Shao, R., & Gentner, D. (2022). Perceptual alignment contributes to referential transparency in indirect learning. Cognition 224.

Cognition 224 (2022) 105061

Contents lists available at ScienceDirect

Cognition

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

Perceptual alignment contributes to referential transparency in indirect learning☆

Ruxue Shao, Dedre Gentner *

Department of Psychology, Northwestern University, USA

ARTICLE INFO

Keywords:
Incidental word learning
Structure-mapping
Adjectives

ABSTRACT

Young children can sometimes acquire new word meanings—even for property terms—through incidental learning (e.g. Carey & Bartlett, 1978). We propose that an important support for this process is spontaneous perceptual comparison processes that lead children to notice key commonalities and differences. Specifically, we hypothesize that when the target property appears as a difference between two highly similar and alignable objects, spontaneous comparison processes operate to highlight the property. The property may then be linked to an accompanying word, even if the child has no prior intention to learn the word. To test this, we revisited the Carey and Bartlett paradigm, varying the perceptual alignability of the objects that 3- and 4-year-olds saw while hearing a novel color word, *chromium*. In Experiments 1 and 2, children in the High Alignment condition were able to identify chromium objects in a subsequent task, whereas those in the Low Alignment condition were not. Experiment 3 showed that direct instructions to learn the word led to a different pattern of results. Experiment 4 showed that the incidental learning persisted over delay and transferred to new objects. We conclude that perceptual alignment contributes to referential transparency and to incidental learning of word meanings. Implications for hypothesis-testing theories of word learning are discussed.

1. Introduction

When thinking about early word acquisition, people often envision a situation in which a caretaker provides a child with a label for some object of interest. For example, a caretaker may point to a rubber duck and say, "Look! This is a duck!" Over the past decades, we have accumulated much knowledge about how children acquire noun labels in this kind of direct naming situation. Researchers have characterized factors that support early word learning, including perceptual salience (e.g. Landau, Smith, & Jones, 1988; Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006; Yu & Smith, 2011), perceptual individuability (Gentner, 1982; Gentner & Boroditsky, 2001); conceptual knowledge (e.g. Booth & Waxman, 2009; Clark, 1987; Fernald, Zangl, Portillo, & Marchman, 2008; Markman, 1989), prosodic cues (e.g. Fernald & Mazzie, 1991; Jusczyk, Houston, & Newsome, 1999; Mattys, Jusczyk, Luce, & Morgan, 1999), joint attention (e.g. Brooks & Meltzoff, 2008; Tomasello & Farrar, 1986), social context (e.g. Akhtar & Tomasello, 2000; Baldwin et al., 1996; Behne, Carpenter, & Tomasello, 2005; Ferguson & Lew-Williams, 2016; Tomasello & Barton, 1994), and syntactic cues (e.g. Bloom & Kelemen, 1995; Fisher, Gertner, Scott, & Yuan, 2010; Gleitman, 1990). Any or all of these cues can contribute to what Cartmill et al. (2013) called "referential transparency"—the degree to which the situation in which a word is uttered is conducive to the child's learning the word.

In this work we investigate another source of referential transparency—available perceptual comparisons in the environment. We suggest that children's spontaneous comparison processes can lead them to grasp word meanings that might otherwise be missed. Our studies test the power of spontaneous comparison in a particularly challenging arena—the acquisition of property words via incidental learning. Before describing the studies, we review the relevant research.

☆ This research was supported by NSF SLC Grant SBE-1041707 awarded to the Spatial Intelligence and Learning Center (SILC), by NSF Grant BCS-1729720 awarded to Susan Hespos and Dedre Gentner, and by ONR Grant N00014–16-1-2613 awarded to Dedre Gentner and Ken Forbus.

* Corresponding author at: Department of Psychology, Northwestern University, Swift Hall 102, 2029 Sheridan Road, Evanston, IL 60208, USA.
E-mail address: gentner@northwestern.edu (D. Gentner).

https://doi.org/10.1016/j.cognition.2022.105061
Received 8 November 2019; Received in revised form 22 August 2021; Accepted 6 February 2022
Available online 18 February 2022

1.1. Incidental learning of word meaning

Although most studies of word learning have involved direct word learning, children can also acquire new words in a wide range of indirect[1] situations. Infants as young as 18-months of age can learn labels for novel objects through overhearing others' conversations (e.g. Akhtar, 2005; Callanan, Akhtar, & Sussman, 2014; Foster & Hund, 2012; Gampe, Liebal, & Tomasello, 2012). At around the same age, they can also learn words for objects through media consumption, especially with the guidance of an adult viewer (e.g. Krcmar, Grela, & Lin, 2007; Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009). By preschool, children can also acquire new words through joint reading activities (e.g. Blewitt, Rump, Shealy, & Cook, 2009; Justice, 2002; Sénéchal, Thomas, & Monker, 1995).

Most of these studies have focused on noun learning—a reasonable first step, since children's early vocabulary is strongly noun-focused (Bornstein et al., 2004; Braginsky, Yurovsky, Marchman, & Frank, 2015; Gentner, 1982, 2006; Gentner & Boroditsky, 2001, 2009; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005). However, there are also studies showing that preschoolers can learn aspects of verb meaning from indirect speech (e.g., Arunachalam & Waxman, 2010; Yuan, Fisher, & Snedeker, 2012). In one study, Yuan and Fisher (2009) showed 2-year-old children videos of two women chatting with each other about unseen events. The speakers used a novel verb several times in either a transitive frame (e.g., "Jane blicked the baby") or an intransitive frame ("Jane blicked"). After this, children were shown two video clips, one depicting a causal event (one girl swinging another girl's leg) and the other depicting a non-causal event (one girl waving her arm). When asked to "Find blicking," children who had heard transitive dialogues looked significantly longer at the causal event than those who had heard intransitive dialogues. Subsequent studies showed that 21-month-olds can succeed at this task when the number of participants is held constant (Arunachalam, Escovar, Hansen, & Waxman, 2013), and that 22-month-olds can retain the information gained indirectly from a dialogue phase over a brief delay (Messenger, Yuan, & Fisher, 2015).

Thus there is evidence that preschoolers can learn about the meanings of nouns and verbs from indirect speech. However, much less work has addressed whether children can acquire property terms through indirect means, and if so, how they are able to do so.

1.2. Learning adjectives

Adjectives are conceptually demanding. Because they denote specific properties of objects, a child must process the object referent before they can comprehend the target property. A child who hears "Look at the bifish dax" needs to be able to connect *dax* with the referent object, as well as to discover the intended dimension (e.g., color) and the intended value along the dimension (Gasser & Smith, 1998; Sandhofer & Smith, 2004). A further challenge is that during early language learning, English-speaking children tend to acquire a focus on nouns (Gentner & Boroditsky, 2001; Hoyos, Shao, & Gentner, 2016). There is evidence that this can temporarily impede the learning of other syntactic categories: Sandhofer and Smith (2007) found that children with fewer nouns in their productive vocabularies learned more adjectives than those with more nouns, unless strong syntactic cues were present. In addition, adjectives have relatively low frequency in parental input (Salerni, Assanelli, D'Odorico, & Rossi, 2007; Sandhofer, Smith, & Luo, 2000; Tribushinina & Gillis, 2012). Not surprisingly, then, corpus studies have found that children are slower to learn property words than concrete nouns (e.g. Bates et al., 1994; Bornstein et al., 2004; Braginsky et al., 2015; Gentner, 1982; Gentner & Boroditsky, 2001, 2009; Gleitman et al., 2005).

The same pattern is found in experimental studies. For example, 14-month-olds who were directly taught a novel noun correctly extended the label to object categories, and not to same-property items (Booth & Waxman, 2003; Waxman & Booth, 2001); but when taught a novel adjective, 14-month-olds, and even 18-month-olds, showed no preference between object-based vs. property-based extensions (Booth & Waxman, 2003, 2009; Waxman & Booth, 2001). Taylor and Gelman (1988) found similar patterns with 2-year-olds in a more complex word-learning task. Even among 3- and 4-year-olds, children often fail to take an adjective as a property term without additional assistance, such as applying the novel adjective to a highly familiar object, using highly similar training and initial test items, or providing rich referential and syntactic information (Hall, Waxman, & Hurwitz, 1993; Klibanoff & Waxman, 2000; Mintz & Gleitman, 2002; Waxman & Klibanoff, 2000). Overall, research suggests that young children learn early to attach words to objects, and only slowly learn to attach words to properties of objects.

The fact that 3- and 4-year-old children have difficulty learning property terms even under direct presentation conditions might lead us to be pessimistic concerning the likelihood of success in indirect learning. Yet, in an important study, Carey and Bartlett (1978) found that a single incidental learning experience enabled many 3–4-year-olds to learn a new color adjective. Because this study is central to our discussion, we describe it in some detail.

1.3. Carey and Bartlett's (1978) study

In Carey and Bartlett's (1978) classic study, 36-to 46-month-olds were led to connect a novel word (*chromium*) with an unfamiliar color (dark olive-green). During snack time at preschool, their teacher asked, "You see those two trays over there? Bring me the chromium one. Not the red one, the chromium one." (Carey & Bartlett, 1978, p.18). This was the full extent of the initial exposure: children's exposure to the new word and color was embedded within a task, without a direct referential label nor any pedagogical cues to signal the interaction as a potential learning experience. Thus, this paradigm serves as an apt method for studying incidental learning.

Seven-to-ten days after the initial exposure, children were assessed on their understanding of the word *chromium* and the color dark olive-green. They were exposed to the word and color again ten weeks later in two encounters that occurred two days apart. Seven-to-ten days after the second encounter, the same battery of tests was once again administered. The assessment included a color-naming task in which children were asked to label colors, including dark olive-green, and a comprehension task in which children were asked to pick *chromium* from an array of nine colors. Carey and Bartlett (1978) found that over half of the children retained knowledge of the word, even after this delay. Heibeck and Markman (1987) corroborated these findings. Using a similar procedure, they found that 2–4-year-olds could learn a novel property words from this indirect method.

The Carey and Bartlett (1978) study is a milestone study for at least three reasons. First, it showed that children can learn a word's meaning from one exposure. This kind of rapid learning from one exposure (*fast mapping*) has inspired many further studies (e.g. Bion, Borovsky, & Fernald, 2013; Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Horst & Samuelson, 2008; Spiegel & Halberda, 2011). Second, this research utilized indirect word learning, rather than direct labeling. Third, it focused on property terms, which, as reviewed above, are typically learned rather late.

Although most of the follow-up studies have focused on the first point—that children can fast-map the meaning of a word from one or a few exposures—the second two points are equally remarkable. Against the general pattern of relatively slow learning of property words, Carey and Bartlett's finding that 3- and 4-year-olds could learn a property term

[1] Throughout this paper, we will use "incidental word learning" for the case in which children learn a word's meaning through an interaction in which the word is used, but with no intention to teach the word. We will use the more general term "indirect word leaning" to subsume both incidental word learning and word learning via overhearing.

from one indirect exposure is striking. After all, in their task, children had no need to learn the meaning of *chromium*; they could satisfy the teacher's request simply by giving her the tray that was not red. Given the double challenge of learning property terms and learning via incidental input, the finding that some of their 3-to-4-year-old participants encoded and retained the new word's meaning is quite striking. How did children accomplish this feat?

1.4. Structure-mapping framework for indirect property-word learning

We propose that a key process in incidental learning of property terms is perceptual comparison. In Carey and Bartlett's (1978) study, the two trays were extremely similar—identical in all but their color—and therefore likely to be compared. There is considerable evidence that comparing two things involves a process of structural alignment that seeks relational correspondences between the two representations (Falkenhainer, Forbus & Gentner, 1989; Forbus, Ferguson, Lovett, & Gentner, 2017; Forbus & Gentner, 2017; Gentner, 1983, Gentner, 2010; Gentner & Markman, 1997). Structural alignment renders common structure more salient (Christie & Gentner, 2010; Graham, Namy, Gentner, & Meagher, 2010; Markman & Gentner, 1993; Medin, Goldstone, & Gentner, 1993). More relevant to the present paper, structural alignment also renders *alignable differences*[2]—differences that play the same role in the common structure—more salient (Gentner & Gunn, 2001; Gentner & Markman, 1994; Hoyos, Horton, Simms, & Gentner, 2020; Markman & Gentner, 1993; Matlen, Gentner, & Franconeri, 2020; Sagi, Gentner, & Lovett, 2012).

Another key point in this discussion is the special role of high-similarity comparison. According to structure-mapping theory, the process of structural alignment should be faster and easier for high-similarity pairs than for low-similarity pairs. Because alignable differences tend to 'pop out'[3] when two representations are structurally aligned, this means that high-similarity comparisons should give rise to rapid noticing of differences (Falkenhainer et al., 1989; Forbus et al., 2017; Gentner & Forbus, 2011; Sagi et al., 2012). This prediction might seem paradoxical, since there are bound to be more differences between low-similarity pairs than between high-similarity pairs. But there is considerable evidence that high-similarity comparison supports rapid, spontaneous difference-detection (Gentner & Gunn, 2001; Sagi et al., 2012), in both conceptual and perceptual comparisons. Gentner and Markman (1994) gave adults a large set of word pairs and asked them to find a difference between as many pairs as possible in a brief time period. Participants identified differences for many more high-similarity pairs than low-similarity pairs, and this surplus was chiefly made up of alignable differences (see also Gentner & Gunn, 2001).

The same pattern—that it is easier to generate differences for high-similarity pairs than between low-similarity pairs—is found for perceptual comparison. Markman and Gentner (1996) gave adults image pairs and asked them to list either differences or commonalities. Again, participants listed more differences for highly similar images than for less similar ones, and again, this surplus was made up of alignable differences. Sagi et al. (2012) directly measured people's response time to state a difference between two images. The results confirmed that people were faster for high-similarity image pairs than for low-similarity pairs. Subjective reports suggested a popout effect—sudden effortless emergence of a difference—for the high-similarity pairs. Indirect evidence for popout of alignable differences in children[4] can be seen in a study conducted in the Chicago Children's Museum (Gentner et al., 2016), in which children constructed toy buildings out of an erector set. The buildings frequently collapsed, because children often used only vertical and horizontal pieces in their models, neglecting the critical role of diagonal bracing in creating stability (Benjamin, Haden, & Wilkerson, 2010). Gentner et al. (2016) gave 8-year-olds a brief comparison experience just prior to the study. They were shown two sample buildings, one stable and one unstable (no bracing). For one group, the buildings were highly alignable; for the other group,[5] the two buildings looked rather different, making them less easy to align. When children were asked "Which is stronger?" and allowed to wiggle the buildings, all children recognized that the braced building was stronger (more stable). Although the brace was never pointed out, over 80% of the children in the high-alignment condition used a diagonal brace in a subsequent "repair the wobbly building" task. Only 50% (chance level) of the children in the low-alignment group did so.

There is also evidence that perceptual comparison is instrumental in direct learning of word meanings—including adjective meanings (Au & Laframboise, 1990; Gelman & Markman, 1985; Sandhofer & Smith, 2001; Tribushinina et al., 2013; Waxman & Klibanoff, 2000). For example, Gelman and Markman (1985) found that 4-year-olds invoked contrast among objects in applying adjectives, but not nouns. More specifically, Waxman and Klibanoff (2000) found evidence for comparison and alignment in direct learning of adjectives for texture. They showed 3-year-olds (for example) a toy horse covered with bumps and said "This is a very blickish horse." They then showed the child two other animals and asked them to find "another one that's blickish." Children performed well when the alternatives were highly similar (and alignable) to the initial standard—for example, two similar horses, one bumpy and one not. However, when the two alternatives were dissimilar to the initial horse standard—for example, two similar rhinoceroses, one bumpy and one not—3-year-olds were at chance. In a second experiment, 3-year-olds who first received the high-alignable alternatives went on to succeed with low-alignable alternatives; but those who began with low-alignable alternatives performed below chance on both the first set and the second (low-alignable) set. This sequence of a close comparison followed by far comparisons (i.e., comparisons that share relational structure but relatively few surface commonalities) is known as *progressive alignment* and has been shown to support relational insight and generalization (Gentner, Anggoro, & Klibanoff, 2011; Gentner & Namy, 2006; Kotovsky & Gentner, 1996; Thompson & Opfer, 2010).

The studies just reviewed (aside from the Carey and Bartlett (1978) study) have used direct learning tasks. In the present work, we advance the hypothesis that spontaneous comparison processes enable incidental learning of property terms. As reviewed above, for high-similarity pairs, (1) the process of structural alignment is fast and essentially effortless (Gentner & Hoyos, 2017; Kurtz & Gentner, 2013; Sagi et al., 2012); and (2) alignable differences tend to pop out as highly salient (Gentner & Gunn, 2001; Gentner & Markman, 1994; Markman & Gentner, 1993; Sagi et al., 2012). Applying these ideas to Carey and Bartlett's study, we hypothesize that, while simply complying with the request to "give me the chromium one, not the red one," children experienced a spontaneous comparison between the two highly similar trays, resulting in pop-out of the alignable difference in color—chromium vs. red. The juxtaposition of this salient new color with the new word, *chromium,* led some children to encode the connection, even without any intention to learn a new word.

In this paper, we present a series of studies testing this account. We adopted a paradigm similar to that used by Carey and Bartlett (1978) but

[2] An example alignable difference is "Motorcycles have two wheels, cars have four." An example nonalignable difference is "A car has room for your dog." (See Markman & Gentner, 1993 for the coding system).

[3] 'Pop-out' in visual processing refers to the phenomena whereby a visual feature is noticed via preattentive processes (Pashler, 1998; Treisman & Gelade, 1980; Treisman, Vieira, & Hayes, 1992). Pop-out typically depends on contrast between simple features: for example, a vertical search target will pop out from a field of horizontal distractors.

[4] Admittedly, this evidence is indirect, as we did not measure response times.

[5] There was also a No-training group, which also failed to exceed chance on the repair task above.

varied the alignability of the pairs. Our account predicts that (1) when the object pairs are highly alignable, as in Carey and Bartlett's experiment, young children will be likely to learn the new color term (*chromium*) even with indirect exposure, and (2) this kind of spontaneous word learning should be much less likely when object pairs are less easy to align. However, an alternate possibility is that the perceptual alignment of items was not a key component in children's learning. Instead, other cues within the paradigm – such as the linguistic contrast with a familiar color term – may have been sufficient for children to focus on the distinguishing property of color and correctly map color to word. In this case, we should see little or no effect of perceptual alignability on children's indirect property-word learning. In the present research, we test this by varying the perceptual alignability of the pair, while holding the linguistic cues constant. The logic is straightforward: first, if perceptual comparison is instrumental in learning the meaning of *chromium*, then children will be less likely to learn the meaning if given a less readily alignable pair (Experiments 1, 2, and 4). Second, if the alignment process in this case is a spontaneous, nonintentional process, then the pattern will be quite different if children are asked to learn the word's meaning (Experiment 3).

2. Experiment 1

We presented 3- and 4-year-olds with a pair of objects—one blue, one dark olive green ("chromium"[6])—and asked the children to "point to the chromium one, not the blue one". There were two conditions: children in the High Alignment (HA) condition saw two objects that were highly similar (identical except for color) and therefore easy to align, whereas those in the Low Alignment (LA) condition saw objects that were less similar and therefore more difficult to align. After a break, we assessed their understanding of the word *chromium* through a yes-no classification task. We predicted that children in the HA condition would be more likely to encode the new chromium color and would therefore show more accurate classification than those in the LA condition.

2.1. Methods

2.1.1. Participants

A total of 128 typically developing children[7] participated in this study, including 65 3-year-olds ($M = 42.89$ months, $SD = 2.89$ months, females = 32) and 63 4-year-olds ($M = 54.05$ months, $SD = 3.28$ months, females = 32). Children were assigned to either the High Alignment (HA) or Low Alignment (LA) condition, with age and gender counterbalanced across the two conditions. An additional 25 children (five 4-year-old and twenty 3-year-olds) were tested but excluded from further analyses, six children due to failing to complete the experiment, four children due to experimenter error, and fifteen children for responding either all-yes or all-no in the Meaning Assessment test (see below for details). In this experiment and all succeeding experiments, children were recruited from the suburban Chicago area through a voluntary participant pool or local preschools; the racial and economic composition of the sample reflected those of the local population (approximately 80% European-American, 20% Hispanic, multiracial or Black; middle- and upper-middle class). Children were given a small gift for participation. In this experiment and all following experiments, parents gave written informed consent for their children to participate and children gave verbal assent. All research activities were reviewed and approved by Northwestern University's Institutional Review Board.

2.1.2. Procedure and materials

The experiment took place in a single session. Children were initially exposed to the novel word, *chromium*, in an indirect learning context. The word *chromium* was used only once, as in Carey and Bartlett's study. After a break of approximately ten minutes, they were assessed on their understanding of the novel word through a yes-no classification task.

2.1.2.1. Initial exposure. The Initial Exposure to the novel word, *chromium,* was embedded within a pointing task in which children were asked to point to one of two options. The task was designed to be quite easy, to preserve the sense of a fun game. For example, on the first trial, children saw a lion and a cow and were requested to "point to the lion—the lion, not the cow" or vice versa. The *chromium* exposure trial was always the third trial and was followed by a fourth trial to prevent recency effects. Appendix A shows the full set of materials used in the pointing game. The first two trials and the final trial were identical for the HA and LA conditions.

On the *chromium* exposure trial, children in both the HA and LA conditions saw geometric shapes made of foam and of approximately the same size. Those in the HA condition saw two highly similar objects—a chromium square and a blue square—whereas those in the LA condition saw two objects that were similar in size but not shape—a chromium circle and a blue square (Fig. 1). All children were asked, "Look at these two! Can you point to the chromium one? The chromium one, not the blue one." Note that, for both groups, this task could be solved simply by pointing to the 'not blue' one; however, our prediction is that, at least in the HA condition, children will register the color contrast. Left/right placement of the objects was counterbalanced in all trials. For the first two and final trials, the target object was also counterbalanced.

2.1.2.2. Break. Between the Initial Exposure and Meaning Assessment, children completed an unrelated filler task for approximately 5 min. Then children went on to a yes-no classification task intended to serve as a warm-up for the key Meaning Assessment test. In the warm-up task, children were given a series of 12 easy classification decisions: e.g., Child was asked "Is this an animal" while being shown a picture of a truck (See Appendix B [online supplemental material]). Children were shown a Yes Box with a smiley face and a No Box with a frowny face and were asked to put the cards in the corresponding box. No corrective feedback was given. To be included in the final analysis, children had to reach a criterion of 8 of 12 trials correct.

2.1.2.3. Meaning assessment. The Meaning Assessment test immediately

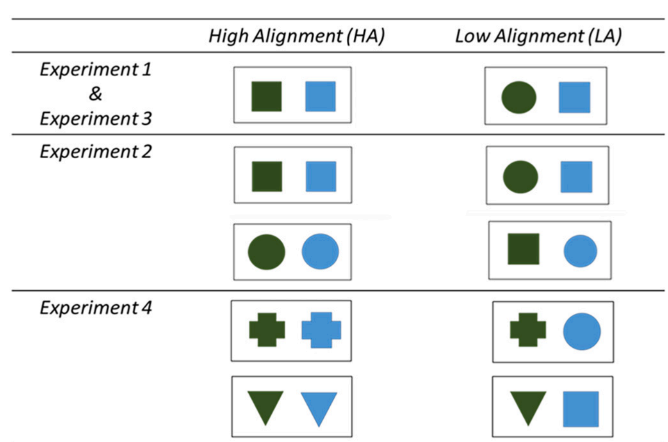

Fig. 1. Initial Exposure objects presented to children in Experiments 1, 2, 3, and 4.

[6] In the following text, chromium will be used for the color and *chromium* to refer to the word.

[7] Because this was a new venture, we set an initial sample size of 32–33 per group. Based on the findings of Study 1, we used GPower to calculate minimal sample sizes for the succeeding studies. This yielded larger desired sample sizes for 3-year-olds than for 4-year-olds. Study 2 (3-year-olds) was slightly underpowered, with 27–28 per group. The remaining studies (Studies 3 and 4, with 4-year-olds) were adequately powered, with sample sizes of 20–22.

followed the warm-up classification task and was continuous with it. Children were shown eight geometric figures of various colors and shapes, one at a time in a pseudo-random order so that no consecutive objects shared the same color or shape (Fig. 2). As each object was presented, the experimenter asked, "Look at this one! Is this a chromium one?" The object was placed in the Yes Box or the No Box based on the child's answer.[8] Children were excluded from further analyses for responding either all-yes or all-no in the Meaning Assessment test.

The eight figures consisted of three chromium figures—a square, a circle, and a hexagon—and five non-chromium figures—a blue square, a blue circle, a maroon triangle, an orange circle, and a purple square. (See Fig. 2a). Among the chromium figures, the chromium square had been seen in the Initial Exposure in the HA condition; likewise, the chromium circle had been seen in the LA condition. The chromium hexagon was new to both groups. The maroon triangle had been seen by all children in one of the filler trials of the Initial Exposure.

2.2. Results

2.2.1. Initial exposure

Three- and four-year-olds in both the HA and LA conditions pointed to the target with high accuracy. Of the 128 children, 124 (96.9%) pointed to the chromium object on the first attempt. All four children who failed to do so were 3-year-olds, with three of them in the LA condition. These children were excluded from further analysis, since we do not expect children to map the word *chromium* to its corresponding color if they failed to identify the correct referent object.

2.2.2. Meaning assessment

The Meaning Assessment test was used to assess whether children had learned the meaning of *chromium* during the Initial Exposure phase. We scored the number of hits (chromium objects correctly endorsed; max. = 3) and false alarms (non-chromium objects incorrectly endorsed; max. = 5) made by each child.

2.2.2.1. Correct identifications/hits. To test the prediction that children in the HA condition would be more likely to derive the meaning of chromium than those in the LA condition, we conducted planned comparisons on the number of hits. Children in the HA condition made more hits than those in the LA condition, both among the 3-year-olds ($M_{HA} = 2.13$, $SD_{HA} = 1.10$; $M_{LA} = 1.48$, $SD_{LA} = 1.02$; $t(59) = 2.36$, $p = .01$, $d = 0.61$, one-tailed *t*-test[9]) and among the 4-year-olds ($M_{HA} = 2.50$, $SD_{HA} = 0.76$; $M_{LA} = 1.77$, $SD_{LA} = 1.15$; $t(61) = 2.97$, $p = .002$, $d = 0.75$, one-tailed *t*-test). In addition, only the HA group performed better than chance (chance = 1.5 hits; both $ps < 0.004$, two-tailed *t*-tests[10]); the LA group did not differ from chance (both $ps > 0.19$, two-tailed *t*-tests). Fig. 3 shows the mean number of hits made by each Age × Alignment group.

2.2.2.2. Generalization check. An important question is whether high alignment helped children generalize *chromium* to new objects. Therefore we looked specifically at the two new chromium objects in the Meaning Assessment test. For 4-year-olds, but not 3-year-olds, those in the HA condition were more likely than those in the LA condition to correctly identify both of the two new chromium objects. Table 1 shows the generalization patterns across studies, showing an ordinal advantage for HA in all four studies.

2.2.2.3. False alarms. An alternate explanation for the greater number of hits in the HA condition than in the LA condition is that, for whatever reason, children had a stronger tendency to respond positively in the HA condition. To assess this possibility, we examined false-alarm rates—rates of saying "yes" to the five non-chromium objects. Regardless of Age or Alignment condition, none of the groups made more false alarms than chance[11] (chance = 2.5 false alarms). Children's low false alarm rates make it unlikely that the results are due to a 'yes' bias.

2.3. Experiment 1 discussion

These findings provide evidence that perceptual alignment can support incidental word learning. Although virtually all children in both conditions accurately pointed to "the chromium one, not the blue one" during the Initial Exposure task, only those in the HA group learned the meaning of *chromium* from this experience. The HA group performed significantly better than the LA group on the Meaning Assessment test; in fact, the LA group (both ages) performed at chance. We suggest that structural alignment led the HA group to learn a new word, even without any intention to do so. Children in the HA group experienced spontaneous structural alignment between the two highly similar objects during the Initial Exposure task, and this led to pop-out of the alignable difference between the familiar color (blue) and the novel color (dark olive green). The salient new color became linked with the salient new word, *chromium*. On this account, children were not trying to learn the new word; rather, a spontaneous alignment process operated to permit a new connection.

However, although the results so far are consistent with this account, they could also be explained in a very different way. Perhaps the performance differences between the HA and LA groups resulted simply

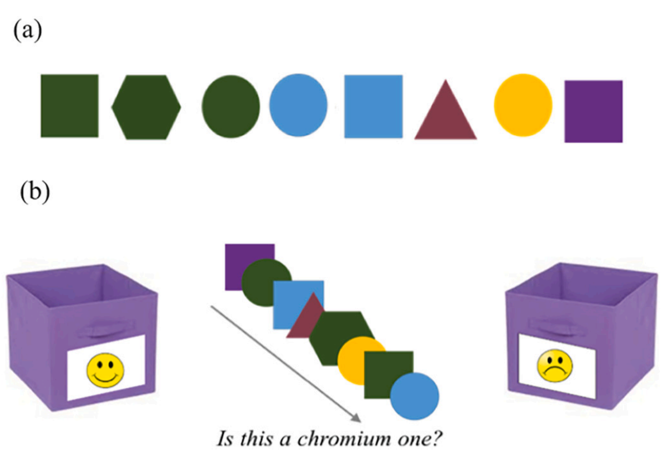

Fig. 2. Schematics of (a) Meaning Assessment objects presented to children in Experiments 1, 2, and 3; (b) Meaning Assessment procedure. Objects were presented in pseudo-random order in all studies.

[8] If the child failed to answer, or indicated that they did not know, the experimenter repeated the question. If the child still did not answer, the experimenter placed the Yes Box and No Box side by side in front of the child and asked, "Is this a chromium one? What do you think? Which box should we put it in?" Across all experiments, no participant required more than three probes to answer a question.

[9] We predicted that children in the HA condition would outperform their peers in the LA condition. Thus, we use one-tailed *t*-tests throughout this paper when comparing the performances of children in the HA and LA conditions.

[10] We did not have directional hypotheses regarding children's performance compared to chance. Thus, we used two-tailed *t*-tests throughout this paper when comparing performance with chance.

[11] The mean number of false alarms made by 4-year-olds in the HA condition ($M = 1.72$, $SD = 1.35$), 4-year-olds in the LA condition ($M = 1.23$, $SD = 0.96$), and 3-year-olds in the LA condition ($M = 2.00$, $SD = 1.13$) were all significantly lower than chance (all $ps < 0.03$, two-tailed *t*-test). The mean number of false alarms made by 3-year-olds in the HA condition was not significantly different from chance ($M = 2.03$, $SD = 1.40$; $p = .07$, two-tailed *t*-test).

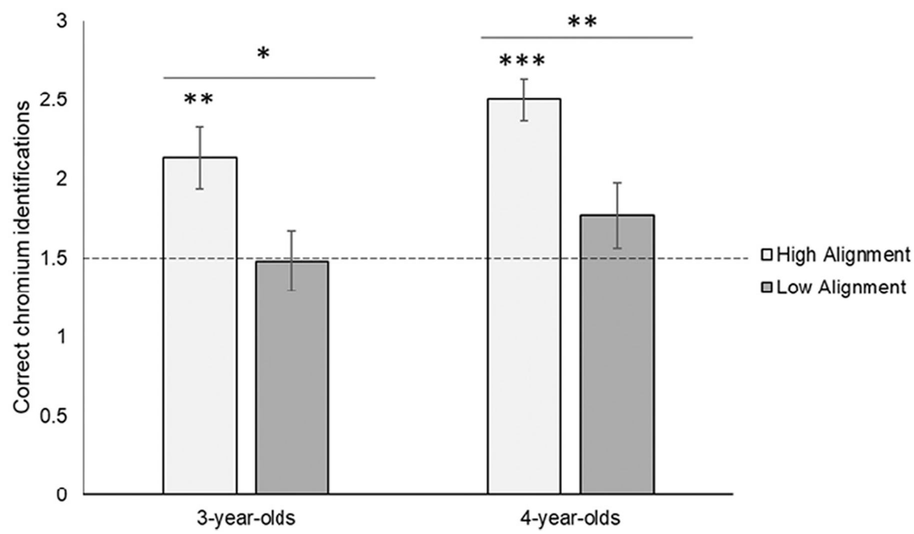

Fig. 3. Experiment 1: Mean number of chromium objects (max = 3) correctly endorsed as *chromium* (hits) in the Meaning Assessment test. Error bars depict +/− one standard error. The dashed line represents chance performance. * $p < .05$, ** $p < .01$, *** $p < .005$.

Table 1
Generalization rate: Across studies, the proportion of children in each condition who endorsed the new chromium object(s) in the Meaning Assessment Task.

		High Alignment		Low Alignment	Chi-square
Experiment 1	3-year-olds	0.50		0.38	$X^2(2, N = 61) = 3.79, p = .15, \varphi = 0.25$
	4-year-olds	0.97	*	0.71	$X^2(2, N = 63) = 7.41, p = .03, \varphi = 0.34$
Experiment 2	3-year-olds	0.81	*	0.50	$X^2(2, N = 52) = 5.44, p = .02, \varphi = 0.32$
Experiment 3	4-year-olds	0.77		0.59	$X^2(2, \mathrm{N} = 44) = 1.93, p = .38, \varphi = 0.21$
Experiment 4	4-year-olds	0.84		0.57	$X^2(1, N = 40) = 3.48, p = .06, \varphi = 0.30$

Note: Proportions of children endorsing all possible new chromium objects. The number possible differed across studies: Experiment 1: 2 new chromium objects; Experiment 2: 1; Experiment 3: 2; Experiment 4: 3. The Experiment 4 results are for the Day 1 task.

from differences in the amount of available information. The two objects shown to the HA group differed in only one dimension – color. In contrast, the LA group saw objects that differed in *both* color and shape. From a hypothesis-testing point of view, the LA participants had to consider two possible hypotheses for the meaning of *chromium*—a color hypothesis (olive green) and a shape hypothesis (round)—whereas the HA participants only had to consider the color hypothesis. Thus, the HA advantage could have resulted from differences in the number of pre-existing hypotheses children had to consider, rather than from their having noticed the color contrast online through a spontaneous alignment process. Experiment 2 was designed to distinguish between these two possibilities.

3. Experiment 2

In Experiment 2 we equated the information that children received by giving all children two Initial Exposure trials. These trials were designed so that after the second trial, all children had sufficient information to logically conclude that color was the relevant dimension for the meaning of the word *chromium*. As in Experiment 1, we manipulated the alignability of the objects presented, with half of the children seeing high-alignment pairs and half of the children seeing low-alignment pairs. If children were engaged in hypothesis-testing or other types of intentional reasoning, then those in both the High Alignment (HA) and Low Alignment (LA) conditions would be able to correctly infer the meaning of *chromium*. However, if the HA advantage in Experiment 1 derived from spontaneous perceptual alignment, we should still see an advantage of HA over LA in Experiment 2's Meaning Assessment test.

3.1. Methods

3.1.1. Participants

Since the 3- and 4-year-olds showed the same pattern of performance in Experiment 1, we examined only the younger age group in Experiment 2. We ran a total of 55 3-year-olds ($M = 42.64$ months, $SD = 2.80$ months, females = 27). Children were randomly assigned to either the HA or the LA condition, with approximately the same number of males and females in each condition. Six additional children participated in the experiment but were excluded due to either caretaker interference (one child), failing to complete the experiment (one child), failing to reach the warm-up task's criterion (one child), or answering all-yes in the Meaning Assessment test (three children). Children were recruited through the same methods as in Experiment 1, and the demographic composition was approximately the same.

3.1.2. Materials and procedure

The materials and procedure for Experiment 2 were similar to those of Experiment 1 but with one crucial change: children received two Initial Exposure trials rather than one (See Fig. 1, middle row). The key trials were embedded in the same simple pointing game as before. Children received a warm-up trial (e.g. "point to the lion—the lion, not the cow."), followed by the first *chromium* exposure trial. Then they received an easy trial (e.g. "point to the long one—the long one, not the short one."), followed by the second exposure trial and then a final easy trial. On both exposure trials, children were requested to "point to the chromium one—the chromium one, not the blue one". The order of the two *chromium* exposure trials and the order of the two easy trials were counterbalanced.

On the first exposure trial, the LA group saw two objects that differed in both shape and color, as in Experiment 1. For example, half of the LA group saw a chromium circle and a blue square. At this point, those children could have two possible hypotheses for the meaning of *chromium*: olive green or circle. On the second exposure trial, they saw a chromium square and a blue circle. Crucially, the target (non-blue) object on Trial 2 was of a different shape from the target object on Trial 1 but retained the same color (see Fig. 1). Thus, by the end of the second trial, the LA group could rule out the shape hypothesis and arrive at the correct meaning of *chromium*—the olive-green color.

Children in the HA condition saw the same four objects as those in the LA condition. However, the objects were presented in two highly alignable pairs: e.g., a chromium circle and a blue circle on the first trial, and a chromium square and a blue square on the second. Thus, children in the HA condition did not receive additional information from the second trial.

3.2. Results

3.2.1. Initial exposure

As in Experiment 1, all children performed well in the Initial Exposure. Among the 55 children, 52 (94.5%) of them correctly pointed to the chromium object on both trials. The three children who did not do so were all in the LA condition and were excluded from further analyses.

3.2.2. Meaning assessment

We compared the number of correctly identified chromium objects (hits, max. = 3) and incorrectly identified chromium objects (false alarms, max. = 5) that children made in the Meaning Assessment test.

3.2.2.1. Correct identifications/hits. Planned comparisons revealed that the HA group ($M = 2.46$, $SD = 0.76$) made significantly more hits than the LA group ($M = 1.88$, $SD = 0.99$), $t(50) = 2.35$, $p = .01$, $d = 0.66$, one-tailed *t*-test. Children in the HA condition made significantly more hits than chance (chance = 1.5 hits), $t(25) = 6.45$, $p < .001$, $d = 1.26$, two-tailed *t*-test. There was a non-significant trend for children in the LA condition to make more hits than chance, $t(25) = 1.98$, $p = .06$, $d = 0.39$, two-tailed *t*-test. Fig. 4 shows the mean number of hits in the two groups; Table 1 shows the proportion of each group who generalized the word *chromium* to the new test objects.

3.2.2.2. False alarms. For both groups, the mean number of false alarms did not exceed chance (2.5 false alarms): $M_{HA} = 2.19$, $SD_{HA} = 1.39$, $p = .27$; $M_{LA} = 1.58$, $SD_{LA} = 1.27$), $p = .001$, two-tailed *t*-tests.

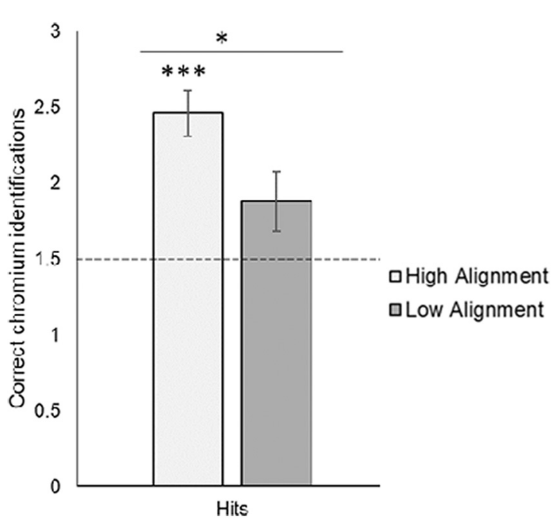

Fig. 4. Experiment 2: Mean number of chromium objects (max = 3) correctly endorsed as *chromium* (hits). Error bars depict +/− one standard error. The dashed line shows chance level. * $p < .05$, ** $p < .01$, *** $p < .005$.

3.3. Experiment 2 discussion

Experiment 2 replicated the high alignment advantage found in Experiment 1 and extended the findings to rule out an information-based explanation. As in Experiment 1, children in both groups correctly picked out the chromium object in the Initial Exposure; however, the HA group was better able to identify chromium instances in the Meaning Assessment test, and was significantly more likely to apply the new word *chromium* to the new instance (Table 1). Thus, the difference in performance in the two groups is not driven by unequal information. Instead, we suggest that the key distinguishing factor is a greater likelihood of spontaneous structural alignment by the HA group than the LA group.

How does alignability promote learning in such a context? Following Gentner and Hoyos (2017), we suggest that high overall similarity of the objects benefitted the HA group in two ways. First, it prompted spontaneous comparison of the two objects. Second, the objects' high similarity triggered a fast and essentially effortless alignment, making the alignable difference of color "pop out" for easy detection (Gentner & Hoyos, 2017; Sagi et al., 2012). Once the color of the target object was highlighted in this way, children often linked it with the word *chromium.* We suggest that this was unintentional learning—that this linkage occurred even without any deliberate attempt to learn the word's meaning—a point to which we return later.

The LA group (like the HA group) was highly accurate at pointing to the chromium object when asked to "point to the chromium one, not the blue one." However, the lack of similarity between the items made them unlikely to trigger spontaneous comparison. Further, even if children had compared the objects, the lack of overall similarity would have made alignment less fluent, and therefore less likely to result in pop-out of the alignable difference. Of course, children could have engaged in deliberate comparison, as an adult might do, in order to derive the meaning of *chromium.* In this case, the LA group would have had two hypotheses as to the meaning of *chromium* after Trial 1 and could have selected the correct meaning after Trial 2. But this apparently did not happen; performance in the LA group did not exceed chance. In sum, the results of Experiment 2 are consistent with our claim that the HA advantage stems from the greater likelihood of spontaneous comparison, and that this process did not require a deliberate intention to learn the meaning of *chromium.*

4. Experiment 3

We have suggested that success in our task occurred via implicit learning, and did not require any conscious intention to learn the meaning of the word *chromium.* Both groups of children complied with the experimenter's request to point to the object that was "not the blue one," and were highly accurate in doing so. But only children who experienced highly alignable comparisons learned the term. We suggest that the HA group learned the meaning of *chromium* because high perceptual similarity triggered a spontaneous comparison process, rendering the chromium (olive green) color more salient as an alignable difference. This new color was implicitly linked with the co-occurring new word, *chromium.* Children in the LA condition, which lacked the spontaneous invitation to compare, mostly failed to learn the term, even with sufficient information to do so.

If this account is correct, there are some important implications. It would mean that the view of language-learning as hypothesis-driven is limited in scope. That is, some potential meanings (or hypotheses) may arise out of children's spontaneous comparison processes. But before drawing conclusions, we must ask–was this learning truly implicit? Our procedures were aimed at promoting unintentional rather than deliberate word learning. Rather than asking children to learn a new word, we simply used the new word in a communicative exchange, during one round of a pointing game. In addition, we inserted an unrelated filler task between the initial exposure and the later Meaning Assessment. Finally, we included other colors in the Meaning Assessment task, so that the novel word *chromium* would not be perceived as the focus of the task. But to test this claim of unintentional learning further, in Experiment 3, we contrast the current incidental task with a direct-learning version of the task. This enables us to test between two possibilities. One possibility is that children in our prior studies were trying to learn the word, and this happened to be easier with high perceptual alignability. In this case, asking children to learn the word should not greatly alter their performance. The second possibility, and the one we propose, is that children in our prior studies were simply playing a pointing game, and the high-alignment advantage arose through spontaneous implicit comparison. In this case, instructions to learn the word should improve children's performance, especially the performance of the LA group (which lacks the hypothesized advantage of spontaneous comparison).

In Experiment 3, at the outset of the study, children were directly instructed to learn the word *chromium*. The rest of the procedure was identical to that of Experiment 1. We then contrast the results with those of the matched age group (i.e., 4-year-olds) in Experiment 1. Our predictions are that (1) overall, children will perform better in the direct condition (Experiment 3) than in the indirect condition (Experiment 1)—suggesting that the intention to learn the word was not already present in the indirect case; and (2) this improvement will be greatest in the LA group. This would be evidence that pedagogical support can compensate for the lack of environmental support.

4.1. Methods

4.1.1. Participants

Forty-four 4-year-olds were included in Experiment 3 (M = 53.17 months, SD = 2.65 months, females = 20). An additional nine children participated but were excluded: one child for failing to complete the experiment, two children due to experimenter error, two children for failing to reach the predetermined criterion in the warm-up task, and four children for answering either all-yes or all-no in the Meaning Assessment test. All children were recruited through the same measures as in Experiment 1, and the demographic composition was approximately the same. Children were randomly assigned to either the HA or LA condition, with approximately equal numbers of males and females in each condition.

4.1.2. Materials and procedure

The materials and procedure for Experiment 3 were similar to those of Experiment 1, except that children were explicitly told that their goal was to learn the meaning of a new word. Before the Initial Exposure, the experimenter told the child, "We are going to play a fun game together! In this game, you have a special mission and that is to learn a new word –*chromium*. Can you say *chromium*?" After the child repeated the word, the experimenter said, "Great! Remember, you need to figure out what *chromium* means by the end of the game, okay?" After the introduction, the experiment was conducted as in Experiment 1. At the end of the study, the experimenter asked, "In this game you were supposed to learn a new word. Do you remember what word it was?" If the child produced the word, the experimenter replied, "Right, the word was *chromium*! What do you think *chromium* means?" If the child indicated that they forgot or simply remained silent, the experimenter asked, "Was the word… *avocado*? Was it…*stegosaurus*? Was it…*chromium*?" If the child picked the correct word, the experimenter said, "Right, the word was *chromium*. What do you think *chromium* means?" If the child picked the wrong word, the experimenter replied, "Oh, I think that the word was *chromium*. What do you think *chromium* means?"

4.2. Results

4.2.1. Initial exposure

Regardless of condition, all children (44/44) pointed to the correct chromium object during the Initial Exposure.

4.2.2. Meaning assessment

We next compared the number of correctly identified chromium objects (hits, max. = 3) and incorrectly identified chromium objects (false alarms; max. = 5) that children made in the Meaning Assessment test.

4.2.2.1. Correct identifications/hits. In contrast to the prior studies, the two groups did not differ in their number of hits ($M_{HA} = 2.59$, $SD_{HA} = 0.73$; $M_{LA} = 2.32$, $SD_{LA} = 0.89$; $t(42) = 1.11$, $p = .14$, $d = 0.33$, one-tailed t-test). Both groups made hits at significantly above-chance rates (chance = 1.5 hits; HA: $t(21) = 6.97$, $p < .001$, $d = 1.49$, two-tailed t-test; LA: $t(21) = 4.29$, $p < .001$, $d = 0.92$, two-tailed t-test). Fig. 5 shows the mean number of hits made by children in the HA and LA conditions in Experiments 1 and 3. The two groups also did not differ in ability to generalize *chromium* to new objects (Table 1).

4.2.2.2. False alarms. Both groups made significantly fewer false alarms than expected by chance (chance = 2.5 false alarms; $M_{HA} = 1.82$, $SD_{HA} = 1.14$; $M_{LA} = 1.59$, $SD_{LA} = 1.26$, both $ps < 0.02$, two-tailed t-tests), suggesting that their high hit rates were not due to an overall bias to respond positively.

4.2.2.3. Cross-study comparison. Our prediction was that children, especially those in the LA condition, would perform better in the direct condition (Experiment 3) than in the indirect condition (Experiment 1). To test this prediction, we conducted planned comparisons of the results of Experiment 3 with those of the 4-year-olds in Experiment 1. Fig. 5 shows the contrast in hit rate between the two studies (which were identical except for the early instruction to learn the new word *chromium* in Experiment 3). To test our prediction that direct instruction should preferentially benefit the LA group, we conducted planned comparisons comparing the HA and LA groups in Experiment 3 (direct task) with those in Experiment 1 (indirect task). As predicted, children in the LA condition did better in the direct task ($M = 2.32$, $SD = 0.89$) than in the indirect task ($M = 1.77$, $SD = 1.15$), $t(51) = 1.86$, $p = .03$, $d = 0.52$, one-tailed t-test.[12] Also as expected, direct instructions did not improve the already-high performance in the HA group; there was no difference in their mean number of hits between the direct ($M = 2.59$, $SD = 0.73$) and indirect tasks ($M = 2.50$, $SD = 0.76$), $t(52) = 0.44$, $p = .33$, $d = 0.12$, one-tailed t-test.

4.2.2.4. Word memory check. We examined how well children were able to remember the word learning instructions that they were given at the beginning of the experiment. Of the 44 children included in the final analysis, 31 of them (70%, p = .01, binomial test) spontaneously produced some form of the word *chromium* when asked "what was the word that you were supposed to learn" (e.g. "chromium", "chromy"). Of the remaining 13 children, all but one picked out the correct word from three possible options. When asked to say what *chromium* means, 17 children provided a color meaning (e.g. "black", "dark"), 4 children provided a shape meaning, (e.g. "circle", "shapes"), 2 children provided a texture meaning (e.g. "hard", "a little soft"), 12 children indicated that they did not know, and 9 children responded in other ways (e.g. "grill", "animals", "we're done with the game"). The number of children who gave a color meaning did not differ between the HA condition (9 out of 22 [40.1%]) and LA condition (8 out of 22 [36.4%]), ($X^2(1, N = 44) = 0.10$, $p = .76$, $\varphi = 0.47$).

[12] One-tailed t-tests were used here because there is a clear directional prediction: performance in the direct task should be as high or higher than that in the indirect task.

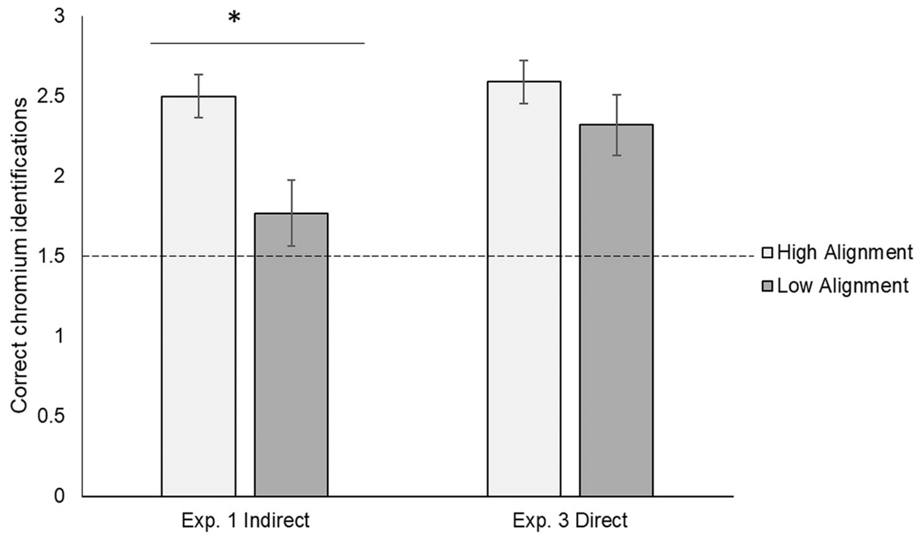

Fig. 5. Mean number of hits (chromium objects correctly endorsed as *chromium*) (max = 3) by children in Experiment 1 (Indirect) and Experiment 3 (Direct). Error bars depict +/− one standard error. The dashed line represents chance performance. * $p < .05$, ** $p < .01$, *** $p < .005$.

4.3. Experiment 3 discussion

Giving children a pedagogical goal to learn the word's meaning improved performance in the LA group. Under incidental learning conditions (Experiments 1 and 2), the LA group performed at chance on the Meaning Assessment test; but when given direct instructions to learn the word *chromium*, they performed well above chance. Further, whereas the LA group performed worse than the HA group in Experiments 1 and 2, there was no difference between the groups in Experiment 3. In contrast, the HA group showed high performance in both direct and indirect conditions—evidence that perceptual alignability can lead to implicit word learning.

Unlike the LA group, the HA group was not affected by the addition of explicit instructions—children performed equally well on the Meaning Assessment test regardless of whether they were given a goal to learn the new word. This is consistent with the idea that children in the HA condition were learning via spontaneous alignment processes, rather than by intentionally seeking to learn the word *chromium*. This pattern[13] suggests that children in our prior studies were not approaching the experiment as a word-learning task.

The contrast between the LA group and the HA group (whose performance was unaffected by instruction) highlights the importance of perceptual alignment in promoting spontaneous, non-intentional word learning. At the same time, the findings show that perceptual alignment is not *necessary* for learning property words. The difference between the LA and HA groups found in the prior studies disappeared when children were given direct task instructions—evidence that pedagogical support can compensate for a lack of ideal perceptual conditions.

The three studies so far show that perceptual alignment can support implicit learning of property terms. But how robust is this learning? Is it a fleeting impression, or will children retain these meanings over time? And is the learning narrowly construed, or can children apply the *chromium* concept to very different materials? In our experiments so far, we have assessed children's understanding of *chromium* ten minutes after the Initial Exposure, using objects similar to those seen in the Initial Exposure. In the next experiment, we ask whether children will retain the meaning of *chromium* over a delay, and whether they can apply it to a new set of objects.

5. Experiment 4

In this study, we examined the strength of children's learning by (1) assessing children's understanding of the novel word's meaning after a longer delay; and (2) increasing the diversity of the transfer items in the Meaning Assessment test. As in Experiment 2, both the high alignment and low alignment groups saw two exposure trials during the Initial Exposure, designed to provide sufficient information to allow children to reject a shape hypothesis and conclude that *chromium* refers to a color. As before, we assessed participants' knowledge of the word *chromium* ten minutes after the first exposure (Day 1). In addition, to investigate robustness under delay, we also assessed their understanding two-to-four days later (Day 2). Day 2 featured two tasks used to assess children's understanding of *chromium*: a Retention Task to assess the durability of the word-meaning mapping and a Transfer Task to assess whether children could extend the word to new kinds of materials. One further change was that, in the interest of generalizability, we developed new stimuli for both the Initial Exposure and Meaning Assessment.

5.1. Participants

There were 41 4-year-olds ($M = 53.54$ months, $SD = 3.61$ months, females = 26) in Experiment 4. Participants were randomly assigned to either the HA condition ($n = 20$, female = 12) or the LA condition, with approximately equal numbers of males and females in each group. An additional four 4-year-olds were tested but excluded from the analyses, one for failing to complete the task and three for adult interference. All children were recruited through the same methods as in Experiment 1, and the demographic composition was approximately the same. A group of 3-year-olds also participated in Experiment 4; however, they performed at chance in both the HA and LA conditions on Day 1, making it difficult to interpret their subsequent performance. We present their Day 1 results separately in Appendix C (online supplemental material).

5.1.1. Day 1 materials and procedure

The procedure for Day 1 was the same as that of Experiment 2. Children were exposed to the novel word and color during a pointing task and assessed on their learning of the word-color mapping approximately ten minutes later through a yes-no sorting task. However, we used new stimuli for both the Initial Exposure and Meaning Assessment.

[13] We cannot rule out the possibility that the lack of difference between direct and indirect tasks for the HA group could be due to a ceiling effect in our studies. With more complex stimuli, we might see an advantage of direct instruction even for the HA group. However, the more important finding is that we *do* see a difference for the LA group.

During the Initial Exposure (Fig. 6a), children in the HA condition saw two crosses (one chromium and one blue) and two upside-down triangles (also one chromium and one blue). Children in the LA condition saw a chromium cross paired with a blue circle, and a chromium upside-down triangle paired with a blue square. During the Meaning Assessment test (Fig. 6b), children saw a set of 14 geometric shapes (as opposed to only 8 shapes in Experiments 1–3) created under the same guidelines as in previous studies. The set included five chromium objects, four blue objects and five objects of various colors. As in previous experiments, the set included all the objects presented across the Initial Exposure trials. In addition, the set also included objects with seen shapes and new colors (e.g. the orange cross), objects with new shapes and seen colors (e.g. the chromium hexagon), completely new objects (e.g. the yellow L shape), and a maroon triangle that had been seen by all children in one of the Initial Exposure's filler trials.

5.1.2. Day 2 materials and procedure

Two-to-four days (M = 2.43, SD = 0.75) after Day 1, children returned to the lab for Day 2 of the experiment. All children went through the same procedure: a Retention Task, a Reminder, and a Transfer Task. The HA and LA groups received identical instructions and materials for all Day 2 tasks.

The Retention task was the same as the Meaning Assessment test on Day 1. After the Retention task, both groups of children received a refresher prompt, analogous to the Low Alignment Initial Exposure on Day 1. They were shown a card depicting a chromium fish, together with another card depicting a different-looking blue fish. (Fig. 7a) The experimenter asked, "Look at these! Can you point to the chromium one? The chromium one, not the blue one!". Afterwards, children completed the Transfer Task. This task used the same format as the Meaning Assessment test (and Retention task), but with new materials: 14 picture cards of differently colored and shaped fish. As shown in Fig. 7b, five of the fish were chromium, four were blue, and the rest were of various colors. All fish in the Transfer task were new; neither of the two fish in the Reminder were shown again in the Transfer task.

5.2. Results

5.2.1. Overall performance

Fig. 8 shows the number of hits across the three tasks. To compare overall performance, we conducted a repeated-measures ANOVA with the 39 children who were included in all three tasks. We used Hits as the dependent variable, Task (Day 1 Meaning Assessment, Day 2 Retention, and Day 2 Transfer) as a within-subject independent variable and Alignment (HA and LA) as a between-subject independent variable. We found that the HA group performed significantly better than the LA group overall across the three tasks, $F(1,37) = 5.02$, $p = .03$. There was no significant main effect of Task ($p = .23$), nor an interaction between Task and Condition ($p = .78$). Thus, the HA group's advantage in learning the meaning of *chromium* persists over time and over transfer to new materials. We next present the detailed results for each task.

5.2.2. Day 1 results

5.2.2.1. Initial exposure. As in prior studies, children performed very well on the Initial Exposure task. 40 of the 41 (97.6%) children successfully pointed to the chromium object on both Initial Exposure trials. Only one child (HA) was incorrect on the first key exposure trial. This child was excluded from further analyses.

5.2.2.2. Meaning assessment. As in Experiment 1, children in the HA group correctly endorsed more chromium objects than those in the LA group ($M_{HA} = 4.63$, $SD_{HA} = 0.96$; $M_{LA} = 3.71$, $SD_{LA} = 1.65$), $t(38) = 2.12$, $p = .02$, $d = 0.64$, one-tailed t-test. However, unlike Experiment 1—in which only children in the HA condition performed above chance—here, both groups made more hits than expected by chance (chance = 2.5 hits) (both $ps < 0.004$, two-tailed t-tests). Fig. 7 shows the mean number of hits made by 4-year-olds in Experiment 4. The proportion of children who generalized to all three new chromium objects was 0.84 for the HA group and 0.57 for the LA group, a nonsignificant difference (See Table1).

5.2.2.2.1. False alarms. Children in the both the HA ($M = 1.32$, $SD = 1.57$) and LA condition ($M = 1.76$, $SD = 2.39$) made significantly fewer false alarms than expected by chance (chance = 4.5 false alarms; both $ps < 0.001$, two-tailed t-tests).

5.2.2.2.2. d prime. With the extended set of objects in the Meaning Assessment, we were able to examine children's sensitivity to the meaning of *chromium*. Combining hits and false alarms, we calculated the d' score for each individual child and compared the mean d' score of the HA group and LA group. There was a nonsignificant trend towards a higher mean d' score for children in the HA condition ($M = 2.20$, $SD = 0.86$) than in the LA condition ($M = 1.57$, $SD = 1.46$), $p = .06$, one-tailed t-test. Fig. 9 shows the mean d' scores for 4-year-olds in the HA and LA conditions.

5.2.3. Day 2 results

5.2.3.1. Retention task. Of the 40 4-year-olds included in Day 1's Meaning Assessment analyses, one child in the LA condition was excluded from the following analyses due to answering yes on all trials, leaving 39 children.

5.2.3.1.1. Correct identifications/hits. There was a non-significant trend for the mean number of hits (max. = 5) to be higher in the HA group ($M = 4.68$, $SD = 1.00$) than in the LA group ($M = 4.00$, $SD = 1.62$), $t(37) = 1.57$, $p = .06$, $d = 0.50$, one-tailed t-test. Children in both conditions made significantly more hits than expected by chance (chance = 2.5 hits, both $ps < 0.007$, two-tailed t-tests). Fig. 8 shows the mean number of hits in the Retention Task.

5.2.3.1.2. False alarms. The mean number of false alarms was significantly lower than chance in both the HA ($M = 1.42$, $SD = 1.84$) and the LA condition ($M = 1.55$, $SD = 2.44$) (chance = 4.5 false alarms, both $ps < 0.001$, two-tailed t-tests).

5.2.3.1.3. d prime. The mean d' scores did not differ significantly between the HA condition ($M = 2.19$, $SD = 0.84$) and the LA condition ($M = 1.79$, $SD = 1.43$), $p = .15$, one-tailed t-test (Fig. 9).

5.2.3.2. Reminder. When shown two fish and asked to "point to the chromium one, not the blue one," 100% of the children chose the correct chromium object on the first attempt, consistent with previous results on the Initial Exposure Task.

5.2.3.3. Transfer task. As before, we excluded children who answered yes to all trials in the Transfer task (n = 1, LA condition, same child excluded in the Retention task). However, we included children who answered no to all trials in the Transfer task (n = 4, LA = 3), since an all-no response could indicate failure to transfer.

5.2.3.3.1. Correct identifications/hits. Children in the HA condition made more hits than those in the LA condition ($M_{HA} = 4.42$, $SD_{HA} = 1.17$; $M_{LA} = 3.45$, $SD_{LA} = 2.11$), $t(37) = 1.76$, $p = .04$, $d = 0.59$, one-tailed t-test. Fig. 8 shows the mean number of hits made by 4-year-olds in Experiment 4. Children in the HA condition made significantly more hits than chance (chance = 2.5 hits), $t(18) = 7.16$, $p < .001$, $d = 1.64$, two-tailed t-test. There was a non-significant trend for children in the LA condition to make more hits than chance, $t(19) = 2.01$, $p = .06$, $d = 0.45$, two-tailed t-test.

5.2.3.3.2. False alarms. The mean number of false alarms made by children in both the HA and LA conditions ($M_{HA} = 1.11$, $SD_{HA} = 2.26$; $M_{LA} = 1.35$, $SD_{LA} = 2.48$) was significantly lower than chance (chance = 4.5 false alarms), both $ps < 0.001$, two-tailed t-tests.

5.2.3.3.3. d prime. Four-year-olds in both the HA and LA conditions

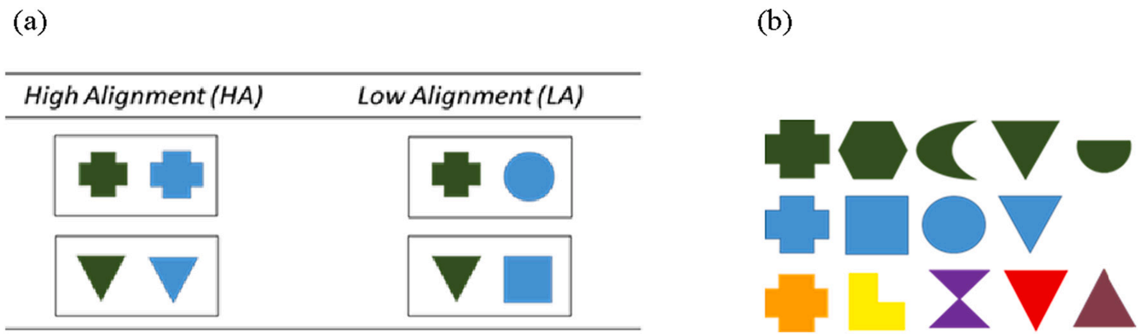

Fig. 6. Experiment 4: (a) Day 1 Initial Exposure objects; (b) Day 1 Meaning Assessment objects (also used in Day 2 Retention task).

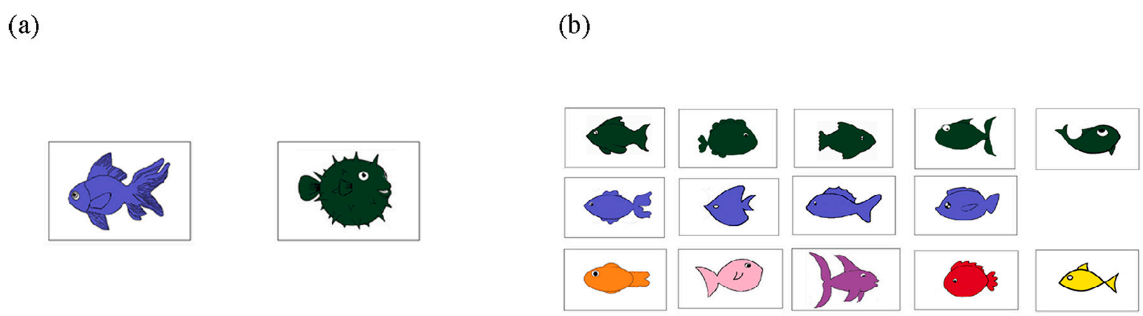

Fig. 7. Experiment 4, Day 2: Materials used in (a) Reminder task; (b) Transfer task.

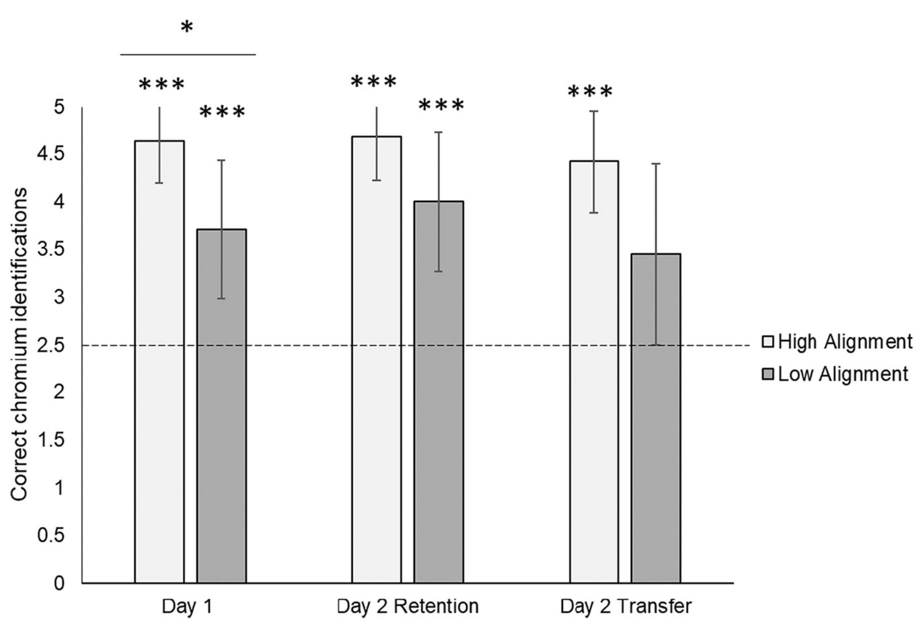

Fig. 8. Experiment 4: Mean number of hits (chromium objects correctly endorsed as *chromium*) (max = 5) by 4-year-olds in the Day 1 Meaning Assessment test, the Day 2 Retention task, and the Day 3 Transfer task. Error bars depict +/− one standard error. The dashed line represents chance performance. * $p < .05$, ** $p < .01$, *** $p < .005$.

showed high sensitivity to the novel word and color in the Transfer task, as evidenced by their high *d'* scores. The two groups did not differ significantly: HA group ($M = 2.22$, $SD = 1.04$) and LA group ($M = 1.61$, $SD = 1.42$), $t(37) = 1.53$, $p = .07$, $d = 0.49$, two-tailed *t*-test. (Fig. 9).

5.3. Experiment 4 discussion

In Experiment 4, we examined the robustness of incidental word learning through spontaneous alignment. As in our prior studies, the HA group outperformed the LA group overall. The HA group retained the meaning of *chromium* over a delay of two-to-four-days and were able to transfer this meaning to completely new objects. Interestingly, the LA group performed comparatively better here than in Experiment 1. These findings show that the insight gained from spontaneous comparison go beyond a momentary advantage—they can confer lasting benefits and inform understanding of new situations.

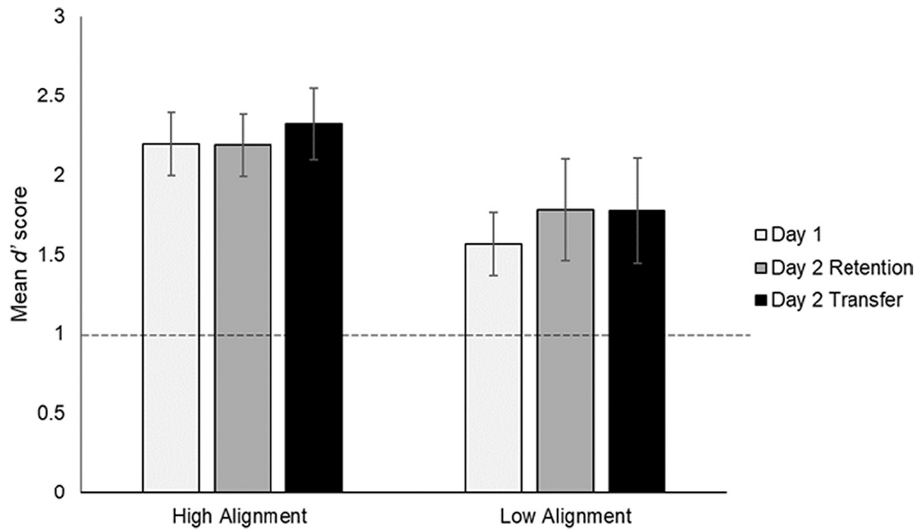

Fig. 9. Experiment 4: Mean *d'* scores across the three tasks. Error bars depict +/− one standard error. The dashed line represents a *d*' score of 1.00.

6. General discussion

Our goal in these studies was to test the claim that spontaneous comparison processes contribute to referential transparency and to children's word learning. Specifically, we proposed that highly similar contrast pairs in the environment invite spontaneous processes of structural alignment, which support rapid noticing of alignable differences. Taking Carey and Bartlett's (1978) fast-mapping study as the paradigm example of this kind of learning, we repeated this study, varying the similarity of the contrast pairs. We predicted that high-similar pairs would lead to spontaneous structural alignment, and therefore to pop-out of the alignable difference in color between the two objects. Thus, children who saw highly similar pairs should be far more likely to notice and learn the color *chromium* than those who saw less similar pairs. A further claim was that this insight arose from a spontaneous comparison process and did not require a prior goal to learn the word's meaning.

Across four experiments, we found evidence for these claims. In our studies, 3- and 4-year-old children were given either highly alignable objects—e.g., two squares, one chromium and one blue—or two objects that were less similar and harder to align—e.g., a chromium circle and a blue square. Children were asked to "point to the chromium one, not the blue one." Across studies, over 96% of the children pointed correctly—yet the high-alignment group was far more likely to extract the meaning of *chromium*, as assessed in a subsequent meaning assessment task. Indeed, in Experiments 1 and 2, the low-alignment group was at chance on identifying chromium objects in this task. In Experiment 4, we found that the new word meaning was retained over a two-to-four-day delay and could be transferred to new kinds of objects.

These findings are consistent with our hypothesis that the high-alignment group learned the meaning of *chromium* via a spontaneous structural alignment process. But there was an alternate interpretation—namely, that children in both conditions were trying to understand the word, and that only the high-alignment group had adequate information to do so. From a hypothesis-testing perspective, the low-alignment group saw two figures that differed in both shape and color, they could have entertained either a shape hypothesis ("*chromium* means square") or a color hypothesis ("*chromium* means dark green"), or both. In contrast, the high-alignment group saw two figures that differed only in color, and would therefore have had only one possible hypothesis—"*chromium* means dark green." To test this possibility, in Experiment 2, we equated the information given to the two groups. 3-year-olds were given two initial exposures, designed so that both groups would have adequate information to rule out the shape hypothesis and arrive at the correct color hypothesis (see Fig. 1). However, as in Experiment 1, only the high-alignment group succeeded—consistent with the claim that the meaning of *chromium* arises out of spontaneous comparison in the high-alignment group, rather than via deliberate hypothesis-testing.

As a further test of our thesis that spontaneous comparisons can yield insights without conscious intention, in Experiment 3 we repeated the same incidental learning procedure as in the other studies, but preceded it by telling children that they should try to learn the word *chromium*. The prediction is that the goal to learn the word's meaning will improve performance in the low-alignment group, but not the high-alignment group. The results bore out this prediction: For the low-alignment group, performance was significantly better in the direct condition than in the prior incidental conditions; for the high-alignment group, there was no difference. These results buttress our claim that under favorable conditions, implicit comparison processes can reveal new word meanings to children.

But there is one more potential concern to consider.[14] Perhaps the 3- and 4-year-olds in our study were confused in the low alignment condition because the objects used were nameable using familiar shape terms ("circles", "squares", "triangles", and "crosses"). When presented with an olive-green circle and a blue square, children might expect an adult to refer to the objects using the known shape terms ("circle" and "square"). The same is not true in the high alignment condition, since the known term "square" would be insufficient to pick out one of the two squares. Thus, perhaps children in the Low Alignment condition would expect the speaker to distinguish the two objects using known terms (e.g., "the square, not the circle") and are wondering why the person is talking about "the chromium one, not the blue one". Although this is certainly a possibility, we note that speakers can refer to objects in different ways depending on their pragmatic focus; a person who is focused on color may refer to objects by color even if the objects could have been referred to in other ways. For example, suppose a child asks their parent "Which one do you want?" (offering a blue truck and a red sportscar). The parent could choose to say "the truck" but they might also just say "the blue one." There is evidence that 3- and 4-year-olds can take account of speakers' focus in interpreting different ways of referring to objects (e.g., E. Clark, 1990, 2014).

6.1. Characterizing high-quality input for word learning

Past research has established a correlation between the amount of

[14] We thank an anonymous reviewer for bringing this concern to our attention.

language a child hears and the child's vocabulary size (e.g. Goodman, Dale, & Li, 2008; Hart & Risley, 2003; Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Weisleder & Fernald, 2013). However, recent research has shown that the *quality* of the input, and not just the sheer quantity, is also a strong predictor of children's vocabulary growth (e.g. Cartmill et al., 2013; Hirsh-Pasek et al., 2015; Rowe, 2012). Of particular relevance here is Cartmill et al.'s (2013) finding that the referential transparency of parents' language is an important contributor to children's language learning. *Referential transparency* refers to how clearly a word's meaning can be inferred from the immediate extralinguistic context. Cartmill et al. assessed referential transparency by using Gleitman's Human Simulation Paradigm (Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1990) to code parent-child dialogues. Referential transparency was gauged by how easily a naïve adult could guess what noun a parent was using by watching (but not hearing) the parent-child interaction. There was wide variation across parents in this measure; and, importantly, the judged referential transparency of the input speech at 14 months correlated (over and above the quantity of speech) with the size of the children's vocabularies at 54 months.

The present research reveals a new contributor to referential transparency—namely, available highly alignable comparisons—and suggests a role for perceptual comparison and structural alignment in deriving the meaning of a new word. Seeing just one highly alignable pair was sufficient for 3- and 4- year-old children to learn the word *chromium* from an incidental exposure. Children given the same linguistic input with a less alignable pair were significantly less likely to learn the word's meaning. Thus, our findings suggest that alignable external comparisons can contribute to referential transparency. Further, having a relevant distinction emerge as an alignable difference should increase its salience, thus contributing to memory for the word's meaning, as found in Experiment 4 (Vlach and Sandhofer, 2012).

Of course, it is likely that many kinds of extralinguistic features support referential transparency. For noun learning, Cartmill et al. (2013) speculated that perceptual factors such as having a visible referent are important. Other factors that may support noun learning are the ease of individuating the object from its context (Gentner & Boroditsky, 2001), the perception of nonarbitrary structure (Prasada, Ferenz, & Haskell, 2002) and (relatedly) the regularity and complexity of the object's shape (Imai & Gentner, 1997).

Our findings extend the notion of high-quality perceptual input to adjective learning. As discussed earlier, for property terms, identifying the referent object is just the beginning; the child must also extract the relevant property. We suggest that situations that permit comparison and contrast of perceptually aligned objects are especially effective for allowing children to extract specific properties. This pattern has been found for direct learning of adjectives (Au & Laframboise, 1990; Gelman & Markman, 1985; Sandhofer & Smith, 2001; Tribushinina et al., 2013; Waxman & Klibanoff, 2000), as reviewed earlier. Gentner, Loewenstein, and Hung (2007) found a parallel pattern in learning of part-names. In their study, 3-year-old children were shown a drawing of an unfamiliar "Martian" animal[15] and were told "This animal has a blick. Can you show me which one of these also has a blick?" There were two alternatives, identical except that only one of them shared a distinctive body part with the standard. When the alternatives were highly similar to the standard, 3-year-olds succeeded in choosing the alternative with the same body-part. When the alternatives were dissimilar to the standard, they performed at chance. In a further study, 3-year-olds who were given a progressive alignment experience (high-similarity triads followed by low-similarity triads) succeeded on the low-similarity triads. Thus, structural alignment processes can support children's learning of names for object parts as well as object properties.

[15] To increase the challenge, the creatures were not named; thus there was no basic-level category to refer to.

More generally, we suggest that for nonobvious word meanings—such as those of verbs, adjectives, and relational nouns—comparison is likely to be important in both direct and indirect learning (Childers, Heard, Ring, Pai, & Sallquist, 2012; Childers & Paik, 2009; Fisher, 1996; Gentner et al., 2011; Graham, Namy, Gentner and Meagher, 2010; Haryu, Imai, & Okada, 2011; Sandhofer & Smith, 2001). For example, Haryu et al. (2011) found evidence for progressive alignment in verb-learning. They taught 4-year-old children a verb for a novel event and asked whether the children could extend the verb to other examples. Children were initially limited to close overall matches (i.e., literally similar events). That is, they extended the verb only when the new event shared similar objects as well as the same action as the initial event; they failed to recognize the same action when the objects were dissimilar. In a second study, Haryu et al. found that progressive alignment from close to far matches enabled a new group of 4-year-olds to extend the verb based on sameness of action, without support from object similarity.

6.2. The role of linguistic cues

A notable feature of our studies is that children heard the term *chromium* only once during the learning experience, rather than several times as in the verb-learning studies described earlier. However, children in our studies (in both the HA and LA conditions) had the advantage of hearing a linguistic cue that could invite comparison. Children were asked, "can you point to the chromium one—the chromium one, not the blue one". This framing ("the chromium one, not the blue one") by itself would probably have allowed an adult to infer that the key dimension of difference was color, regardless of the similarity between the two objects. However, studies by Heibeck and Markman (1987) suggest that preschoolers may not understand this convention. When given a highly alignable pair of objects, 2–4-year-olds were equally able to learn a novel property term whether they received an explicit contrast along the color dimension (e.g., "Bring me the chartreuse one, not the red one.") or a less informative frame (e.g., "Bring me the chartreuse one, not the other one."). This suggests that children were primarily responding to the perceptual cues rather than to the linguistic contrast. This conclusion is consistent with our finding that children wo saw highly alignable pairs were more likely to learn the meaning of chromium than those who saw low-alignable pairs*,* despite their having received the same linguistic contrast. Future research should explore how the relative importance of linguistic framing changes over development.

7. Conclusion

Incidental learning experiences are an important part of children's language learning, but they are almost invisible in nature. Such experiences can occur at any time and place, with no necessary learning intention on the part of the child nor pedagogical intention on the part of the adult. Yet during serendipitous moments, high-quality input can allow a child to acquire new word meanings even without a prior intention to do so. In this paper we show that an important contributor to incidental learning is children's own spontaneous comparison processes, invited by highly alignable available comparisons.

Credit authors statement

This paper is based on the doctoral thesis of Ruxue Shao and the work was done in collaboration with thesis advisor Dr. Dedre Gentner.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the children, families, and preschools that participated in the study and Emma Bulzoni and Katherine Cha for their help in conducting the experiments. We also thank Ellen Markman, Sue Hespos, Nina Simms, and members of the Cognition and Language laboratory for insightful suggestions. This research was supported by NSF SLC Grant SBE-0541957 awarded to the Spatial Intelligence and Learning Center (SILC) and ONR Grant N00014-16-1-2613 awarded to Dedre Gentner.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2022.105061.

References

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science, 8*(2), 199–209.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, & K. Hirsch-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (p. 115).

Arunachalam, S., Escovar, E., Hansen, M. A., & Waxman, S. R. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language & Cognitive Processes, 28*, 417–425.

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition, 114*(3), 442–446.

Au, T. K. F., & Laframboise, D. E. (1990). Acquiring color names via linguistic contrast: The influence of contrasting terms. *Child Development, 61*(6), 1808–1823.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development, 67*(6), 3135–3153.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., … Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*(1), 85–123.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8* (6), 492–499.

Benjamin, N., Haden, C. A., & Wilkerson, E. (2010). Enhancing building, conversation, and learning through caregiver–child interactions in a children's museum. *Developmental Psychology,, 46*(2), 502–515.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition, 126*(1), 39–53.

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology, 101*(2), 294.

Bloom, P., & Kelemen, D. (1995). Syntactic cues in the acquisition of collective nouns. *Cognition, 56*(1), 1–30.

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infants' expectations for count nouns and adjectives. *Journal of Cognition and Development, 4* (3), 357–381.

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development, 80*(1), 15–22.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., … Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75*(4), 1115–1139.

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2015). Developmental changes in the relationship between grammar and the lexicon. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the cognitive science society* (pp. 1–6). Austin, TX: Cognitive Science Society.

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language, 35*(1), 207–220.

Callanan, M. A., Akhtar, N., & Sussman, L. (2014). Learning words from labeling and directive speech. *First Language, 34*(5), 450–461.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference, 15*, 17–29.

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences, 110*(28), 11278–11283.

Childers, J. B., Heard, M. E., Ring, K., Pai, A., & Sallquist, J. (2012). Children use different cues to guide noun and verb extensions. *Language Learning and Development, 8*(3), 233–254.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language, 36* (1), 201–224.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. J. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hover, England: Psychology Press.

Clark, E. V. (1990). Speaker perspective in language acquisition. *Linguistics, 28*, 1201–1220.

Clark, E. V. (2016). *First language acquisition* (3rd). Cambridge: Cambridge University Press.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*, 1–63.

Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports, 6*, 25434.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*(2), 209.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental psycholinguistics: On-line methods in children's language processing, 44*, 97.

Fisher, C. (1996). Structural limits on verb meaning: The role of analogy in children's interpretation of sentences. *Cognitive Psychology, 31*, 41–81.

Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(2), 143–149.

Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). Extending SME to handle large-scale cognitive modeling. *Cognitive Science, 41*, 1152–1201.

Forbus, K. D., & Gentner, D. (2017). Evidence from machines that learn and think like people. *Behavioral and Brain Sciences, 40*, 35–36.

Foster, E. K., & Hund, A. M. (2012). The impact of scaffolding and overhearing on young children's use of the spatial terms between and middle. *Journal of Child Language, 39* (2), 338–364.

Gampe, A., Liebal, K., & Tomasello, M. (2012). Eighteen-month-olds learn novel words through overhearing. *First Language, 32*(3), 385–397.

Gasser, M., & Smith, L. B. (1998). Learning nouns and adjectives: A connectionist account. *Language & Cognitive Processes, 13*(2–3), 269–306.

Gelman, S. A., & Markman, E. M. (1985). Implicit contrast in adjectives vs. nouns: Implications for word-learning in preschoolers. *Journal of Child Language, 12*(1), 125–143. https://doi.org/10.1017/s0305000900006279

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek, & R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544–564). Oxford University Press.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775.

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.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman, & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge, UK: Cambridge University Press.

Gentner, D., & Boroditsky, L. (2009). Early acquisition of nouns and verbs: Evidence from Navajo. In V. Gathercole (Ed.), *Routes to language: Studies in honor of Melissa Bowerman* (pp. 5–36). New York, NY: Taylor & Francis.

Gentner, D., & Forbus, K. (2011). Computational models of analogy. *WIREs Cognitive Science, 2*, 266–276.

Gentner, D., & Gunn, V. (2001). Structural alignment facilitates the noticing of differences. *Memory and Cognition, 29*(4), 565–577.

Gentner, D., & Hoyos, C. (2017). Analogy and abstraction. *Topics in Cognitive Science, 9* (3), 672–693.

Gentner, D., Levine, S. C., Dhillon, S., Ping, R., Bradley, C., Isaia, A., & Honke, G. (2016). Rapid learning in a children's museum via analogical comparison. *Cognitive Science, 40*, 224–240.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*(3), 285–307.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *Psychological Science, 5*(3), 152–158.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity, *52*, 45-56. *American Psychologist, 52*, 45–56.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*, 297–301.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*(2), 135–176.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*(1), 23–64.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28*(1), 99.

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35*(3), 515–531.

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*, 280–290.

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How two-and four-year-old children interpret adjectives and count nouns. *Child Development, 64*(6), 1651–1664.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator, 27*(1), 4–9.

Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development, 82*(2), 674–686.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 1021–1034.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., … Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science, 26*(7), 1071–1083.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*(2), 128–157.

Hoyos, C., Horton, W. S., Simms, N. K., & Gentner, D. (2020). Analogical comparison promotes theory-of-mind development. *Cognitive Science, 44*(9), Article e12891. https://doi.org/10.1111/cogs.12891

Hoyos, C., Shao, R., & Gentner, D. (2016). The paradox of relational development: Could language learning be (temporarily) harmful? In D. Grodner, D. Mirman, A. Papafragou, J. Trueswell, J. Novick, S. Arunachalam, S. Christie, & C. Norris (Eds.), *Proceedings of the 38th annual conference of the cognitive science society* (pp. 2507–2512). Austin, TX: Cognitive Science Society.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*(2), 236.

Imai, M., & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62*, 169–200.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*(3–4), 159–207.

Justice, L. M. (2002). Word exposure conditions and preschoolers' novel word learning during shared storybook reading. *Reading Psychology, 23*(2), 87–106.

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development, 71*(3), 649–659.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.

Krcmar, M., Grela, B., & Lin, K. (2007). Can toddlers learn vocabulary from television? An experimental approach. *Media Psychology, 10*(1), 41–63.

Kurtz, K. J., & Gentner, D. (2013). Detecting anomalous features in complex stimuli: The role of structured comparison. *Journal of Experimental Psychology: Applied, 19*(3), 219.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*(4), 431–467.

Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition, 24*(2), 235–249.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Matlen, B. J., Gentner, D., & Franconeri, S. L. (2020). Spatial alignment facilitates visual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 46*(5), 443–457.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38*(4), 465–494.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254–278.

Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development, 11*(4), 356–368.

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition, 84*(3), 267–293.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition, 83*, 141–165.

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development, 77*(2), 266–280.

Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development, 80*(5), 1360–1375.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762–1774.

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36*(6), 1019–1050.

Salerni, N., Assanelli, A., D'Odorico, L., & Rossi, G. (2007). Qualitative aspects of productive vocabulary at the 200-and 500-word stages: A comparison between spontaneous speech and parental report data. *First Language, 27*(1), 75–87.

Sandhofer, C., & Smith, L. B. (2007). Learning adjectives in the real world: How learning nouns impedes learning adjectives. *Language Learning and Development, 3*(3), 233–267.

Sandhofer, C. M., & Smith, L. B. (2001). Why children learn color and size words so differently: Evidence from adults' learning of artificial terms. *Journal of Experimental Psychology: General, 130*, 600–620.

Sandhofer, C. M., & Smith, L. B. (2004). Counting the nouns: Simple structural cues to verb meaning. *Developmental Science, 7*(3), 378–388.

Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language, 27*(3), 561–585.

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87*(2), 218.

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology, 109*(1), 132–140.

Taylor, M., & Gelman, S. A. (1988). Adjectives and nouns: Children's strategies for learning new words. *Child Development*, 411–419.

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development, 81*(6), 1768–1786.

Tomasello, M., & Barton, M. E. (1994). Learning words in non-ostensive contexts. *Developmental Psychology, 30*(5), 639.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 1454–1463.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treisman, A. M., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *American Journal of Psychology, 105*, 2.

Tribushinina, E., & Gillis, S. (2012). The acquisition of scalar structures: Production of adjectives and degree markers by Dutch-speaking children and their caregivers. *Linguistics, 50*(2). https://doi.org/10.1515/ling-2012-0009

Tribushinina, E., Van Den Bergh, H., Kilani-Schoch, M., Aksu-Koç, A., Dabašinskienė, I., Hrzica, G., & Dressler, W. (2013). The role of explicit contrast in adjective acquisition: A cross-linguistic longitudinal study of adjective production in spontaneous child speech and parental input. *First Language, 33*(6), 594–616.

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: memory processes support children's retention of learned words. *Frontiers in Psychology: Developmental Psychology, 3*(46).

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.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*(11), 2143–2152.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science, 14*, 165–180.

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*(5), 619–626.

yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development, 83*, 1382–1399.