4- and 5-Year-Olds Use Mental Models of Events in Online Reference Resolution

Amanda Rose Yuile and Cynthia Fisher

Many referential terms are ambiguous. For example, in (1) the pronoun '*he*' refers back to an entity mentioned in the prior sentence, but the prior sentence has two potential referents for this pronoun. Similarly, in a context in which two girls have been mentioned, the phrase '*the girl*' in (2) is ambiguous.

(1) John was passing the book to Bill. He…
(2) Sarah asked the girl to pass the salt.

As adults, we can resolve many of these ambiguities by generating inferences based on a linguistic representation of the discourse context. For instance, pronouns and definite expressions can often be resolved via straightforward inferences based on information such as grammatical position (Arnold et al., 2018; Gordon, Grosz, & Gilliom, 1993), order of mention (Hartshorne, Nappa, & Snedeker, 2015), and the distance separating the expression from its antecedent (Gernsbacher & Hargreaves, 1988). In other words, listeners can resolve many ambiguities by maintaining representations of the linguistic properties of the text. In fact, computational models have used such heuristics to good effect in modeling referential processing (see Lee et al., 2013 for an example).

However, this is only part of the story. Adults also generate inferences about reference based on pragmatic information. Many researchers have proposed that adults understand sentences relative to a semantically interpreted *situation model* (e.g., Richmond & Zachs, 2017; Zwaan & Radvansky, 1998). The situation model is a representation of the discourse that includes the entities mentioned in the discourse and the relationships among them, the properties of those entities, as well as relevant world knowledge. This representation is created incrementally, as comprehenders integrate each incoming sentence with the current state of the model. This integration allows listeners to resolve many ambiguities and to generate predictions about upcoming linguistic information. In reference comprehension, adult listeners make inferences about which referents are likely to be talked about next, given their incrementally created conceptual

* Department of Psychology, University of Illinois at Urbana-Champaign. Corresponding author: Amanda Rose Yuile (amandah3@illinois.edu). This research was supported by NICHD Grant 2-ROI-HD-5448 to Cynthia Fisher.

 Proceedings of the 45th annual Boston University Conference on Language Development, ed. Danielle Dionne and Lee-Ann Vidal Covas, 829-844. Somerville, MA: Cascadilla Press.

representation of the situation and their estimate of the speaker's goals (e.g., Hartshorne, O'Donnell, & Tenenbaum, 2015).

Nieuwland and colleagues (2007) reported particularly clear evidence of situation-based reasoning in adult reference comprehension. In an ERP (event-related brain potential) study, definite expressions such as "the nephew," in a discourse context involving two nephews, elicited a negative, frontally distributed ERP component that is thought to index the processing of referential ambiguity (NRef). However, in a discourse context in which two nephews were introduced, but one left the scene ("the nephew who was into history left early") before the critical sentence, the NRef component was greatly reduced, suggesting that participants found "the nephew" unambiguous. The researchers carefully controlled the number and order of prior mentions for each referent, so the effect could not be explained by use of linguistic heuristics. Instead, in Nieuwland et al.'s experiment, adults created a conceptual representation that included the two nephews in the same location. When one nephew was described as leaving early, listeners then removed the referent from that location in their model of the situation. Thus, this inference was generated as a natural consequence of creating a structured event representation that included the bindings between predicates and their arguments and the causal relations within the event. In the current work, we ask whether children are capable of using the situation model to constrain reference during online language comprehension, as adults do.

Like adults, children use linguistic features of the discourse to constrain the interpretation of ambiguous noun phrases. For instance, in a simple context in which only one referent has been mentioned, 18-month-olds used an expectation of discourse continuity to disambiguate an anaphoric expression such as "Do you see another one?" (Lidz, Waxman, & Freedman, 2003). Similarly, 12-month-olds can use person-specific language ("my ball") to determine which of two balls a speaker is referring to (Saylor, Ganea, & Vasquez, 2011). Children also link pronouns with prominently established discourse topics—referents that have been mentioned repeatedly, and in subject position (e.g., Hartshorne, Nappa, & Snedeker, 2015; Song & Fisher, 2005, 2007).

Children can also learn new words through their introduction in the discourse. In an experiment by Horowitz and Frank (2015), 3- to 6-year-old children assigned a novel label to a toy when the label was embedded between sentences describing the same toy's features. Akhtar (2002) found that 2-year-olds linked a novel word to the dimension of an object (shape or color) that had been the focus of the preceding discourse. Similarly, Bohn and colleagues (2020) found that 3- and 4-year-olds assigned an ambiguous pronoun ("it") to an object belonging to the same category that had been the topic of the speaker's previous utterances (e.g., animals). Thus, children appear to use linguistic information in reference comprehension.

The evidence on children's use of pragmatic reasoning is less clear, especially as it relates to online referential processing as sentences unfold. In comprehension studies using offline measures, recent reports suggest that children can use pragmatic inference to guide reference comprehension. For instance, Sullivan and

Barner (2016) found that children as young as 2 years old could make a bridging inference to identify the implied topic of a stretch of discourse, and in turn to identify the referent of a novel word. That is, children united utterances such as, "I'm *thirsty*", and "There's a *pliff* on the table!", to pick out a drink.

Children, like adults (Arnold, 2001; Stevenson, Crawley, & Kleinman, 1994), also use the semantic properties of verbs, and discourse connectives, to guide pronoun interpretation. For instance, Kehler, Hayes, and Barner (2011) found that when 5- and 6-year-old children were prompted with passages containing motion verbs (e.g., "Mrs. Horse skipped to Mrs. Sheep"), they were more likely to continue the story with reference to the subject (Mrs. Horse). In contrast, when prompted with passages containing transfer-of-possession verbs (e.g., "Mr. Seal handed a gift to Mr. Penguin"), they were more likely to continue the story with reference to the goal (Mr. Penguin). In another study, 4-year-olds used discourse connectives to infer the referent of a novel noun (Sullivan et al., 2019). Finally, in an online comprehension task, Pyykkönnen, Matthews, and Jarvikivi (2010) presented 3- and 4-year-old children with stories that began with a sentence containing either a high transitivity (e.g., *hit*) or low transitivity (e.g., *saw*) verb, followed by a sentence with an ambiguous pronoun. In this visual world task, children's looking patterns revealed a stronger subject bias for the pronoun following sentences with low transitivity verbs. Together, these studies indicate that semantic properties of sentences can change the prominence of each participant in the discourse for later reference. However, it is difficult to tell whether these findings reflect inferences over a situation model.

On the one hand, it is possible that these effects arise from causal reasoning over event knowledge. In Sullivan et al.'s (2019) experiment, children could have inferred the referent of the novel noun based on the causal and temporal relationship indicated by the discourse connective. In the study, participants listened to sentences such as "One animal gave the cat to the other animal *and/because* the cat scratched the *zilfa*." When children encountered *and* they could infer that the following clause spelled out what happened next in the event and therefore that the cat scratched the goal participant (who now held the cat). When children encountered *because* they could infer that the following clause explained why the preceding event happened and therefore that the cat scratched the agent (causing him to hand it over). Thus, *and* and *because* help establish causal and temporal relationships between the two clauses; these relationships guide inferences about which event participants might be mentioned next.

On the other hand, effects of semantic prominence are also sometimes attributed to use of linguistic discourse representations. Such representations may include probabilistic semantic knowledge. For example, it is possible that children in Kehler et al.'s (2011) experiment had learned through experience with language that subjects are likely to be mentioned again following motion verbs, and goals are likely to be mentioned again following transfer-of-possession verbs (see Hartshorne, O'Donnell & Tenenbaum, 2015, for a review). In other words, a listener might learn to use event type to predict upcoming reference. Therefore, it

is not clear whether these effects should be attributed to causal reasoning over a situation model as opposed to linguistic knowledge.

However, prior research demonstrates that children *do* construct situation models during comprehension. For instance, Lynch and colleagues (2008) found in a narrative recall task that 4- and 6-year-old children were more likely to recall events that had a greater number of causal connections within the narrative. This indicates that children treated causally connected events as more central to the narrative and used causal links to structure their representation of the narrative. Rall and Harris (2000) found that 3- and 4-year-old children more accurately recalled sentences in a story if the verb was consistent with the main protagonist's spatial perspective. For example, in *Cinderella*, children more accurately recalled, "the fairy godmother *came* into the cottage" than "the fairy godmother *went* into the cottage." These findings indicate that children create a causally structured situation model while comprehending language.

Can children use their situation model during online reference comprehension? We know of only one prior experiment that examines the use of discourse-provided event information in children's online reference comprehension: Borovsky and colleagues (2014) investigated whether children could use recently encoded event information to anticipate an upcoming word. In the experiment, school-aged children were presented with several facts, such as "The monkey eats candy" and "The monkey rides the bus". Next, in a visual world task, children were presented with sentences repeating the recently learned facts. The children readily used what they had been told to anticipate the final word of a sentence ("The monkey rides…").

The present work builds on these findings to ask whether children can use the situation model to generate an inference about unmentioned events, affecting their interpretation of a referentially ambiguous noun. To address this question, we conducted a visual world eye-tracking experiment with 4- and 5-year-old children and a group of adult participants. In our experiment, participants listened to stories modeled after Nieuwland et al. (2007). In the critical sentence, the story's protagonist asked an animal to join her in an activity, e.g. *She asked the bunny with the crown to splash in the water with her*. We varied whether or not the story had introduced two potential referents for the target noun (two bunnies), and whether the story stated that one had left the scene before the critical sentence. We tracked participants' gaze to the pictured animals to measure the time-course of reference resolution. If preschoolers exploit the constraints of the situation model, then in a two-referent context in which one referent has left the scene, they should identify the target animal in the critical sentence before the disambiguating information, because only the target animal is still present in the story (though both remain visible in the display). We included a group of adult participants to ensure that we could replicate the effect found by Nieuwland et al. (2007) using a visual world paradigm.

1. Methods

1.1. Participants

Forty-two undergraduate students from the University of Illinois at Urbana-Champaign participated in exchange for course credit. All participants were native English speakers. Data from six additional participants were excluded following data cleaning for excessive data loss (see procedure below).

Eighty 4- and 5-year-old children, all native English speakers, participated (*M* age = 57.7 months, *range* = 48.03-71.72, 45 female). Our target sample size was 90 participants. We chose to end data collection prior to reaching this target due to an indefinite lab closure caused by the COVID-19 global pandemic. Data from 14 additional children were tested but excluded because they wanted to stop listening to the stories early (10) or due to excessive data loss (4).

1.2. Materials

Participants listened to spoken stories modeled after Nieuwland et al. (2007). We constructed twenty-four stories, each involving a protagonist named Maisy and two animals (see Table 1 for an example). In each story, the first sentence introduced Maisy. The second sentence introduced the two animals, and the third sentence referenced only Maisy, to foreground her. The fourth sentence described the actions of the two animal characters, and the fifth sentence made explicit reference to Maisy to foreground her again. Then, in the critical sentence of each story (bolded in Table 1), Maisy asked one animal to join her in an activity: "She asked the bunny..." Each story ended with a sentence that described Maisy and the target animal doing the activity. All references to animal characters were modified phrases, such as “the bunny with the crown”, as shown in Table 1. Stories were recorded by a female native speaker of English (the first author).

Table 1. Example Story

One afternoon Maisy was at the beach.		
She was with her friends the bunny (One-referent unambiguous: turtle) with the necklace and the bunny with the crown.		
Maisy was wearing her polka dot shorts.		
Ambiguous	**Two-referent Unambiguous**	**One-referent Unambiguous**
The bunny with the necklace was flying a kite and so was the bunny with the crown.	*The bunny with the necklace had to leave early*, but the bunny with the crown was flying a kite.	*The turtle with the necklace had to leave early*, but the bunny with the crown was flying a kite.
Maisy didn’t like flying kites; she liked splashing in the water.		
She asked the bunny with the crown to splash in the water with her.		
And then they splashed in the water together.		

As shown in Table 1, we created 3 versions of each story. These versions differed on two dimensions. First, in the Ambiguous and the Two-referent

Unambiguous conditions, the two animals introduced in the story were both of the same kind (e.g., both bunnies), differentiated by modifiers (e.g., "the bunny with the crown"). In the One-referent Unambiguous condition, the two animals were of different kinds (e.g., a bunny and a turtle). Second, in the fourth sentence of each story, shown separately by condition in Table 1, the stories differed in whether one of the animals left the scene. In the Ambiguous condition, neither animal left the scene; in the Two-referent and One-referent Unambiguous conditions, one animal left the scene. These differences created the key manipulation: In the critical sentence shown bolded in Table 1, the phrase "the bunny with the" is temporarily ambiguous in the Ambiguous condition, because two bunnies are still present. To identify the intended referent, listeners must await the modifier ("the bunny with the *crown*"). In contrast, in both Unambiguous conditions, this phrase might be treated as unambiguous before the modifier. Returning to the Table 1 examples, in the One-referent Unambiguous condition, only one bunny was introduced (and the other animal, the turtle, left the scene). In this case, the noun itself ("the bunny") is sufficient to identify the target. In the Two-referent Unambiguous condition, two bunnies were introduced, but because one has left the scene, listeners might infer, on the basis of the discourse-provided semantic information, which bunny is being referred to before hearing the disambiguating attribute at the end of the phrase. We counterbalanced whether the target animal character was introduced first or second across stories.

The stories were accompanied by clip-art style pictures, as shown in Figure 1a. A picture of Maisy (pictured in different outfits across stories) always appeared in the center of the screen. A picture of each animal character, and pictures associated with each activity mentioned in the story, appeared in the four corners of the screen. For example, Figure 1a shows a bunny with a crown and a bunny with a necklace (the displays for the Ambiguous and the Two-referent Unambiguous conditions), a splash of water, and a kite. In each display, the animal images always appeared on the same horizontal line, i.e. both on the top half of the screen or both on the bottom half of the screen. Across stories, we counterbalanced the left/right and top/bottom position of the animal images.

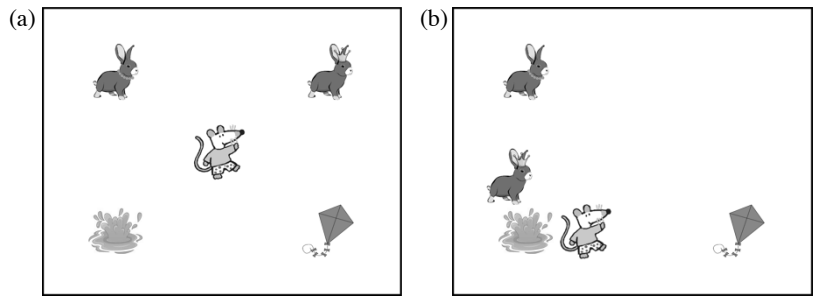

Figure 1. Example trial screens

To ensure that our materials had the intended structure, we tested a group of 30 native English-speaking adults (*M* age = 22.5 years, 20 female) in a sentence completion norming experiment. In the norming task, we presented participants with stories that ended at the first determiner in the critical sentence (*"She asked the"*) and asked participants to finish the sentence. We found that the adults' sentence completions mentioned two animals (e.g., two bunnies, or a bunny and a turtle) more often in the Ambiguous condition than in the two Unambiguous conditions, and that participants correctly remembered who had left the scene. When only one referent remained in the scene (in the Two-referent and One-referent Unambiguous conditions), participants used modifiers more often if the two animals mentioned were the same kind (in the Two-referent Unambiguous condition) than if they were of different kinds (in the One-referent Unambiguous condition). Therefore, our norming experiment confirmed that in our stories referential ambiguity varied appropriately with how many potential referents the target noun had — the bunnies versus the bunny and the turtle — and whether one referent left the scene prior to the critical sentence.

1.3. Procedure

The procedure varied slightly for adult and child participants. In this section we will describe the procedure for children first, and then note differences in the adult version of the task.

Children were seated in a comfortable child-sized chair in front of a computer monitor. The experimenter explained that the child would listen to stories, and that the child and the experimenter would work together to act out each story. Before presenting the first story, the experimenter told the child that they were missing some cactuses and if the child found any during the stories, they would give the child a sticker. This cactus game was added to make the experiment more fun and engaging for children. A picture of a cactus appeared on the screen after every 4 trials. Whenever this occurred, the experimenter handed the child a sticker to put on a sticker worksheet given to them at the start of the experiment.

Each trial began with a blank white screen. When the child was ready to begin, the experimenter pressed the space bar on the display computer to begin the story. The images of the story characters and activities were displayed (see Figure 1a) for 500 ms before the story audio played. The average story duration was 40.96 seconds (range = 36.40 – 44.44 seconds). The images remained visible throughout the trial. When the audio finished playing, the experimenter asked the child several questions to probe recall for the story. In all story conditions, the experimenter asked what Maisy did at the end of the story and who she did it with. In Two-referent and One-referent Unambiguous trials, the experimenter also asked the child which animal character left the scene. If the child answered incorrectly or told the experimenter that they did not remember, the experimenter provided the correct answer. After these memory questions, the experimenter acted out the story by clicking and dragging the images of Maisy and the target animal to the target activity image as shown in Figure 1b, while providing a

narration (e.g., “Maisy and the bunny with the crown splashed in the water.”). Each time the experimenter referred to the animal characters, they used the full, modified noun phrase. Visual fixations were recorded from the onset of the images until the experimenter pressed the right arrow key to indicate that the trial was finished. Participants were offered a brief break after trial 12 (halfway through the story task). During this break, children were allowed to get up from the chair and were offered a juice box.

Children saw 24 trials, 8 in each story condition. Trial order was blocked by story condition. Within each block, stories were presented in a fixed semi-random order, with the constraint that the target character image did not appear on the same side of the screen more than twice in a row. We created 16 total lists to counterbalance the order of stories and conditions across blocks.

After the task was complete, parents or guardians were asked to complete a short survey that asked whether children understood or understood and said each animal name mentioned in the stories. On average, parents reported that their children understood and said 22 out of 24 animal names used in the stories.

The procedure was the same for the adults, except that: (1) The stories were presented in a random order for adults, with story conditions intermixed rather than blocked. (2) Adult participants pressed the space bar themselves to begin each story trial. The white screen at the start of each trial had black text that read “Press the space bar to play the story.” (3) The cactus finding game was omitted. (4) No memory questions were asked. (5) Participants were instructed to “act out” the last sentence of each story themselves, by clicking on and moving the images on the screen. This act-out task was included to engage participants in the stories. During the first story trial, the experimenter demonstrated the act-out task by clicking and dragging the image of Maisy and the target animal character to the target activity (see Figure 1b).

1.4. Eye-movement Recording

Stimuli were presented in Matlab using the Psychophysics Toolbox extension (Brainard, 1997) and the SR Research EyeLink API. We used an EyeLink 1000 eye-tracker (SR Research, Ltd) in a remote arm configuration to automatically record participants’ eye gaze. Eye gaze was recorded at a sampling rate of 500Hz. Before the experiment, an experimenter manually adjusted the display so that the distance between the participant’s eye and the eye-tracking camera was approximately 533mm. A target sticker was placed approximately 25mm above the participant’s left eye. The experimenter then performed a manual 9-point calibration and validation routine using a built-in black-and-white bull’s-eye image as the calibration target. This bulls-eye image was also used to perform a ‘drift-check’ after every 6 trials throughout the experiment.

1.5. Data Processing

We grouped visual fixations into 10ms bins for further analysis. Fixations were categorized into areas of interest (AOIs: target animal, competitor animal, first-mentioned activity, second-mentioned activity, and Maisy). Because we had the impression that it was possible for participants to gaze idly between the images, the areas captured by our AOIs did not encompass the entire monitor display. Any on-screen fixations that fell outside the AOIs were coded as looks away. Trials in which participants failed to attend to any AOI for at least two-thirds of the recorded story were removed (Adults: 7%, Children: 2%). Next, to ensure that there would not be excessive data loss in our planned analysis windows, we calculated the proportion of away looks within a time window starting at the onset of the pronoun and ending at the offset of the modifier in the critical sentence. Trials in which participants failed to attend to any AOI for at least two-thirds of this region were removed (Adults: an additional 7% of trials, Children: 7%). Following this two-step data processing procedure, we removed participants who did not retain a minimum of 4 trials (50%) within each story condition (Adults: 6 participants, Children: 4 participants). For adults, this resulted in the retention of 42 subjects and 938 trials; we removed 7% of trials for retained adult participants. For children, this resulted in the retention of 80 total participants and 1,770 total trials; we excluded 9% of trials for the retained child participants. In all analyses below, we offset measurement time windows by 200ms for adults and by 300ms for children, to account for the time it takes to plan and execute an eye movement.

2. Results

2.1. Fixations to the Target Referent

Figure 2 shows the time-course of fixations to the target and competitor animals from the onset of the pronoun to the average offset of the modifier in the critical sentence ("She asked [the ANIMAL] with the [MODIFIER]"). As Figure 2 shows, in the Ambiguous condition, adult and child participants looked roughly equally to the target and competitor images until they heard the disambiguating modifier. This reflects the expected temporary ambiguity of the phrase in this condition. However, in both the Two-referent and One-referent Unambiguous conditions, adults and children looked more at the target animal even *before* hearing the modifier. This early target preference reflects the informativeness of the animal name itself in the One-referent condition. However, in the Two-referent condition, any early preference for the target animal suggests a situation-derived inference about who Maisy could ask to play with her, given who remains on the scene.

These patterns were tested in a planned analysis window that began at the onset of the first determiner in the critical sentence and ended at the average onset of the modifier. We call this window the 'noun window'. We averaged across the noun window to create a target advantage score for each participant and trial. The

target advantage score was calculated by subtracting the proportion of looking toward the competitor from the proportion of looking toward the target (see Table 2 for condition means).

Table 2. Average target advantage score in the noun window in each condition for child and adult participant groups

	Adults		Children	
Condition	Mean	SD	Mean	SD
Ambiguous	.03	.46	.06	.56
Two-referent	.10	.48	.11	.54
One-referent	.15	.47	.13	.53

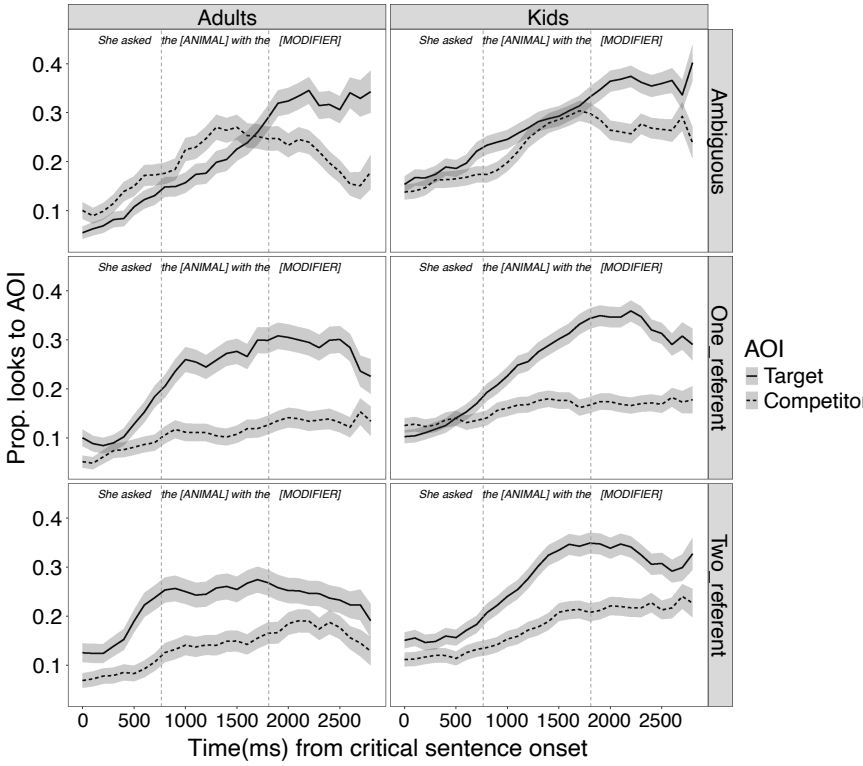

Figure 2. Mean proportion fixations to target and competitor animals; shading shows by-subject standard error. Vertical lines represent the average onset of the first determiner and the disambiguating modifier. Each figure reads, "She asked the [ANIMAL] with the [MODIFIER]."

We ran two mixed effects regressions, one for adults and one for children, in R (R Core Team, 2020) using the *lmer4* package (Bates et al., 2015). P-values

were calculated using the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2016). The models included target advantage score as the outcome measure and condition as a fixed effect. We used weighted effects codes for the condition variable to compare the critical Two-referent Unambiguous trials both to the Ambiguous and One-referent Unambiguous trials.[1] The first comparison tested whether listeners showed a larger target advantage in Two-referent Unambiguous trials (in which only one candidate referent remained on the scene) than in Ambiguous trials. The second comparison tested whether listeners showed a smaller target advantage in Two-referent Unambiguous trials than in One-referent Unambiguous trials (in which the head noun was sufficient to identify the target). The models also included subject and item random intercepts.[2] This represents the maximal random effects structure permitted by the data.

Both adults and children showed a reliably greater target advantage in the Two-referent Unambiguous than in the Ambiguous trials (Adults: $b = .15$, SE = .03, $t = 4.38$, $p < .001$; Children: $b = .07$, SE = .03, $t = 2.52$, $p = .02$). Adults and children also showed a reliably larger target advantage in One-referent Unambiguous than Two-referent Unambiguous trials (Adults: $b = .13$, SE = .03, $t = 4.06$, $p < .001$; Children: $b = .07$, SE = .03, $t = 2.34$, $p = .02$). These findings show that as children and adults listen to a story, they determine reference using not only linguistic information (e.g., how often, and how recently each candidate referent has been mentioned), but also their current model of the situation.

2.2. Memory Question Accuracy

After each story, children were asked 2-3 questions to probe recall for the activity Maisy did, which animal character she played with, and (in Two-referent and One-referent Unambiguous trials) which animal character had left the scene. Children answered by pointing to an image or speaking their answer; see Table 3 for proportion of correct answers.[3] Children generally remembered what Maisy did at the end of each story. However, accuracy in remembering who Maisy did the activity with varied by story condition. Children were most accurate at remembering who Maisy played with in the One-referent Unambiguous condition and least accurate in the Ambiguous condition. In Two-referent Unambiguous trials, children could use their knowledge of which animal had left the scene to help them remember who Maisy played with, providing additional support for the conclusion that children determine reference using their model of the situation.

---

[1] A second model was run for each participant group to compare One-referent to Ambiguous trials. Target advantage score was greater for One-referent trials than Ambiguous trials in both adults, $b = .07$, SE = .03, $t = 2.23$, $p = .03$, and children, $b = .06$, SE = .03, $t = 2.34$, $p = .02$.

[2] Random slopes were removed because they were estimated to be near 0.

[3] Memory question data for 4 participants is missing and not included in these proportions because the child did not answer the questions (1), appeared to purposefully give incorrect answers (1), or the experimenter forgot to video-record the session (2).

Finally, participants were accurate in remembering who left the scene during unambiguous trials. However, we note that the experimenter offered a correction for the second question (*Who did Maisy do it with?*) if the child answered incorrectly. Thus, this question may reflect memory for what the experimenter recently told the child instead of or in addition to memory for the story.

Table 3. Mean proportion and standard deviation of correct answers to memory questions by story condition

	Ambiguous		Two-referent		One-referent	
Question	M	SD	M	SD	M	SD
Q1. What did Maisy do?	.92	.15	.92	.15	.92	.14
Q2. Who did Maisy do it with?	.81	.19	.88	.18	.92	.13
Q3. Who had to leave early?	--	--	.90	.20	.94	.18

3. Discussion

The present work asked whether children could use their model of a situation to make inferences about unmentioned events, thus generating expectations about upcoming reference. We found that adults and preschoolers used their current model of the situation to constrain the interpretation of a temporarily ambiguous phrase during online reference comprehension. When two potential referents were introduced in the story, but one left the scene prior to the critical sentence, participants looked to the target animal before hearing the disambiguating modifier. In contrast, during ambiguous stories in which two potential referents were introduced and no one left the scene, participants could not make anticipatory fixations to the target referent. Note that the linguistic prominence of event participants was equated across stories. Each animal character was mentioned the same number of times, and we counterbalanced whether the target animal was introduced first or second. This suggests that participants generated a situation-based inference about the likely referent based on the event information in the story—that one animal left early and therefore couldn't be asked to play.

This finding complements evidence, discussed in the introduction, that children appear to make pragmatic inferences in reference comprehension. In offline tasks, children generate inferences based on the topic of the discourse to guide the interpretation of a novel word (Sullivan & Barner, 2016). Children also use the event structure, as indicated by the verb type or a discourse connective, to guide referential processing (Pyykkönnen, Matthews, & Jarvikivi, 2010; Kehler, Hayes, & Barner, 2011; Sullivan et al., 2019). Children construct situation models during narrative comprehension (Lynch et al., 2008; Rall & Harris, 2000), and activate recently provided event information to predict an upcoming linguistic argument (Borovsky et al., 2014). Building on this work, our findings demonstrate that children are capable of performing causal reasoning over their representation of a situation to generate inferences about unmentioned events during online language processing.

Together, the prior work and our results suggest that children create and incrementally update mental representations of events as the discourse describes

them. These representations include spatial, temporal, and causal information. Our data show for the first time that preschoolers can use these representations in online language comprehension to reason about likely referents.

The present work raises several new questions regarding the nature of children's pragmatic understanding. First, why do children succeed in this task when they struggle to draw other pragmatic inferences during online comprehension? Lee and colleagues (2017) found that 7-year-old children struggle to maintain predictions from a fictional narrative when the information conflicts with their world knowledge. For example, children listened to a story about a fairy that liked to eat snow. In a critical story sentence, children struggled to maintain anticipatory fixations to the target snack (snow), if a more typical patient of the verb *eat* (cake) was also visible in the display. While children initially looked to the discourse-congruent referent (snow), this anticipatory effect diminished across the verb phrase. Additionally, children with larger receptive vocabularies spent less time considering the discourse-congruent referent than those with smaller receptive vocabularies. This suggests that children in this task had difficulty overriding expectations formed on the basis of their existing lexical-semantic relationships and/or world knowledge. One potential explanation for the discrepancy between this result and the present work is that in our experiment, the discourse-provided event information did not compete with stored semantic and world knowledge. Children only needed to generate an inference based on who remained in the scene and therefore who was available to ask.

Another question that arises from this work is under what circumstances children can integrate discourse-provided facts with presuppositions of verbs in referent prediction. The verb *ask* strongly implies that someone is present to ask; however, not all verbs carry the same presupposition. For instance, listeners should not anticipate a referent who is still present in the story, given the verb *remember* in (4).

(4) She remembered that the bunny with the crown…

Adults integrate situation-derived expectations with verb meanings during comprehension. For example, adults can use the semantic characteristics of the verb *return* in "return the block to area 3" to constrain their attention to blocks that had been previously moved. And, while playing a referential communication game, adults attended to objects that the speaker had seen previously moved (Chambers & San Juan, 2008). This suggests that adults' referential predictions consider not only linguistic statistics and situation-specific properties of referents, but also the verb-relevant properties of potential referents (i.e., which object had been moved). As children learn the presuppositions of verbs, we might expect their referential predictions to reflect them in the same manner. That is, children might consider relevant background knowledge (e.g., which objects have been moved) when anticipating what objects could be *put back,* as opposed to *put*.

Finally, we can ask to what degree a situation model is involved in earlier language comprehension and learning. Situation models are a workspace that

permits uniform causal reasoning over multiple sources of knowledge in order to generate online predictive inferences. In fact, causal reasoning over event knowledge is also involved in non-linguistic domains. For example, 25-month-olds look toward a location where they expect an agent to search for her toy, even when the agent has a false belief about the toy's location (Southgate, Senju, & Csibra, 2007). In the current study and in the false-belief task, children incrementally build an event representation and reason over that representation to make predictive inferences. For this reason, we speculate that a situation model might be involved in earlier language comprehension and learning.

References

Akhtar, Nameera. (2002). Relevance and early word learning. *Journal of Child Language, 29*, 677-686. doi: 10.1017/S0305000902005214

Arnold, Jennifer E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes, 31*(2), 137-162. doi: 10.1207/S15326950DP3102_02

Arnold, Jennifer E., Eisenband, Janet G., Brown-Schmidt, Sarah & Trueswell, John C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76*, B13-B26. doi: 10.1016/S0010-0277(00)00073-1

Arnold, Jennifer E., Strangmann, Iris M., Hwang, Heeju, Zerkle, Sandra & Nappa, Rebecca. (2018). Linguistic experience affects pronoun interpretation. *Journal of Memory and Language, 102*, 41-54. doi: 10.1016/j.jml.2018.05.002

Bates, Douglas, Maechler, Martin, Bolker, Ben & Walker, Steve. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi: 10.18637/jss.v067.i01.

Bohn, Manuel, Le, Khuyen Nha, Peloquin, Benjamin, Köymen, Bahar & Frank, Michael C. (2020). Children's interpretation of ambiguous pronouns based on prior discourse. *Developmental Science*. doi: 10.1111/desc.13049

Borovsky, Arielle, Sweeney, Kim, Elman, Jeffery L. & Fernald, Anne. (2014). Real-time interpretation of novel events across childhood. *Journal of Memory and Language, 73*, 1-14. doi: 10.1016/j.jml.2014.02.001

Brainard, David. H. (1997). The Psychophysics Toolbox, *Spatial Vision*, *10*, 433-436.

Chambers, Craig G. & San Juan, Valerie. (2008). Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition, 108*, 26-50. doi: 10.1016/j.cognition.2007.12.009

Ehrlich, Kate & Rayner, Keith. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22*(1), 75-87. doi: 10.1016/S0022-5371(83)80007-3

Gernbacher, Morton Ann. & Hargreaves, David J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27*(6), 699-717. doi: 10.1016/0749-596X(88)90016-2

Hartshorne, Joshua K., Nappa, Rebecca & Snedeker, Jesse. (2015). The development of the first-mention bias. *Journal of Child Language, 42*(2), 423-446. doi: 10.1017/S0305000914000075

Hartshorne, Joshua K., O'Donnell, Timothy J. & Tenenbaum, Joshua B. (2015). The causes and consequences explicit in verbs. *Language, Cognition, and Neuroscience, 30*(6), 716-734. doi: 10.1080/23273798.2015.1008524

Horowitz, Alexandra C., & Frank, Michael C. (2015). Young children's developing sensitivity to discourse continuity as a cue for inferring reference. *Journal of Experimental Child Psychology, 129,* 84-97. doi: 10.1016/j.jecp.2014.08.003

Kehler, Andrew & Rohde, Hannah. (2013). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics, 39*(1-2), 1-37. doi: 10.1515/tl-2013-0001

Kuznetsova, Alexandra, Brockhoff, Per B. & Christensen, Rune H.B. (2017). lmerTest: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1-26. doi: 10.18637/jss.v082.i13

Lee, Heeyoung, Chang, Angel, Chambers, Nathanael & Jurafsky, Dan. (2013). Deterministic coreference resolution based on entity-centric, precision-ranked rules. *Computational Linguistics, 39*(4), 885-916. doi: 10.1162/COLI a 00152

Lee, Ruth, Chambers, Craig G., Huettig, Falk & Ganea, Patricia A. (2017). Children's semantic and world knowledge overrides fictional information during anticipatory linguistic processing. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society (CogSci 2017)*

Lidz, Jeffrey, Waxman, Sandra & Freedman, Jennifer. (2003). What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition, 89,* B65-B73. doi: 10.1016/S0010-0277(03)00116-1

Lynch, Julie S., Van Den Brock, Paul, Kremer, Kathleen, Kendeou, Panayiota, White, Mary Jane & Lorch, Elizabeth P. (2008). The development of narrative comprehension and its relation to other early reading skills. *Reading Psychology, 29*, 327-365. doi: 10.1080/02702710802165416

Nieuwland, Mante S., Otten, Marte & Van Berkum, Jos J.A. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of Cognitive Neuroscience, 19*(2), 228-236. doi: 10.1162/jocn.2007.19.2.228

Pyykkönen, Pirita, Matthews, Danielle & Järvikivi, Juhani. (2010). Three-year- olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution, *Language and Cognitive Processes*, *25*(1), 115-129. doi: 10.1080/01690960902944014

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rall, Jaime & Harris, Paul L. (2000). In Cinderella's slippers? Story comprehension from the protagonist's point of view. *Developmental Psychology, 36*(2), 202-208. doi: 10.LO37//0O12-1649.36.2.202

Richmond, Lauren L. & Zacks, Jeffrey M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences, 21*(12), 962-980. doi: 10.1016/j.tics.2017.08.005

Saylor, Megan M., Ganea, Patricia A. & Vázquez. (2011). What's mine is mine: Twelve-month-olds use possessive pronouns to identify referents. *Developmental Science, 14*(4), 859-864. doi: 10.1111/j.1467-7687.2010.01034.x

Song, Hyun-joo & Fisher, Cynthia. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language, 52*, 29-57. doi: 10.1016/j.jml.2004.06.012

Song, Hyun-joo & Fisher, Cynthia. (2007). Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua, 117*(11), 1959-1987. doi: 10.1016/j.lingua.2006.11.011

Southgate, Victoria, Senju, Atsushi & Csibra, Gergely. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7). doi: 10.1111/j.1467-9280.2007.01944.x

Sullivan, Jessica & Barner, David. (201). Discourse bootstrapping: Preschoolers use linguistic discourse to learn new words. *Developmental Science, 19*(1), 63-75. doi: 10.1111/desc.12289

Sullivan, Jessica, Boucher, Juliana, Kiefer, Reina J., Wiliams, Katherine & Barner, David. (2019). Discourse coherence as a cue to reference in word learning: Evidence for discourse bootstrapping. *Cognitive Science, 43*. doi: 10.1111/cogs.12702

Zwaan, Rolf A. & Radvansky, Gabriel A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 122*(2), 162-185. doi: 10.1037/0033-2909.123.2.162

Proceedings of the 45th annual Boston University Conference on Language Development

edited by Danielle Dionne
and Lee-Ann Vidal Covas

Cascadilla Press Somerville, MA 2021

ISSN 1080-692X
ISBN 978-1-57473-067-8 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com