TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 9 (2017) 776–799

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12272

This article is part of the topic "2016 Rumelhart Prize Issue Honoring Dedre Gentner," Jeffrey Loewenstein and Arthur Markman (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/doi/10.1111/tops.2017.9.issue-2/issuetoc

Grammatical Constructions as Relational Categories

Micah B. Goldwater

School of Psychology, The University of Sydney

Received 17 May 2016; received in revised form 4 January 2017; accepted 11 January 2017

Abstract

This paper argues that grammatical constructions, specifically argument structure constructions that determine the "who did what to whom" part of sentence meaning and how this meaning is expressed syntactically, can be considered a kind of relational category. That is, grammatical constructions are represented as the abstraction of the syntactic and semantic relations of the exemplar utterances that are expressed in that construction, and it enables the generation of novel exemplars. To support this argument, I review evidence that there are parallel behavioral patterns between how children learn relational categories generally and how they learn grammatical constructions specifically. Then, I discuss computational simulations of how grammatical constructions are abstracted from exemplar sentences using a domain-general relational cognitive architecture. Last, I review evidence from adult language processing that shows parallel behavioral patterns with expert behavior from other cognitive domains. After reviewing the evidence, I consider how to integrate this account with other theories of language development.

Keywords: Grammatical constructions; Relational categories; Analogy; Language development; Construction grammar; Structural priming

1. Introduction

A central question in cognitive science has concerned whether language is learned, represented, and processed via mechanisms common across domains of cognition or whether language is better characterized with mechanisms unique to language (e.g., the famous debate between Jean Piaget, who argued for the former, and Noam Chomsky, who argued

Correspondence should be sent to Micah B. Goldwater, School of Psychology, University of Sydney, Room 336, Brennan MacCallum (A18), NSW 2006, Australia. E-mail: micahbg@gmail.com

for the latter [from 1975]). In this paper, I consider the evidence for a specific piece of the domain-general hypothesis: Those grammatical constructions are a form of relational category.[1] I begin by defining these key terms.

Grammatical constructions are units of language that unify form and function, structure and meaning, syntax and semantics (see e.g., Goldberg, 1995; Jackendoff, 2002). I specifically focus on argument structure constructions that capture the "who did what to whom" aspect of sentence meaning and govern the syntax of how that meaning is expressed. Relational categories are categories that are defined by the extrinsic relations among objects, not the intrinsic features of objects (e.g., Genter & Kurtz, 2005; Goldwater, Markman, & Stilwell, 2011; Goldwater & Schalk, 2016; Markman & Stilwell, 2001). For example, a barrier is any object, situation, or idea that can prevent an agent from achieving its goal (such as a giant rock or poverty; Genter & Kurtz, 2005), or a positive feedback system is any phenomenon that shows non-linear growth due to the increasing output of the system being fed back into the system leading to further increase (such as polar icecap melting or economic pricing bubbles; see Rottman, Gentner, & Goldwater, 2012).

This paper argues that grammatical constructions are relational categories in that any given construction is represented by the generalization of the exemplar utterances that fit its form and express its meaning, and fosters the production of novel exemplar utterances (building on and reiterating arguments from Goldberg, 2006; Goldberg, Casenhiser, & White, 2007). The idea is that argument structure constructions categorize sentences via common semantic and syntactic relations among their constituent words, just as a *positive feedback system* categorizes phenomena via common causal relations among their constituent events. For example, consider the double-object dative, which has the surface syntactic sequence of a noun phrase (NP) followed by a verb phrase (VP), and then followed by two more NPs. Furthermore, this syntax reliably expresses a semantic relation among the argument NPs that the first NP transferred the third NP to the second NP (that is, they describe *transfer events*) as in:

(1) John gave Bob a cookie.
(2) Doris told Fran the news.

The argument that grammatical constructions are relational categories rests on the growing body of evidence that (a) the same relational learning mechanisms that help children learn abstract relational categories such as *barrier* help children learn abstract construction semantics such as *transfer events* and their corresponding syntax, and that (b) adults' processing of constructions is supported by the same mechanisms underlying the use of relational categories wherein one is highly familiar or expert.[2] The paper now reviews this evidence in turn and concludes by examining how this proposal fits in with domain-general accounts of language learning and processing more generally.

2. Behavioral patterns of relational category learning

There is considerable evidence that central to learning relational categories is the process of structural alignment, the same cognitive process underlying analogical thinking and more “simple” similarity judgments (Gentner, 1983; Gentner & Markman, 1997). As discussed throughout this special issue, structural alignment involves the placement of two representations into correspondence based on matching relations, such as the oft-cited Rutherford model of the atom that was an analogy with the solar system wherein the nucleus corresponded to the sun, and the electrons to the planets. When the relational structures of two exemplars are put into correspondence, this can form the basis of a relational category by fostering the abstraction of their relational commonalities. While two exemplars is technically the minimum needed for a category, any such abstraction would be more greatly supported by the continued alignment of more exemplars sharing these relations. For example, aligning a series of exemplar sets of entities bound by a central force could create an abstract “central force” concept that could classify together an atom, the solar system, the milky-way galaxy, etc. The number of experimental and computational investigations of relational category learning has been rapidly growing over the past few years, (e.g., Corral & Jones, 2014; Doumas, Hummel, & Sandhofer, 2008; Goldwater & Markman, 2011; Goldwater & Gentner, 2015; Goldwater & Schalk, 2016; Jung & Hummel, 2015; Kurtz, Boukrina, & Gentner, 2013; Tomlinson & Love, 2010 for review).[3]

In addition to highlighting commonalities, structural alignment aids learning via highlighting differences related to the common structure (called “alignable differences,” Markman & Gentner, 1993). In the Rutherford atomic model, an alignable difference would be the kind of central force, gravity for the solar system, and electromagnetic for the atom. The recognition of alignable differences is key to relational categories in at least two ways: (a) They can form the basis of a distinction between concepts or functional elements within a broader system, and (b) help show how common relations are realized in the distinct sets of features of different category exemplars, which is key to expert recognition of novel disparate category members (Chi & VanLehn, 2012).

Relational learning in childhood (and in older learners) shows a consistent pattern. At the onset of learning in any particular domain, initial alignment is supported by a correlation between superficial similarity and similarity in relational structure (see Gentner, 2010, for review). That is, superficial similarity both seems to motivate a reason to attempt the alignment in the first place and help guide the process of finding the correspondence across two representations. The increased superficial similarity has aided children in discovering both relational commonalities (e.g., Kotovsky & Gentner, 1996; Loewenstein & Gentner, 2001) and alignable differences (Gentner, Loewenstein, & Hung, 2007). For example, in Gentner et al. (2015) comparing highly superficially similar stable and unstable buildings allowed children to discover the key difference: The stable buildings featured triangular bracing. After these initial alignments are made, the “training wheels” of superficial similarity can often be discarded and more pure relational matches can be discovered (e.g., Braithwaite & Goldstone, 2015; Kotovsky & Gentner, 1996). For example, consider Gentner, Anggoro, and Klibanoff (2011) wherein 3-year-old children had trouble learning relational concepts such as “food for” when tasked to extend the concept to moderately dissimilar exemplars (e.g., from food for a rabbit to food for a horse), but then were able to generalize the concept when supported by initial high-similarity alignments (e.g., comparing food for rabbits with food for mice).

In addition to this pattern characterizing learners' experience in specific domains, development of domain-general cognitive processes improves children's ability to align relational representations. Improvement in response-inhibition increases children's ability to resist responding based on superficial similarity devoid of underlying relational commonality (e.g., Morrison, Doumas, & Richland, 2011), and increase in working-memory capacity improves children's ability to consider more complex relations among multiple objects (e.g., Andrews & Halford, 2002).

While executive resources critically support learning novel relational categories, experience in a domain is marked by the reduced need to expend executive resources to align novel exemplars with prior knowledge, and instead experts can more directly apply acquired relational knowledge as holistic chunks. This pattern is seen across domains of experience. For example, chess experts can rapidly build a winning strategy from recognizing arrangements of chess pieces as exemplars of familiar categories defined by their spatial relations (such as a "fork" configuration; Chase & Simon, 1973). For a second disparate example, this pattern is similar to "the career of metaphors," wherein to comprehend a novel metaphor, an alignment is needed between the compared concepts to discover which common relations and properties the metaphor is highlighting (e.g., in *science is a glacier,* to mean science progresses slowly and steadily), but after a metaphor is well known, it can be applied to novel instances directly without a full alignment (e.g., *fear is a roadblock to success*; Bowdle & Gentner, 2005).[4]

In sum, relational category learning sees a developmental trajectory wherein structural alignment is central to the learning process, initial alignments are boosted by superficial similarity, and the early use of relational knowledge is effortful and fragile. After extensive experience (and the continued alignment of a large number of novel exemplars with existing category knowledge), robust abstractions are formed that are simply a part of an internal representational vocabulary that allows for direct recognition of familiar relational structure in novel category members with relatively little cognitive effort. Next, the paper reviews the evidence that the developmental trajectory of construction learning and processing follows this same pattern.

3. Parallels between child language and relational learning

Why would patterns of learning grammatical constructions mirror relational learning more generally? First, consider the relationship between verb and construction meaning. Verb meaning is relational in that verbs describe semantic relations between nouns (Gentner, 1982; Jackendoff, 1990, among many others). Furthermore, as exemplified by the verbs in sentences (1) and (2), the relational meaning of verbs and the relational meaning of constructions are often in lockstep. The verbs of (1) and (2) intrinsically express the semantic relation of transfer, which is why they are so frequently used in a dative form. While the construction constrains the semantic relations among the nouns, the main verb is the single best predictor of the construction as a whole, and the best predictor of any single word to the overall sentence meaning (Chomsky & Miller, 1968; Healy & Miller, 1970). The verb provides richer semantic detail than the construction, fleshing out the skeletal construction meaning (Goldberg, 2006). However, while the meaning may be more skeletal, the construction still provides semantic information and predictive value on top of the verb's contribution motivating that constructions are worthwhile generalizations to form (Bencini & Goldberg, 2000; Goldberg, Casenhiser, & Sethuraman, 2005). One clear way to see that is when verbs appear in constructions they typically do not, they seem to inherit the construction meaning for that use. For example, consider that "sneeze," typically an intransitive verb, expresses the means of moving an object to a new location in (3) consistent with the semantics of the prepositional dative (example from Goldberg, 1995). Furthermore, (4) shows that construction semantics can even lend their meaning to novel verbs derived from nouns (i.e., "novel denominal verbs"; example from Kaschak & Glenberg, 2000) in that "to crutch" clearly means to move an object to another location using a crutch. I discuss examples like (4) in greater detail below.

(3) Linda sneezed the envelope off the table.
(4) Lisa crutched her apple over to Jen so she would not starve.

These links between grammatical construction and verb meaning are critical in child language learning. There is substantial evidence that the grammatical form of the sentence (or at least the number of noun arguments) constrains verb interpretation from the onset of verb learning (e.g., Yuan, Fisher, & Snedeker, 2012). Furthermore, this paper is arguing that the relationship between verb and construction learning is bidirectional. That is, central to discovering the relational meaning of grammatical constructions (beyond the constraints of just argument number) is generalizing the common semantic relations across the verbs that frequently appear in the construction, similar to how any given verb meaning must be abstracted across instances of its use (see Goldberg, 2006; Tomasello, 2003; and below). Fittingly, there is now substantial evidence that structural alignment is critical in verb generalization. Childers and colleagues (e.g., Childers et al., 2016) have shown that aligning events described by a novel verb helps young children generalize the verb to events with novel objects, and Haryu, Imai, and Okada (2011) show that initial extension of verbs to events with highly similar objects better allows the verb to then be extended further to events with more disparate objects, consistent with the progressive alignment pattern.

There are several experiments examining artificial grammar learning that are also inline with a relational learning account, for example, the work of Waterfall and colleagues on "variation sets" (e.g., Onnis, Waterfall, & Edelman, 2008). Variation sets are consecutive utterances that overlap in some but not all words. Variation sets aid grammar learning because what the consecutive utterances share and what they do not highlight what constitutes a functional unit, fostering the discovering of phrasal categories. A relational learning account predicts benefit from such sequences because the overlap aids initial alignment, and then what varies are alignable differences which highlight key structural properties of the grammar. This pattern is similar to the research on children learning about triangular bracing from highly similar stable and unstable building pairs discussed above (Gentner et al., 2015).

Onnis et al. (2008) showed the utility of variation sets using both corpus analyses of naturally occurring child-directed speech and artificial grammar learning studies. The corpus analysis shows the rate of natural occurrence of variation sets in child-directed speech predict rates of grammar learning (Onnis et al., 2008). The work on artificial grammars showed the benefits of variation sets by presenting two sets of learners with the identical set of input (artificial) utterances, but it varied the sequence to either contain variation sets or to sequentially separate the utterances with overlapping content. The learners with the variation sets outperformed the other group. Many statistical learning accounts cannot explain this kind of order advantage because they rely on inferring grammatical structure from global distributional information (see discussion in Goldstein et al., 2010), while a relational learning account predicts that variation sets are powerful opportunities for structural insights that go beyond global information (also see Kolodny, Lotem, & Edelman, 2015).

While the pattern of Onnis et al. (2008) is consistent with the current argument, it examined artificial grammar learning devoid of semantic content. Casenhiser and Goldberg (2005) taught children novel grammatical constructions using nonce verbs, familiar nouns, a novel word order, and novel semantics (events wherein an objects suddenly appears). Across two papers, a pattern consistent with progressive alignment emerged (Casenhiser & Goldberg, 2005; Goldberg et al., 2007). When the construction appeared with a single high-frequency verb early in learning (aiding initial alignment), this enabled generalization of the word order and semantics to additional verbs later on.

Perhaps the most compelling evidence for structural alignment processes in child language comes from work on structural priming. Structural priming is shown when a grammatical construction is repeated across utterances despite alternative felicitous ways to communicate, for example continuing to use the double-object dative as in (1) and (2) instead of switching to the prepositional dative as in (3) and (4) (Bock, 1986). Structural priming has been seen as a window in the representations of children's grammatical constructions, and specifically to whether constructions are represented abstractly or whether they are tied to individual words. The logic is that if children show priming across utterances that show no lexical overlap, then they are represented abstractly (e.g., see contrasting evidence in Savage, Lieven, Theakston, & Tomasello, 2003; Shimpi, Gámez, Huttenlocher, & Vasilyeva, 2007).

Goldwater, Tomlinson, Echols, and Love (2011) examined whether structural alignment and mapping between utterances was the mechanism underlying structural priming in young children (i.e., whether structural priming was another example of domain-general analogical processes). That is, they proposed that children in dialog will attempt to use their interlocutor's sentence structure as the basis of their own (see Pickering & Garrod, 2004 for a related but distinct idea). To investigate this proposal, Goldwater et al., assessed 4- and 5-year-old children's tendency to repeat grammatical constructions using a turn-taking scene description task. The experimenter would describe two scenes, such as a family at a dinner table or children in a classroom, and then the child would describe a third scene. All the scenes depicted transfer and could be felicitously described with either dative as in (5) or (6). In addition, they could be described with other constructions as in (7).

(5) The father handed his son a piece of cake.
(6) The father handed a piece of cake to his son.
(7) The father and his son shared some cake.

The scenes came from three scene categories: food sharing scenes, sports scenes, and classroom scenes, with three examples each. The key manipulation was their order. In the high-similarity condition, the two scenes the experimenter described and the third scene the child described came from the same category. In the low-similarity condition, the scenes were intermixed across categories (see Fig. 1).

The same six scenes were described by the experimenters, and the same three scenes were described by the children in each condition, only the order differed. There was a third condition wherein the children described all nine scenes. Both the 4- and the 5-year-old children produced more datives, such as (5) and (6), and fewer other constructions such as (7), in both the high- and low-similarity conditions compared to when the children described all the scenes on their own. The 5-year-olds showed no effect of the similarity manipulation as they were more likely to repeat the dative that the experimenter used throughout the experiment in both conditions. However, while the 4-year-olds were more likely to repeat the experimenter dative in the high-similarity condition, they were just as likely to use the alternate dative form as the experimenter's dative in the low-similarity condition.

Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011) reasoned that children in the low- and high-similarity conditions were attempting to produce their utterances via analogy from the experimenter's utterances. How was this pattern of construction repetition predicted by such a structural alignment account? There are two relevant sets of relations to map when repeating a grammatical construction: syntactic and semantic. Producing either dative (in contrast to a non-dative) showed that the semantic relations of a three-argument transfer event were successfully mapped, while producing the specific dative alternate the experimenter used showed that both semantic and syntactic relations were successfully mapped.[5] The latter case reflects a more relationally complex mapping. The pattern in Goldwater et al. reflected that all children could successfully use the experimenter's semantic relations to guide their sentence production, but the 4-year-olds needed the aid of the correlated superficial scene similarity to achieve the more complex mapping.

The early requirement for high levels of more superficial semantic similarity to achieve complex mappings was generally consistent with the pattern of relational development outlined above (e.g., Kotovsky & Gentner, 1996). Furthermore, the increased use of both datives equally (when comparing the low-similarity 4-year-olds to the 4-year-olds who described all of the scenes themselves) cannot be explained by other current models of

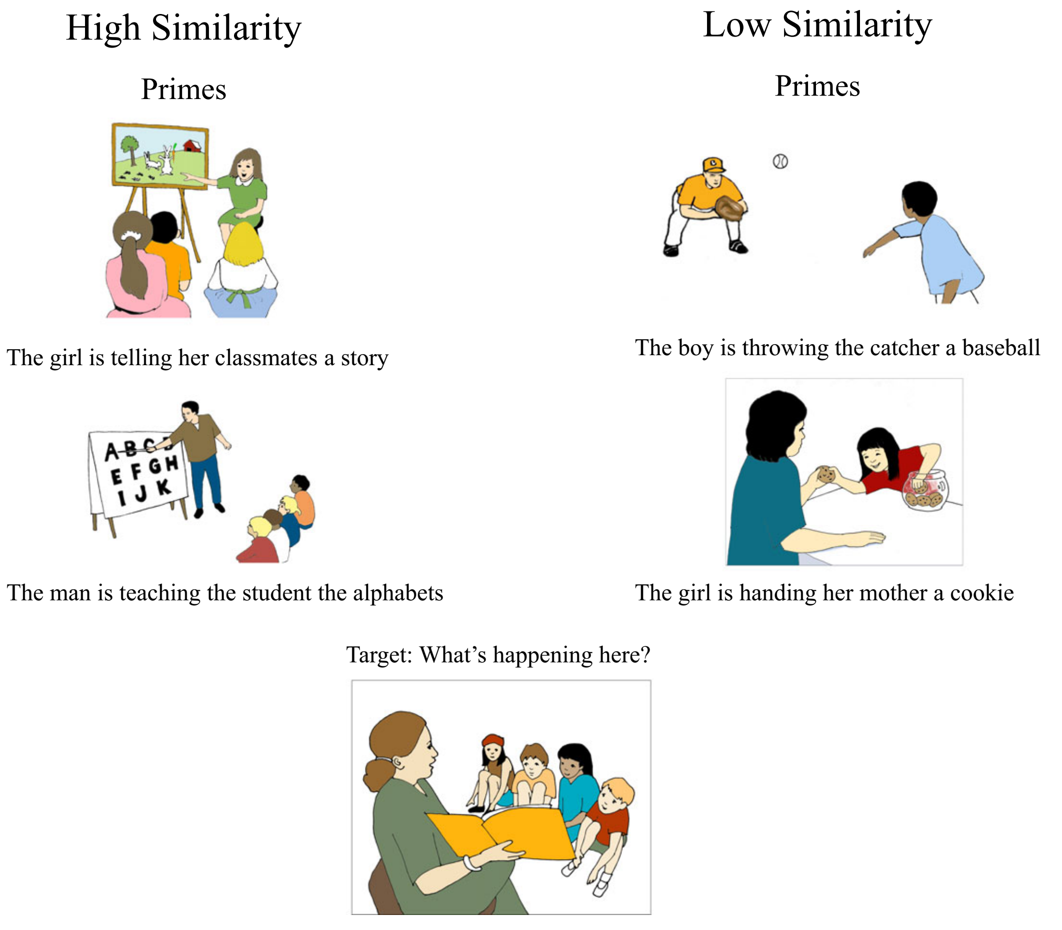

Fig. 1. High- and low-similarity sets of scenes from Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011). The first two were described by experimenters, the third by the child.

structural priming because priming in these models is a result of competition between two alternates (e.g., Chang, Dell, & Bock, 2006; Pickering & Garrod, 2004). That is, priming one alternate comes at the cost of the other. These models simulate priming via a competition between word orders given pre-established semantics. The structural alignment account posits the same mechanisms can lead to mapping semantic and syntactic relations across utterances, not that syntactic persistence is a separate process from construction sentence meaning.

While only the structural alignment account seems capable of predicting the pattern of Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011), it does not distinguish between two possibilities for the cause of difference between the 4- and 5-year-olds. It was possible the 5-year-olds have more robust knowledge/more abstract representation of the dative construction, but it was also possible the change was largely due to increases in domain-general processing capacity. Andrews and Halford (2002) show the transition from 3- to 5-years-old is a critical time in children's growing ability to handle relational complexity. The 5-year-olds increased working-memory capacity could explain the improved syntactic mapping.

Goldwater and Echols (in progress; download preprint here: https://osf.io/tw87z/) are examining whether the cause of the 4-year-olds' reliance on high superficial similarity in the previous work was rooted in a less robust representation of the dative construction by examining if that knowledge could be improved (at least temporarily) in the course of the turn-taking scene description task itself. If the children's previous deficit was general cognitive capacity, one would not predict performance could be so readily improved. Inspired by research on progressive alignment, this experiment presented two blocks of three pairs of scenes (making six total scene pairs) to 4-year-old children. The experimenter described the first scene and the child described the second of every pair. That is, they alternated for twelve scenes. Across conditions, the children described the same scenes in the same order. In addition, the experimenter described the same scenes in the same order for the second block of three scene pairs. This second block of scene pairs that were identical for all children and shared little superficial similarity (i.e., were from different scene categories). The key between-condition manipulation was in the first three scenes the experimenter described. In one condition, the three scenes the experimenter described were of high similarity to the following scene the child described. In the other condition, they were of low similarity. Taken together, there were two conditions: (a) high-to-low similarity, which first presented a block of high-similarity scene pairs, and then a second block of low-similarity scene pairs; (b) low-to-low similarity, which presented two blocks of low-similarity scene pairs. Analyzing the first block alone replicated Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011) because the high-similarity pairs elicited significantly more repetitions of the matching dative alternate by the children than in the low-similarity condition. Critically, this matching dative advantage in the high-to-low condition persisted into the second block, despite all children now engaging with the identical low-similarity scene pairs. That is, the progressive alignment pattern of how alignment supported by high similarity can act as training wheels to support later low-similarity alignment was clearly demonstrated (see Fig. 2).

Summarizing the prior and ongoing work together, 4-year-olds increase dative use overall with low- or high-similarity scene pairs, but they are only more likely to match the specific dative of their interlocutor with high-similarity pairs. However, following a block of high-similarity pairs, 4-year-olds can match the specific dative alternate with less support from superficial similarity. Five-year-olds show the mature pattern of matching the specific dative alternate regardless of similarity relations between scene pairs.

Looking at this pattern of the decreasing dependence on superficial similarity, a structural alignment account seems to supply the best explanation. The pattern suggests that the relational knowledge of the dative construction gets increasingly robust and the child is better able to align scenes and sentences with these representations. Currently, it is unclear if other current accounts of structural priming can explain any key piece of the

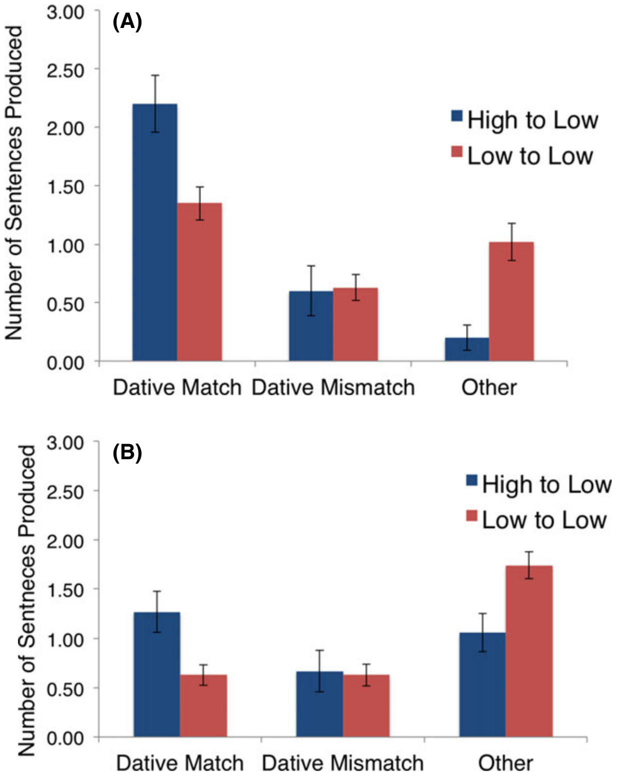

Fig. 2. Results from 4-year-olds from Goldwater and Echols (2016). (A) First block of three pairs of trials (above). (B) Second block of three pairs of trials (below). High-to-low condition elicited significantly more matching datives in both blocks.

pattern: the effects of similarity, progressive alignment, and the ability to boost both dative alternates at once.

In sum, this section supports the argument that grammatical constructions are relational categories by reviewing the experimental evidence that structural alignment is a key process in verb learning, artificial grammar learning, and structural priming. Next, the paper presents some initial attempts to simulate construction learning and structural priming using a computational architecture for analogical thinking and relational abstraction.

4. Simulating learning and producing grammatical constructions with a computational architecture for analogical thinking

With Scott Friedman, Dedre Gentner, Ken Forbus, and Jason Taylor, we used a group of interrelated models all centered on the Structure-Mapping Engine (1989), the computational implementation of Gentner's structure-mapping-theory (1983), to simulate the development of structural priming. This simulation involves two sets of processes, the generalization of construction event semantics across verbs, and then the structural priming process itself to test the effects of semantic abstraction. To set up both simulations, I describe how sentences are represented in the model.

The representations and simulations are grounded in the work of Chang et al. (2006), particularly how the training corpus is composed of a small set of lexical items, concepts that the words refer to, and sentence constructions. Here, we focus on the dative constructions. Every sentence has three levels of representation. The first was "sequential," where it is just the linear order of lexical forms (that is, words connected by a *follows* relation). The second was "semantic," which contains the conceptual referents of the words, the semantic roles they are linked to, and some additional semantic relations specific to the events the sentences describe (see below). The third level was "referential," which bound the concepts of the semantic level to the word-forms of the syntactic level.

To further detail the semantic level, there were three kinds of dative events distinguished by three kinds of additional semantic relations. There were events of exchange wherein one object changed possession between two others, for example, (8). There were events of creation wherein a new object came into existence (9). There were events of exchange where two objects changed possession (10).

(8) John gave Bill the book.
(9) Bob baked Dave a cake.
(10) Sally sold Tim some shoes.

To simulate the abstraction of event semantics, the verbs at the beginning of the simulation governed roles specific to them. That is, instead of *agent*, there was *giver* and *baker*. These representations are consistent with the "verb-island" stage of development in Tomasello's constructivist account of grammar learning (2003) when construction representations are tied to individual words (i.e., they are not fully abstract). This simulation had 10 blocks of 10 training trials each. During these training trials, the model would attempt to align pairs of sentences. If lexical structure, event semantics, and word order were sufficiently alignable, but the semantic roles were non-identical (because they were specific to different verbs), then the model would infer that they were both subtypes of a minimally more abstract role that subsumed those two. Then, new sentences with those verbs would be represented with these more abstract roles. When one of these sentences could be aligned with a different verb-specific or multi-verb role, new even more abstract roles would be formed in the same manner, subsuming the previous roles. With more training, the roles became more and more abstract.

In between every training block, we tested the roles' abstractness with structural priming sequences. To simulate structural priming, the model was fed pairs of dative sentences into a short-term memory buffer. The first in the pair was a fully formed sentence with all three levels of representation. The second of the pair was just the semantic level. The job of the model was to express the meaning by producing referential and sequential representations. The model did this with SME. That is, if SME could align the semantics with the semantics of the previous sentence, then it would project the referential and sequential structure of the previous sentence as analogical inferences to the current semantics. More specifically, when semantic roles aligned across sentences, they would be expressed in the same word order. The result is a novel sentence with novel words, but expressed in the same construction as the previous sentence. That is, structural priming was shown. If the semantics of the pairs of sentences could not be aligned, then MAC-FAC, a model of similarity-based retrieval (Forbus, Gentner, & Law, 1995) would find an appropriate match in the long-term memory store (which contained previous exemplar sentences and sentence-generalizations; see Taylor, Friedman, Forbus, Goldwater, & Gentner, 2011 and below for more details) to serve as the basis of the syntactic form of the new sentence. In this case, the previous sentence did not influence the new sentence, and thus structural priming would not be shown (see Fig. 3).

At the start of the simulation, priming was only shown when pairs of sentences shared verbs because sharing verb-specific roles were necessary for alignment. This was tested with a block of 50 pairs of sentences. Fig. 4 shows the results of each new block of 50 test-pairs, each following a block of 10 training trials. This pattern fits a key component of Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011), that across-verb priming of specific dative alternates increased from 4 to 5 years of age.

While, we believe this simulation demonstrates the plausibility of a structure-mapping account of construction learning by formally specifying possible developmental mechanisms, caveats need noting. First, critical to the specific pattern of semantic role abstraction was the particular event semantics reflected in the classes of (8)–(10). That is, the particular abstract roles those were formed reflected these classes because the semantic

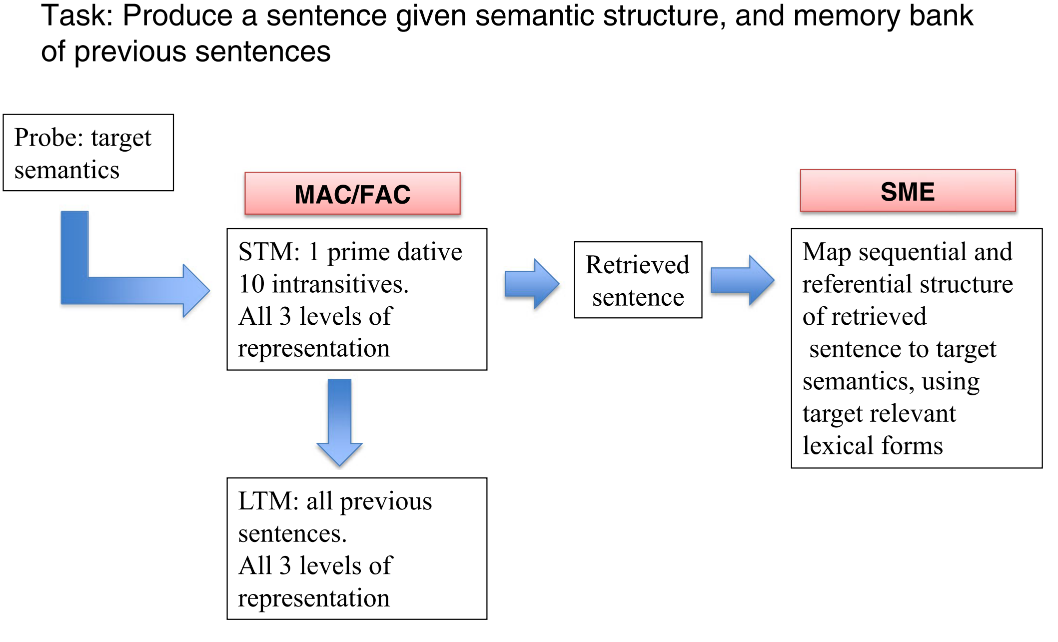

Fig. 3. Simulating structural priming with Mac-Fac & SME. See Taylor et al. (2011) for more details.

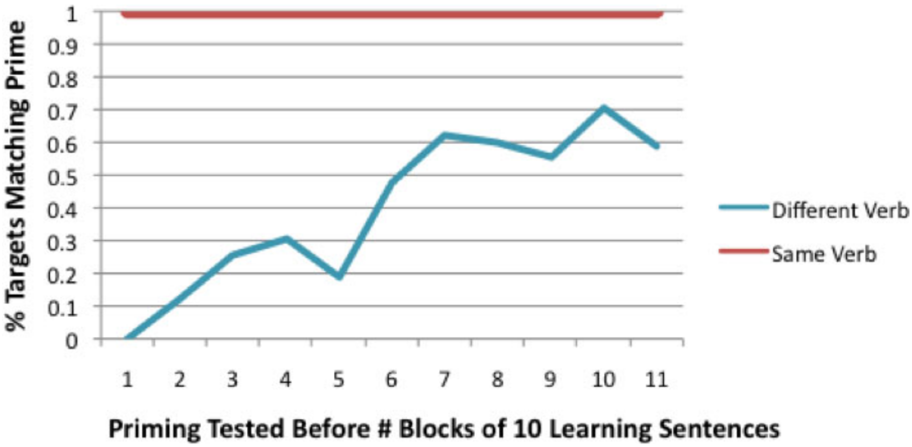

Fig. 4. Results from simulating structural priming. There were 50 pairs of full sentences and probe semantics to produce sequential and referential forms for. Overall, the ability of the model to produce sentences that match the previous sentence's construction increased across the 10 blocks.

relations of each class fostered alignment of within-category sentences. Although to be clear, we are not committed to those semantic representations specifically, just that some form of overlapping semantic representations across events would aid the alignment and abstraction process (e.g., see Pinker, 1989; Ambridge, Pine, Rowland, & Chang, 2012, for analyses of dative semantic classes). In addition to overlapping event semantics aiding abstraction, Chambers and Jurafsky (2009) developed an algorithm to find abstractions of semantic roles by using coreference in a discourse to guide alignment (e.g., the patient of *arrest* is likely to also be the patient of *convict*, suggesting the two verbs share a more general patient role.) This process is specifically helpful for computational applications that can only learn from language input without reference to perceivable events, but we see no reason why children would not exploit both coreference and event representations to find generalizations.

The second caveat to note is how incomplete these simulations were. That is, much like Chang et al. (2006), these simulations did not demonstrate boosts to both dative alternates at once (as demonstrated in Goldwater, Markman et al., 2011; Goldwater, Tomlinson et al., 2011; Goldwater & Echols, 2016) because the semantics of any given sentence were predetermined to serve as the basis for alignment to then project a word sequence. More computational work will be needed to implement the current account that the event semantics themselves can be based on the previous sentence in the same manner that the sequential and referential structure can. Last, clearly mature syntax is more than the linear word order represented here. But we take it as a strength of the model that structural priming could be demonstrated with such a stripped-down syntactic representation. Still, an account of forming hierarchical syntactic structure is also critical for this modeling approach (see discussion below).

In sum, this section discussed an initial formal attempt to implement a relational learning account of grammatical construction learning using a computational architecture with

SME at its heart. This model aligns input sentences to find generalizations, and like a cluster model of categorization (e.g., Anderson, 1990; Love et al., 2004) its memory store consists of exemplars and abstractions (much like our representation of language; Jackendoff, 2002; Goldberg, 2006) While this formal implementation is clearly far from complete, it captures a key qualitative pattern of a shift from verb-specific to verb-general structural priming, as demonstrated in Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011), showing the promise of this approach for further development.

5. Grammatical constructions as relational categories in adult language processing

To further support the central argument of this paper, I now turn to evidence from adult language processing. First, I discuss the evidence that grammatical constructions are relational categories like any other by showing how construction semantics are an inherent and productive piece of our conceptual system more generally. This research (by Goldwater, Markman, and colleagues) has focused on the inter-relations between two kinds of relational categories: schema-governed categories, which classify entire relational systems (e.g., *marriage*, *catalysis*), and role-governed categories, which are represented by the roles within such systems (e.g., *husband and wife*, *reagent and solvent*). The working hypotheses were because construction semantics are schemas that refer to relations among multiple argument-roles, that (a) when interpreting a novel denominal verb, that novel verb inherits the relational semantics of the construction in which it appears (Kaschak & Glenberg, 2000) and (b) lexicalizing the inherited relational structure with a novel verb would license the creation of role-governed categories specific to that novel verb (such as the agent of the action the verb referred to).

To investigate hypothesis 1, Goldwater and Markman (2009) introduced novel denominal verbs (in italics) in one of two constructions, the passive as in (11), and the middle as in (12).

(11) The tomatoes were *sauced* quickly to prepare for the banquet.
(12) The tomatoes had *sauced* quickly to prepare for the banquet.

The subjects read the first five words (i.e., the main clause) as a whole, and then read each next word one at a time, progressing to the next word by making a sensicality judgment. That is, if the sentence continued to make sense with each new word, the subject would deem the sentence sensical, and the next word was presented. If the sentence was deemed non-sensical, then the trial was over and the next sentence would appear. The critical trials of the experiment all had a rationale clause, for example, “in order to X” following the main clause. During the rational clause, when the preceding main clause was in the middle construction more sentences were rejected as non-sensical compared to when the main clause was in the passive construction. Goldwater and Markman reasoned (building on Mauner & Koenig, 2000) that this reflected how the passive construction semantics provided an agent of the event (though not explicitly mentioned in the sentence) responsible for the action described by the verb and as the basis of the rational clause, while the middle construction did not provide an agent.

To investigate hypothesis 2, Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011) then showed that after the construction provides an agent for a novel demoninal verb, this licenses the use of a novel noun derived from that verb to refer to that novel agent. That is, novel lexicalized relational structures license novel role-governed categories. Goldwater et al. showed that agent terms, as in (16), were most easily understood following novel denonimal verbs (as in 13), than when following either a paraphrase of the novel verb (as in 14) or a novel adjective (as in 15) wherein there were no novel lexicalized schema concepts endowed with the semantic relations of a grammatical construction.

(13) Paul went to Mardi Gras and *whiskied* himself stupid.
(14) Paul went to Mardi Gras and drank himself stupid with *whiskey*.
(15) Paul went to Mardi Gras and had a *whiskeylicious* time.
(16) The next day, the *whiskier* felt terrible.

Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al.'s (2011) interpretation of the novel verb condition's advantage was that it reflected general conceptual processes. However, there's an alternative that the reading time data alone could not rule out: that the advantage for the novel verb condition was purely morphosyntactic in nature. That is, it is possible that the agent terms were easier to understand because the *-er* derivational morpheme only operates on verbs, and the other two conditions did not provide a verb from which to directly derive the agent term. Goldwater, Markman, Trujillo, and Schnyer (2015) used ERP methodology to rule out this possibility. If the advantage for the novel verb condition was because the agent term was seen as a error in derivational morphology in the other conditions, then following the paraphrase, the novel agent term should elicit a more positive ERP around 600 ms post-stimulus-onset than when following the novel verb (the P600, see Kaan, Harris, Gibson, & Holcomb, 2000). However, there was no such increased P600, and instead, the paraphrase elicited an increased negativity around 400 ms post-stimulus-onset, consistent with the N400, a marker of the ease of semantic integration of a word into the representation of a sentence, and with ERP's marking differences in working-memory load and cognitive conflict more generally (e.g., Kutas & Federmeier, 2011; Rugg & Allan, 2000). This suggested that the same mechanisms underlying relational semantics more generally explained the novel verb advantage. That is, taking these three lines of experiments on adult language processing together, there is strong evidence that construction semantics are a productive part of the mature conceptual system.

While Goldwater and colleagues' work on adult sentence processes focused on the agent role specifically, there is evidence from adult structural priming that construction representations are comprised of full sets of semantic roles and their sequences. This is important evidence that constructions are represented as coherent and holistic relational concepts (similar to how experts represent their domain of expertise; see discussion above). Chang, Bock, and Goldberg (2003) showed that even when syntax was held constant, semantic role sequences could be primed across utterances. That is, (17) and (18) would selectively prime their respective semantic role sequences in new sentences despite sharing the same phrase order of NP-VP-NP-PP.

(17) The maid rubbed polish onto the table.
(18) The maid rubbed the table with polish.

Popov and Hristova (2014) gave further evidence that structural priming effects are not about individual roles, but full sets. Consider the following example sentences:

(19) The hunter watched the alpinist with binoculars.
(20) The doctor watched the patient by using glasses.
(21) The doctor watched the patient who wore glasses.
(22) The doctor and the patient watched by using glasses.

"For a target sentences, such as (19) participants interpreted the ambiguous role 'with binoculars' as an instrument of the action more often when the corresponding role in the previous sentence was also an instrument, as in (20), compared to when it was an attribute of the preceding noun-phrase as in (21). The effect was present only when the whole structure of the base was analogous to the target – the nonanalogical base (22) did not increase the amount of instrumental interpretations of the target, although its key role was instrumental. Therefore the effect was not due to the activation and priming of the key concept instrument alone . . . but the result of systematic mapping" (p. 1196) of multiple roles bound by relations.

Furthermore, Popov and Hristrova argue that priming the entire set of thematic roles was automatic. Consistent with the claim that construction meaning as a whole is automatically accessed by adults, Johnson and Goldberg (2013) presented evidence that constructions with nonce words (e.g., *He daxed her the norp*) primed verbs in a lexical decision task with semantically congruent meanings (e.g., *give*, *transfer*).

In sum, evidence from adult language processing suggests that the meanings of mature grammatical constructions are accessed automatically, retrieved in their entirety, and are a productive part of our conceptual system more generally. This pattern is similar to learning relational concepts in other domains wherein learned relational schemas are retrieved from memory in entirety and can be the basis of generating novel uses (e.g., see Chase & Simon, 1973; Anderson & Fincham, 2014, for examples from chess and mathematical problem solving, respectively).

6. How relational learning fits into domain-general accounts of language development more generally

This paper has argued for the role of domain-general relational cognitive processes underlying child language development and adult language processing. However, most domain-general accounts of language learning have focussed on how syntax is learned from the sequential and distributional statistics of words and phrases (e.g., Chang et al., 2006; Gómez & Gerken, 2000). There are of course a large variety of models within this umbrella. For example, some accounts have focused on how basic predictive-learning mechanisms (rooted in the animal learning literature; Rescorla & Wagner, 1972) explain learning basic rules of morpho-syntax (e.g., Arnon & Ramscar, 2012), and others in the tradition of computational linguistics and machine learning have developed unsupervised learning algorithms (with no built-in language knowledge) to learn grammar from natural language corpora (e.g., Kolodny et al., 2015; though of course animal and machine learning traditions inform each other; see Sutton & Barto, 1998).

Some argue that purely distributional learning mechanisms could discover these high-order abstractions precluding the need for analogical abstraction mechanisms (e.g., Twomey, Chang, & Ambridge, 2014; see Ambridge & Lieven, 2015; Ambridge, Goldwater, & Lieven, unpublished data); however, they need not be mutually exclusive as sequential-statistical learning mechanisms, and mechanisms of analogical abstraction work together in other domains. For example, one could independently learn via the co-occurrence statistics of causes and effects that two different natural phenomena are characterized by multiple causes causing a single effect (e.g., Fernbach & Sloman, 2009; Holyoak & Cheng, 2011), and then structural alignment can support recognizing this commonality in the causal relations between the two phenomena (Goldwater & Gentner, 2015).

Likewise, recent advances in computational models of syntax learning combine statistical and analogical learning. For example, the computational models presented in Kolodny et al. (2015) and Bod (2009) infer hierarchically structured syntactic representations from the sequential statistics of individual words and phrases, but also each use forms analogical comparison to recognize commonalities in the sequential relations among words and phrases, which in turn is central for the models' ability to generate and interpret novel utterances, and create novel syntactic abstractions (though these models do so in distinct ways).

How the models of Kolodny et al. (2015), Bod (2009), and others learn hierarchical syntactic representations would also nicely complement the semantic abstraction mechanisms of the Friedman and colleagues simulations discussed above. In those simulations, while semantics were represented by hierarchical relational structures, syntax was merely a linear sequence. Integrating hierarchical syntactic representations with the hierarchical semantic representations should only increase the system's computational power, and importantly improve the ability of the model to scale beyond toy corpora to natural language (see discussion in Kolodny et al., 2015).[6]

In sum, this paper has tried to build the case that analogical learning across many exemplar utterances forms utterance-categories represented as abstract semantic and syntactic relations. This section has tried to make clear that these analogical processes are not to the exclusion of other forms of learning mechanisms. Indeed, this is of course the basic idea behind domain-general accounts of language learning: The child has the potential to apply his or her general cognitive capacities, whatever they may be, to language learning. Thus, there is no reason to presume that if associative, sequential, distributional, and analogical learning mechanisms are *all* at children's disposal that they would not *all* play some role in learning grammatical constructions.

7. Argument summary

This paper has argued that grammatical constructions are relational categories in that they are represented by the abstraction of exemplar sentences' semantic and syntactic relations. Like learning relational categories in other domains, they show a characteristic behavioral pattern suggesting that the structural alignment of exemplars is key to discovering their relational commonalities. This characteristic pattern of how superficial similarity acts as "training wheels" to help discover the deeper commonalities has been shown in both artificial language learning studies, and in studies of children's language production. Furthermore, similar to how experts can directly apply their relational knowledge to novel contexts with a relatively effortless mapping process, adult sentence processing and structural priming show the use of abstract grammatical semantics and syntax seemingly automatically.

The paper also presented computational simulations using the domain-general cognitive architecture of Forbus and Gentner with the Structure-Mapping Engine at its heart. More work will be required to fully adapt this system to learn constructions from natural corpora, and closely simulate both relational category learning and language production experiments. However, the computational system is already in good position because it is intrinsically aligned with multiple theoretical linguistic approaches (e.g., Goldberg, 2006; Jackendoff, 2002) in that it both stores exemplar utterances and forms abstractions from these exemplars, akin to cluster models of categorization (e.g., Anderson, 1990; Love et al., 2004), and captures key qualitative developmental patterns of children's sentence production (from Goldwater, Markman et al., 2011; Goldwater, Tomlinson et al., 2011). Future computational and experimental research should focus on how structural alignment works with associative and distributional learning mechanisms to form a more complete domain-general model of grammar learning.

8. Conclusion

I started this paper affecting to situate the current argument in the context of the famous Chomsky and Piaget debate about whether language was distinct from the rest of cognitive development. Chomsky carefully considered Piaget's case and then famously rejected the argument that any "general developmental mechanisms" could explain the uniqueness of human language. Many decades later, thanks to (in no small part) Dedre Gentner's work on analogy and relational learning, we have a better understanding of what these general developmental mechanisms really are, and can thus better consider whether they can explain how children learn language.

Acknowledgments

Of course, I would like to first thank Dedre Gentner for her mentorship and influence of the research described here. In addition, I'd like to thank Art Markman and Jeff Loewenstein for editing this volume, inviting me to contribute, and their intellectual influence on the work as well. Last, I would like to acknowledge that Australian Research Council grant no. DP150104267 funds my research.

Notes

1. Unfortunately, considering the evidence in favor of domain-specific theories is beyond the scope of the paper.
2. Categories, such as *barrier*, or *mammal*, typically have labels that, among other things, support them as the direct object of explicit reasoning (e.g., by a biologist developing a taxonomy). On the one hand, this could suggest that constructions are only categories for (psycho-, computational, and theoretical) linguists, but for typical language users, there are just collections of words, phrases, and compositional rules. On the other hand, it is quite clear speakers have formed robust abstractions of sentence-sized pieces of semantics and syntax, and there is evidence that when directly asked to categorize sentences, speakers do use their shared grammatical constructions just as much as their shared words (Goldberg, 2006). Does this imply there may be relational categories from other domains, such as vision, that do not have labels? Most likely, but in domains outside of language, child development, and education, the field of cognitive science (in the name of experimental control) has done relatively little work analyzing the representations that people actually have (see Ray Jackendoff's 2014 Rumelhart address for an elaborated argument along those lines https://www.youtube.com/watch?v=M8W4tn8rso4).
3. While the current account assumes relational abstractions go beyond just a collection of exemplars (see Goldwater & Gentner, 2015, for experimental evidence), these abstractions are graded. In contrast, the reader should refer to Tomlinson and Love (2006) and Davis, Goldwater, and Giron (2016) for exemplar-models of relational categorization. Most likely, extending cluster models of categorization that store abstractions and exemplars based on their respective utility (e.g., Love, Medin, & Gureckis, 2004) to relational categories will provide the best explanation.
4. To be clear, the use of familiar relational knowledge by experts is not entirely effort-free, particularly when it needs to be adjusted or amended. While applying familiar relational knowledge holistically is a relatively effortless process compared to initial learning, recent neuroimaging studies show that the neurosignature of structural alignment is re-engaged when relational knowledge needs to be adjusted to fit new exemplars that do not perfectly align with past ones (e.g., Anderson &

Fincham, 2014; Davis et al., 2016; Fiebach, Schlesewsky, Lohmann, Von Cramon, & Friederici, 2005).

5. To be clear, the two dative alternates have subtly distinct semantics, as the PP form does not entail a true recipient that possesses the theme object, just an end location for caused motion of the object (see Goldberg, 2006; Pinker, 1989). However, as all the events described in the experiment were of transfer, this distinction may not be important for the discussed pattern of results, and at a minimum, one could say the semantics mapped were of a three-role-relation, while the non-datives the children produced had only one or two semantic roles.
6. In addition to improving simulations of language learning, integrating these distributional syntactic learning mechanisms with structural alignment could solve a current empirical mystery in the structural priming literature. In contrast to Goldwater and colleagues' findings reviewed above, Rowland, Chang, Ambridge, Pine, and Lieven (2012) showed that children as young as 3-years old showed dative priming without specific support from superficial semantic similarity or lexical overlap, arguing that they learned these syntactic patterns from distributional learning alone. However, this structural priming task was perhaps less demanding as the experimenters began the sentences for the child, for example, "The students handed..." and the children then just needed to complete the sentence, for example, "their homework to the teacher" or "the teacher their homework." The claim of the structural alignment account of structural priming is that children will take advantage of recent sentences as the basis of their sentence construction. In Goldwater, Markman et al. (2011) and Goldwater, Tomlinson et al. (2011) it was easier for the 4-year-olds to build their sentences from the high-similarity adult sentences, allowing for the dative syntax to be more precisely mapped. In the case of Rowland et al. (2012) the children just had to select the two post-verbal arguments and put them in some order, perhaps reducing their "relational load" in an analogous manner. That is, in both the high-similarity condition, and in this partial-sentence completion task, the need for an abstract event-semantics-to-syntax mapping to guide sentence production was reduced, allowing their syntactic knowledge alone to guide mapping between sentences (though further research is clearly needed).

References

Ambridge, B., & Lieven, E. (2015). A constructivist account of child language acquisition. *The Handbook of Language Emergence*, *87*, 478.

Ambridge, B., Pine, J. M., Rowland, C. F., & Chang, F. (2012). The roles of verb semantics, entrenchment, and morphophonology in the retreat from dative argument-structure overgeneralization errors. *Language*, *88*(1), 45–81.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Psychology Press.

Anderson, J. R., & Fincham, J. M. (2014). Extending problem-solving procedures through reflection. *Cognitive Psychology*, *74*, 1–34.

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, *45*(2), 153–219.

Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, *122*(3), 292–305.

Bencini, G. M., & Goldberg, A. E. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language*, *43*(4), 640–651.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355–387.

Bod, R. (2009). From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science*, *33*(5), 752–793.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, *112*(1), 193.

Braithwaite, D. W., & Goldstone, R. L. (2015). Effects of variation and prior knowledge on abstract concept learning. *Cognition and Instruction*, *33*(3), 226–256.

Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, *8*(6), 500–508.

Chambers, N., & Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In *Proceedings of the joint conference of the 47th annual meeting of the ACL and the 4th international joint conference on Natural Language Processing of the AFNLP: Vol. 2* (pp. 602–610). Association for Computational Linguistics.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*(1), 29–49.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81.

Chi, M. T., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177–188.

Childers, J. B., Parrish, R., Olson, C. V., Burch, C., Fung, G., & McIntyre, K. (2016). Early verb learning: How do children learn how to compare events? *Journal of Cognition and Development*, *17*(1), 41–66.

Chomsky, N., & Miller, G. A. (1968). Introduction to the formal analysis of natural languages. *Journal of Symbolic Logic*, *33*(2), 299–300.

Corral, D., & Jones, M. (2014). The effects of relational structure on analogical learning. *Cognition*, *132*(3), 280–300.

Davis, T., Goldwater, M., & Giron, J. (2016). From concrete examples to abstract relations: The rostrolateral prefrontal cortex integrates novel examples into relational categories. *Cerebral Cortex*.

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, *115*(1), 1.

Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 678.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., Von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca’s area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, *24*(2), 79–91.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, *19*(2), 141–205.

Genter, D., & Kurtz, K. J. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the lab* (pp. 151–175). Washington, DC: American Psychological Association.

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

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.

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., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2015). Rapid learning in a children's museum via analogical comparison. *Cognitive Science*, *40*, 224–240.

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. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45.

Goldberg, A. (1995). *Constructions: A construction grammar account of argument structure*. Chicago, IL: University of Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford, UK: Oxford University Press on Demand.

Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2005). The role of prediction in construction-learning. *Journal of Child Language*, *32*(02), 407–426.

Goldberg, A. E., Casenhiser, D., & White, T. R. (2007). Constructions as categories of language. *New Ideas in Psychology*, *25*(2), 70–86.

Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences*, *14*(6), 249–258.

Goldwater, M. (2017). Progressive alignment in children's sentence production. Retrieved from osf.io/tw87z

Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition*, *137*, 137–153.

Goldwater, M. B., & Markman, A. B. (2009). Constructional sources of implicit agents in sentence comprehension. *Cognitive Linguistics*, *20*(4), 675–702.

Goldwater, M. B., & Markman, A. B. (2011). Categorizing entities by common role. *Psychonomic Bulletin & Review*, *18*(2), 406–413.

Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition*, *118*(3), 359–376.

Goldwater, M. B., Markman, A. B., Trujillo, L. T., & Schnyer, D. M. (2015). Licensing novel role-governed categories: An ERP analysis. *Frontiers in Human Neuroscience*, *9*.

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin*, *142*(7), 729–757. Advance online publication. https://doi.org/10.1037/bul0000043.

Goldwater, M. B., Tomlinson, M. T., Echols, C. H., & Love, B. C. (2011). Structural priming as structure-Mapping: Children use analogies from previous utterances to guide sentence production. *Cognitive Science*, *35*(1), 156–170.

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*(5), 178–186.

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

Healy, A. F., & Miller, G. A. (1970). Verb as main determinant of sentence meaning. *Psychonomic Science*, *20*(6), 372–372.

Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology*, *62*, 135–163.

Jackendoff, R. (1992). *Semantic structures* (Vol. *18*). MIT press.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press.

Johnson, M. A., & Goldberg, A. E. (2013). Evidence for automatic accessing of constructional meaning: Jabberwocky sentences prime associated verbs. *Language and Cognitive Processes*, *28*(10), 1439–1452.

Jung, W., & Hummel, J. E. (2015). Making probabilistic relational categories learnable. *Cognitive Science*, *39*(6), 1259–1291.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, *43*(3), 508–529.

Kolodny, O., Lotem, A., & Edelman, S. (2015). Learning a generative probabilistic grammar of experience: A process-level model of language acquisition. *Cognitive Science*, *39*(2), 227–267.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*(6), 2797–2822.

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1303.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development*, *2*(2), 189–219.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*(2), 309.

Markman, A. B., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language*, *32*(4), 517–535.

Markman, A. B., & Stilwell, C. H. (2001). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence*, *13*(4), 329–358.

Mauner, G., & Koenig, J. P. (2000). Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language*, *43*(1), 110–134.

Morrison, R. G., Doumas, L. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. *Developmental Science*, *14*(3), 516–529.

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*(3), 423–430.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(02), 169–190.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Popov, V., & Hristova, P. (2014). Automatic analogical reasoning underlies structural priming in comprehension of ambiguous sentences. In P. Bello et al. (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1192–1197). Austin, TX: Cognitive Science Society.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.

Rottman, B. M., Gentner, D., & Goldwater, M. B. (2012). Causal systems categories: Differences in novice and expert categorization of causal phenomena. *Cognitive Science*, *36*(5), 919–932.

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, *125*(1), 49–63.

Rugg, M. D., & Allan, K. (2000). Event-related potential studies of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 521–537). Oxford, UK: Oxford University Press.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, *6*(5), 557–567.

Shimpi, P. M., Gámez, P. B., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3-and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, *43*(6), 1334.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1). Cambridge, MA: MIT Press.

Taylor, J. L., Friedman, S. E., Forbus, K. D., Goldwater, M., & Gentner, D. (2011). Modeling structural priming in sentence production via analogical processes. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2916–2921).

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomlinson, M. T., & Love, B. C. (2006). From pigeons to humans: Grounding relational learning in concrete examples. In *Proceedings of the national conference on artificial intelligence* (Vol. 21, No. 1, p. 199). Menlo Park, CA; Cambridge, MA; London: AAAI Press; MIT Press; 1999.

Tomlinson, M. T., & Love, B. C. (2010). When learning to classify by relations is easier than by features. *Thinking & Reasoning*, *16*(4), 372–401.

Twomey, K. E., Chang, F., & Ambridge, B. (2014). Do as I say, not as I do: A lexical distributional account of English locative verb class acquisition. *Cognitive Psychology*, *73*, 41–71.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, *83*(4), 1382–1399.