A Look "Inside" Children's Real-time Processing of Spatial Prepositions

Zoe Ovans[1], Barbara Landau[2], Heesu Yun[1], Sarah Yi[1], John Trueswell[1]
[1]University of Pennsylvania, [2]Johns Hopkins University

Abstract

A wealth of evidence indicates that children use their developing linguistic knowledge to incrementally interpret speech and predict upcoming reference to objects. For verbs, determiners, case-markers, and adjectives, hearing linguistic information that sufficiently constrains referent choice leads to anticipatory eye-movements. There is, however, limited evidence about whether children also use spatial prepositions predictively. This is surprising and theoretically important: spatial prepositions provide abstract semantic information that must interface with spatial properties of, and relations between, objects in the world. Making this connection may develop late because of the complex mapping required. In a visual-world eye-tracking task, we find that adults and 4-year-olds hearing 'inside' (but not 'near') look predictively to objects that afford the property of *containment*. We conclude that children make predictions about the geometric properties of objects from spatial terms that specify these properties, suggesting real-time use of language to guide analysis of objects in the visual world.

Keywords: Spatial prepositions; Sentence processing; Language development; Eye-tracking

Introduction

The task of real-time language comprehension is immense. The speech that we hear is rapid and capricious: adults speak at over 3 words per second (Chermak & Schneiderman, 1985), and listeners must parse (and sometimes re-parse) this input at the same rate, transforming the surface structure they hear into representations of meaningful sentences in a format that interfaces seamlessly with nonlinguistic representations of their environment. For young learners, whose executive function systems are still developing (Diamond, 2020), this task is even more daunting, yet typically-developing school-aged children are able to accomplish it.

The last few decades of psycholinguistic research have provided a great deal of evidence that children (and even toddlers) can interpret speech incrementally and can use their developing linguistic systems to predict upcoming words and their reference to objects. For example, preschool-aged children have been shown to make rapid use of their sentence context to constrain real-time comprehension of nouns and verbs (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Mani & Huettig, 2012; Borovsky et al., 2012, *inter alia*). In a recent example, Gambi, Pickering & Rabagliati (2016) presented children with sentences such as "Pingu will ride the horse" and found that 3-6 year-old children were above chance at looking to nouns that were predicted by the verb before noun onset. For example, upon hearing the verb "ride," they looked to a horse, as it is a predictable object of riding (importantly, children did not look to a cowboy, which is also semantically related to the verb but not a predictable object). Other examples of children's predictive online processing capabilities include use of gender-marked determiners to predict noun reference in Spanish (Lew-Williams & Fernald, 2007), case markers in German (Özge, Kornfilt, Maquate, Küntay & Snedeker, 2022), number morphology on verbs in Italian (Bosch & Foppolo, 2023) and gender morphology on adjectives in Russian (Aumeistere, Bultena, & Brouwer, 2022). In each case, children are able to use linguistic information to constrain the set of possible referents they consider, as revealed by anticipatory eye-movements to those referents in eye-tracking studies.

As this literature shows, children are readily capable of anticipating reference in many cases. However, the processing of spatial prepositions has received much less attention in the literature on developmental sentence processing (Pierre & Johnson, 2021; Christou, Sanz-Torrent, Coloma, Guerra, Araya & Andreu, 2021). Adult comprehenders can use spatial prepositions predictively: Chambers, Tanenhaus, Eberhard, Filip & Carlson (2002) presented the first evidence that adults make online use of the information provided by spatial prepositions in order to constrain their domain of reference quite narrowly, to objects with particular spatial properties. Participants were presented with a display that contained eight objects: four large and four small. In one condition, one of the large objects was a container large enough to fit each of the small objects inside it, and in another condition three of the large objects were such containers. Participants heard sentences like "Put the whistle [a small object] inside/below the can [a container]," while their eye-movements were recorded. Chambers et al. found that in the "inside" condition, participants began looking to the can in the 1-container (but not the 3-container) condition before the offset of the preposition, indicating that they quickly integrated the information conveyed by the preposition with their visual context. Conversely, looks to the target (can) emerged late (after the onset of the noun) in the "below" condition for both 1-container and 3-container contexts, as the preposition did not allow participants to make a prediction about the upcoming referent. This experiment demonstrated that adults are able to use information provided by spatial prepositions in order to constrain their hypotheses about potential referents based on fairly specific visual information (e.g. features like +CONTAINER) during real-time comprehension.

Children's real-time processing of spatial prepositions

As discussed below, there is a significant gap in the literature regarding children's real-time use of spatial prepositions, with some of the existing work suggesting that these terms

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

may be difficult for learners to use predictively. This is surprising, as prepositions present a good candidate for a class of abstract terms that children may process in an adult-like manner: they are frequent in children's input (Meintis, Plunkett, Harris & Dimmock., 2002) and are produced early by learners (Tomasello, 1987). Additionally, offline tasks reveal that children appear to comprehend them in an adult-like manner from an early age (i.e. like adults, they are more likely to look to a typical instance of "on" than an atypical one, showing a fairly sophisticated understanding of the boundaries of on-ness, Meintis et al., 2002). These results raise the possibility that children's online comprehension of spatial prepositions may be similarly adult-like, however current evidence is mixed.

Pierre & Johnson (2021) investigated preposition processing with 2-year-old children, and found varying results. Children looked to appropriate referents on first mention for the prepositions *on* and *under* but not *in* or *next to.* However, as they presented children with a word-learning task in addition to the sentence processing task, the chance-performance results may be the result of a failure to map the novel word to the novel object, and not a failure to process the prepositions themselves. That is, in their task children not only had to interpret spatial prepositions in real time, they also had to use these representations to map novel words to novel objects. As such, failure to succeed at this task might be the result of either failure to process the prepositions or a more downstream failure at the word-referent mapping stage.

In another recent study, Christou et al., (2021) investigated real-time comprehension of spatial prepositions in typically developing children (as a control for testing children with Developmental Language Disorder (DLD)). They presented children with sentences like "El gato está bajo la mesa" (*The cat is below the table)* and found that as typically-developing children heard the preposition (and before the noun onset) they looked to appropriate referents (e.g. a cat underneath a table), indicating that they used information from the preposition to make their selections. It should be noted though that the children in their study were older (average ages of 7-8 years), and children by this age have been shown to be largely adultlike in many aspects of sentence processing, including even online use of executive function (Kidd & Bavin, 2005). The auditory stimuli children heard were also quite unnatural, as 1000 milliseconds of silence was inserted between each word, limiting the extent to which conclusions about real-time processing can be made. To the best of our knowledge, these studies constitute the only evidence as to whether children use spatial prepositions to constrain reference in online processing, yet neither were set-up to test this question directly.

The overall lack of evidence regarding children's processing of spatial prepositions is surprising, as the question of whether children do so in an adult-like manner is an important one. Spatial prepositions provide abstract information about the spatial properties of objects and the relations between them in the world. A ball is only *near* a cup by virtue of an abstract spatial relationship between the two objects, and can only be *inside* a cup when the latter has a specific set of geometric properties (e.g. a particular type of concavity and size). It is possible that the real-time mapping of these abstract relations to children's linguistic system occurs late, as it requires children to quickly integrate information from two very disparate mental systems. In addition, studying children's anticipatory processing as a result of hearing spatial prepositions provides a crucial testbed for questions of developmental language comprehension more generally, as measuring children's understanding of the preposition itself (prior to hearing information about the object noun) provides a relatively pure measure of children's understanding of spatial terms.

Notably, all of the child work on real-time processing of spatial prepositions has been done with pictorial stimuli, which only indirectly depict spatial information, whereas Chambers et al. tested adults acting on physically co-present objects. Given that spatial information is intimately connected to perception-action interfaces in complex ways (Bertenthal, 1996; Fajen & Phillips, 2013; Gottlieb, 2007), the study of real-time use of spatial prepositions may be most successfully examined in the context of children acting on physical objects. This may be crucial for revealing and fully understanding how space and language are connected and deployed developmentally.

In the two studies below, we test whether 4-year-old children are adult-like in their real-time processing of spatial prepositions (that is, whether they integrate them into a sentence representation and a representation of their referential context as they hear them). We ask whether children are able to use the information provided by the preposition, along with the information provided by their spatial reasoning system (e.g. the knowledge that containers have to possess a particular set of properties) to guide their real-time hypotheses about upcoming words and their subsequent referents. This work was done with participants acting on physically co-present objects, using head-mounted eye-tracking. In Experiment 1, we first carry out a conceptual replication of Chambers et al.,'s (2002) findings on adults using an experimental paradigm that was created to also be appropriate for testing children. Upon finding that the results for adult participants confirm those of the original Chambers et al., study, Experiment 2 tests this updated paradigm on 4-year-old children.

Experiment 1: Conceptual replication of Chambers et al., (2002)

Participants

Sixteen adult participants recruited from the University of Pennsylvania Psychology subject pool participated for course credit. Three additional participants were tested, however their data were not included in the analysis either because they reported themselves to be bilingual and not English-dominant (2 participants) or because of equipment failure leading to data loss (1 participant). The remaining 16

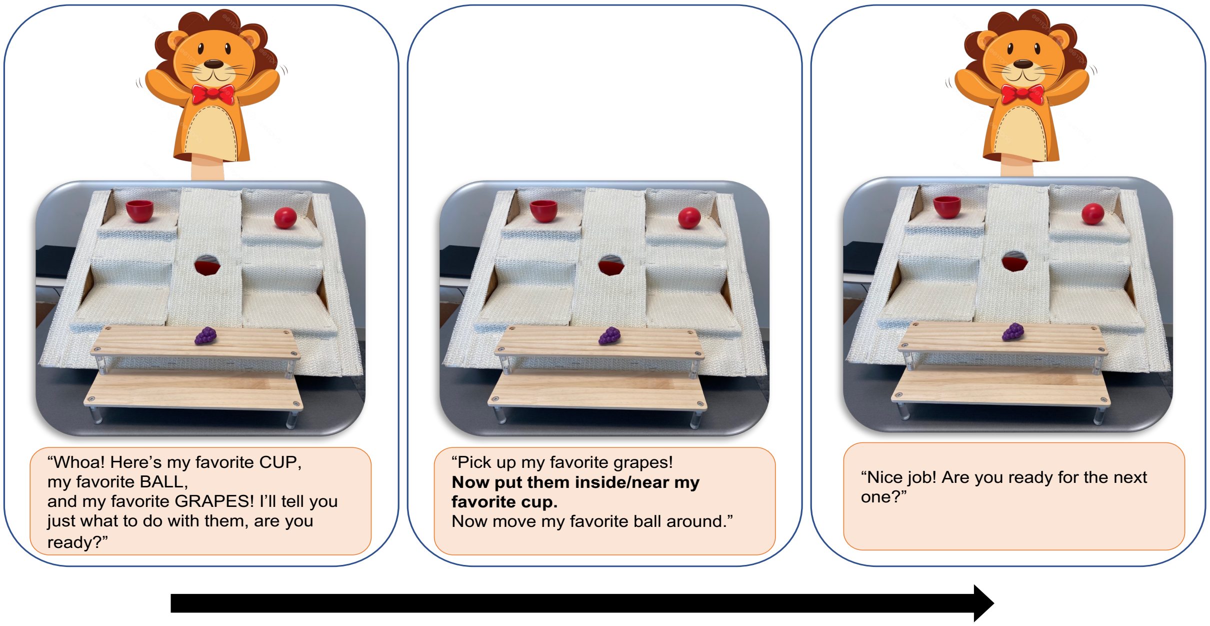

Figure 1: Schematic of a typical target trial in Experiments 1 & 2, with target sentence bolded.

participants reported themselves to be monolingual English speakers with no known cognitive disabilities and were provided with course credit for their participation in the experiment.

Design & Stimuli

As we aimed to compare results from adult participants directly to those of children, adults were tested in a manner appropriate for children. In order to make the experiment suitable for child participants, several changes from the initial Chambers et al., (2002) set-up were made. First, the contrast between 1 vs. 3-container trials was eliminated, and only 1-container trials were included in the present studies. The two conditions, informative ("inside") and uninformative ("near") were instead compared directly.

The number of items on each trial was also reduced: while Chambers et al. had 4 potential small items and 4 large items that served as either containers or distractors, the present study paired this down to three items per trial: one container, one distractor, and one small item below the container and distractor. Reducing the number of items was done in order to reduce distractions and increase target looks overall Containers and distractors were always painted the same color in order to set them apart from the small objects, to distract from the true purpose of the experiment, and to ensure that objects looked relatively uniform, with no distractingly shiny surfaces, visible writing, etc. While the container and distractor objects were sometimes semantically related (e.g. a bowl and a plate), the small objects were always semantically unrelated to them on a given trial. Container objects all had either a visible cavity or a top that was shown to be removeable. Distractor objects were flat or otherwise visibly not suitable containers. Containers and distractors were generally similar in size, and small objects were sized such that it was clear that they would fit into the containers without effort.

All objects were chosen such that 4-year-olds could be expected to know their labels, and these labels were confirmed by MCDI norms where possible (Fenson et al., 2007). The objects used were always inanimate to avoid animate objects being particularly tempting targets for children (some semi-animates such as plants and fruits were included as these were not deemed overly distracting).

Trials consisted of either 3 or 4 sentences, with target trials always containing 3 sentences and fillers containing 3 or 4 sentences. Regardless of trial type, the first sentence was always an instruction to pick up the small object (e.g. "pick up my favorite grapes"). For target trials, the second sentence was always the target sentence (e.g. "Now put them inside/near my favorite cup"). Then, the final sentence for target trials always targeted the distractor so that all items were mentioned on every trial (e.g. "Now roll the ball around"). Preposition condition ("inside" vs. "near") varied within subjects, but for a given trial was counterbalanced between subjects along with trial order. On target trials, the second sentence always mentioned the container regardless of preposition condition (otherwise, participants might have learned over the course the experiment that "near" could refer to the distractor object while "inside" always referred to the container). The addition of the adjective phrase "my favorite" was included in every trial in order to extend the window of analysis beyond the preposition itself.

Filler trials were visually identical to targets, and differed in that the distractor was always mentioned in the second sentence and the container in the third. This was done so that participants could not learn over the course of the experiment

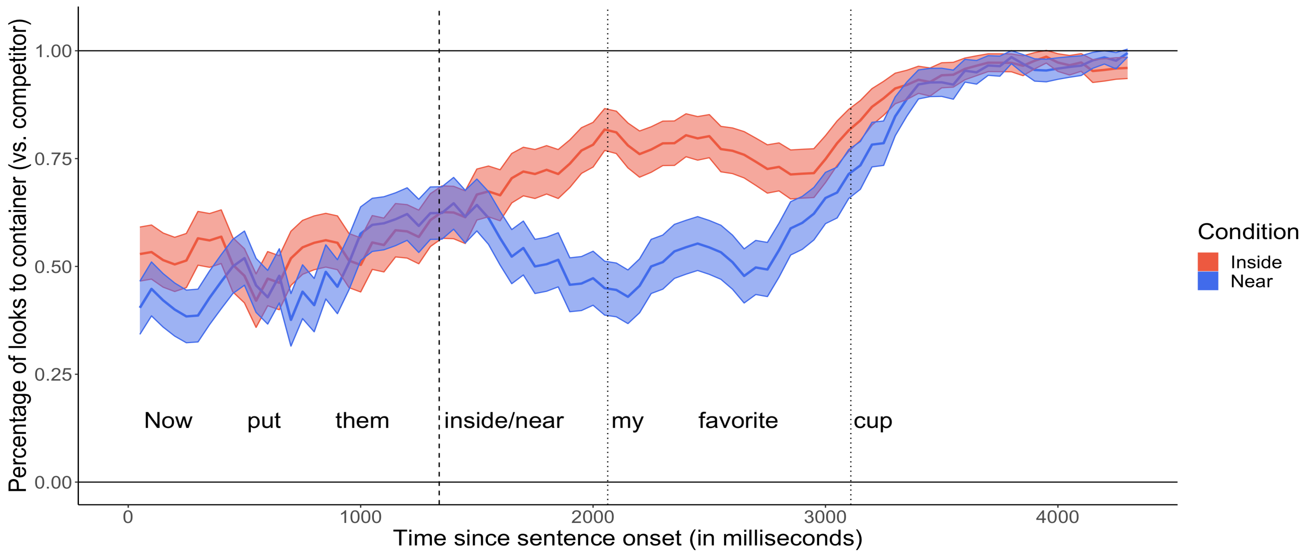

Figure 2: Eye-tracking results from Experiment 1 (Adults), showing looks to the container objects vs the competitor. Vertical lines indicate average word onsets of the preposition, adjective phrase, and noun.

that the container was likely to be referred to second (e.g. "Pick up my favorite triangle"/"Now put it under my favorite leaf"/"Now put my favorite pot on top of them both"). Filler sentences sometimes used other spatial prepositions, but did not contain inside or near. A fourth sentence was also included on some filler trials to make the experimental design less predictable. As with the target trials, all items were eventually mentioned on filler trials.

Overall, the experiment contained 10 target trials and 10 filler trials, pseudorandomly interspersed such that no more than three test or filler trials occurred in a row, and the experiment always began and ended on a filler trial to reduce introductory or wrap-up effects.

Procedure

After providing informed consent, participants were seated in front of a sloped display that contained cut-outs where objects could be placed (see Figure 1). Participants wore Tobii Glasses 2 eye-tracking glasses during the course of the experimental session (Tobii Technology, Danderyd, Sweden). Adults were first informed that they were participating in a study that was ultimately designed to be appropriate for children. Following a short calibration, participants were introduced to a hand puppet who appeared behind the display (see Figure 1). The puppet introduced the game with the following dialogue: "Hi! I'm Sally, and today, we're going to play with all of my favorite things! I'll show you just how I like to play with them. Can you help me?" After assent from the participant, the puppet continued "Great! I have small hands, so my assistant's going to help me out. Are you ready for the first one?" After further assent, the assistant (a confederate experimenter) added items to the display always in the order Top left (container or distractor), Top right (container or distractor), bottom middle (small object). Container and distractor position was counterbalanced throughout the experiment, but was always consistent for a given trial (e.g. the cup was always on left, and ball was always on the right). The order of trial presentation was varied across lists such that half of the participants saw the trials in the reversed order.

As the confederate introduced each item, the puppet labeled it, saying e.g. "Wow, that's my favorite cup! And wow, that's my favorite ball. And wow, those are my favorite grapes!" Once the three items were introduced, the puppet then said "I'll tell you just what to do with these. Are you ready?" Once the participant agreed again, the puppet would duck behind the display and a series of pre-recorded sentences (described in the design section above) were played in the puppet's voice (from a laptop hidden behind the display board). Sentences were played one at a time, with the next one being played once the participant had finished an instruction. As such the timing of sentence onsets differed slightly across participants, however the timing of words presented within each sentence remained consistent.

After every trial, the puppet provided the participants with positive feedback and asked if they were ready for the subsequent trial (e.g. "Great job! Are you ready for the next one?"). At the conclusion of the experiment adults were asked what they thought the study was testing, to determine whether the preposition manipulation was transparent to participants. The entire procedure took approximately 15 minutes to complete.

Results

Data analysis and coding. Eye-tracking data was analyzed using Tobii Pro Lab (Tobii Pro, 2014) software, with area of interest boundary boxes (AOIs) placed around each the container, distractor, and small objects. AOIs completely covered the objects with a few inches of buffer space surrounding them, and were manually coded every few frames to ensure they followed the objects, compensating for participants' head movements and movements of the objects themselves. Audio recordings from the video files were analyzed in Praat (Boersma, 2001) to determine sentence and

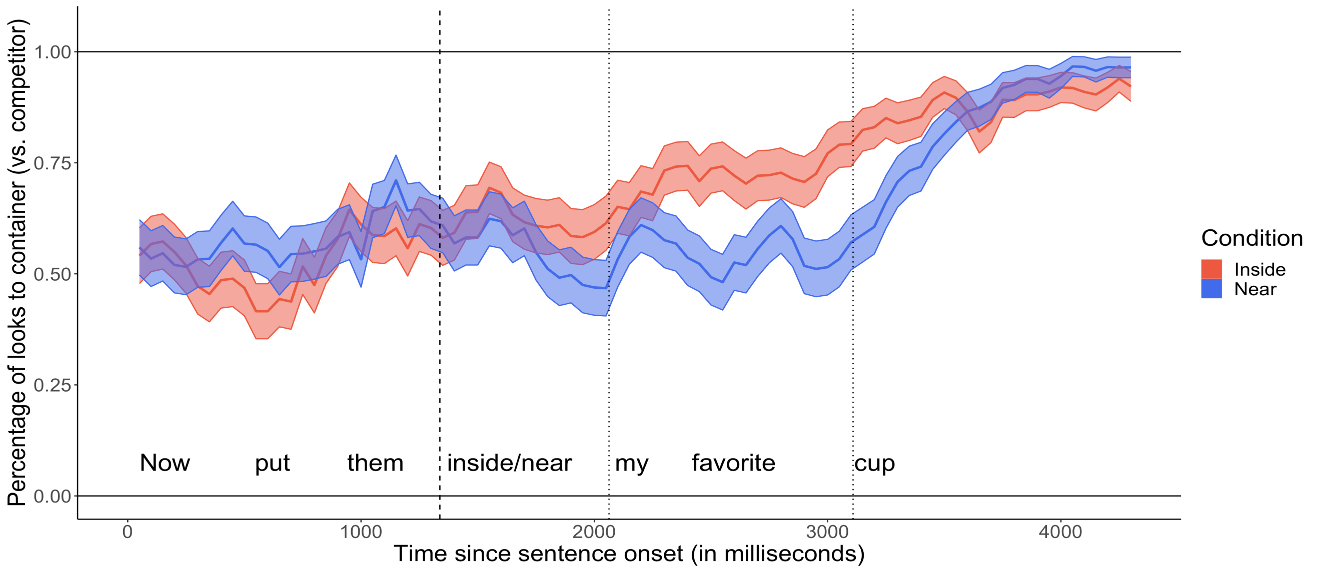

Figure 3: Eye-tracking results from Experiment 2 (Children), showing looks to the container objects vs the competitor. Vertical lines indicate average word onsets of the preposition, adjective phrase, and noun.

word onsets.

An informal analysis of participants' guesses regarding what the experiment was testing indicated that participants were naïve to the preposition manipulation, with most participants reporting that they believed the study to be on color processing or general ability to follow directions.

Eye-tracking results. We analyzed participants' looks to the container object vs. the competitor object in a preregistered time window beginning at the onset of the preposition and ending at the onset of the noun. For this analysis, looks elsewhere in the visual scene were excluded. A logistic mixed effects analysis with random intercepts and slopes for participants and items revealed that adults were significantly more likely to look to the container object in the *inside* condition than in the *near* condition during this time window (β= 1.07; SE = 0.32; z = 3.39; $p < .001$). These model results were further confirmed with a cluster-based permutation test (Maris & Oostenveld, 2007) using the jlmerclusterperm R package (Choe, 2023), on the first 5 seconds of the test trials. With 1000 simulations, this test revealed a significant condition difference between 1800 and 3200 milliseconds after sentence onset ($t > 1.5$, $p < .05$).

Discussion

The results of Experiment 1 constitute a successful conceptual replication of the results of Chambers et al., (2002). Additionally, several changes from the 2002 study add corroboration to the original findings. Our analysis did not rely on cumulative looks to the target, and thus provides a more detailed view of the time-course of processing (analyses of cumulative looks, on the other hand, ignore the information present in participants' switching their gaze targets over time). Our analysis reveals that adults were near chance in looking to the distractor object in our time window, indicating that our containers were not generally more interesting or salient than our distractor objects. Finally, extending the window of analysis with an adjective phrase ("my favorite") allows us to more clearly see the effect of the preposition over time.

These results indicate that adults are able to integrate the information provided by spatial prepositions with the information provided by their spatial reasoning system. We infer that increased looks to the container objects in the informative "inside" condition are the result of participants' using information from the preposition as well as object geometry to guide their real-time interpretation of the sentence. In order to determine that a cup is likely to be mentioned, adults had to evaluate the geometry of the objects, including the curvature of the sides of the containers and the relative sizes of their openings. Experiment 2 extends these results by testing 4-year-old children on the same paradigm.

Experiment 2: Processing in 4-year-olds

Following the result of Experiment 1, that adults' online sensitivity to preposition choice is readily observable in our experimental set-up, Experiment 2 applied the same method to the testing of 4-year-old children. The stimuli for Experiment 2 were identical to those used for Experiment 1. Pilot testing revealed that the materials and length of the experiment was suitable for children in our age range.

Participants

16 four year-old children (4;0-5;0, average age 4;7) were recruited from area preschools. One additional child was tested, but was excluded due to high trackloss (over 30% of gaze samples were untracked). Children's caregivers provided informed consent, and children provided assent to participate before beginning the study. Following their participation, children were provided with a small book to thank them for their time.

Procedure

Children were tested in a quiet room in their preschool, following the procedures of Experiment 1. In the event that the glasses were too large to reliably stay on children's heads, a small headband was used to secure the eye-tracking apparatus. As with adults, the experimental session lasted about 15-20 minutes, and children reported generally enjoying the game.

Results

Data analysis and coding. While in our preregistration we noted that children would be excluded if they failed to perform the act-out actions for more than 50% of filler trials, no participants had to be excluded for this metric as children generally understood the instructions. AOI and audio coding was done in an identical manner to Experiment 1.

Eye-tracking results. We analyzed children's eye-movement data in an identical way to the adult data in Experiment 1. A logistic mixed effects model with random intercepts and slopes for items (the maximal model that converged) revealed that children, like adults, were significantly more likely to look to the container object in the *inside* condition than in the *near* condition during our preregistered time window (β= .686; SE = 0.25; z = 2.73; $p < .01$). These model results were further confirmed with a cluster-based permutation test on the first 5 seconds of the test trials, which revealed a significant condition difference between 2400 and 3400 milliseconds from sentence onset ($t > 1.5$, $p < .05$). When combined with adult data from Experiment 1, no effect of age-group or interaction between age-group and condition was found (*p*s > .1).

General Discussion

The results presented here support an account in which children, like adults, leverage the linguistic information available to them in processing spatial prepositions to restrain their referential context, and make anticipatory looks to appropriate referents. To do this, children and adults must integrate the preposition into a sentence context, e.g. to know the subject noun must fit inside the object noun, and not the other way around. This is notably a feat that even the most sophisticated artificial language systems have difficulty with (Conwell & Ullman, 2022). Children then had to leverage information from their visual system (e.g. that they were looking at objects with container-like properties), and combine both sources of information to constrain their predictions about potential referents, before launching an anticipatory eye-movement, all within a fraction of a second.It should be noted that children's looks were likely not guided solely by the likelihood of lexical co-occurrence between our prepositions and nouns. If that were the case, one might expect that children would make more anticipatory looks to containers when the container nouns were more frequent words, as they would have had more opportunity to learn these co-occurrences. A post-hoc analysis revealed no correlation between container looks in our study and Google Ngram (Michel et al., 2010) word frequency (β= -.005 [-0.47 to 0.46], p =.98). This indicates that children's semantic interpretations of the prepositions they heard, and not simply their likelihood to co-occur with particular objects, guided their interpretation of reference in our task.

One caveat to this claim that children were adult-like in our study is that the condition effect for children was approximately 600 milliseconds delayed compared to that of adults. While it is well-known that saccade latencies decrease with age (e.g. Bucci & Seassau, 2012), this delay is still longer than expected by general age differences in eye-movement planning. While we leave open the possibility that this delay reflects a mechanistic difference in children's processing strategies from that of adults, we cannot rule out a simpler explanation: children were tested in school settings that were, by their nature, noisier than the quiet lab setting in which adults were tested. The delay we observe may be similar to those often found with hearing in noise (e.g. Ben-David et al., 2011).

Future work will further specify the extent to which children make use of their spatial reasoning system during online comprehension. To even more definitively rule out a low-level lexical co-occurrence interpretation, (e.g. that children simply have stronger associations between prepositions like inside and the nouns denoting containers), a planned follow-up version of the current study will display container objects upside-down (or otherwise visually un-openable). The predictions are as follows: If children are indeed integrating the information provided by their visuo-spatial processing system when interpreting prepositions in real time, and this system is hampered from making inferences about containers, we should no longer see a large effect of preposition choice on container looks. If however children are merely relying on word associations instead of the semantics of the preposition, the results should look identical to the current ones.

Other planned future work will determine whether children make even more fine-grained distinctions between types of spatial configurations, such as more canonical or "core" spatial relations (e.g. a mug *on* a plate, where the figure object is fully supported by the ground object) vs. less canonical representations (e.g. a mug *on* a hook, where the majority of the figure object is below the ground). Finding that even these fine-grained distinctions are considered would further support the notion that children's real-time processing of spatial prepositions occurs in a deep and adult-like manner.

The current work serves to replicate and extend the results of Chambers et al., (2002) by demonstrating that adults and children both make use of the meaning of spatial prepositions and the visuo-spatial information in their referential context to update their hypotheses about reference in real-time. This conclusion provides a new way in which children are adult-like in their sentence processing abilities, despite resource limitations such as memory and attention. Furthermore, it indicates that children are able to leverage the abstract non-linguistic representations from their visuo-spatial reasoning system with real-time linguistic processing more generally.

References

Aumeistere, A., Bultena, S., & Brouwer, S. (2022). Wisdom comes with age? The role of grammatical gender in predictive processing in Russian children and adults. *Applied Psycholinguistics, 43*(4), 867-887.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research, 54*, 243-262.

Bertenthal, B. I. (1996). Origins and early development of perception, action, and representation. *Annual review of psychology*, *47*(1), 431-459.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot. Int.*, *5*(9), 341-345.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of experimental child psychology, 112*(4), 417-436.

Bosch, J. E., & Foppolo, F. (2023). Prediction during spoken language processing in monolingual and multilingual children: Investigating the role of literacy. *Linguistic Approaches to Bilingualism.*

Bucci, M. P., & Seassau, M. (2012). Saccadic eye movements in children: a developmental study. *Experimental brain research, 222*, 21-30.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of memory and language, 47*(1), 30-49.

Chermak, G. D., & Schneiderman, C. R. (1985). Speech timing variability of children and adults. *Journal of Phonetics.*

Choe, J. (2023). jlmerclusterperm: Cluster-Based Permutation Analysis for Densely Sampled Time Data. R package version 1.1.0. https://cran.r-project.org/package=jlmerclusterperm.

Christou, S., Sanz-Torrent, M., Coloma, C. J., Guerra, E., Araya, C., & Andreu, L. (2021). Real-time comprehension of Spanish prepositions and prepositional locutions in bilingual children with developmental language disorder: A study based on eye-movement evidence. *International Journal of Language & Communication Disorders, 56*(1), 51-71.

Conwell, C., & Ullman, T. (2022). Testing relational understanding in text-guided image generation. *arXiv preprint arXiv:2208.00005*.

Diamond, A. (2020). Executive functions. In *Handbook of clinical neurology 173,* 225-240.

Fajen, B. R., & Phillips, F. (2013). Spatial perception and action.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. *Cognition, 157,* 340-351.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of child language, 14*(1), 23-45.

Gottlieb, J. (2007). From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron*, *53*(1), 9-16.

Kidd, E., & Bavin, E. L. (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretations of ambiguous sentences. *Journal of Child Language*, *32*(4), 855-876.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological science, 18*(3), 193-198.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods, 164*(1), 177-190

Meints, K., Plunkett, K., Harris, P. L., & Dimmock, D. (2002). What is 'on'and 'under'for 15-, 18-and 24-month-olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology, 20*(1), 113-130.

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Google Books Team, ... & Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *science*, *331*(6014), 176-182.

Özge, D., Kornfilt, J., Maquate, K., Küntay, A. C., & Snedeker, J. (2022). German-speaking children use sentence-initial case marking for predictive language processing at age four. *Cognition, 221,* 104988.

St. Pierre, T., & Johnson, E. K. (2021). Looking for Wugs in all the Right Places: Children's Use of Prepositions in Word Learning. *Cognitive Science, 45*(8), e13028.

Tomasello, M. (1987). Learning to use prepositions: A case study. *Journal of child Language*, *14*(1), 79-98.