novel “fast-mapped” higher-order contingencies among agents, actions and objects to generate expectancies in real-time speech comprehension, paralleling findings from studies of sentence processing in highly familiar contexts (Borovsky, Elman, & Fernald, 2012; Kamide, Altmann, & Haywood, 2003).

Can listeners extend their knowledge from a similar situation to generate real-time referential predictions in a novel one? For example, after hearing a story about a monkey who rides in a car, as described above, would a listener then generalise this knowledge to generate real-time linguistic predictions about similar agents who might then participate in a similar event (e.g. by expecting a gorilla to ride in a car)? Although prior research has demonstrated that learners can predictively activate recently trained event knowledge to a separate instance of an identical situation, we might not expect listeners to generate “promiscuous” extensions of novel event knowledge after minimal training. Instead, a sensible strategy may be a conservative one, where listeners would require additional evidence before extending their linguistic predictions to novel situations. A central question in this case is how the structure of novel event experiences may influence generalisation.

These questions highlight the focus of the current investigation. The current studies examine the conditions where adult listeners extend previously learned event relations to novel agents during real-time language comprehension using a visual world eye-tracking task.

1.1 Using visual world paradigms to explore real-time language interpretation

In the past two decades, there has been an explosion of research that has used the visual world paradigm to explore mechanisms of real-time language comprehension (Huettig, Rommers, & Meyer, 2011; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). This technique, which measures eye-movements towards visual referents in response to spoken language, has revealed that listeners attend to a host of linguistic and non-linguistic cues while actively and dynamically generating predictions about unspoken referents.

One finding that is of particular relevance for the current investigation is that listeners can generate appropriate predictions for upcoming spoken words by combining cues that extend across multiple lexical items (Borovsky, et al., 2012; Kamide, et al., 2003; Kukona, Cho, Magnuson, & Tabor, 2014). For example, Borovsky, Elman and Fernald (2012) demonstrated this phenomenon in an experiment where participants listened to sentences containing an informative agent and action (*The pirate chases the...*) while viewing a visual array that contained items related to both the agent and action (SHIP), only the agent (TREASURE), only the action (CAT), and unrelated to both cues. They found that adults and children generated anticipatory fixations towards the Target item (SHIP) before it was spoken. This finding is taken as evidence that listeners can rapidly activate well-known event relationships among agents, actions and objects in spoken language interpretation. An additional study (described above; Borovsky, et al., 2014), found that recently acquired event knowledge could also be interpreted in a similar predictive fashion in real-time. However, prior work on combinatorial linguistic processing using this method has only explored cases where the knowledge was already well established, or previously trained. No work has yet explored whether and how similar predictive effects emerge when listeners interpret entirely novel linguistic combinations on the fly. This question is explored in the current investigation using a visual world eye-tracking paradigm.

1.2 Event knowledge activation and generalisation in language processing and acquisition

There is converging evidence from a variety of experimental techniques that language comprehension involves the real-time activation of prior event knowledge (Bicknell et al., 2010; Metusalem et al., 2012). This ability requires listeners to dynamically activate meanings that situate the linguistic message within the broader world (Zwaan & Radvansky, 1998). The evidence suggests that language comprehension involves the real-time activation of a broad web of knowledge, and requires that listeners activate not only the immediately mentioned referential context, but also associated locations, object, instruments, patients and agents (McRae & Matsuki, 2009). By extension, these findings suggest that listeners are

similarly encoding this information when they initially learn about novel events and the language that accompanies them.

However, when encountering an entirely novel sentence, no research has explored whether and how listeners activate referential predictions by generalizing from prior (similar) experiences. Experiment 1 addresses this question by asking whether listeners not only generate predictions for events involving repeated agents (replicating prior work) but, most importantly, also asks whether this case can be extended to novel events performed by similar (semantically-related) agents. For example, if a listener has previously learned about a novel event such as, *an ant wearing sunglasses*, and, *a rabbit wearing a hat*, then one might expect the listener to also generate a prediction for sunglasses (vs. hat) when they hear an agent who shares some similarity to the ant perform the same action (e.g. *The ladybug wears the...*).

While it was expected that this investigation would replicate prior findings that listeners can use the previous experience to interpret the same event predictively (e.g. looking to the sunglasses when hearing *The ant wears the...*), there was some uncertainty with respect to whether participants would generalise this information to a novel agent after a single experience. In either case (generalisation or no), it is likely that additional experience with events would facilitate novel extensions to new agents. In this case, the structure of this experience may be crucial.

For example, if one has encountered multiple instantiations of the same event with a variety of similar agents (e.g. a bus riding event completed by a number of zoo animals, like a monkey and zebra) then this may facilitate extension to a novel, similar agent (e.g. a lion). Such a result would be consistent with findings that variable learning conditions support linguistic generalisation (Gerken & Bolt, 2008; Gómez & Maye, 2005; Gómez, 2002; Wonnacott, Boyd, Thomson, & Goldberg, 2012). Alternatively, repeated experience with the same event may reinforce encoding of the event representation, potentially facilitating generalisation to other situations. This finding would be consistent with reports that having a highly frequent or prototypical exemplar may support linguistic generalisation of argument structures (Goldberg, Casenhiser, & Sethuraman, 2004). These alternate possibilities are explored in Experiment 2, by exposing learners to multiple instances of an event that is either simply repeated with the same agent, or repeated with multiple, related agents.

Methods

Participants

Forty-one college student participants (35 F, 6 M) at Florida State University participated in return for course credit. Participants reported that they had normal or corrected-to-normal hearing, no prior diagnosis or treatment for any cognitive, attentional, speech or language disorders, and exposure to English as their primary language in childhood. An additional 8 students participated but were not included in the analyses for the following reasons: 6 reported extensive exposure to a language other than English in childhood, one reported a hearing issue, and 2 other participants were excluded due to experimental error that resulted in a failure to collect eye-tracking data.

Materials

Story materials—Following the procedure from Borovsky et al., 2014, all participants heard eight stories that consisted of novel events that established relationships among two novel agents, two actions, and four thematic objects. Stories were depicted in colourful cartoon images (400 ×400 pixels in size) that were created using a comic strip creation website, Toondoo (<http://www.toondoo.com>). Story images were accompanied by narrations that reinforced the visually depicted events in the story. These narrations were recorded by a native-English speaking female at 44.1 kHz on a mono channel and normalised to mean 70 dB intensity offline.

Each story followed the identical narrative structure established in prior work by Borovsky and colleagues (2014). Each story included two agents, who performed the same two actions on two different objects (see Figure 1). The event sequence was presented in identical order across stories. First, the narrator introduced one of the agents (e.g. RABBIT, *Hey there's a rabbit*), and the event relationships of the story are established as the agent performs different actions (e.g. wearing a hat, eating cake) across the story. Next, a second agent is introduced (e.g. ANT, *Hey there's an Ant!*) and this agent then performed identical actions with different objects (e.g. wearing sunglasses, eating ice cream). All Agent-Action-Theme relationships in stories are listed in Appendix A.

This story structure presented four novel event relationships among the depicted agents, actions and objects in the story. In the above example, these pairings are: rabbit – wears – hat, rabbit – eats – cake, ant – wears – sunglasses, ant – eats – ice cream. The primary goal was to create very low frequency event relationships that were unlikely to have been previously experienced by the study participants. Additional counterbalancing of the pairings of agents and themes across versions was done to control for the possibility that participants may have had particular biases regarding the likelihood of various combinations among the event constituents. For example, in one version, the rabbit may have worn a hat and eaten cake, in another, it would have worn sunglasses and eaten ice cream.

It is also important to note that the verbal narrative description of these events never mentioned these elements of the event within the same sentence. That is, for the rabbit – wears – hat relation, the narrator named the agent as a rabbit in one piece of the story, and then later mentioned that *He's wearing a hat*, but never explicitly stated, *The rabbit is wearing a hat*. The goal of this structure was to test recognition and extension of this relationship without the participant previously hearing this same acoustic stimulus in the sentence recognition portion of the study. As every event relation was equally likely to occur with the agents across versions, this control further ensured that the identical acoustic signal conveyed the same critical event relationship across versions, irrespective of the agent who participated in the event.

Agent selection—The main questions of this study concern how participants extend the event information from the initial stories to novel agents participating in the same events during real-time language comprehension. Therefore, the inter-relationships among the agents within the stories were carefully controlled such that the two story agents were distinct, and that each extension agents was relatively similar to one of the story agents.

First, to keep the agents within each story maximally distinct, all story agents corresponded to one mammal and one non-mammal category. Mammals and non-mammals were drawn from four subcategories. Mammal categories were: woodland creatures, zoo animals, pets, and farm animals. Non-mammal categories were: bugs, birds, reptiles and sea creatures. Generalisation agents were selected from the same subcategories as those that appeared in the story. For instance, in the story example illustrated in Figure 1 with the rabbit and ant, the items that were selected for generalisation were mouse and ladybug, respectively. Appendix B lists all training and extension agents along with their associated category.

Agent generalisation norming—The primary interest in this study was whether and how participants would extend their interpretation of the inter-relationships among the agents in the stories to other, similar agents during sentence recognition. Therefore, a norming task was carried out to ensure that participants would find the relationships between the story and extension agents to be more similar than that of the story and distractor agents. Seventy-one adult participants from Amazon’s Mechanical Turk were presented with two story agents (e.g. RABBIT and ANT) and asked to select which item was most similar to the two possible associated test sentences in the generalise condition (e.g. *mouse* or *ladybug*). Because each participant rated half of all potential combinations, the size of the norming sample ensured that at least 35 participants provided ratings for each item. The ratings indicated that participants agreed that the story agents were highly similar to the extension agents and not similar to the distractor agents. Participants matched the extension agent to the appropriate story agent of the same category 97.5% of the time overall, with the ratings ranging from 77.8% – 100% for each individual item.

Sentence recognition stimuli—When each story was completed, the participants completed a visual-world sentence comprehension task that was designed to measure the real-time recognition and extension of the event relationships that had been previously mentioned in the story. Participants viewed four thematic objects from the prior story in isolation on a screen as they heard a sentence that described one of the depicted items. Sentences corresponded to one of two conditions: (1) Repeat condition, where the agent in the sentence was identical to the one that had appeared in the story (ANT -> ANT), (2) Generalise condition, where the agent in the sentence was similar, but not identical to the story agent (e.g. ANT -> LADYBUG). As described above, the generalisation agents were previously normed for similarity. Each sentence was constructed with a standard Article-Agent-Action-Article-Theme structure and normalised to be identical in duration and intensity. For example, in the story and test sentence example, illustrated in Figures 1 and 2, participants heard a sentence that repeated either the prior relationship that they encountered in the story in the repeat condition (e.g. *The ant wears the sunglasses*), or, they would have heard a sentence containing a similar agent from the same animal subcategory performing the same action (e.g. *The ladybug wears the sunglasses*).

The accompanying images corresponded to the thematic objects that initially appeared in the prior story block (e.g. CAKE, ICE CREAM, SUNGLASSES, HAT), and each image was assigned to one of four target or distractor conditions: Target, Agent-Related, Action-Related, or Unrelated. In the repeat condition, these relationships were assigned according to

the information that was presented in the prior story. In the example depicted in Figure 1 and 2, for the sentence, *The ant wears the sunglasses*, the Target object was the SUNGLASSES, the Agent-Related item the CAKE, the Action-related item was the HAT, and the Unrelated item, the ICE CREAM. In the generalise condition (e.g. when the ladybug was mentioned as the agent of the sentence *The ladybug wears the hat* rather than the ant), the same items were assigned to the corresponding Target and Distractor images as in the Repeat condition (see Figure 2). Eight sentences were associated with each image quartet (four each in the Repeat and Generalise conditions). This arrangement therefore allowed each image to serve in all Target and Distractor roles across all conditions equally, yielding a completely balanced, within-subjects design.

Visual images of the thematic objects were isolated from their original cartoon depiction in the story and placed on a white background in a 400×400 pixel square. Auditory stimuli were recorded by the same native English-speaking female as for the stories, in a child-directed voice. All spoken sentences were edited in Praat software (Boersma & Weenink, 2012) and were normalised to a standard mean intensity of 70 dB. The onsets and durations of all words were additionally aligned in Praat. This alignment procedure ensured the fixation time course across all experimental items was identical with respect to the auditory stimulus.

Procedure

Experimental task—Participants sat a stationary armchair in front of a 17” LCD monitor with an attached eye-tracking video camera (SR Research Eyelink 1000+ eye tracker). Participants initially read instructions for the task, which indicated they would hear a short story, which they should attend to and do their best to understand. Participants were then instructed that, after each story, they would hear some additional sentences accompanied by four images. They were asked to use the mouse to select the image that “goes with the spoken sentence.” Participants then completed a single practice trial before completing a five-point calibration routine with a standard black-and-white 20-pixel bull’s-eye image. After the calibration was complete, the experimental task began.

For each block in the study, participants initially heard a short story that was intended to establish novel relationships among agents, actions and objects. These spoken stories were accompanied by colourful cartoon images that supported the spoken story descriptions. Each image initially appeared on the screen for 2000 ms before the start of the narration for that image. The image remained on the screen until the participant clicked the mouse to indicate that they were ready to hear the next frame in the story.

After the story was complete, the participants then completed an eye-tracked sentence comprehension task. Each trial in this task began with the presentation of a centrally located 20-point black and white bull’s-eye image (identical to the image presented during the calibration procedure). This image served as a drift-check prior to the onset of the trial. If it was discovered that the calibration had drifted, the procedure was stopped for re-calibration, though this was not often necessary. Once the participant completed the drift-check procedure, four objects from the preceding stories reappeared on the screen. The images appeared for 2000 ms in silence before the onset of the spoken sentence. Eye-movements

towards each of the objects were recorded in response to the spoken sentence. The images remained on the screen until the participant clicked on an image after the sentence was spoken. Participants were offered a break halfway through the experimental task, which lasted approximately 10 minutes.

Eye-movement recording—Eye-movements were recorded using a SR Research Eyelink 1000+ eye-tracker in remote mode at 500 Hz sampling rate. Eye-movements were automatically classified as blinks, fixations and saccades using the tracker's default settings. Data were binned into 50 ms samples offline for further analysis.

Approach to analysis—We analyzed the dataset comprising eye-movements in response to the spoken sentence. The primary goal was to determine whether participants' eye-movements indicated anticipatory fixations towards the sentence-final thematic object. As in prior work (e.g., Borovsky, et al., 2012; Kamide, et al., 2003) fixations towards the Target item that exceed each of the other three distractor interest areas (Agent-Related, Action-Related, Unrelated) before the onset of the Target were taken to as strong evidence for prediction of the final item. While this target divergence criteria itself may seem logically straightforward, there has been some variability in the field regarding the appropriate metric of target divergence. In some work, researchers have directly compared fixations towards the Target vs. other competitors over relatively broad time windows, such as windows that span entire spoken words within the sentence (e.g., Kamide, et al., 2003). This broad time window approach is advantageous in that it is tied specifically to events that occur across the sentence that span often several hundred ms. However, the coarse-grained approach sacrifices fine-grained precision that is available with this methodology, and can mask more rapidly changing patterns that occur within the span of a single word. In an attempt to capture these real-time dynamics within a sentence, we implemented a non-parametric statistical approach that can identify fine-grained time windows of Target divergence while controlling for Type I error (see Groppe, Urbach, & Kutas, 2011; Maris & Oostenveld, 2007 for more statistical details about this procedure). This procedure has been recently adopted by a number of groups (Barr, Jackson, & Phillips, 2014; Borovsky, Ellis, Evans, & Elman, in press; Von Holzen & Mani, 2012). Specifically, a cluster-based permutation analysis is used to identify precise time windows across the sentence period where fixations towards the Target significantly exceed fixations towards each of the visual competitor items, individually. Target fixations are taken to be anticipatory when this procedure identifies a time window that begins before the onset of the sentence-final object where fixations towards the Target exceed fixations towards each of the other items, individually.

Results

Behavioral accuracy

To ensure that participants were attending to and understanding the sentences, participants were asked to select the target picture that corresponds to the spoken sentence. Participants completed this task with a high degree of accuracy (98.6%), missing only 9 out of 656 trials. Trials with incorrect responses were removed from subsequent analyses.

Time course data and analyses

The time course of fixations towards the Target and distractor items in the Repeat and Generalise conditions is illustrated in Figure 3. These plots illustrate apparent differences between the Repeat and Generalise condition with respect to anticipation of the Target object. In the repeat condition, participants appeared to view target item before the onset of the spoken object. These fixations to the target suggest that participants rapidly encoded information from the initial story and predictive re-activated this information during a real-time sentence processing task, replicating findings from prior work (Borovsky, et al., 2014). On the other hand, fixations towards the Target in the Generalise condition did not appear to diverge from the other distractors until the sentence-final object is spoken. This pattern would appear to suggest that participants did not use information from the initial story to interpret a novel sentence predictively.

Cluster-based permutation analyses were used to investigate whether there is statistical evidence for anticipatory fixations in the Repeat and Generalise condition, separately. The timing of differences between the Target and distractors in each condition is illustrated in Table 1. In the repeat condition, fixations to the Target exceeded fixations to all other distractors starting from 1300 ms and continued to the end of the sentence. Importantly, this divergence in looks to the Target vs. other distractors occurred well before the onset of the sentence-final noun at 2036 ms post sentence onset. This pattern contrasts with the Generalise condition, where the fixations to the Target vs. the Action-related distractor do not diverge until the 2250 ms time bin, approximately 200 ms after the onset of the sentence final noun. The findings suggest that participants generated real-time predictions for the appropriate target in the Repeat condition, whereas gaze patterns in the Generalise condition did not reflect anticipatory processing.

Discussion

There are two important findings from experiment 1. First, the results in the repeat condition indicated that participants predictively interpreted novel event information that was directly repeated from an immediately preceding story context, replicating prior work (Borovsky et al., 2014). Despite this finding, the same participants did not show robust real-time generalisation of this novel relationship to a new, but similar agent.

One potential explanation for these findings is that the listeners could not encode or retrieve the original event information in the generalise condition. This possibility seems unlikely because the same participants did engage in predictive processing of these event relations in the condition where the event information was simply repeated. An ability to encode and retain the prior mapping has also been demonstrated in prior study with adults and children using similarly structured stories (Borovsky et al, 2014). Therefore, it seems unlikely that an encoding or retrieval failure can fully account for these findings. Instead, the results suggest that the novel event representation may not yet support generalisation after a single experience.

Instead, listeners may require additional experience to support the extension of recently established event knowledge during real-time language comprehension. The need for

repetition has some commonsense appeal; it would not necessarily be desirable for a single odd or “one-off” experience to dramatically alter the interpretation of other related events. Listeners are likely to require additional evidence to facilitate extension of recently acquired event knowledge to novel agents in real-time language processing. Prior work on grammatical construction learning reveals that adults and young children can generalise relatively abstract syntactic patterns to novel forms after only a few exposures (e.g., Gomez & Gerken, 1999; Kaschak, 2006; Wonnacott, et al., 2012). Simultaneously, the distribution of the input is likely to influence generalisation, though there is contradictory evidence about what kinds of distributions may best support extension. While some work suggests that increased variability in the input may support linguistic generalisation (Gerken & Bolt, 2008; Gómez & Maye, 2005; Gómez, 2002; Wonnacott, et al., 2012), other accounts suggest that a repeated prototypical exemplar facilitates generalisation of argument structures (Goldberg, et al., 2004). These alternative accounts generate different predictions regarding how additional novel event experience may support generalisation within the current experimental paradigm. Namely, the variability account suggests that listeners may need additional experience with related, but non-identical event situations to support generalisation. On the other hand, the prototype account suggests that more experience with a single example should facilitate subsequent generalisation. The next study directly tests these accounts by augmenting the novel event input in two ways: (1) By simply repeating the same novel event (i.e. increasing the frequency using a single exemplar) and (2) By adding a second agent to each event who is similar to the initial agent (i.e. increasing the variability). If it is the case that listeners simply need additional exposure, and that a more frequent single exemplar supports extension of the event input in real-time language comprehension, then we should expect to see greater anticipatory looks towards the target in the first, Repeated Agent condition, but not the latter, Multiple Agent, condition. If variability supports generalisation, then we would expect instead to see the opposite pattern.

Experiment 2

Experiment 1 did not provide strong evidence for generalisation from a single experience, thereby suggesting that learners need additional exposure to support this process. Experiment 2 was therefore designed to explore how the structure of a single additional experience with a particular event might influence generalisation. This experiment is similar to Experiment 1, except that listeners hear an additional story sequence that either (1) simply repeats the event with the same agents (Repeated Agent), or (2) repeats the same event with similar, but novel agents from the same animal subcategory. For example, after learning about an event where a rabbit wears a hat, the listener would either, in the Repeated Agent condition, hear the same story repeated, or hear a new story where a similar agent (e.g. another woodland creature, a hamster) carries out the same event (wearing a hat). In both cases, the eye-tracked sentence comprehension task would precede in an identical fashion as in Experiment 1.

The goal of manipulating these learning conditions was to allow for experimental testing of novel event extension in real-time language learning when participants had the opportunity to increase exposure to the event either by boosting the frequency of the event in the

Repeated Agent condition or increasing the variability of the event in the Multiple Agent condition.

If we find evidence for generalisation in the Repeated Agent condition (as measured by evidence of anticipatory processing in an eye-tracked auditory sentence comprehension task), then this finding would support exemplar learning accounts of language, where generalisation is facilitated when there are highly frequent prototypical exemplars. Alternatively, prediction in the Multiple Agent condition would be more consistent with variability accounts of language learning.

Method

Participants—Forty college student participants (31F, 9 M) from Florida State University participated in exchange for course credit. All participants reported English to be their primary and native language and an absence of hearing, vision, language, speech or learning issues. An additional nine students participated but were excluded from the analysis for the following reasons: eight were non-native speakers of English, and an experimenter error resulted in the loss of one participant's eye-tracking data.

Materials

Story materials: The same auditory and visual stimuli from Experiment 1 were used, along with new story material depicting the same events with different (related) agents in the Multiple Agent condition. In the new stories that were added to create this condition, new agents were selected who performed the same actions with the same objects as in Experiment 1. These novel agents were selected to correspond to the same animal subcategory as in Experiment 1. For instance, if a story agent in Experiment 1 was a bug (e.g. an ant who wears sunglasses and eats ice cream), then another bug who performed the same actions on identical objects was also included (e.g. a butterfly who wears sunglasses and eats ice cream).

Sentence materials: The auditory and visual materials were identical to those used in Experiment 2, except that a new set of stories with two new agents performing identical actions with the same objects as in Experiment 1. The agents were selected to belong to the same animal subcategory as in experiment 1. This arrangement yielded stories that contained two animals from two different subcategories. As before, story agents were selected to be both maximally distinct between categories and similar within categories. Therefore two agents from two subcategories were selected from one mammal subcategory (either woodland creatures, zoo animals, or farm animals), and two agents were selected from a non-mammal category (either insects, reptiles, or sea creatures). The structure of the agents in the stories is outlined in Appendix B.

Sentence recognition materials: The eye-tracked sentence comprehension task in Experiment 2 included the generalisation sentence materials from Experiment 1 (See Figure 2).

Procedure

Experimental task: The procedure was identical to that of Experiment 1, except that each event training phase was extended to now include either an additional repetition of each agent (Repeated Agent condition) or a repetition of a similar, but not identical agent (Multiple Agent condition).

Eye-tracking procedures and Data Analysis: Details regarding eye movement recording and data analysis are identical to those of Experiment 1.

Results

Behavioral Accuracy—Participants were asked to select the target picture that corresponds to the sentence to ensure that they understood and attended to the task. As in Experiment 1, participants were highly accurate on this task (98.1%), missing only 12 out of 640 trials. Trials with incorrect responses were not included in subsequent analyses.

Time course data and analyses—Figure 4 plots fixation proportions towards the Target and distractor items across the time course of the spoken sentence in the Repeat Agent and Multiple Agent conditions. Visual inspection of these plots suggests that there are differences between the Repeat and Multiple agent condition with respect to anticipation of the sentence-final Target object. In the Repeat condition, fixations towards the target item appear to diverge from looks to all other items before the onset of the Target item, whereas in the Multiple Agent condition, fixations towards the Target do not diverge from those to other distractors until after the onset of the spoken object. This illustrated difference between conditions seems to support the possibility that participants were able to generalise their prior event experience in real-time to interpret a novel event containing a new agent, but only in the repeated agent condition. Next, this visually-apparent pattern is statistically analyzed using cluster permutation analyses as in experiment one.

Non-parametric cluster analyses of the time course were used to investigate whether there was statistical evidence for anticipatory fixations for the Repeated Agent and Multiple Agent conditions individually. The timing of the differences between the Target and distractors in each condition is illustrated in Table 2. This analysis indicated that, in the Repeated Agent condition, fixations to the Target diverged from all other distractors at 1900 ms, before the onset of the sentence-final object at 2036 ms. However, in the Multiple Agent condition, the fixations to the Target vs. the Action-related distractor do not significantly diverge until the 2100 ms time bin, after the onset of the sentence-final noun. These analyses indicate that participants did anticipate the sentence-final Target in the Repeated Agent condition, but did not show evidence for anticipatory fixations towards the Target in the Multiple Agent condition.

General Discussion

This research was motivated by two simple observations. First, many utterances in everyday language are novel and describe events that are unfamiliar to the listener, yet are nevertheless easily understood. Second, the majority of research on psycholinguistic mechanisms of

speech comprehension focuses on the processing of familiar events. Much of this prior research has characterised speech comprehension as a rapid, incremental and dynamic process where listeners actively generate predictions about upcoming referents as language unfolds (Pickering & Garrod, 2013). Little work had explored whether and how predictive mechanisms of language interpretation operate in speech that describes novel situations. The current work represents an initial foray into this topic, by probing how listeners might leverage their prior experience when interpreting spoken language about a novel event. Broadly, the findings suggest that incremental linguistic interpretation is sensitive to the amount and structure of the listener's prior experience.

The first study sought to replicate and extend prior work in this area. We measured real-time sentence interpretation in two conditions: (1) Where the event information underlying the sentence had been recently learned (in a single exposure), and (2) Where sentence consisted of information that required the listener to generalise from a recently learned event-mapping. The first condition had been shown in a prior study (Borovsky, et al., 2014) to support real-time predictive processing in adults and school-aged children, even though the event had only been experienced a single time. Our results replicate these prior findings: listeners were quickly able to re-activate and predict the appropriate coordinating thematic object that cohered with their prior event experience. This result alone suggests that the language system is exquisitely capable of immediately integrating, and activating recently acquired event information in real-time spoken language processing. On the other hand, in the generalisation condition, listeners did not show strong evidence that they extended this novel event information in real-time to a similar, but not identical situation performed by a related agent. Together these findings demonstrate that there are important limitations to which a single event experience might influence language processing in related situations. In general, this result suggests that real-time language processing system is at once sensitive to the changing event dynamics that coordinate with an unfolding sentence context and is also conservative in the degree to which it will initially extend from the listener's evolving experience. This behavior seems highly optimal: the listener is simultaneously able to adjust their expectations in response to new experiences, and is limited in the degree to which a single, odd experience may extend to other situations. The listener is likely to need additional exposure to a particular event structure before this information can support generalisation.

Therefore, the second study explored whether and how additional experience with the initial event might influence real-time processing in a related event. Prior theoretical accounts of the mechanisms underlying abstraction in grammar learning suggested that the distribution of experience should matter – but in contrasting ways. While some accounts proposed that variability in experience should support generalisation (e.g. Gomez, 2002), others suggested that exposure to a frequently repeated prototype is optimal (Goldberg, 2006). The findings in experiment 2 are most consistent with this latter account: listeners generalised during a language-processing task when events were previously encountered twice with the same agent, but did not generalise in a condition with different agents. At face value, this finding may seem surprising. One might expect that hearing a second related agent might pragmatically indicate that many different agents could participate in the same event. Instead, these findings indicate that, at least in the earliest stages of learning about event

structure, listeners must first reinforce their initial representation via repetition to support extension.

One reason for this pattern is that event mappings were only encountered a couple of times in the study, and testing occurred after a short delay. The events in the training paradigm were only presented one (Experiment 1) or two times (Experiment 2). Although a bevy of findings in the word learning literature suggest that learners can integrate much about novel mappings after even only a single exposure, a fuller representation of words and events is likely to emerge after more extended training (Carey & Bartlett, 1978). It is possible that event variability and repetition would differentially support generalisation in more extensive training conditions. Indeed, the majority of work that explores generalisation of linguistic structure uses more extensive training examples compared to the relatively limited training examples in the current experiment. Future work will be necessary to explore how the distribution of the input over extended learning affects real-time mechanisms of language processing.

The connection between semantic memory and real-time generalization is another key factor to consider in these experiments. Generalization agents and training agents belonged to the same super-ordinate category domain, such as zoo creatures and birds. This design therefore relied on the listener's pre-existing category knowledge to license some generalizations between agents but not others. This experimental structure raises a host of questions regarding how listeners leverage their categorical knowledge to support generalization during real-time linguistic processing. Prior research indicates that listeners are sensitive to semantic structure in the lexicon as they understand individual spoken words in the visual world paradigm (Huettig & Altmann, 2005; Yee & Sedivy, 2006). Some research suggests that even very young children may be sensitive to the pre-existing semantic structure to support processing of known relations between individual words, as well as between verbs and objects (Borovsky, Ellis, Evans, Elman, 2016). Adult listeners also probabilistically activate an array of sentential outcomes that vary in semantic distance from a highly expected object in a familiar event context (e.g. Federmeier & Kutas, 1999). The current experiments add to this literature by suggesting that category knowledge can also support real-time (pre)activation of sentential outcomes for relatively unfamiliar/novel events by allowing the listener to recognize semantic correspondences between a novel agent, and one who has previously engaged in an event.

More generally, these findings suggest that the cognitive mechanisms that support categorization and semantic memory are crucial for real-time linguistic processing, and may be particularly important in supporting generalization in real-time. Identifying the learning conditions and mechanisms that support category groupings and generalization among similar concepts is a core enterprise within the semantic memory literature. This field has identified a multitude of conceptual relations in semantic memory, including among perceptual/sensorimotor, associative and taxonomic dimensions (McRae & Jones, 2014). Recent findings have also highlighted the importance of event relations in semantic memory. A number of studies have clearly demonstrated priming among verbs and object concepts that share connections via generalized event knowledge (e.g. Ferretti, McRae & Hatherell, 2001; Hare, Jones, Thomson, Kelly & McRae, 2009). Event knowledge is also activated

during sentence comprehension (Metusalem et al., 2012) and exposure to events in sentences can modify connections between concepts in semantic memory (Jones and Love, 2007). For example, Jones and Love (2007) found that similarity ratings between items like POLAR BEAR and COLLIE increased when participants recently read about them participating in chasing events like *The polar bear chases the seal* and *The collie chases the cat*. These results indicate that connections between objects in semantic memory could be driven by computing similarity across events (such as when POLAR BEARS and COLLIES both chase other animate entities). The current study extends this work by suggesting that the relation between semantic and event knowledge is not unidirectional, and that semantic knowledge may also drive event interpretation and generalization. Specifically, the findings in this investigation indicate that listeners activate their semantic knowledge (such as their understanding of similarities among animal subcategories) to support real-time inferences about event relations in sentences.

The framing of this work within a categorization/semantic memory approach also suggests a number of promising areas for future research. Specifically, the connection between event generalization and semantic memory suggests a need to identify whether and how the mechanisms that support conceptual relations in semantic memory may similarly support event generalization in sentence processing. In particular, future work is needed to systematically manipulate the extent to which event generalization during real-time sentence processing is possible as a function of semantic distance between category members. One productive direction for a future study would be to compare whether generalization over repeated exposures might emerge more readily when the novel agent shares a greater number of semantic features with that of the training agent.

It is also important to consider that the novel event mappings occurred in a fictional, “cartoon-world” context where animal agents participated in anthropomorphic activities. This cartoon world method is highly advantageous from a learning perspective because it is possible to precisely control and limit effects due to uncontrolled variability in listener’s prior experience. There is also a long tradition of using cartoon-like illustrations, line-drawings, and “clip-art” materials to explore fundamental mechanisms of psycholinguistic processing (e.g. Snodgrass & Vanderwart, 1980). Fictional contexts may even support some elements of language learning, such as learning vocabulary in book-reading contexts (Weisberg et al., 2015). Nevertheless, it is still possible that the artificiality of the novel event structures could have promoted different generalisations from initial mappings than might otherwise occur in everyday situations involving real-agents embedded in a natural world (Filik, 2008; Filik & Leuthold, 2008). Much of language learning is scaffolded by rich information from the physical world and everyday events. Therefore, a full account of the memory and linguistic mechanisms that support extension of event knowledge in real-time language processing will also need to account for potential differences between fictional and real worlds. Future work will be necessary to explore extent to which these mechanisms extend to learning and prediction in everyday real-world contexts.

The current findings are also unable to determine whether and how prediction and extension of event knowledge would occur over longer delays between the initial mapping and sentence processing. Because listeners deploy generalised event knowledge during

incremental processing (McRae & Matsuki, 2009), it seems uncontroversial that listeners recruit knowledge gained over their entire lives to shape their ongoing interpretation of speech. However, there may be important differences in the activation and implementation of knowledge that was acquired within the last several minutes compared to that which was initially learnt several hours or days ago. For example, there are some findings that indicate that the abstraction of structural regularities in artificial language, and lexical consolidation is supported by sleep consolidation in infancy and adulthood (Gaskell & Dumay, 2003; Gómez, Bootzin, & Nadel, 2006; Hupbach, Gomez, Bootzin, & Nadel, 2009). Memory consolidation processes are likely to interact with predictive language processing skills, and this would be an exciting avenue for future research. Nevertheless, the current findings suggest that even with very minimal training, listeners are able to generalise from their recent experience to dynamically update their ongoing interpretation of language.

In sum, these experiments represent an initial venture toward uncovering fundamental processes in how adults predictively interpret unfamiliar spoken language contexts in real-time. While much more work is needed, these initial findings begin to chip away at foundational questions regarding how memory and processing skills interact when listeners interpret novel utterances in everyday spoken language contexts. Importantly, these results reinforce the importance of prior experience in predictive processing. Not only does the *amount* of experience matter, but the *structure* of prior experience is fundamental as well.

Acknowledgments

I wish to thank Julie Carranza for enthusiastically allowing me to record her voice for the spoken stimuli. I am also grateful to: Courtney Yehert, Angele Yazbec, Virginia Prestwood, Laura Hamant, Megan Blom, Ashley Candler, Ireney Duval, Elizabeth Gentry, Emanuel Boutzoukas, Brooke Nolan, Ashley Candler for their assistance with stimuli development, data collection and data entry in this study. Michael Kaschak and Gerry Altmann provided invaluable input on prior versions of this manuscript. This work was partially supported by R03 DC013638 to A.B.

References

- Altmann G, Mirkovi J. Incrementality and prediction in human sentence processing. *Cognitive Science*. 2009; 33(4):583–609. DOI: 10.1111/j.1551-6709.2009.01022.x [PubMed: 20396405]
- Amato MS, MacDonald MC. Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition*. 2010; 116(1):143–148. DOI: 10.1016/j.cognition.2010.04.001 [PubMed: 20430372]
- Barr DJ, Jackson L, Phillips I. Using a voice to put a name to a face: the psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General*. 2014; 143(1):404–413. DOI: 10.1037/a0031813 [PubMed: 23398179]
- Bicknell K, Elman JL, Hare M, McRae K, Kutas M. Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language*. 2010; 63(4):489–505. DOI: 10.1016/j.jml.2010.08.004 [PubMed: 21076629]
- Boersma P, Weenink D. Praat: doing phonetics by computer (Version 5.3.22). 2012. Retrieved from <http://www.praat.org/>
- Borovsky A, Ellis EM, Evans JL, Elman JL. Semantic density interacts with lexical and sentence processing in infancy. *Child Development*. (in press).
- Borovsky A, Elman JL, Fernald A. Knowing a lot for one's age: Vocabulary and not age is associated with incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*. 2012; 112(4):417–436. DOI: 10.1016/j.jecp.2012.01.005 [PubMed: 22632758]

- Borovsky A, Sweeney K, Elman JL, Fernald A. Real-time interpretation of novel events across childhood. *Journal of Memory & Language*. 2014; 73:1–14. doi: DOI: 10.1016/j.jml.2014.02.001 [PubMed: 24976677]
- Carey S, Bartlett E. Acquiring a single new word. *Papers and Reports on Child Language Development*. 1978; 15:17–29.
- Chang F, Kidd E, Rowland CF. Prediction in processing is a by-product of language learning. *Behavioral and Brain Sciences*. 2013; 36(4):350–351. DOI: 10.1017/S0140525X12002518 [PubMed: 23789743]
- Chomsky N. *Aspects of the theory of syntax*. Cambridge: M.I.T. Press; 1965.
- DeLong KA, Troyer M, Kutas M. Pre-Processing in Sentence Comprehension: Sensitivity to Likely Upcoming Meaning and Structure. *Language and Linguistics Compass*. 2014; 8(12):631–645. DOI: 10.1111/lnc3.12093 [PubMed: 27525035]
- Federmeier KD, Kutas M. A rose by any other name: Long-term memory structure and sentence processing. *Journal of memory and Language*. 1999; 41(4):469–495. DOI: 10.1006/jmla.1999.2660
- Ferretti TR, McRae K, Hatherell A. Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*. 2001; 44(4):516–547.
- Filik R. Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition*. 2008; 106(2):1038–1046. DOI: 10.1016/j.cognition.2007.04.006 [PubMed: 17524387]
- Filik R, Leuthold H. Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology*. 2008; 45(4):554–558. DOI: 10.1111/j.1469-8986.2008.00656.x [PubMed: 18282200]
- Gaskell MG, Dumay N. Lexical competition and the acquisition of novel words. *Cognition*. 2003; 89(2):105–132. DOI: 10.1016/S0010-0277(03)00070-2 [PubMed: 12915296]
- Gerken L, Bollt A. Three Exemplars Allow at Least Some Linguistic Generalizations: Implications for Generalization Mechanisms and Constraints. *Language Learning and Development*. 2008; 4(3): 228–248. DOI: 10.1080/15475440802143117
- Goldberg AE. *Constructions at work: The nature of generalization in language*. Oxford University Press; 2006.
- Goldberg AE, Casenhiser D, Sethuraman N. Learning argument structure generalizations. *Cognitive Linguistics*. 2004; 14:289–316. DOI: 10.1515/cogl.2004.011
- Gómez R, Maye J. The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy*. 2005; 7(2):183–206. DOI: 10.1207/s15327078in0702_4
- Gómez RL. Variability and Detection of Invariant Structure. *Psychological Science*. 2002; 13(5):431–436. DOI: 10.1111/1467-9280.00476 [PubMed: 12219809]
- Gómez RL, Bootzin RR, Nadel L. Naps Promote Abstraction in Language-Learning Infants. *Psychological Science*. 2006; 17(8):670–674. DOI: 10.1111/j.1467-9280.2006.01764.x [PubMed: 16913948]
- Gomez RL, Gerken L. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*. 1999; 70(2):109–135. DOI: 10.1016/S0010-0277(99)00003-7 [PubMed: 10349760]
- Grice HP. *Logic and Conversation*. In: Cole P, Morgan JL, editors *Syntax and Semantics*. Vol. 3. New York: Academic Press; 1975.
- Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*. 2011; 48(12):1711–1725. DOI: 10.1111/j.1469-8986.2011.01273.x [PubMed: 21895683]
- Hare M, Jones M, Thomson C, Kelly S, McRae K. Activating event knowledge. *Cognition*. 2009; 111(2):151–167. [PubMed: 19298961]
- Hockett CF. The Origin of Speech. *Scientific American*. 1960; 203(3):88–96. [PubMed: 13683472]
- Huetting F, Altmann GT. Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*. 2005; 96(1):B23–B32. DOI: 10.1016/j.cognition.2004.10.003 [PubMed: 15833303]

- Huetig F, Mani N. Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*. 2015; :1–13. DOI: 10.1080/23273798.2015.1072223
- Huetig F, Rommers J, Meyer AS. Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta Psychologica*. 2011; 137(2):151–171. DOI: 10.1016/j.actpsy.2010.11.003 [PubMed: 21288498]
- Hupbach A, Gomez RL, Bootzin RR, Nadel L. Nap-dependent learning in infants. *Developmental Science*. 2009; 12(6):1007–1012. DOI: 10.1111/j.1467-7687.2009.00837.x [PubMed: 19840054]
- Jones M, Love BC. Beyond common features: The role of roles in determining similarity. *Cognitive Psychology*. 2007; 55(3):196–231. [PubMed: 17094958]
- Kamide Y, Altmann GTM, Haywood S. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory & Language*. 2003; 49:133–156. DOI: 10.1016/S0749-596X(03)00023-8
- Kaschak M. What this construction needs is generalized. *Memory & Cognition*. 2006; 34(2):368–379. DOI: 10.3758/bf03193414 [PubMed: 16752600]
- Kukona A, Cho PW, Magnuson JS, Tabor W. Lexical interference effects in sentence processing: Evidence from the visual world paradigm and self-organizing models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2014; 40(2):326–247. DOI: 10.1037/a0034903
- Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*. 2007; 164(1):177–190. DOI: 10.1016/j.jneumeth.2007.03.024 [PubMed: 17517438]
- McRae K, Jones MN. Semantic memory. In: Reisberg D, editor *The Oxford handbook of cognitive psychology*. Oxford, UK: 2013. 206–219.
- McRae K, Matsuki K. People Use their Knowledge of Common Events to Understand Language, and Do So as Quickly as Possible. *Language and Linguistics Compass*. 2009; 3(6):1417–1429. DOI: 10.1111/j.1749-818X.2009.00174.x [PubMed: 22125574]
- Metusalem R, Kutas M, Urbach TP, Hare M, McRae K, Elman J. Generalized event knowledge activation during online sentence comprehension. *Journal of Memory & Language*. 2012; 66(4): 545–567. DOI: 10.1016/j.jml.2012.01.001 [PubMed: 22711976]
- Pickering MJ, Garrod S. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*. 2013; 36(4):329–347. DOI: 10.1017/S0140525X12001495 [PubMed: 23789620]
- Rabagliati H, Gambi C, Pickering MJ. Learning to predict or predicting to learn? *Language, Cognition and Neuroscience*. 2015; 31(1):94–105. DOI: 10.1080/23273798.2015.1077979
- Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*. 1980; 6(2)doi: 10.1037/0278-7393.6.2.174
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. *Science*. 1995; 268:1632–1634. DOI: 10.1126/science.7777863 [PubMed: 7777863]
- Von Holzen K, Mani N. Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*. 2012; 113(4):569–586. DOI: 10.1016/j.jecp.2012.08.001 [PubMed: 22980955]
- Weisberg DS, Ilgaz H, Hirsh-Pasek K, Golinkoff R, Nicolopoulou A, Dickinson DK. Shovels and swords: How realistic and fantastical themes affect children's word learning. *Cognitive Development*. 2015; 35:1–14. doi: DOI: 10.1016/j.cogdev.2014.11.001
- Wonnacott E, Boyd JK, Thomson J, Goldberg AE. Input effects on the acquisition of a novel phrasal construction in 5 year olds. *Journal of Memory and Language*. 2012; 66(3):458–478. doi: DOI: 10.1016/j.jml.2011.11.004
- Yee E, Sedivy JC. Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2006; 32(1): 1–14. DOI: 10.1037/0278-7393.32.1.1
- Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. *Psychological Bulletin*. 1998; 123(2):162–185. DOI: 10.1037/0033-2909.123.2.162 [PubMed: 9522683]

Appendix A

Outline of the agents, actions and thematic objects presented in Experiment 1 and 2. In each story set, each agent performed both actions on a unique set of objects. In Experiment 2, The first Agent in each pair performed the same set of actions (e.g. if Rabbit eats ice cream and wears hat, so does the Hamster). All possible combinations of Agents, Actions and Objects occurred across lists in the study.

Experiment 1				
Agents	Action1	Object1	Action2	Object2
Rabbit, Ant	Eat	Cake, Ice Cream	Wears	Hat, Sunglasses
Bee, Squirrel	Holds	Microphone, Teddy bear	Turns on	Lamp, Flashlight
Fox, Fish	Rides	Skateboard, Bike	Eats	Grapes, Banana
Seahorse, Lion	Opens	Bag, Present	Reads	Letter, Book
Monkey, Duck	Flies	Airplane, Kite	Sits on	Rock, Fence
Parrot, Horse	Sits on	Bench, Chair	Smells	Flower, Popcorn
Pig, Turtle	Turns on	TV, Computer	Cuts	Paper, Bread
Frog, Dog	Eats	Candy, Apple	Rides	Bus, Car
Experiment 2				
Agent (Category 1, Category 2)	Action1	Object1	Action2	Object2
Rabbit, Ant Hamster, Butterfly	Eat	Cake, Ice Cream	Wears	Hat, Sunglasses
Bee, Squirrel Spider, Skunk	Holds	Microphone, Teddy bear	Turns on	Lamp, Flashlight
Fox, Fish Raccoon, Dolphin	Rides	Skateboard, Bike	Eats	Grapes, Banana
Seahorse, Lion Starfish, Elephant	Opens	Bag, Present	Reads	Letter, Book
Monkey, Duck Tiger, Pigeon	Flies	Airplane, Kite	Sits on	Rock, Fence
Parrot, Horse Pelican, Cow	Sits on	Bench, Chair	Smells	Flower, Popcorn
Pig, Turtle Sheep, Snake	Turns on	TV, Computer	Cuts	Paper, Bread
Frog, Dog Alligator, Cat	Eats	Candy, Apple	Rides	Bus, Car

Appendix B

Outline of the Story and Extension agents in Experiment 1 and 2

Story Agent(s)	Extension agent	Category
Dog, Cat	Guinea Pig	Pets
Rabbit, Hamster	Mouse	Pets
Ant, Butterfly	Ladybug	Bugs
Bee, Spider	Fly	Bugs
Squirrel, Skunk	Deer	Woodland creatures
Fox, Raccoon	Wolf	Woodland creatures
Fish, Dolphin	Shark	Sea creatures
Seahorse, Starfish	Crab	Sea creatures
Lion, Elephant	Zebra	Zoo animals

Story Agent(s)	Extension agent	Category
Monkey, Tiger	Giraffe	Zoo animals
Duck, Pigeon	Bluebird	Birds
Parrot, Pelican	Robin	Birds
Pig, Sheep	Goat	Farm animals
Horse, Cow	Chicken	Farm animals
Turtle, Snake	Lizard	Reptiles
Frog, Alligator	Iguana	Reptiles

Story A

Hey there's a rabbit!



Hey there's an ant



Oh look! He's wearing a hat!



Oh look! He's wearing the sunglasses!



Where's he going now?

Blank screen

Where's he going now?

Blank screen

He's eating the cake! What a silly rabbit!



He's eating the ice cream! What a silly ant!

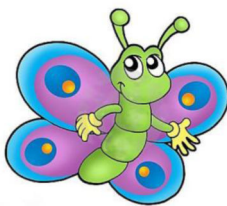


Story B

Hey there's a hamster!



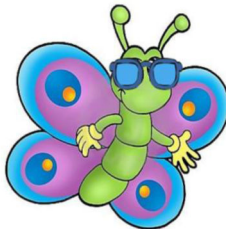
Hey there's a butterfly



Oh look! He's wearing a hat!



Oh look! He's wearing the sunglasses!



Where's he going now?

Blank screen

Where's he going now?

Blank screen

He's eating the cake! What a silly hamster!

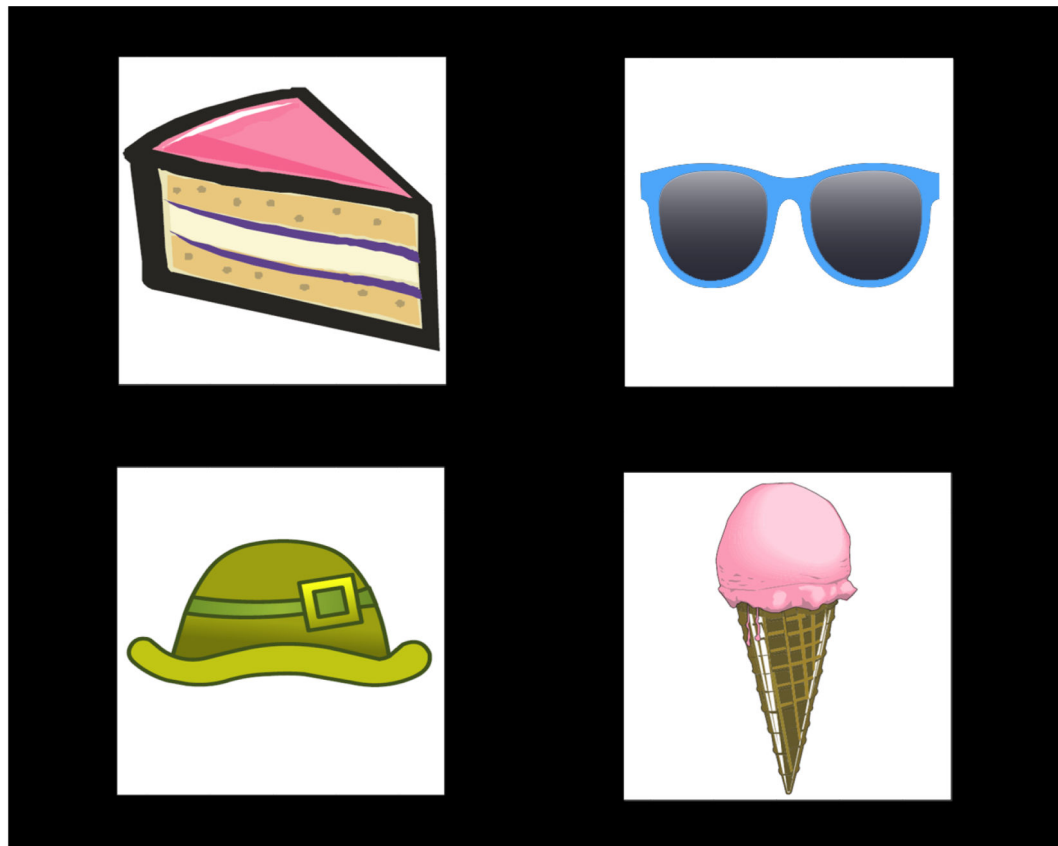


He's eating the ice cream! What a silly butterfly!



Figure 1.

Illustration of one version of a story in Experiment 1 and 2. Panel A represents the story as it appeared in Experiment 1. In Experiment 2, participants either saw Panel A repeated twice (Repeated Agent condition) or saw Panel A and Panel B a single time (Multiple Agent Condition). Each image was accompanied by spoken narration (listed above the image). The stories were self-paced, and participants clicked on a mouse when they were ready to hear the next image and narrative sentence of the story. Story versions were counterbalanced such that all possible combinations of agents, actions and objects occurred.



Spoken Sentence:

Repeat Condition

The rabbit wears the hat.
The rabbit eats the ice cream.
The ant wears the sunglasses.
The ant eats the cake.

Generalize Condition

The mouse wears the hat.
The mouse eats the ice cream.
The ladybug wears the sunglasses.
The ladybug eats the cake.

Art1 Agent Action Art2 Target

Target	Agent-Related	Action-Related	Unrelated
HAT	ICE CREAM	SUNGLASSES	CAKE
ICE CREAM	HAT	CAKE	SUNGLASSES
SUNGLASSES	CAKE	HAT	ICE CREAM
CAKE	SUNGLASSES	ICE CREAM	HAT
HAT	ICE CREAM	SUNGLASSES	CAKE
ICE CREAM	HAT	CAKE	SUNGLASSES
SUNGLASSES	CAKE	HAT	ICE CREAM
CAKE	SUNGLASSES	ICE CREAM	HAT

Figure 2.

Example of image array and possible sentences in Experiment 1 Repeat and Generalise conditions. In Experiment 2, only sentences from the Generalise condition were presented.

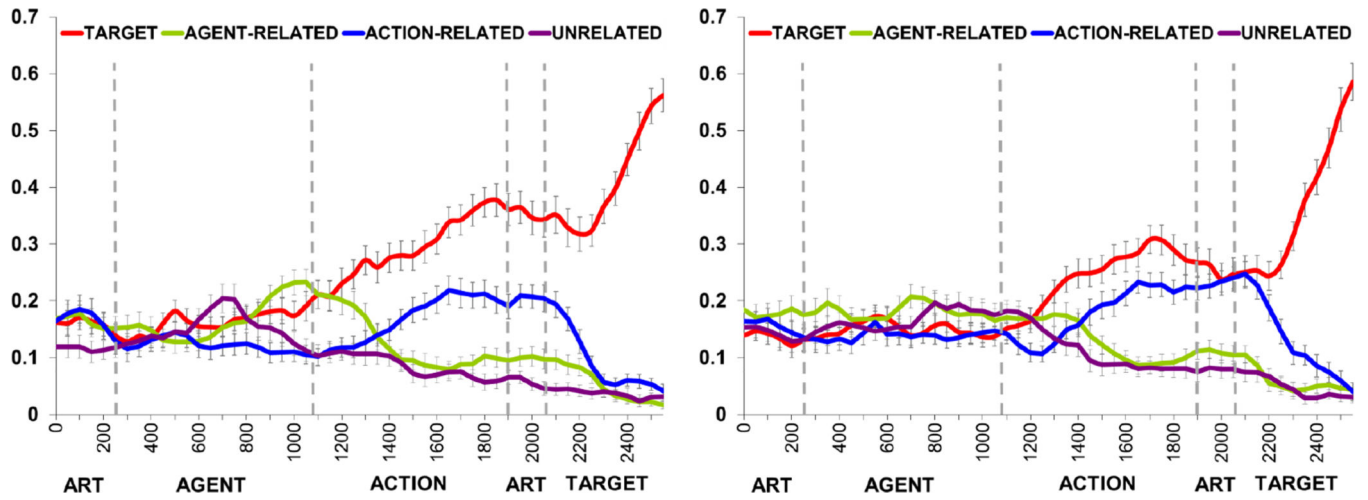


Figure 3. Timecourse of looks to interest areas plotted from sentence onset to offset in 50 ms time bins in the A) Repeat condition and B) Generalise condition in Experiment 1. Error bars represent ± 1 SE of mean fixation proportions in each time bin.

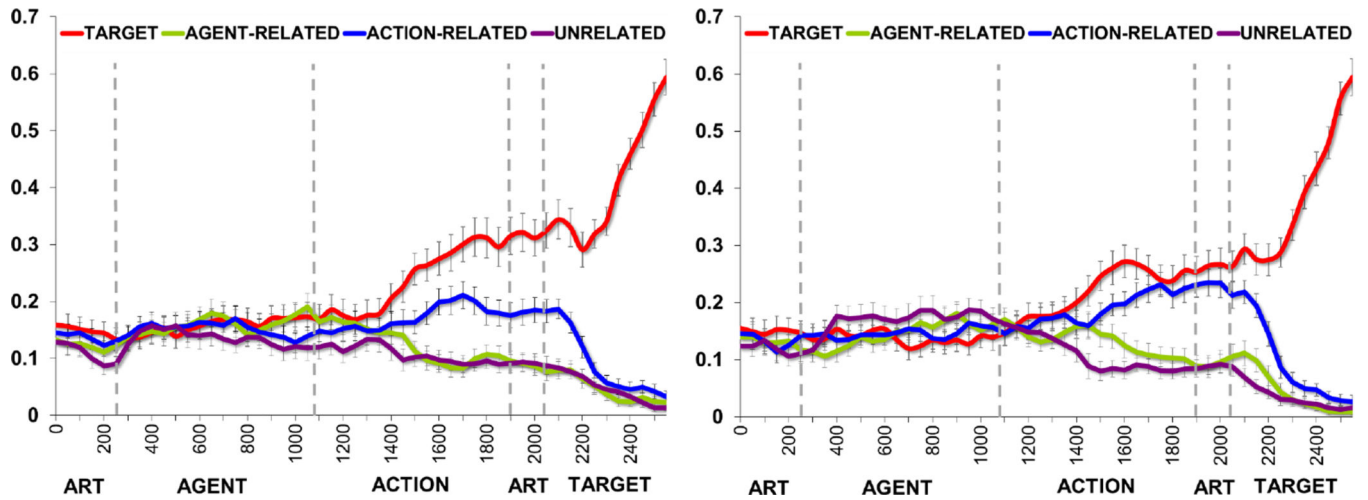


Figure 4. Timecourse of looks to the interest areas from sentence onset to sentence offset in 50 ms time bins in the A) Repeat Agent condition and B) Multiple Agent condition of Experiment 2. Error bars represent ± 1 SE of mean fixation proportions in each time bin.

Table 1

Results of cluster based permutation analyses in the all conditions in Experiment 1. The time window where the Target significantly exceeded that of each distractor is reported in ms, followed by the Cluster t statistic and Monte Carlo p value for that window, in parentheses. Time window is measured from the onset of the spoken sentence, which was 2550 ms in duration. The onset of the sentence final target was at 2036 ms post sentence onset.

	Target vs. Agent-Related	Target vs. Action-Related	Target vs. Unrelated
Repeat Condition	1300 – 2550 ms ($t = 183.3, p = .001$)	1050 – 2550 ms ($t = 142.7, p = .0009$)	1050 – 2550 ms ($t = 232.4, p = .001$)
Generalise Condition	1400 – 2550 ms ($t = 143.0, p = .0008$)	2250 – 2550 ms ($t = 51.6, p = .0008$)	1250 – 2550 ms ($t = 179.3, p = .0008$)

Table 2

Timing of cluster based permutation analyses in the all conditions in Experiment 2. Timing is measured from the onset of the spoken sentence, which was 2550 ms in duration. The time window where the Target significantly exceeded that of each distractor is reported in ms, followed by the Cluster t statistic and Monte Carlo p value for that window, in parentheses. The onset of the sentence final target was at 2036 ms post sentence onset.

	Target vs. Agent-Related	Target vs. Action-Related	Target vs. Unrelated
Repeated agent condition	1500 – 2550 ($t = 164.4, p = .0008$)	1900 – 2550 ($t = 93.7, p = .0008$)	1400 – 2550 $t = 159.0, p = .001$
Multiple agent condition	1500 – 2550 ($t = 159.2, p = .0008$)	2100 – 2550 ($t = 77.0, p = .0008$)	1400 – 2550 ($t = 177.4, p = .0009$)