Cognition 243 (2024) 105671

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Full Length Article

Language systematizes attention: How relational language enhances relational representation by guiding attention

Lei Yuan [a,*], Miriam Novack [b], David Uttal [c], Steven Franconeri [c]

[a] Department of Psychology and Neuroscience, University of Colorado Boulder, USA
[b] Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, USA
[c] Department of Psychology, Northwestern University, USA

ARTICLE INFO

Keywords:
Language and thought
Relational language
Relational representation
Attention
Relational systems

ABSTRACT

Language can affect cognition, but through what mechanism? Substantial past research has focused on how labeling can elicit categorical representation during online processing. We focus here on a particularly powerful type of language—relational language—and show that relational language can enhance relational representation in children through an embodied attention mechanism. Four-year-old children were given a color-location conjunction task, in which they were asked to encode a two-color square, split either vertically or horizontally (e.g., red on the left, blue on the right), and later recall the same configuration from its mirror reflection. During the encoding phase, children in the experimental condition heard relational language (e.g., "Red is on the left of blue"), while those in the control condition heard generic non-relational language (e.g., "Look at this one, look at it closely"). At recall, children in the experimental condition were more successful at choosing the correct relational representation between the two colors compared to the control group. Moreover, they exhibited different attention patterns as predicted by the attention shift account of relational representation (Franconeri et al., 2012). To test the sustained effect of language and the role of attention, during the second half of the study, the experimental condition was given generic non-relational language. There was a sustained advantage in the experimental condition for both behavioral accuracies and signature attention patterns. Overall, our findings suggest that relational language enhances relational representation by guiding learners' attention, and this facilitative effect persists over time even in the absence of language. Implications for the mechanism of how relational language can enhance the learning of relational systems (e.g., mathematics, spatial cognition) by guiding attention will be discussed.

1. Introduction

The Whorfian hypothesis has spurred one of the most prolific yet contentious debates in the history of cognitive science (Bohannon, Landau, & Gleitman, 1986; Boroditsky, 2001; Brown & Lenneberg, 1954; Choi & Bowerman, 1991; Clark & Clark, 1977; Evans & Levinson, 2009; Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2012; Hermer-Vazquez, Spelke, & Katsnelson, 1999; Hunt & Agnoli, 1991; Lupyan, 2012; Pinker, 1994; Ratliff & Newcombe, 2008; Rosch, 1973; Talmy, 1983). While the nuances of this effect are still debated, there is broad agreement that language ***can*** change thought in certain contexts, such as when language is actively used during a task (Dessalegn & Landau, 2013; Lupyan, Abdel Rahman, Boroditsky, & Clark, 2020). Although this "online" effect of language is a concession from a strong anti-Whorfian view, it stands in stark contrast with accumulating evidence from cognitive development demonstrating that the acquisition of language fundamentally ***alters*** cognitive capacities in domains as diverse as categorization (Jones, Smith, & Landau, 1991; Lupyan, 2008; Samuelson & Smith, 1999; Sloutsky & Fisher, 2012; Waxman & Booth, 2001), mathematics (Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson, 2010; Parrish, 2014; Purpura, Logan, Hassinger-Das, & Napoli, 2017; Purpura, Napoli, Wehrspann, & Gold, 2016), and spatial cognition (Dessalegn & Landau, 2008; Loewenstein & Gentner, 2005; Miller & Simmering, 2018; Pruden, Levine, & Huttenlocher, 2011; Spencer, Simmering, Schutte, & Schöner, 2007). Accumulating evidence has highlighted the role that labeling plays in eliciting categorical thinking in both adults and children (Lupyan, 2008, 2012; Lupyan et al., 2020;

* Corresponding author.
E-mail address: lei.yuan@colorado.edu (L. Yuan).

https://doi.org/10.1016/j.cognition.2023.105671
Received 1 February 2023; Received in revised form 16 November 2023; Accepted 20 November 2023
Available online 30 November 2023

Sloutsky & Fisher, 2012; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Vales & Smith, 2015). However, many domains—e.g., math, space, visualizations—with systems of relational concepts require the learner to not just represent a single item, but the relations among them. In the current study, we test one mechanism through which relational language changes relational representation by guiding children's visual attention. The change in visual attention—and the corresponding facilitative effect of language on mental representation—persisted during the study, even when relational language was no longer provided. These results have implications for the mechanisms underlying relational representation and the role of language in teaching systems of relational concepts (e.g., mathematics, spatial cognition).

1.1. How language changes thought: the labeling effect and beyond

Much research on the effects of language on thought has centered on the effect of labeling. One study (Lupyan & Thompson-Schill, 2012) has shown that hearing verbal labels (e.g., "cow") facilitated visual search efficiency (e.g., searching for a cow amid an array of other animals), compared to seeing a visual preview of the object/animal, or the typical sound associated with it (e.g., the "mooing" sound of a cow). Based on this and other similar results, Lupyan proposed the Label Feedback Hypothesis (Lupyan, 2012), which states that language modulates ongoing cognitive and perceptual processes. For example, hearing the label "cow" evokes a categorical representation of a cow, causing cow-like creatures to be perceived as more similar to each other, and as more distinct from non-cow-like creatures. In other words, labels exert their influence on perception by changing the nature of mental representation evoked by language. Vales and Smith (2015) extended this paradigm to 3-year-old children and showed that hearing object labels together with a visual preview of an object increased children's efficiency in searching for the object (e.g., an ice cream cone) among a group of distractors, compared to only seeing a visual preview of the target object alone. While the Label Feedback Hypothesis emphasizes how language changes the internal representation of objects, the explanation offered by Vales and Smith (2015) focuses on how language changes the allocation of ***visual attention***: They argue that providing basic-level category labels biases children's attention to encode and represent object shape (Landau, Smith, & Jones, 1988; Smith et al., 2002; Smith, Jones, & Landau, 1996; Smith & Samuelson, 2006), and the resulting object shape representation leads to more efficient visual search in an array. Regardless, both theories were meant to explain how hearing object labels can change ongoing perceptual processes.

However, many real-life learning tasks—especially those involving learning systems of relational concepts (e.g., taxonomy, math, spatial representations)—require the learner to not just represent a single item, but the relations between multiple items. It has been proposed that relational language is particularly suited for promoting relational representation because linguistic structures give learners a means through which they can latch on to represent to-be-learned non-linguistic relations (Gentner, 2003). For example, Loewenstein and Gentner (2005) gave preschool children a challenging relational mapping task in which they searched for hidden toys on a three-tiered shelf after seeing where the toys were hidden in another three-tiered shelf. Hearing experimenter-provided spatial terms like "top, middle, bottom" significantly improved children's search success, compared to not hearing any spatial terms and having to rely on the perceptual mapping between the hidden and search shelves alone. Likewise, the relational structure embedded in number names may play an instrumental role in children's learning of relative numerical magnitudes (Gentner, 2010). Most children first learn number words as part of a rote courting routine without understanding the precise magnitudes represented by those numbers. But by mapping one object to the number name "one", two objects to "two", and three objects to "three", children may start to extract the critical relational mappings between the counting sequence (e.g., "one, two, three") and the increasing magnitudes (e.g., from one to two to three objects). This magnitude-increasing-one maps onto counting-increasing-one relation can then be recruited to learn the magnitude for the remaining single-digit numbers. Empirical evidence showing strong correlations between parents' early number words in child-directed speech and children's later numeracy understanding (Gunderson & Levine, 2011; Levine et al., 2010) supports this "linguistic structures bootstrap conceptual structures" proposal.

But critical questions remain as to how exactly linguistic bootstrapping takes place during learning. For example, how do children select the individual items, how do they extract the various relations, and how exactly do linguistic structures work to facilitate relational learning in this process? Answers to these questions are critical for understanding the mechanism through which language can shape cognition and have direct implications for education interventions that are aimed at using language to facilitate learning and development.

1.2. Proposals on how relational language works to change relational representation

Representing novel spatial relations—such as an arbitrary spatial relation between two randomly chosen colors—is a surprisingly difficult task. As an illustration, try to remember the configuration in Fig. 1 (a). Now try to identify the two targets with the same configuration in Fig. 1 (b). This is a difficult task because it requires the viewer to represent the precise relative spatial location as left (red, blue) instead of left (blue, red). In contrast, try to search for the same target in Fig. 1 (c). The task becomes significantly easier; instead of representing the precise spatial relation, one can simply perform a feature search—vertical split lines rather than diagonal lines. To answer the question of how relational language works to change relational representation such as the one shown in Fig. 1 (a), Dessalegn and Landau (2008) gave 4-year-old children a challenging color-location conjunction task, in which they were first shown spatial relations such as a red bar on the left of a blue bar and later had to recognize the same configuration from its mirror reflections (e.g., a red bar on the right of a blue bar). Children who were presented with relational language describing the spatial relationship during the learning phase—e.g., "Look, red is on the left of blue"—performed significantly better at testing, compared to children who were not presented with any relational language—e.g., "Look at this". Interestingly, guiding children's attention to inspect each object within the relation—e.g., flashing the red bar and then the blue bar or similar manipulations that had been shown to aid in relational representation in adults and older children (Gleitman, January, Nappa, & Trueswell, 2007; Grant & Spivey, 2003; Livins, Doumas, & Spivey, 2015; Yuan, Uttal, & Gentner, 2017)—didn't improve 4-year-olds' performance in this study. Furthermore, using both a comprehension task and a production task, the researchers found that the facilitative effect of language did not depend on children's understanding of the precise meaning of spatial terms (e.g., picking out a yellow ducky on the left—as opposed to the right—of a dot); in other words, children's performance benefited from hearing relational language, even when they did not show ostensible understanding of the precise meanings represented by those spatial terms (an interesting result that we will return to later).

Based on the above results, Dessalegn and Landau (2008) concluded that language has both an "abstract" and a "momentary" effect on relational representation. First, the authors reasoned that, since attention manipulation didn't improve the encoding and maintenance of spatial relations, the facilitative effect of language must rely on augmenting some internal representations, not manifested in discernable external sensory processes. Perhaps this may be accomplished through the activation of a spatial template (e.g., two horizontally arranged objects)—much like labeling can evoke a categorical (e.g., cow) representation. This conclusion is consequential, because if it is all "in the head" (i.e., internal representation), then this proposal precludes the utility of sensory experiences in shaping relational thinking. Second,

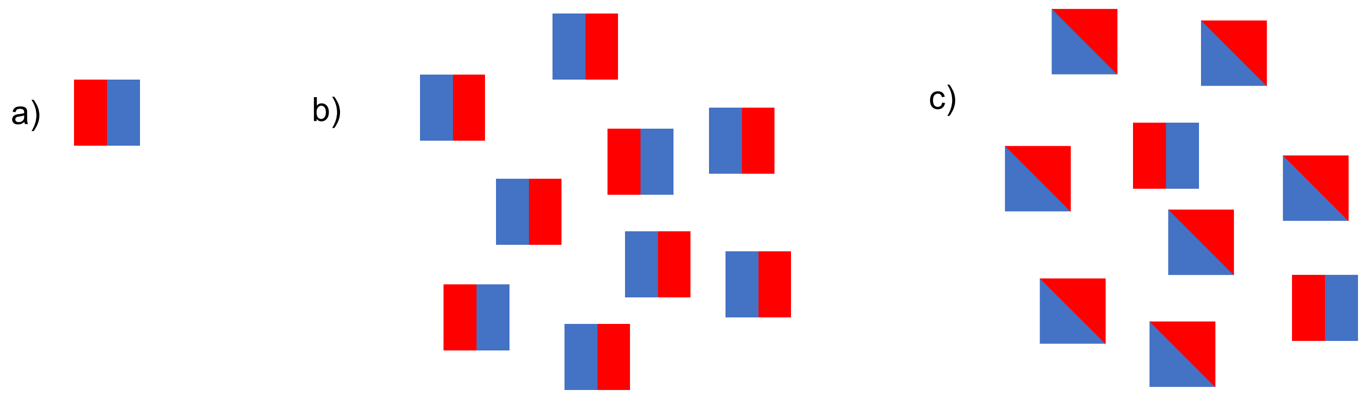

Fig. 1. Illustration of the difficulty of representing spatial relation.
Note. a) a possible spatial relational representation as left (red, blue). b) search for the target in a) in the middle of its mirror reflections red (blue, red) is difficult, because one must represent the precise spatial relation. c) search for the target in a) in the middle of totally different spatial relations (i.e., diagonal splits); the task becomes easier because one can rely on feature matching (e.g., vertical lines) rather than spatial relational representation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

since performance did not rely on children's understanding of the spatial terms, the effect of language seems "momentary"—only manifests during the task, not to be changed by anything stored in long-term memory and not changing any subsequent long-term representations. Overall, this argument is similar to the online labeling effect of language frequently cited in adult literature: language has a temporary effect on cognition through evoking a categorical representation of visual stimuli.

1.3. Is the effect of language entirely "abstract"?

In addition to language, a separate line of research has considered the powerful ways in which guiding attention can also induce relational representation in both children and adults. For example, Yuan et al. (2017) asked 3-year-old children to use a map to find hidden objects in a miniature room. This is a challenging task for 3-year-olds, especially when the icons on the map and the locations in the room lack high-level perceptual similarity (e.g., using a dot to indicate the location of a chair) (Marzolf & DeLoache, 1994) or when there are strong distracting perceptual cues that may lead to mismatch (e.g., using a picture of a chair to indicate the location of a table) (Rattermann, Gentner, & Deloache, 1990). Providing relational language improved children's performance, but so did providing back-and-forth pointing gestures that highlighted the one-to-one correspondence of icons on the map and referent locations in the room. Gestures were even more helpful than relational language in cases of low-similarity or distracting mismatch icons, suggesting that gestures can provide strong visual cues—perhaps stronger than verbal cues in certain contexts—to draw learners' attention to the relational correspondences between two representations. This result is consistent with studies showing that gestures guide visual attention in learning relational tasks such as mathematics (Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018) and can "spatialize" important conceptual information, including analogical structures (Cooperrider, Gentner, & Goldin-Meadow, 2016; Goldin-Meadow, 2011; Goldin-Meadow, Alibali, & Church, 1993; Richland & McDonough, 2010). Overall, these results suggest that guiding attention via external sensory processes (e.g., gesture) may present direct guidance for establishing relational mapping, whereas the way in which relational language guides attention may vary depending on how strongly the language can evoke the intended relational representation in children.

Franconeri, Scimeca, Roth, Helseth, and Kahn, L. E. (2012) provided

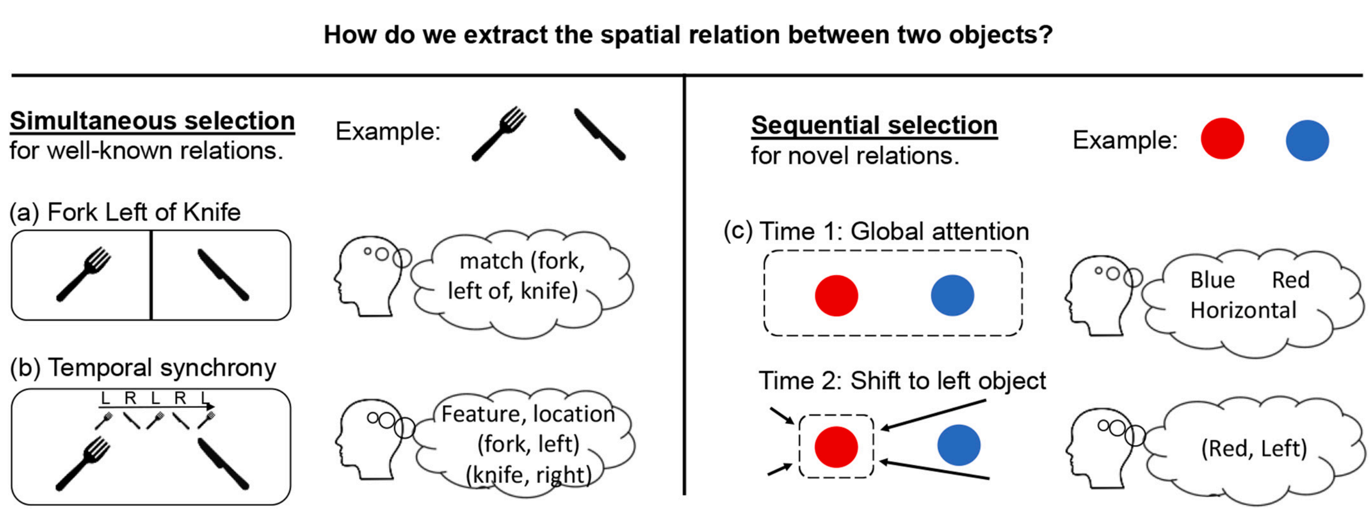

Fig. 2. Two forms of spatial relational representation.
Note. Left: Simultaneous selection of well-known relations, such as a fork on the left of a knife on a dinner plate, through two possible ways of template matching (adapted from (Franconeri et al., 2012). (a) matching the entire relation: Left (fork, knife). (b) matching through temporal synchrony of component and relational role: (fork, left) – (knife, right) – (fork, left) – (knife, right). Right: Sequential selection of novel relations, such as a red dot on the left of a blue dot, through an attention shift from a global window to one side of that window, with the direction of the shift providing the relative spatial position of one object (adapted from (Roth & Franconeri, 2012). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

a formal model through which attention plays a direct role in how adults extract novel spatial relational representations. As shown in Fig. 2 (left), frequent and well-encoded spatial relations—e.g., a fork on the left of a knife on a dining table—might be extracted through a simultaneous selection mechanism with various proposed implementations—such as (a) and (b)—akin to matching stimuli to long-term stored mental templates. In contrast, extracting novel spatial relations—e.g., a red bar is higher/bigger/left of a blue bar—requires the perceiver to shift attention. For example, as shown in Fig. 2 (right) people may start by attending to the global structure of the scene (horizontal arrangement) and shift (or focus) attention to the object on the left (c). Under this account, it is obligatory that people focus their attention on one object at a time in order to extract their relationship, as opposed to attending to the two objects simultaneously or splitting attention equally between them. Various empirical studies using different methodologies have largely supported this attention shift account of relational representation in adults (Franconeri et al., 2012; Roth & Franconeri, 2012; Yuan, Uttal, & Franconeri, 2016). If the same mechanism applies to young children, then we may expect that successful encoding of relational representation must be preceded by differential attention allocation to the two objects in a relation. Note, the reverse may not necessarily be true—inducing different attention allocation to the two objects may not necessarily lead to successful encoding of spatial relations; some cues—such as gestures that carry conventionalized communicative meanings (Goldin-Meadow, 2011)—may be more useful than other cues—such as sequential visual highlighting, which may present little communicative meaning to very young children.

1.4. *Is the effect of language just "momentary"?*

The second argument in Dessalegn and Landau (2008)'s proposal of how relational language changes relational representation—and indeed in the effect of labeling on adults' cognition in general—is that the effect of language is momentary, only exerting an online effect when language is used during a task. While there have been demonstrations that teaching adults novel words can elicit categorical representation (Zhou et al., 2010), there seems to be a disconnection between the online and offline effects of language on thought. If language has such a powerful effect on cognition, why is it so fragile and short-lived? Lupyan et al. (2020) proposed that language may exert an offline effect on thought through a permanent change in mental representation. Prior to learning a label, the representation of a stimulus may be primarily driven by the continuous perceptual property (e.g., the wavelength of light). Learning a new label pushes within-category members to be closer together and between-category members to be further apart in the mental representation space. Through many usages of the label, this altered categorical representation may replace the initial continuous perceptually driven representation, "biasing" behavioral and neural activities even in the absence of language. However, this account is hard to reconcile with findings showing that the labeling effect can easily be wiped out in tasks involving verbal interference (but see (Nedergaard, Wallentin, & Lupyan, 2022).

Here, we test one potential hypothesis for how relational language can cause long-term changes in thoughts or thought processes via changing learners' attention patterns. We suggest that relational language can guide learners to attend to the visual stimuli in systematic ways, which in turn aid in the construction of robust relational representation. We refer to this as the "language systematizes attention" hypothesis. The changes in attention pattern may initially only occur in the context in which language was used, but habitual use of the language systematically recruits attention in those specific contexts, and eventually leads to the same recruitment of attention and constructed mental representation when language is no longer present. In the case of relational representation, hearing relational terms, such as "red is on the left of blue" or "the toy is hidden on top of the shelf", may prompt children to attend to and encode the relative spatial location of the target/referent object—by isolating one object at a time as suggested in the attention shift account of relational representation (Franconeri et al., 2012). Children may not initially know the precise meaning of spatial terms such as left or right, but they may know that words like "left" and "right" denote relative spatial locations (as opposed to color or shape) (Dessalegn & Landau, 2008). This partial knowledge (Byrge, Smith, & Mix, 2014; Wilkinson, 1982; Yurovsky, Fricker, Yu, & Smith, 2014) may be sufficient to guide their attention to the correct input in the right sequence and support them to encode the correct relationship. Through multiple encounters with the relational words and the corresponding scenes, children reinforce a particular attention pattern. Eventually, this attention pattern and its corresponding mental representation become so tightly paired in the context of a particular task that they subsequently get automatically executed, even in the absence of the language. We believe that this is a parsimonious explanation that connects the online and offline effects of language and represents an important pathway through which relational language changes relational representation. By this account, systems of linguistic symbols help to organize otherwise overwhelming perceptual experiences; in other words, language recruits and brings order to perception.

1.5. *Current study*

In the current study, we tested two predictions with respect to how relational language works to change relational representation in children. First, we predicted that the effect of language is not purely "abstract" (or internal) but is driven by changes in external sensory processes—particularly, the recruitment of visual attention. We directly measured children's visual attention using eye trackers while they performed a challenging task that required them to encode the spatial relationship between two colors (e.g., a red bar on the left of a blue bar), either with (the experimental group) or without (the control group) hearing relational language. If the effect of language on thought relies purely on changing internal mental representations without discernable effects on external sensory processes, then we would not expect to see differential attention patterns between the experimental and control groups. However, if relational language enhances relational representation by altering visual attention patterns, then the two groups should exhibit distinct attention patterns. Second, we tested the prediction that the effect of relational language on relational representation is not just "momentary" but can persist over time, even in the absence of language. We deployed a within-subject design, in which children in the experimental group (Labels first) were presented with relational language in the first half of the study but were not presented with relational language in the second half of the study, while the control group (Labels last) received the opposite order (no relational language during the first half of the study and relational language during the second half). At no point during the study were the participants given the correct answer, and thus could not simply learn from feedback. If language only has an online effect on cognition, then performance for the experimental group should worsen in the second half of the study, once they are no longer hearing relational language. However, if repeated exposure to relational language changes how learners attend and their corresponding mental representation, even in the absence of language, we might expect the facilitative effect of language to carry over to the second half of the study when language was no longer present for the experimental group. As shown below, our results support both of our predictions—language enhances relational representation by guiding learners' attention, and this facilitative effect persists over time even in the absence of language.

2. Methods

2.1. *Participants*

Fifty-six 4-year-old children from the greater Chicagoland area were recruited through a university database. Twelve children were excluded due to various reasons (8 failed to complete the whole experiment, 2 didn't understand the instructions, and 2 due to technical difficulty with the sound system of the experiment computer). The final sample consisted of 44 children, who were randomly assigned either to the experimental group (N = 22, 15 females, M age = 54.7 months, range = 49.23–60.66) or the control group (N = 22, 11 females, M age = 54.4 months, range = 48.04–62.66). Three children provided no eye movement data due to calibration failure (two children wore smudged glasses that made it difficult to track their eyes, and one child refused to wear a head sticker that was required for tracking). Thus, the final sample in the eye movement analysis includes 41 children (22 from the experimental group and 19 from the control group). A power analysis conducted in G*power (Faul, Erdfelder, Buchner, & Lang, 2009) indicates that, for a mixed factorial design with two between-subject groups (experimental vs. control) and two within-subjects measures (first vs. second half of the study), a minimum sample of 36 is needed to achieve 0.85 power with a median effect size of 0.25 and a significance level of 0.05. For completeness, we also conducted a simulation-based post hoc power analysis using the "simr" package in R (Green & Macleod, 2016) with 1000 iterations. This analysis indicated an estimated power of 0.77, based on the observed effect sizes and variance components from our study. These analyses suggest that our sample size is adequate for addressing the research questions.

Children were tested individually in a quiet laboratory room. Informed consent was obtained from the legal guardian and assent was obtained from the participants prior to the study. Children received an envelope of stickers for their participation and each family was given a $10 Amazon gift certificate as compensation.

2.2. Stimuli and apparatus

Stimuli were presented using SR-Research Experiment Builder on a 17-in. monitor with 1024 × 768 resolution and 75 Hz refresh rate. Eye movements were recorded by an EyeLink 1000 Desktop Mount eye tracker with a 1000 Hz sampling rate. As shown in Fig. 3 (a), the stimulus squares were composed of two parts, colored in red and blue respectively (colors were chosen to have equal luminance and be color-blind safe), and were 240 × 240 pixels large. There were four unique arrangements of the stimuli: two horizontal and two vertical splits with the red side on the left/right/top/bottom of the blue side as shown in Fig. 3 (a). The stimulus squares were positioned in one of 4 equidistant locations on the screen, such that the squares were always positioned 136 pixels away from the closest vertical edge of the screen and 72 pixels away from the closest horizontal age of the screen.

Children were positioned 550–600 mm from the eye tracker on a child-sized chair. One experimenter monitored the eye-tracking computer (conducted calibration and ensured that tracking was going well) while a second experimenter sat next to the child. The second experimenter helped the child sit, explained the tasks, and recorded the child's responses on each trial into the computer program.

2.3. Design and procedures

The study deployed a mixed factorial design. Children were randomly assigned to either the Labels first (experimental group) or the Labels last (control group). For both groups, there were two phases with 16 trials for each phase and 32 trials in total. Each phase (16 trials) included four arrangements (see Fig. 3) appearing at each of the four possible locations. Children in the Labels first (experimental) group heard pre-recorded relational language (16 trials) during the first phase of the study, followed by non-relational language (16 trials) during the second phase of the study. The relational language described the relative location of the ***target side*** of the stimuli in relation to the ***foil side***, in the form of "Red is on the left/right/top/bottom of blue". The non-relational language simply provided encouraging words (e.g., "Look at this one, look at it carefully"). Children in the Labels last (control) group received the opposite order: 16 trials with non-relational language in the first phase, followed by 16 trials with relational language manipulation in the second phase of the study. We chose to use the red side as the target side for all trials, due to past research showing strong trial-to-trial carry-over effects in young children in which they cannot readily switch task assignments (Chevalier & Blaye, 2009; Garon, Bryson, & Smith, 2008), and due to the current research goal of helping children to build a consistent and strong attentional pattern and internal representation.

Prior to the main experiment, all children heard the same instructions and completed warm-up trials. During the instructions, children were told that they would see a picture, it would disappear, and then they would see two pictures, and their job was to choose the picture that matched the one that disappeared. There were 4 warm-up trials with simple images (e.g., a rabbit, a purple rectangle) to ensure that they understood the task. We also provided instructions about looking at the smiley or winking faces when they appeared on the screen; these stimuli served as fixation markers before each trial during the main experiment.

During the main experiment, each trial included a learning phase, a delay, and a test phase. As shown in Fig. 3 (b), the **learning phase** started when the child fixated on a smiley or winking face (fixation trigger) positioned at the center but perpendicular to where the target would appear; for example, if a left/right split target appeared at the top left quadrant, then the fixation trigger would appear at the bottom left quadrant. This fixation trigger ensured that the child was not already fixating on the target when it appeared and there was no bias as to which side of the target the child's gaze would be closer. After the target appeared 50 milliseconds later, an audio file was played that either described the relation between the two colors "Look, the red is on the top/bottom/left/right of the blue" (***relational language***) or provided generic, encouraging audio, e.g., "Look at this one, look at it carefully" (***non-relational language***). The target image stayed on the screen for a total of 6 s, while the audio lasted for between 2240 and 2230 milliseconds depending on the stimuli (e.g., left/right/top/bottom). The target then disappeared. During the **delay phase**, another fixation smiley face appeared for 1000 ms, which was positioned at the midpoint between the locations where the two test items would appear, as shown in Fig. 3 (b). The **test phase** presented two images—one that matched the target and one that was its mirror reflection. The experimenter asked the child "Do you remember which one you just saw?" and recorded the child's answer in the computer by pressing a designated key. Children were not given feedback as to whether they were correct or incorrect. For every 4 trials, children had a break in which they either saw a photo encouraging them to keep going, or a screen showing 9 smiley or winking faces that provided confirmation of eye-tracking accuracy. We recalibrated periodically if the child failed to trigger the target image during the fixation trigger, or if there were other indications that tracking was not working.

3. Results

3.1. Accuracy analysis

We first examined how participants' overall success in completing the task was affected by their conditions and the presence of relational language using a Linear Mixed Effect Model conducted in the R environment (R Core Team, 2021). The final model[1] included condition

[1] In this and all following LMMs, for specifying the random effect structure, we took a model selection approach (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017) start with the maximal model that include all random intercept and slope, and gradually reduce the complexity if the model failed to converge. To determine the final model, we performed LRT test favoring the model that significantly explain additional variance in the data. If there was no significant difference between two models, we reported the simpler model based on the principle of parsimony (Burnham & Anderson, 2004).

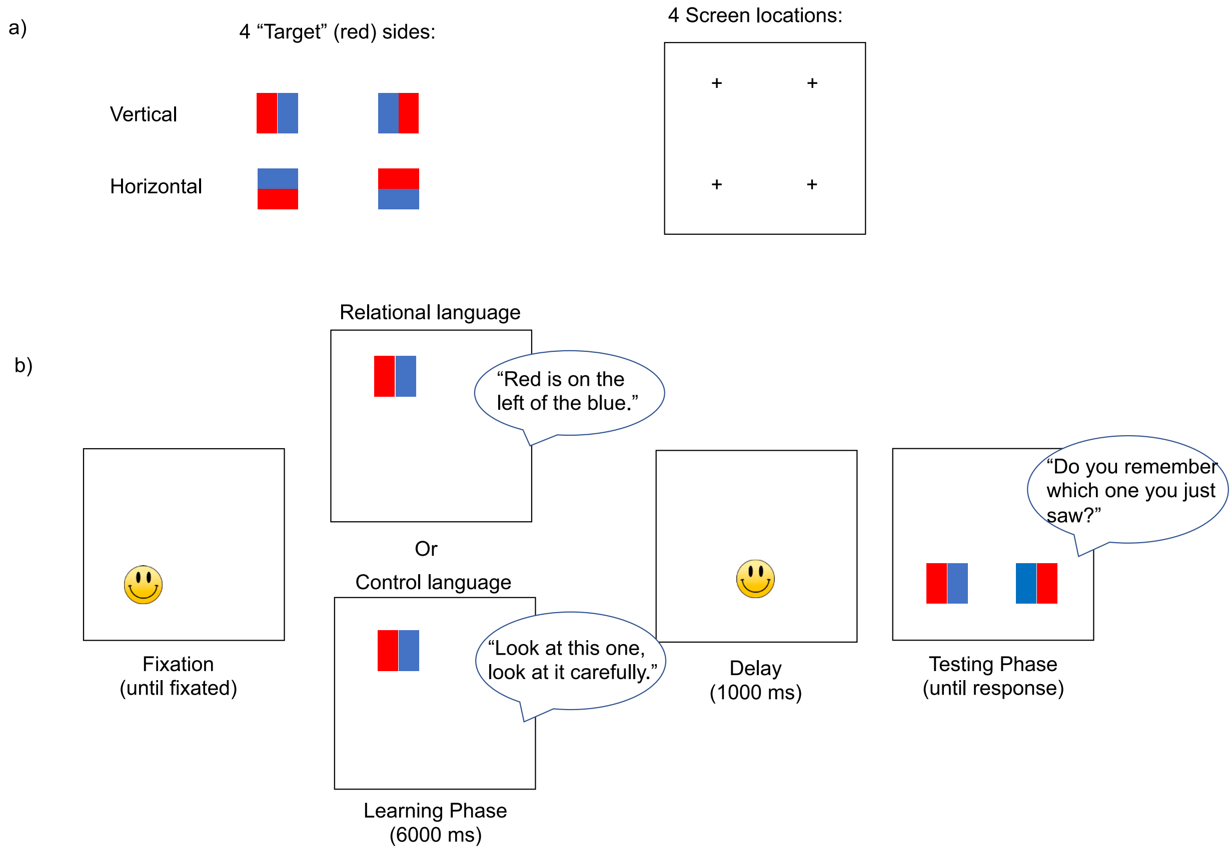

Fig. 3. Illustration of experimental materials and design.
Note. (a) Left: Four stimuli square configurations: vertical and horizontal splits with the target (red) sides on the left, right, bottom, and top of the squares. Right: four quadrant locations where the stimulus squares can appear. (b) Illustration of the experimental procedures with the relational language or the non-relational language. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(experimental vs. control) and the presence of relational language (yes vs. no) as fixed effects, random slope for the presence of relational language over subject and random intercept for trial. Trial accuracy (i.e., 0 or 1) was entered as the dependent variable. Significance values were obtained using the afex package (Singmann et al., 2015) with the likelihood ratio test method, and logit was used as the linking function. Results showed a main effect of condition: children in the Labels first (experimental) condition outperformed those in the Labels last (control) condition, χ^2 (1) = 6.23, p = .013. There was also a significant main effect of language: children overall performed better on trials with relational language than on trials with non-relational language, χ^2 (1) = 13.72, $p < .001$. There was no significant interaction between the condition and the presence of the relational language, χ^2 (1) = 1.00, p = .317.

We next directly test our predictions through specific planned comparisons. Since during the first phase of the study, the Labels first (experimental) condition completed the task with the aid of relational language, while the Labels last (control) condition did not, the contrast between the two groups is indicative of an ***Online effect*** of relational language. To test this effect, a Linear Mixed Effect Model was conducted in which condition (experimental vs. control) was entered as a fixed effect, subject and trial were entered as random effects, and trial accuracy (i.e., 0 or 1) was entered as the dependent variable. As shown in Fig. 4, there was a main effect of condition: children in the Labels first (experimental) condition outperformed those in the Labels last (control) condition, χ^2 (1) = 17.85, $p < .001$, suggesting that hearing relational language enhances relational representation.

To test the ***Sustained effect*** of language on relational representation, we compared the performance of the Labels first (experimental) condition during the second phase of the study, in which they were no longer provided with any relational language, to the performance of Label last (control) condition during the first phase of the study, in which they also did not hear (and had not previously heard) any relational language. A Linear Mixed Effect Model was conducted in which condition (experimental vs. control) was entered as a fixed effect, subject and trial were entered as random effects, and trial accuracy (i.e., 0 or 1) was entered as the dependent variable. As shown in Fig. 4, there was a main effect of condition, children in the Labels first (experimental) condition outperformed those in the control condition, χ^2 (1) = 8.06, p = .005, suggesting that the facilitative effect of relational language on relational representation persisted over time even when language was no longer present.

3.2. Eye movement analysis

To investigate possible differential visual attention patterns between the experimental and control conditions, we first examined the overall fixation duration[2] to the target sides of the stimuli. We then consider how gaze allocation dynamically changes over time according to time-dependent linguistic input. Lastly, we directly linked accuracy with gaze patterns to further examine how differential attention patterns contribute to performance differences. For all analyses, our focus is on the initial 3 s of the learning phase in each trial, during which language

[2] Although attention should not be confused with raw location of eye gaze as people can fixate on a location but attend to another, particularly in experimental tasks in which they were explicitly asked to do so. In many naturalistic settings and certainly with young children, this departure is rare, and many prior studies have used gaze pattern as indicative of the focus of attention.

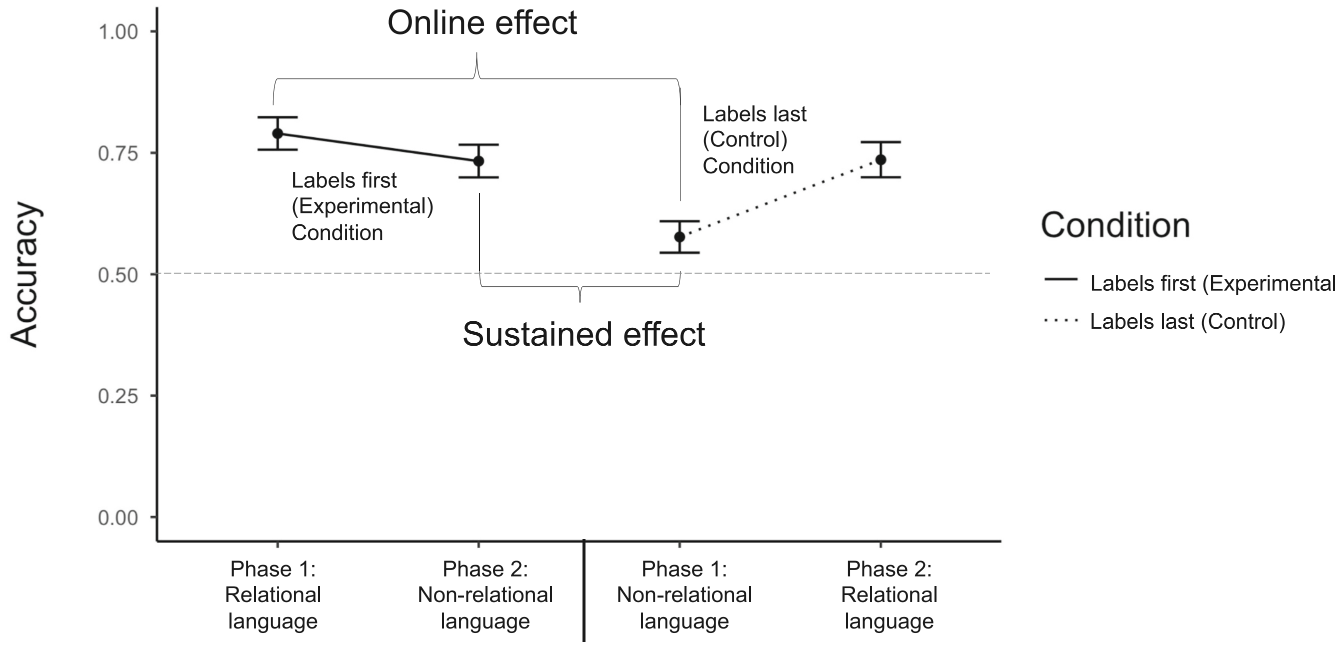

Fig. 4. Participants' behavioral accuracies from different experimental conditions.
Note. Accuracy by condition and the presence of relational language. There was a significant Online effect of language: children who heard relational language during the first half of the study (the experimental condition) outperformed those who heard non-relational language during the first half of the study (the control condition). There was also a significant Sustained effect of language: children who did not hear relational language during the second half of the study but heard language during the first half of the study (the experimental condition) outperformed those who heard non-relational language during the first half of the study (the control condition).

was presented to the participants. Additional analyses based on the entire duration of the trial (6 s) showed the same overall patterns of results. The fixation analysis was based on fixations extracted from the SR Research Data Viewer software. The gaze analysis was based on raw data that was sampled at a 1000 Hz rate. For gaze analysis, 10.5% of the trials (138 out of 1312 trials) were excluded due to missing >50% of the raw gaze data.

3.2.1. Differential fixation patterns

First, to rule out the possibility that relational language affects general attention to tasks—such that children who heard relational language were simply better able to concentrate on the task at hand—we compared children's overall fixation duration to the stimulus square between conditions during the first phase when the Labels first (experimental) condition heard relational language and the Labels last (control condition) did not. Fixations were extracted using SR Research Data Viewer software. A Linear Mixed Effect Model was conducted in which condition was entered as a fixed effect, subject and trial were entered as random effects and total fixation duration was the outcome variable. There was no significant difference between conditions in the amount of time spent looking at the square, χ^2 (1) = 1.53, p = .216. This suggests that relational language is not a general attention enhancement and confirms that both groups were equally attentive during the study.

If the attentional shift account of relational representation applies to young children, we would expect children across conditions to differentially allocate their attention to the two sides of the square. Specifically, we expect children in the Labels first (experimental) condition to allocate more attention to the target side of the square compared to those in the Labels last (control) condition during the first phase of the study—an indication of the ***Online effect*** of relational language. Fig. 5 (a) shows a histogram of fixation duration to the target sides[3] of the stimuli for children in the two conditions during the first half of the study. Given the non-normal distribution of fixation durations, nonparametric tests (Kolmogorov-Smirnov tests) were used to determine group differences. Results confirmed that children in the experimental condition (who heard relational language) had significantly more fixation with longer durations (D = 0.06, p < .001) to the target side of the stimuli, compared to children in the control condition (who did not hear any relational language). As shown in Fig. 5 (a), children in the control group tended to have many short fixations to the target, whereas there was a greater proportion of longer target fixations for children in the experimental condition. This result supports the hypothesis that relational language alters participants' attention patterns.

Another Kolmogorov-Smirnov (K—S) test was conducted to test the ***Sustained effect*** of relational language on relational representation—by comparing the distribution of fixation duration for the Labels first (experimental) condition during the second half of the study (in which they didn't hear any relational language, but had previously heard relational language in the first half of the study) to the Labels last (control) condition during the first half of the study (in which they had never heard relational language during the study). As shown in Fig. 5 (b), results showed that children in the experimental condition had significantly more fixations with longer durations (D = 0.04, p = .016) to the target sides of the stimuli, compared to children in the control condition, supporting the hypothesis that the effect that relational language had in altering participants' attention patterns persists over time even when relational language was no longer present.

[3] Due to the high variability that is common in children's eye tracking data, we enforced a 50-pixel grace region around the square, such that the allowable region for gaze to be classified as within the target or foil was increased by 50 units in all directions except for the split between the two sides.

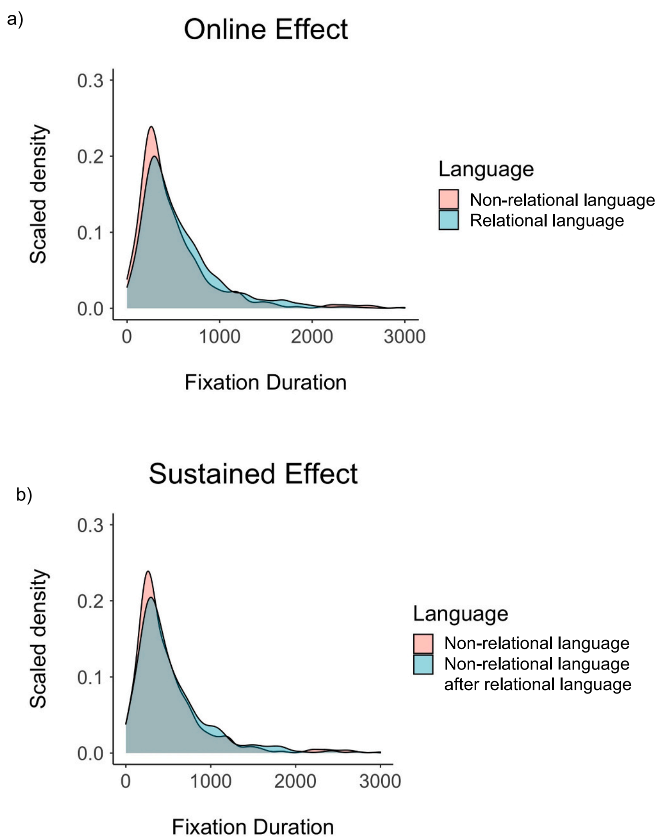

Fig. 5. Distribution of fixations duration to target for different language manipulations.
Note. a) During the first phase of the study, compared to those who heard non-relational language in the control group (shown in red), children who heard relational language (shown in turquoise) had more long fixations to the target side of the square. b) Compared to the control group who had never heard non-relational language (shown in red), children who had *previously* heard relational language during the first half of the study (shown in turquoise) had more long fixations to the target side of the square. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.2. *Differential time-dependent gaze patterns*

Next, we examined how differences in children's attention patterns from the two conditions might rise over time during the study based on moment-by-moment raw gaze data. We aggregated the raw gaze data into 50-millisecond bins and, for each bin, calculated the proportion of samples during which the child was looking to a given interest area (i.e., target side, foil side, or outside the square, or missing) on each trial. Because this measure was not normally distributed, a log transformation of proportion looking was used as the dependent variable for plotting and analysis.

To examine the ***online effect*** of relational language, Linear Mixed Effect Model was conducted for the experimental condition when the participants were presented with relational language during the first phase of the study. The final model included interest area as a fixed effect and participants as a random effect. There was a significant main effect of interest area, χ^2 (1) = 19.69, $p < .001$; children were more likely to attend to the target side than the foil side when provided with relational language. To visualize this difference over time, we used Growth Curve Analyses (Mirman, Dixon, & Magnuson, 2008) conducted using the EyeTrackingR package (Dink & Ferguson, 2015) in the R environment (R Core Team, 2021). Fig. 6 (left) shows GCA results for the online effect of language; the smooth curved line represents the model fit function for log-transformed proportion looking.[4] In contrast to the asymmetrical attention patterns exhibited by children in the experimental condition during phase 1 of the study, children in the control condition attended to the target and foil sides non-differentially, as shown in Fig. 6 (middle). A Linear Mixed Effect Model with interest area as a fixed effect and participants as a random effect failed to detect any significant main effect of interest area, χ^2 (1) = 0.09, p = .764.

We next examined the ***Sustained effect*** of relational language—looking to the target and foil sides for the experimental condition in the second phase of the study (during which participants were not presented with relational language but had previously heard relational language). A Linear Mixed Effect Model—in which interest area was included as a fixed effect and participants was included as random effect—revealed a significant main effect of interest area, χ^2 (1) = 5.08, p = .024. Children were more likely to attend to the target side than the foil side when provided with relational language during the first half of the study, supporting the idea that there is a sustained effect of relational language—the asymmetry of children's attention pattern to the target persists even when relational language was no longer provided. Fig. 6 (right) shows the best fitting function for the Growth Curve Analysis between the target and foil sides for each condition, which revealed a similar pattern as the ***Online effect***—children who had previously heard relational language showed asymmetry in their attention to the target and foil sides while children in the control condition did not demonstrate such asymmetry.

3.3. *Linking accuracy to eye movement data*

To more directly test the hypothesis that relational language enhances the relational representation by guiding learners' attention, we next link individual children's accuracy to their eye movement data on a trial-by-trial basis. The rationale for this analysis is that if changing attention is indeed one mechanism through which relational language enhances relational representation in human learners, then success in the individual trials should depend not only on the exposure to language but also on whether the language successfully elicited asymmetrical attention patterns. Thus, we examined differential attention patterns during the online and sustained effect of relational language separated by whether the participants have successfully completed the current trial.

For the online effect of relational language, a Linear Mixed Effect Model in which interest area and accuracy were included as fixed effects and participant was included as random effect revealed a significant main effect of interest area during the first phase of the study, χ^2 (1) = 17.29, $p < .001$. However, a close examination of the time course revealed different attention patterns that characterized successful versus unsuccessful trials. As shown in Fig. 7 (a), on trials in which children successfully chose the correct configuration, the largest difference between the target side and foil side emerged after the onset of relational language (e.g., "Red is on the left") at around 1.7 s. In other words, the syntactical structure of relational language directs children's attention as the language unfolds over time. In contrast, on trials in which children answered incorrectly, the difference emerged early during the trial before the onset of the target side ("red").

At first glance, this seems puzzling; how do children know where to look even before the utterance of the target side? However, this makes sense in the context of the current experimental design in which the target side was always the red side. This design decision was chosen in an effort to build a consistent attention pattern as past research has shown strong trial-to-trial carry-over effects in young children,

[4] The 7th polynomial of time term was used for the GCA analysis, which provides a good fit for the empirical data (Mirman, Dixon, & Magnuson, 2008).

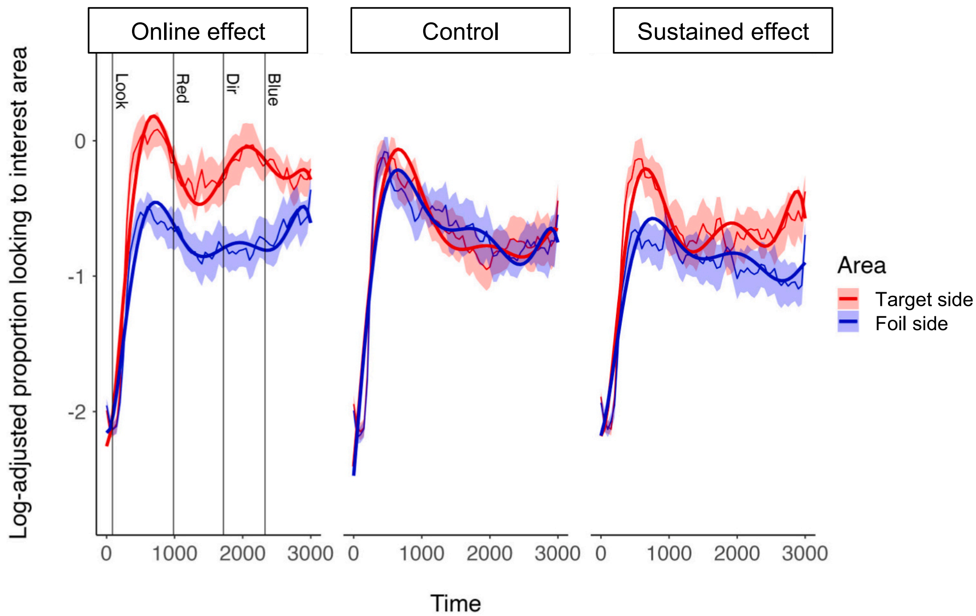

Fig. 6. Time Series Analysis of Gaze to Target and Foil in the different conditions:
Note. Time series analysis for the experimental condition phase 1 hearing relational language (the online effect) (left), for the control condition phase 1 not hearing relational language (middle), and the experimental condition phase 2 not hearing relational language (the sustained effect) (right). The y-axis represents the log-transformed proportion of gaze directed to the target (red) or to the foil (blue) at each time point (x-axis). Solid lines represent the best-fit growth curve analysis (GCA) model (which enters the 7th-order polynomials of time as the predictor). Relative locations of the target/red side (i.e., left, right, top, or bottom) are aggregated. Vertical lines represent the mean time at which keywords (i.e., "Look", "Red", "[Direction]", "Blue") were spoken. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

suggesting that they cannot readily switch task assignments (Chevalier & Blaye, 2009; Garon et al., 2008). Thus, in our study, children learned to visually prioritize the target side over trials. However, specifically during these incorrect trials, children seem to have "lost" the sustained attentional pattern. They started out attending to the target side of the stimuli, but then switched over to the foil side rather than following the relational language to bind the target side (i.e., "red") to the relational term (e.g., "left"). This result suggests that in order for relational language to enhance relational representation online, children's attention had to follow the temporal and syntactical structure of the language. In other words, just having relational language is not enough, it is the changes in attention, guided by language, that promote and support the encoding of relational information.

In contrast to the differential attention patterns exhibited by children in the relational language condition, children in the control condition attended to the two sides of the stimuli similarly regardless of whether they were correct in their answers or not. A Linear Mixed Effect Model in which interest area and accuracy were included as fixed effects and participants was included as random effect failed to reveal any main effect or interaction between interest area and accuracy, *ps* > 0.546. As can be seen in Fig. 7 (b), the time course of attending to the target and foil sides of the stimuli essentially overlapped for the control condition regardless of whether children answered correctly or not. Thus, without the help of relational language, children do not show asymmetrical attention patterns. And even though as a group they did perform significantly above the chance level of 50% ($t\,(21) = 2.37$, $p = .03$), their average performance of 57% was still relatively low in this two-alternative-forced-choice task.

To examine the sustained effect of relational language, a Linear Mixed Effect Model was conducted in which interest area and accuracy were included as fixed effects and participant was included as random effect during the second phase of the study. The result failed to reveal any significant main effect or interaction between interest area and accuracy, *ps* > 0.151. However, as shown in Fig. 7 (c), Growth Curve Analysis shows potential differences in children's attention patterns depending on whether they were successful at the task. As can be seen in Fig. 7 (c) right, when children answered incorrectly, the time course of attending to the target and foil sides of the stimuli largely overlap. However, as shown in Fig. 7 (c) left, when children answered correctly, it appeared that they differentially attended to the target than the foil side of the stimuli. This was confirmed with a Linear Mixed Effect Model in which the interest area was included as fixed effects and participant was included as a random effect. The result revealed a main effect of interest area, χ^2 (1) = 7.13, p = .008. This result suggests that when relational language was no longer present during the second phase of the study, children in the experimental condition were overall more successful than those in the control at identifying the correct spatial relation, but only when they exhibited asymmetrical attention patterns after having been exposed to relational language during the first phase of the study. This effect appears to be small and future studies with longer training sessions should further test the robustness of this result. Nevertheless, this result provides proof of concept that relational language can cause long-term changes in relational representation by changing learners' attention patterns.

3.4. Summary

Children were more successful at encoding and recalling a correct spatial relational representation when they heard relational language, compared to those who did not hear relational language. Moreover, children who heard relational language during the first half of the study sustained their advantage during the second half of the study, when relational language was not provided. Attentional patterns, as measured by eye tracking, indicated that children in the experimental condition had proportionally more longer fixations to the target side of the square compared to the control condition without relational language and that this effect occurred both when relational language was directly presented during the first half of the study and during the second half of the study in which relational language was not presented. Time-series analysis further revealed that children in the experimental condition

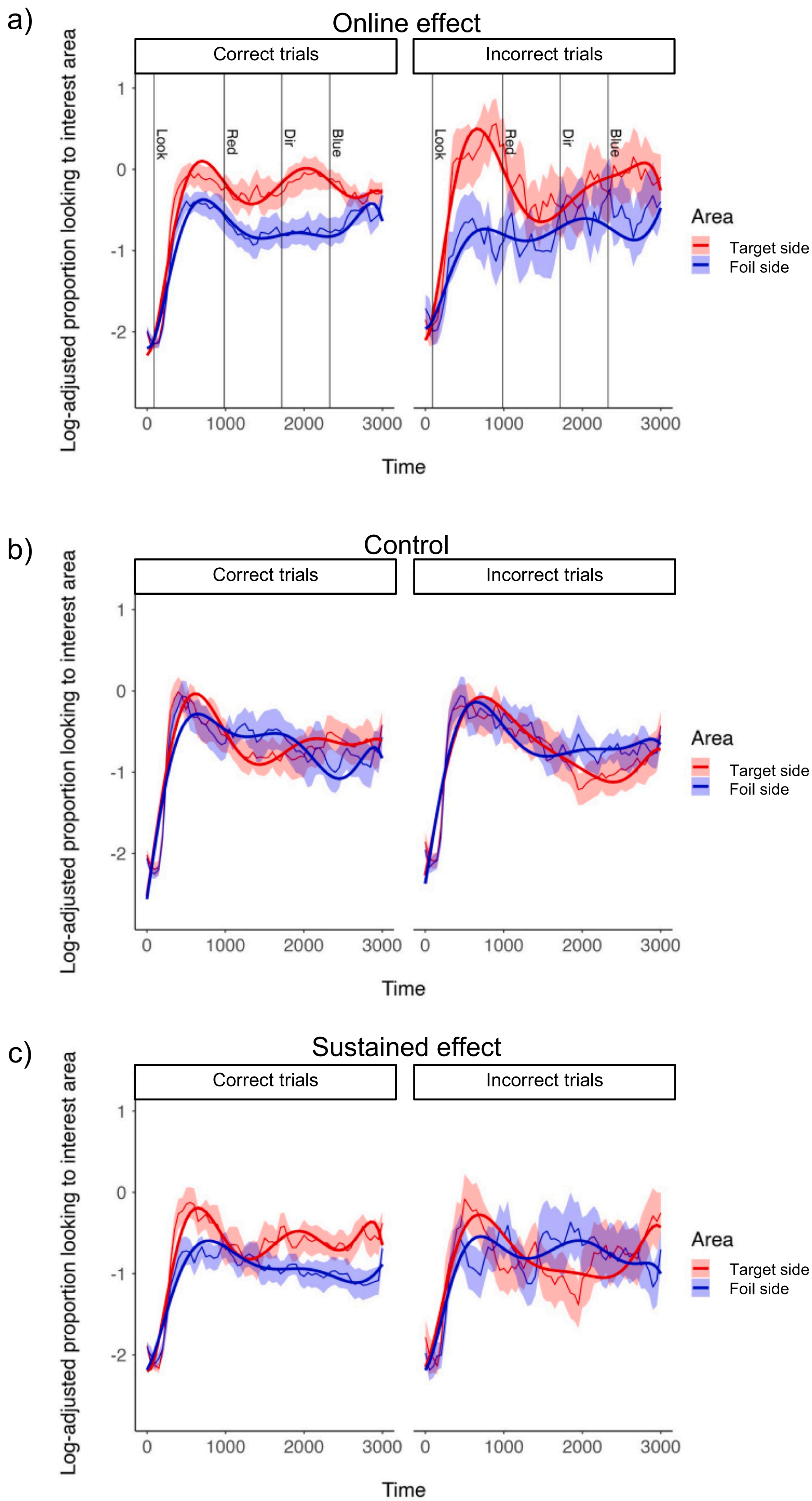

Fig. 7. Time Series Analysis of Gaze to Target and Foil for different effects separated by accuracy.

(both during the presentation of relational language and later in the absence of relational language) differentially attended to the target compared to the foil side of the square over time, while this differential attentional pattern was absent for children in the control condition without relational language. Finally, by linking gaze patterns with behavioral accuracy, it was further found that when children heard relational language online during encoding, they were more successful if their attention followed the relational language to attend to the target side and stayed there when the relational term was uttered, compared to if they just attended to the target side and then moved their attention away before the relational term was presented. When relational language was no longer presented during the second half of the study, children were only successful if they differentially attended to the target than the foil side. These results support the proposal that attention patterns underlie successful encoding of spatial relations, and that relational language enhances relational representation by guiding learners' attention.

4. Discussion

Given the important role of relational representation in human cognition and the formidable force of language in changing cognition, the current study sought to investigate the mechanism through which relational language plays a role in young children's representation of spatial relations. Overall, our results showed that: 1) hearing relational language during a task improved children's accuracy in representing the spatial relationship between two colors, 2) that this enhancement persisted over time during the study in the absence of relational language, and 3) that hearing relational language changes children's visual attention patterns both in the moment and subsequently when the language was no longer present. These results support our two central hypotheses. First, the facilitative effect of relational language on relational representation is not purely "abstract" (or internal) but is driven by changes in external sensory processes—particularly, the recruitment of visual attention. Second, the effect of relational language on relational representation is not just "momentary" but can persist over time in the absence of language. Importantly, the sustained effect of language on performance seems to be supported by the sustained effects of language on visual attention patterns.

How exactly does heard relational language guide children's attention to extract and encode relational representation in the current study? Using a similar paradigm, previous studies (Dessalegn & Landau, 2008, 2013) suggest that most 4-year-olds were successful at deploying an initial global attention to the image and rapidly extracting the overall configuration (e.g., horizontal or vertical arrangement) since most of the errors were mirror reflections, rather than the completely wrong configuration (e.g., diagonal arrangement). The current results suggest that, after this global attention, children in the language condition were able to shift their attention and selectively attend to the target side of the square guided by heard relational language (e.g., "Red is on the left of blue"), enabling them to correctly bind the featural information of the target side (e.g., red) to its geometric information (e.g., left). Further analysis revealed that this successful binding is only achieved when children differentially attended to the target side and maintained attention to the target side when the spatial term (e.g., "left") was uttered, as opposed to shifting attention away from the target side before the presentation of the spatial term. Thus, relational language enhances relational representation by guiding learners' attention; in other words, just being exposed to relational language is not enough, children's attention has to follow the syntactical structure of the language as it unfolds over time. In contrast to the asymmetrical attention pattern shown by children in the experimental condition, children in the control condition who were not presented with relational language lack such selective attention—attending non-differentially to either side of the stimuli. Because according to the attention shift account of relational representation, asymmetric attention allocation is obligatory for extracting spatial relations, children in the control condition overall were much worse than the experimental condition at extracting and encoding the relationship between the two sides of the stimuli.

In addition to the above "online" effective language, the current results suggest that hearing relational language can also change how children subsequently attend to and mentally represent relational information. Based on previous studies showing strong trial-by-trial carry-over effect in children (Chevalier & Blaye, 2009; Garon et al., 2008) and the goal of building consistent and persistent attention shift patterns, the target side of our stimulus in the current study was always the red side of the square. Thus, the relational language that children in the experimental condition heard during the first phase of the study was always "Red is on the left/right/top/bottom of blue". Consequently, children learned that "Red" is the linguistic target during encoding. This attention bias was carried over to the second phase of the study, allowing children to consistently and selectively attend to the "Red" side of the square even when relational language was no longer presented during encoding. This asymmetric allocation of attention in turn allowed children to successfully bind the featural information (e.g., red) to its geometric information (e.g., left) without the immediate aid of relational language. Importantly, children were only successful in trials in which they exhibited this asymmetrical attention pattern. This result provides proof of concept that relational language can cause sustained changes in mental representation via changing learners' attention patterns. Future studies that introduce longer delays are needed to test the robustness of this effect.

To our knowledge, the present study is the first to provide evidence for the attention shift account of relational representation in young children. It represents a crucial mechanism through which young learners extract and represent relational information. In a series of behavioral, EEG, and eye-tracking studies, Franconeri and colleagues have suggested that 1) a shift of attention is obligatory for extracting novel spatial relations between two objects by noting the direction of the shift (Franconeri et al., 2012), 2) participants are faster to verify the correctness of a sentence if they first shift their attention to the linguistic target object (Roth & Franconeri, 2012), and 3) attention shifts uniquely affect spatial relational memory between two objects independent of their identity memory (Michal & Franconeri, 2017; Michal, Uttal, Shah, & Franconeri, 2016; Yuan et al., 2016). The current study provides a plausible link between how young children and adults deploy effective visual routines to extract and build robust relational representations (Ullman, 1984), with implications for understanding how children construct and learn about relational representations in general.

For example, the current results provide a mechanistic account of how children (and even infants) might integrate featural and geometric information in spatial relational representation. Strong claims have been made about the role of language in spatial relational representation (Hermer & Spelke, 1994; Hermer-Vazquez et al., 1999; Newcombe & Uttal, 2006; Ratliff & Newcombe, 2008). Theorists from the nativistic tradition have argued that featural and location (or geometric) information are represented by different "modules" of the human cognitive system and language is the necessary "glue" that helps to bind the different sources of information (Hermer & Spelke, 1994). However, studies have (Lourenco, Addy, & Huttenlocher, 2009) shown that even infants could bind featural and location information when the features present scalar information—e.g., smaller dots on the left and bigger dots on the right—but not when the features present categorical information—e.g., red cross on the left and blue dots on the right. What is undetermined is why infants can successfully integrate featural and geometric information under some conditions but not others. The current results suggest that visual attention pattern evoked by different stimuli is a likely determining factor. Scaler information presents comparative relations, such as bigger vs. smaller dots, or higher vs. lower luminance. Such comparative relations might prompt infants to attend to the stimulus asymmetrically—e.g., prioritizing and shifting attention to the bigger vs. smaller dots. This hypothesis is undetermined based on current data but warrants future investigation.

More broadly, the current results have implications for understanding the mechanism of how relational language promotes learning in domains with systems of relational concepts. There exists substantial evidence that hearing and learning various relational linguistic structures—such as number words or spatial terms—is associated with better conceptual development in relevant domains (Levine et al., 2010; Mix & Cheng, 2012; Newcombe, Levine, & Mix, 2015; Parrish, 2014). For example, children who know more mathematical and spatial language early in life tend to do better at mathematical and spatial tasks later in school, even when the target tasks do not require the use of language (Miller & Simmering, 2018; Purpura et al., 2016). But it's unclear what learning mechanism could explain this correlational evidence. Some have argued that verbal encoding is essential for integrating different sources of perceptual information that are processed through encapsulated modules of the mind (Shusterman & Spelke, 2007). Others suggest that language may prompt children to encode critical aspects of the world—such as spatial language may encourage the encoding of spatial information (Pruden et al., 2011). Still, others argue that the development of selective attention—selectively attending to task-relevant information—itself is the driving force in conceptual development in domains such as number and space (Miller & Simmering, 2018).

The current results suggest an alternative possibility in which relational language serves as "training wheels" for developing effective selective attention that is an integral part of representing relational concepts. As demonstrated here, relational language may guide learners' attention to inspect different parts of the visual world in a systematic way, leading to more robust encoding of relational information and more accurate and consistent behavioral patterns. Day in and day out, learners experience multiple pairings of this language-guided attention, robust mental representation, and successful behavioral outcomes. Eventually, learners can automatically recruit the appropriate attention pattern for solving the task at hand without engaging in overt language use.[5] This could explain the puzzle of why children's performance in certain tasks may initially benefit from linguistic input, but linguistic skills are not predictive of success in those tasks later on (Miller & Simmering, 2018; Pruden et al., 2011). Such long-term changes in perceptual and cognitive processes likely require extensive time and many exposures. How much language exposure is needed and how extensive one can expect the effect of relational language on relational learning are undetermined in the current study but bear significant importance for understanding how language and attention can augment cognition in the long term.

In summary, the main question addressed in this study is this: What mechanism underlies the facilitation effect of relational language on young children's relation thinking, given the plethora of studies showing its benefits (Dessalegn & Landau, 2008; Gentner, Anggoro, & Klibanoff, 2011; Loewenstein & Gentner, 2005; Rattermann & Gentner, 1998; Yuan et al., 2017)? The current result provided proof of concept that relational language improves relational representation by organizing learners' attention both online when language is present and offline when language is absent. However, there are many unanswered questions regarding the larger theoretical question of how relational language affects relational thinking. For example, is it the relational as opposed to other aspects of language that engage children's attention? One past study (Dessalegn & Landau, 2013) has reported that sequentially mentioning the two sides as in the sentence "Red is touching Blue" did not improve children's performance. However, general relational language (not spatial language specifically) such as "Red is prettier than blue" did improve children's performance. The word "prettier" is a comparative relational term, not a spatial relational term per se. One hypothesis is that relational language—because of its comparative nature—is uniquely suited for eliciting asymmetrical attention patterns which in turn aids in the extraction of relational representation. If proven correct, this would refine our understanding of the mechanism of how relational language benefits learning in a host of different domains. Future studies that investigate the relationship between different linguistic inputs and moment-by-moment attention patterns of learners are critical for answering fundamental questions about the mechanisms through which language impacts cognition. While language is often viewed as enhancing cognition, research on multisensory integration in young children has shown that linguistic information, when in conflict with other critical dimensions of a task (e.g., vision, motion), can lead to poorer task performance (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2019). Thus, future studies should also examine the boundary conditions in which linguistic information at times may negatively impact cognition and learning.

5. Conclusion

The debate of language and thought has dominated the field of cognitive science for over half a century. While it is now generally agreed that language can change thought in certain contexts, there is still much to be learned about the mechanism through which language manages to do so, with important theoretical and practical implications for learning, cognitive development, and education in general. The current study highlights the embodied effect of language on thought by transforming what and how we attend in the world. Particular visual attention patterns induced by linguistic structures can have powerful effects in the encoding of new information and constructions of mental representation of the world, even when that attentional guidance or "training wheels" are no longer present during a particular task. This "language systematizes attention" mechanisms can potentially explain both why and how the acquisition of relational language can enhance the learning of systems of relations in crucial domains such as mathematics and spatial cognition.

CRediT authorship contribution statement

Lei Yuan: Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Miriam Novack:** Conceptualization, Methodology, Data curation, Project administration, Writing – review & editing. **David Uttal:** Conceptualization, Supervision, Writing – review & editing. **Steven Franconeri:** Conceptualization, Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Data will be made available on request.

Acknowledgements

The research reported here was supported by the Institute of Education Sciences (https://ies.ed.gov), U.S. Department of Education, Grant R305A120531 to Steven Franconeri, the National Science Foundation Grant IIS-1901485 to Steven Franconeri, and the National Science Foundation Grant DRK12 2200781 to Lei Yuan. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the Institute of Education Sciences. We thank Lily Zihui Zhu and Abraham Ybarra for assistance with data

[5] Although it is possible that children could in principle engage in covert language use, research has shown that spontaneous use of covert speech in problem-solving is rare for kindergarten and early elementary school aged children (Winsler & Naglieri, 2003).

analysis.

References

Bohannon, J. N., Landau, B., & Gleitman, L. (1986). Language and experience: Evidence from the blind child. *Language, 62*(2), 446. https://doi.org/10.2307/414686

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English Speakers' conceptions of time. *Cognitive Psychology, 43*, 1–22. https://doi.org/10.1006/cogp.2001.0748

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49*(3), 454–462. https://doi.org/10.1037/H0057814

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Byrge, L., Smith, L. B., & Mix, K. S. (2014). Beginnings of place value: How preschoolers write three-digit numbers. *Child Development, 85*(2), 437–443. https://doi.org/10.1111/cdev.12162

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of Cue transparency on children's cognitive flexibility. *Developmental Psychology, 45*(3), 782–797. https://doi.org/10.1037/A0015409

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41* (1–3), 83–121. https://doi.org/10.1016/0010-0277(91)90033-Z

Clark, H. H., & Clark, E. V. (1977). In J. Kagan (Ed.), *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Cooperrider, K., Gentner, D., & Goldin-Meadow, S. (2016). Spatial analogies pervade complex relational reasoning: Evidence from spontaneous gestures. *Cognitive Research: Principles and Implications, 1*(1), 28. https://doi.org/10.1186/s41235-016-0024-5

Dessalegn, B., & Landau, B. (2008). More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological Science, 19*(2), 189–195. https://doi.org/10.1111/j.1467-9280.2008.02066.x

Dessalegn, B., & Landau, B. (2013). Interaction between language and vision: It's momentary, abstract, and it develops. *Cognition, 127*(3), 331–344. https://doi.org/10.1016/j.cognition.2013.02.003

Dink, J., & Ferguson, B. (2015). eyetrackingR: An R library for eye-tracking data analysis. http://www.eyetracking-r.com.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*(5), 429–448. https://doi.org/10.1017/S0140525X0999094X

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Franconeri, S. L., Scimeca, J. M., Roth, J. C., Helseth, S.a., & Kahn, L. E.. (2012). Flexible visual processing of spatial relationships. *Cognition, 122*(2), 210–227. https://doi.org/10.1016/j.cognition.2011.11.002

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gentner, D. (2003). Why we're so smart. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). MIT Press.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82* (4), 1173–1188. https://doi.org/10.1111/j.1467-8624.2011.01599.x

Gentner, D., & Goldin-Meadow, S. (2003). Whither Whorf. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind*. The MIT Press. https://doi.org/10.7551/mitpress/4117.003.0004.

Gleitman, L., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language, 57*(4), 544–569. https://doi.org/10.1016/j.jml.2007.01.007

Gleitman, L., & Papafragou, A. (2012). New perspectives on language and thought. In *The Oxford Handbook of Thinking and Reasoning*. https://doi.org/10.1093/oxfordhb/9780199734689.013.0028

Goldin-Meadow, S. (2011). Learning through gesture. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(6), 595–607. https://doi.org/10.1002/wcs.132

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100*(2), 279–297. https://doi.org/10.1037/0033-295X.100.2.279

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462–466. https://doi.org/10.1111/1467-9280.02454

Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Gunderson, E. A., & Levine, S. C. (2011). Some types of parent number talk count more than others: Relations between parents' input and children's cardinal-number knowledge. *Developmental Science, 14*(5), 1021–1032. https://doi.org/10.1111/j.1467-7687.2011.01050.x

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370*(6484), 57–59. https://doi.org/10.1038/370057a0

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology, 39* (1), 3–36. https://doi.org/10.1006/cogp.1998.0713

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review, 98*(3), 377–389.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*(3), 499–516. https://doi.org/10.1111/J.1467-8624.1991.TB01547.X

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk". *Developmental Psychology, 42*(1), 59–69. https://doi.org/10.1037/0012-1649.42.1.59

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321. https://doi.org/10.1016/0885-2014(88)90014-7

Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology, 46*(5), 1309–1319. https://doi.org/10.1037/a0019671

Livins, K. A., Doumas, L. A. A., & Spivey, M. J. (2015). Tracking relations: The effects of visual attention on relational recognition. *Cognitive Science Conference Proceedings, 2001*, 1416–1421.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50*(4), 315–353. https://doi.org/10.1016/j.cogpsych.2004.09.004

Lourenco, S. F., Addy, D., & Huttenlocher, J. (2009). Location representation in enclosed spaces: What types of information afford young children an advantage? *Journal of Experimental Child Psychology, 104*(3), 313–325. https://doi.org/10.1016/j.jecp.2009.05.007

Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition, 108*(2), 566–577. https://doi.org/10.1016/J.COGNITION.2008.03.009

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 3*, 54. https://doi.org/10.3389/fpsyg.2012.00054

Lupyan, G., Abdel Rahman, R., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in Cognitive Sciences, 24*(11), 930–944. https://doi.org/10.1016/J.TICS.2020.08.005

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology. General, 141*(1), 170–186. https://doi.org/10.1037/a0024904

Marzolf, D. P., & DeLoache, J. S. (1994). Transfer in young children's understanding of spatial representations. *Child Development, 65*(1), 1–15. https://doi.org/10.1111/j.1467-8624.1994.tb00730.x

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Michal, A. L., & Franconeri, S. L. (2017). Visual routines are associated with specific graph interpretations. *Cognitive Research: Principles and Implications, 2*(1), 1–10. https://doi.org/10.1186/S41235-017-0059-2/FIGURES/4

Michal, A. L., Uttal, D., Shah, P., & Franconeri, S. L. (2016). Visual routines for extracting magnitude relations. *Psychonomic Bulletin & Review, 23*(6), 1802–1809. https://doi.org/10.3758/s13423-016-1047-0

Miller, H. E., & Simmering, V. R. (2018). Children's attention to task-relevant information accounts for relations between language and spatial cognition. *Journal of Experimental Child Psychology, 172*, 107–129. https://doi.org/10.1016/J.JECP.2018.02.006

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of memory and language, 59*(4), 475–494.

Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math. Developmental and educational implications. *Advances in Child Development and Behavior, 42*, 197–243. https://doi.org/10.1016/B978-0-12-394388-0.00006-X

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*(6), 1850–1870. https://doi.org/10.1111/j.1467-8624.2004.00821.x

Nedergaard, J. S. K., Wallentin, M., & Lupyan, G. (2022). Verbal interference paradigms: A systematic review investigating the role of language in cognition. *Psychonomic Bulletin & Review*. https://doi.org/10.31234/osf.io/mxbpe

Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *WIREs Cognitive Science, 6*, 491–505. https://doi.org/10.1002/wcs.1369

Newcombe, N. S., & Uttal, D. H. (2006). Whorf versus Socrates, round 10. *Trends in Cognitive Sciences, 10*(9), 394–396. https://doi.org/10.1016/j.tics.2006.07.008

Parrish, S. (2014). Number talks: Helping children build mental math and computation strategies, grades K-5. In *Math Solutions*. Sausalito, CA: Math Solutions. https://books.google.com/books?hl=en&lr=&id=p4B9F1u2T4kC&oi=fnd&pg=PR15&dq=Parrish,+S.+(2014).+Number+talks:+Helping+children+build+mental+math+and+computation+strategies,+Grades+K-5.+Sausalito,+CA:+Math+Solutions.&ots=8MFCo2yT9g&sig=AEqnVzyq49EuAd_iq7Wu0.

Pinker, S. (1994). On language. *Journal of Cognitive Neuroscience, 6*(1), 92–98. https://doi.org/10.1162/jocn.1994.6.1.92

Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science, 14*(6), 1417–1430. https://doi.org/10.1111/j.1467-7687.2011.01088.x

Purpura, D. J., Logan, J. A. R., Hassinger-Das, B., & Napoli, A. R. (2017). Why do early mathematics skills predict later reading? The role of mathematical language. *Developmental Psychology, 53*(9), 1633–1642. https://doi.org/10.1037/dev0000375

Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2016). Causal connections between mathematical language and mathematical knowledge: A dialogic reading intervention. *Journal of Research on Educational Effectiveness, 5747*(June). https://doi.org/10.1080/19345747.2016.1204639

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Ratliff, K. R., & Newcombe, N. S. (2008). Is language necessary for human spatial reorientation? Reconsidering evidence from dual task paradigms. *Cognitive Psychology, 56*(2), 142–163. https://doi.org/10.1016/j.cogpsych.2007.06.002

Rattermann, M., & Gentner, D. (1998). The effect of language on similarity: The use of relational labels improves young children's performance in a mapping task. In *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 274–282).

Rattermann, M., Gentner, D., & Deloache, J. (1990). *The effects of familiar labels on young children's performance in an analogical mapping task*. Proceedings of the ….

Richland, L. E., & McDonough, I. M. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology, 35*(1), 28–43. https://doi.org/10.1016/j.cedpsych.2009.09.001

Robinson, C. W., & Sloutsky, V. M. (2019). Two mechanisms underlying auditory dominance: Overshadowing and response competition. *Journal of Experimental Child Psychology, 178*, 317–340. https://doi.org/10.1016/j.jecp.2018.10.001

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In *Cognitive development and Acquisition of Language* (pp. 111–144). Elsevier. https://doi.org/10.1016/B978-0-12-505850-6.50010-4.

Roth, J. C., & Franconeri, S. L. (2012). Asymmetric coding of categorical spatial relations in both language and vision. *Frontiers in Psychology, 3*(November), 464. https://doi.org/10.3389/fpsyg.2012.00464

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition, 73*(1), 1–33. https://doi.org/10.1016/S0010-0277(99)00034-7

Shusterman, A., & Spelke, E. (2007). Language and the development of spatial reasoning *. *The Innate Mind: Structure and Contents*. https://doi.org/10.1093/ACPROF:OSO/9780195179675.003.0006

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. S., & Højsgaard, S. (2015). *Package 'afex'*. http://afex.singmann.science/.

Sloutsky, V. M., & Fisher, A. V. (2012). Linguistic labels: Conceptual markers or object features? *Journal of Experimental Child Psychology, 111*(1), 65–86. https://doi.org/10.1016/J.JECP.2011.07.007

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition, 60*(2), 143–171. https://doi.org/10.1016/0010-0277(96)00709-3

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*(1), 13–19. https://doi.org/10.1111/1467-9280.00403

Smith, L. B., & Samuelson, L. (2006). An attentional learning account of the shape bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology, 42*(6), 1339–1343. https://doi.org/10.1037/0012-1649.42.6.1339

Spencer, J. P., Simmering, V. R., Schutte, A. R., & Schöner, G. (2007). What does theoretical neuroscience have to offer the study of behavioral development? *The Emerging Spatial Mind*, 320–361. https://doi.org/10.1093/acprof:oso/9780195189223.001.0001

Talmy, L. (1983). How language structures space. In H. L. Pick, & L. P. Acredolo (Eds.), *Spatial orientation* (pp. 225–282). Springer US. https://doi.org/10.1007/978-1-4615-9325-6.

Ullman, S. (1984). Visual routines. In *Cognition* (pp. 97–159). Elsevier. https://doi.org/10.1016/0010-0277(84)90023-4. Vol. 18, Issues 1–3.

Vales, C., & Smith, L. (2015). Words, shape, visual search and visual working memory in 3-year-old children. *Developmental Science, 1*, 1–15. https://doi.org/10.1111/desc.12179

Wakefield, E., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science, 21*(6), Article e12664. https://doi.org/10.1111/desc.12664

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*(3), 217–242.

Wilkinson, A. C. (1982). Theoretical and methodological analysis of partial knowledge. *Developmental Review, 2*(3), 274–304. https://doi.org/10.1016/0273-2297(82)90015-6

Winsler, A., & Naglieri, J. (2003). Overt and covert verbal problem-solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5 to 17. *Child Development, 74*(3), 659–678. https://doi.org/10.1111/1467-8624.00561

Yuan, L., Uttal, D., & Franconeri, S. (2016). Are categorical spatial relations encoded by shifting visual attention between objects? *PLoS One, 11*(10), 1–23. https://doi.org/10.1371/journal.pone.0163141

Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. *Developmental Psychology, 53*(6), 1098–1114. https://doi.org/10.1037/dev0000302

Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review, 21*(1), 1–22. https://doi.org/10.3758/s13423-013-0443-y

Zhou, K., Mo, L., Kay, P., Kwok, V. P. Y., Ip, T. N. M., & Tan, L. H. (2010). Newly trained lexical categories produce lateralized categorical perception of color. *Proceedings of the National Academy of Sciences of the United States of America, 107*(22), 9974–9978. https://doi.org/10.1073/PNAS.1005669107/ASSET/E3917755-7B71-4138-8617-7EF97FAB1C52/ASSETS/GRAPHIC/PNAS.1005669107FIG 04.JPEG