

Young children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners*

JEREMY K. BOYD

University of Illinois at Urbana-Champaign

AND

ADELE E. GOLDBERG

Princeton University

(Received 27 July 2009 – Revised 28 April 2010 – Accepted 3 February 2011 –
First published online 22 June 2011)

ABSTRACT

The present study¹ exposed five-year-olds ($M=5;2$), seven-year-olds ($M=7;6$) and adults ($M=22;4$) to instances of a novel phrasal construction, then used a forced choice comprehension task to evaluate their learning of the construction. The abstractness of participants' acquired representations of the novel construction was evaluated by varying the degree of lexical overlap that test items had with exposure items. We found that both child groups were less proficient than adults, but seven-year-olds showed evidence of across-the-board generalization whereas five-year-olds were sensitive to lexical overlap at test. This outcome is consistent with more conservative, item-based learning of syntactic patterns in younger children. Additionally, unlike adults and seven-year-olds, five-year-olds showed no evidence of having mastered the novel construction's linking rules. Thus, younger learners

[*] Addresses for correspondence: Jeremy K. Boyd, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 N. Mathews Ave., Urbana, IL 61801. e-mail: jkboyd@illinois.edu; Adele E. Goldberg, Department of Psychology, Green Hall, Princeton University, Princeton, NJ 08544. e-mail: adele@princeton.edu.

[1] We are grateful to all of the participants who took part in this study, as well as the parents and teachers of children in the local preschool and afterschool programs that we worked at for their cooperation. Many thanks to Erica Wojcik and Christina Farah for their assistance in running participants, and to Sam Glucksberg, Phil Johnson-Laird, Elizabeth Wonnacott, three anonymous JCL reviewers, and Philip Dale for helpful feedback on earlier drafts of this paper. This research was supported by an NICHD postdoctoral fellowship awarded to the first author (2-T32-HD055272A), an Einstein fellowship from the Freie Universität Berlin and an NSF grant awarded to the second author (0613227).

are less likely to generalize abstract argument structure constructions when exposed to the SAME systematic input as older learners.

An ongoing debate in the literature on children's syntactic competence concerns the abstractness of their early constructional representations. Researchers in this area have traditionally focused on the behavior of children younger than about three years old as they process constructions from their ambient language, with the point of contention being whether their representations for these constructions are couched in terms of the same abstract semantic and syntactic categories that adults' are (e.g. agent, patient, subject and object), or whether their representations are less abstract, or item-based.

A number of studies – using mostly production methodologies – have demonstrated that early representations are item-based, but that they become increasingly abstract over time (Braine, 1976; Baker, 1979; Bates & MacWhinney, 1982; MacWhinney, 1982; Schlesinger, 1982; Tomasello, 1992; Ingram & Thompson, 1996; Akhtar & Tomasello, 1997; Lieven, Pine & Baldwin, 1997; Brooks & Tomasello, 1999; Tomasello, 2000; Savage, Lieven, Theakston & Tomasello, 2003; Tomasello, 2003; Dittmar, Abbot-Smith, Lieven & Tomasello, 2008). For example, if two-year-olds learn a novel verb by hearing it used in an intransitive construction and are then given opportunities to use it transitively, they are highly likely to resist creatively causativizing it, and will instead conservatively and exclusively use it only in the exposure construction. As children get older, however, they are increasingly more likely to do what adults do: creatively turning *It pilked* into *He pilked it*. This sort of phenomenon has been taken to suggest that early learners' linguistic knowledge is represented only in lexically specific terms.

At the same time, however, recent work using preferential looking methodologies has suggested that two-year-olds may have some abstract constructional knowledge. Given, for example, an exemplar of the English transitive construction with a novel verb, participants are able to match it to the correct meaning in a two-alternative forced choice task (Fernandes, Marcus, Di Nubila & Vouloumanos, 2006; Gertner, Fisher & Eisengart, 2006). This has been argued to show that early learners do in fact have abstract syntactic knowledge, since above chance performance in these studies was not possible given only item-specific representations.

The comprehension results merit a somewhat more nuanced interpretation though. In Gertner *et al.* (2006), test trials were immediately preceded by practice trials that were designed to familiarize children with the task. All practice trials made use of transitive utterances – the same construction type examined at test – so it is possible that good performance at test was due to scaffolding provided during the practice. This interpretation is suggested by Dittmar *et al.* (2008), who replicated the Gertner *et al.* result, but only

when a practice session was included in the procedure. Without preceding practice, two-year-olds showed no evidence of sensitivity to a general transitive construction. This indicates that while it may be possible for young children to act in a manner consistent with having acquired abstract representations, generalization appears to be quite tentative, which again suggests a less than fully abstract and robust understanding of the transitive construction.

More generally, the overall pattern of results from production and comprehension studies combined seems to be that evidence for early abstract knowledge is easier to find in comprehension, and that evidence for early lexically based knowledge is easier to find in production. This generalization is only partially accurate, however, since act-out comprehension tasks pattern like production tasks (Akhtar & Tomasello, 1997). Thus, early learners' syntactic representations are functionally item-based whenever the task at hand requires them to PRODUCE SOMETHING – be it utterances or actions. Moreover, it is arguably the case that item-based performance on ANY task suggests constructional representations that are less than fully abstract. Adult syntactic knowledge is characterized by even generalization across tasks, so whenever children show uneven performance – as in production versus comprehension tasks – this emphasizes potentially important differences between child and adult competence.

Regardless of the extent to which children's early syntactic representations are lexically based or abstract, an issue that has not been addressed in detail by either camp is how constructional generalization occurs at all. One proposal, known as the CRITICAL MASS HYPOTHESIS (Marchman & Bates, 1994; Tomasello, 2000), suggests that better generalization by older children and adults may simply be due to the fact that they have had more exposure to the requisite data than younger children. This assumes a strong relationship between type frequency and generalization such that as learners experience an increasing number of lexical items in a constructional slot – different causative verbs, for example, in the transitive construction – they have increasingly strong evidence that the slot can be populated by causative verbs that did not co-occur with the transitive construction in the input.

Since two-year-olds have witnessed massive numbers of transitive utterances, it is not surprising that they may have formed at least a tentative generalization of the transitive construction (Goldberg, 2006; Abbot-Smith, Lieven & Tomasello, 2008). Indeed, one might wonder why their representation is still fragile after so much input. This raises the possibility that, in addition to the sheer amount of input experienced, other factors might play a pivotal role in constructional generalization. A prediction of this view is that young children may be more tentative with their generalizations than older children and adults, even when exposed to the

SAME INPUT. In order to investigate this possibility, it is necessary to do something that is not possible when studying the acquisition of constructions from the ambient language – unconfound learners' age with the amount of exposure they have to a construction.

In the present study this was achieved in the context of a novel construction learning paradigm, which allows one to hold the input constant across different groups of learners. Given this design, if older learners generalize better than younger learners, then their advantage has to be due to factors other than the input. Additionally, if younger learners fail to generalize, then this would provide additional evidence in favor of early item-based constructional representations.

Our method revolved around giving learners limited exposure to a wholly novel argument structure construction, involving both a novel form AND a novel abstract meaning. Only novel verbs appeared in the construction, so knowledge of verb semantics could not be used to learn the construction's relational meaning. After exposure, participants performed a forced choice comprehension task to determine whether they were able to correctly interpret new instances of the construction, and distinguish it from a known construction type. Note that since our measure was practically identical to that used by researchers who have argued for early abstract constructional representations (e.g. Fernandes *et al.*, 2006; Gertner *et al.*, 2006), it should, if anything, favor finding generalization in young children.

In previous work, use of the novel construction learning paradigm has provided evidence that older children ($M=6;4$) and adults readily generalize on the basis of remarkably little data (Goldberg, Casenhiser & Sethuraman, 2004; Casenhiser & Goldberg, 2005). In fact, the learning by adults is quite robust. Both SOV and OSV orders can be learned and distinguished from one another. Moreover, undergraduates are willing to use novel constructions in production tasks, with evidence of retention over a seven-day delay (Boyd, Gottschalk & Goldberg, 2009). Additionally, an array of different control conditions have been employed to rule out spurious explanations for above chance performance at test (Goldberg, Casenhiser & White, 2007). These have established that participants' ability to accurately map new exemplars of a novel construction to its correct meaning is due to learning that occurs during exposure.

Overall, previous results from the novel construction learning paradigm indicate that for learners above the age of six, minimal exposure to a novel construction is sufficient for the formation of generalizations that go well beyond the specific exemplars encountered in the input. These older learners readily form abstract constructional representations. What is unclear, however, is whether younger children do the same as easily. As a means of testing for item-specific as well as general constructional representations, the present experiment manipulated the novelty of

the items that participants were tested on. Based on previous results (Casenhiser & Goldberg, 2005; Boyd *et al.*, 2009), we predict that adults and older children will perform at high levels, even when test items are high in novelty. Younger children, on the other hand, may show decrements in performance as novelty increases.

METHOD

Participants

Eighteen five-year-olds (age range 4;6–5;9; $M = 5;2$), 18 seven-year-olds (age range 6;9–8;1; $M = 7;6$) and 18 undergraduates (age range 18;5–43;0; $M = 22;4$) took part in the experiment. An additional 8 five-year-olds and 1 seven-year-old were tested, but their data were not included for reasons explained in the results section. Five-year-olds were recruited from summer recreational programs in the Princeton area, and seven-year-olds were recruited from local after-school programs. Each child was given a children's book in exchange for participation. Adults were either recruited through the use of online advertisements aimed at Princeton students, in which case they received a \$12 payment for participation, or through the Princeton psychology subject pool, in which case they received course credit.

Novel construction

We created a novel construction that describes APPROACH events in which one person moves towards another. The construction has the form NP₁NP₂V, where NP₁ denotes the individual who is moving (the AGENT), NP₂ denotes the individual who is being moved towards (the GOAL), and V is a novel verb that can be construed to encode a manner of motion. For example, *The doctor the construction worker feigos* denotes an event in which a doctor crawls towards a construction worker.

All instantiations of the construction contained two definite English NPs (e.g. *the doctor* and *the construction worker*) followed by a verb with an *-o* suffix, and a present or past tense marker (e.g. *feigos*, *feigoed*). The presence of determiners and inflectional markers served to make NPs and verbs simpler to identify. Our main interest was not in how difficult or easy it was to recognize constructional constituents as nouns or verbs, but in whether participants would be able to learn the form and function of the clausal construction as a whole.

Exposure trials

Participants were familiarized with sixteen different exemplars of the construction during an approximately three-minute EXPOSURE block. Since a



“The doctor the construction worker vakos.”

“The doctor the construction worker vakoed.”

Fig. 1. An example exposure trial. The above still frames represent time slices in a short movie that associates the NP₁NP₂V form with approach semantics.

number of studies have suggested that children have an easier time initially learning an abstract category when they are given input samples with less variability (Goldberg *et al.*, 2004; Casasola, 2005; Casenhiser & Goldberg, 2005; Maguire, Hirsh-Pasek, Golinkoff & Brandone, 2008), we sought to reduce variability during exposure in two ways. First, all of the exposure sentences featured just two NPs – *the doctor* and *the construction worker*. These were balanced so that each one occurred as NP₁ half of the time, and as NP₂ the other half. Second, the frequency distribution of the five verbs that appeared in the exposure sentences was skewed to favor a single verb, since this has been found to enhance initial generalization of novel constructions (Goldberg *et al.*, 2004; Casenhiser & Goldberg, 2005). Eight of the sixteen exposure sentences featured the verb *moopo*, while the remaining eight were evenly divided among four other verbs: *keybo*, *feigo*, *suuto* and *vako*.

The sixteen exposure sentences were embedded in short, 10-second movies that depicted their meanings. Each movie showed an agent approaching a goal, and used a voiceover that featured an exposure sentence with a present tense verb, followed by the same sentence with a past tense verb. For example, if a movie showed a doctor hopping towards a construction worker (see Figure 1), then the sentence *The doctor the construction worker vakos* would be heard simultaneously with the hopping motion. After the hopping had ended, the sentence *The doctor the construction worker vakoed* would play.

Test trials

In order to determine whether participants had learned anything during exposure, we administered a forced-choice comprehension task involving six APPROACH trials, six INTRANSITIVE trials and four LINKING trials. All trials



Fig. 2. Movies depicting approach and intransitive events were combined with different voiceovers to create either approach trials or intransitive trials. For instance, in an approach trial using the two movies shown above, participants would hear *The princess the basketball player pookos*, with the target movie to the right. In an intransitive trial they would hear *The basketball player and the princess are zorping*, with the target movie to the left.

required participants to listen to a voiceover sentence, then pick which of two simultaneously played movies depicted its meaning. The characteristics of the different trial types are described in more detail below.

Approach and intransitive trials. Approach and intransitive trials were designed to determine whether participants recognized that the NP_1NP_2V construction had a general APPROACH meaning that was disjoint from the meaning of a more common construction type – the intransitive. The movie pairs that made up approach and intransitive trials depicted the same two types of events: one movie showed an agent character approaching a goal character (as in exposure trials), while the other showed the same two characters performing a repetitive intransitive motion in synchrony, e.g. clapping. These movie pairs were combined with different kinds of voiceover sentences to produce either approach or intransitive trials. Approach trials featured NP_1NP_2V voiceovers that described the approach movie in a movie pair. The correct answer on an approach trial was therefore a point to the approach movie. Intransitive trials, on the other hand, featured intransitive sentences with novel verbs that described the intransitive movie in a movie pair. This means that the correct answer on an intransitive trial was a point to the movie showing synchronous repetitive intransitive motion. Figure 2 gives an example of the way in which different types of voiceover sentences were combined with movies to create either approach trials or intransitive trials.

Novelty. In addition to testing whether participants could identify approach scenes if and only if they heard instances of the NP_1NP_2V construction, we also wanted to assess the effects of novelty on participants' ability to correctly interpret the construction. To this end, we manipulated the amount of novelty that was present in approach trials.

Three levels of novelty were used – NO NOVELTY, NOVEL VERB and ALL NOVEL. For trials with no novelty, the target movie and its associated



“The construction worker the doctor gippos.”

Fig. 3. A linking trial. Because both movies show approach actions, above-chance performance is only possible if linking rules have been learned.

voiceover sentence were simply repeated from exposure. For trials with a novel verb, the target movie showed the same characters from the exposure block – i.e. the construction worker and the doctor – enacting a novel manner of approach, which was labeled in the voiceover with a new novel verb form. In trials that were all novel, the target movie involved characters that had not been seen during exposure enacting a novel manner of approach. These trials were accompanied by voiceovers in which none of the constituents in the NP_1NP_2V exemplar had ever been heard during the course of the experiment. We hypothesized that increased novelty would generally be associated with worse performance, such that participants would do best on trials in the no novelty condition, and would show increasing decrements to performance in the novel verb and all novel conditions.

Overall, participants were tested on six approach trials and six intransitive trials. The six approach trials were further subdivided according to novelty, with two no novelty trials, two novel verb trials and two all novel trials.

Linking trials. Good performance is possible on approach trials given only a general understanding of the NP_1NP_2V construction. For example, a participant who knows only that two initial NPs are associated with the meaning APPROACH could correctly pick out approach movies. But do participants know more specifically that the novel construction’s first NP encodes the agent, and that its second NP encodes the goal? In order to answer this question, participants were tested on four LINKING TRIALS immediately following the twelve approach and intransitive trials. In linking trials, participants chose between two approach movies in which the agent and goal roles were reversed. For example, if the voiceover sentence was *The construction worker the doctor gippos* (see Figure 3), then the trial featured one movie in which the construction worker approached the doctor (the target movie), and another in which the doctor approached the construction worker (a distractor).

Procedure

Participants were tested individually. A PRETEST block was followed by an EXPOSURE block and then a TEST block. The pretest block consisted of six trials that familiarized participants with the testing procedure by having them match exemplars of known construction types (involving familiar verbs) to one of two simultaneously played movies. The exposure block consisted of sixteen exposure trials, played in different random orders for each participant. The instructions for the exposure block simply asked participants to pay attention to what they would see and hear. The test block was broken down into four sub-blocks. The first three sub-blocks each contained two approach trials randomly intermixed with two intransitive trials. In sub-block one, the approach trials had no novelty; in sub-block two they featured novel verbs; in sub-block three all constituents were novel. The fourth sub-block consisted of four randomly ordered linking trials.

All test trials looped indefinitely until a response was made. Participants were instructed to watch the pair of movies as many times as necessary before picking one. Children pointed to their choice, which the experimenter then recorded with a button press. Adults pressed the button themselves. Test materials were balanced for target side, with half of the target movies appearing on the left, and half on the right. Test trials were additionally counterbalanced across participants for target movie. For the trial shown in Figure 2, for example, half of all participants viewed the two movies while hearing *The princess the basketball player pookos*, while the other half heard *The basketball player and the princess are pooking*. Use of the first voiceover sentence makes the approach movie the target; use of the second makes the intransitive movie the target. This sort of counterbalancing served to reduce the possibility that some specific combination of voiceover sentences with movie pairs might skew our results.

RESULTS

There were eight child participants who showed no response variability in the first three test sub-blocks. In particular, six five-year-olds picked the intransitive movie on every trial, and one five-year-old and one seven-year-old picked the approach movie on every trial. This sort of behavior suggests strategic responding in which children decided to consistently pick a certain type of movie without taking into account a semantic interpretation of the voiceover sentences, as instructed. These participants were consequently excluded from further analysis, along with another five-year-old whose poor performance in the pretest block suggested an inadequate understanding of the task. This left us with analyzable data from eighteen participants in each of the three groups. These data are summarized as proportional accuracy

TABLE 1. Mean accuracy (SE) on approach, intransitive and linking trials (for intransitive trials, novelty refers to the novelty associated with the distracter (approach) movie)

	5-year-old	7-year-old	Adult
No novelty			
Approach	0.83 (0.07)	0.97 (0.03)	1.00 (0.00)
Intransitive	0.81 (0.07)	0.94 (0.04)	1.00 (0.00)
Novel V			
Approach	0.69 (0.10)	0.78 (0.08)	0.97 (0.03)
Intransitive	0.75 (0.08)	0.92 (0.05)	1.00 (0.00)
All novel			
Approach	0.50 (0.10)	0.86 (0.07)	1.00 (0.00)
Intransitive	0.75 (0.08)	0.86 (0.08)	1.00 (0.00)
Linking	0.53 (0.07)	0.88 (0.05)	0.97 (0.03)

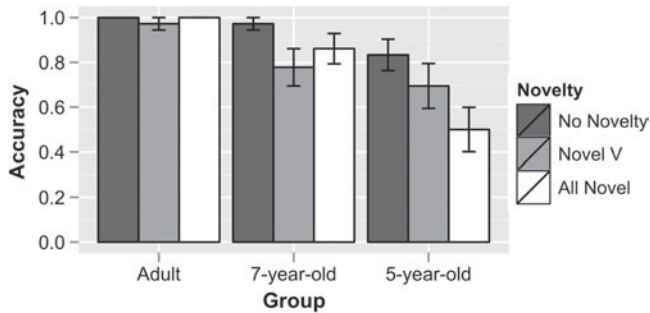


Fig. 4. Approach trial results. Both child groups were significantly less accurate than adults, but only five-year-olds showed reliable novelty effects. Error bars indicate standard error.

scores on the different trial types – approach, intransitive and linking – in Table 1. Figure 4 provides a graphical summary of performance on approach trials.

Statistical analysis of these results was carried out in the R computing language and environment (R Development Core Team, 2010) using logit mixed models (Baayen, 2008; Jaeger, 2008; Quené & van den Bergh, 2008).² For all mixed models appearing in the present work, the reported parameters

[2] Logit mixed models offer several advantages over the use of ANOVAs to analyze categorical data. Because the analysis is conducted in unbounded log-odds space rather than proportion space, logit mixed models avoid generating uninterpretable predictions (i.e., proportions less than zero or greater than one). Mixed models additionally allow for the combination of by-participants and by-items ANOVAs into a single unified analysis, and are robust in the presence of several problems known to affect ANOVA validity: missing data, non-normality, and heteroscedasticity.

TABLE 2. *Mixed model parameters for intransitive trials (parameter estimates indicate the likelihood of a correct response)*

Fixed effects				
	Estimate	SE	Z	p
Intercept	8.53	765.02	0.011	0.99
Group (5-year-olds vs. Adults)	-13.43	1530.05	-0.009	0.99
Group (7-year-olds vs. Adults)	-9.93	1530.05	-0.006	0.99
Novelty (Novel V vs. No Novelty)	-0.13	0.62	-0.21	0.83
Novelty (All Novel vs. No Novelty)	-0.67	0.60	-1.11	0.27
Random effects				
SD				
Participant (intercept)	1.79			

were arrived at via model comparison: beginning with the simplest possible model, additional parameters were added one at a time to identify the model that provided the best possible balance between fit and complexity.

Intransitive trials

We focus first on the analysis of intransitive trials. Intransitive trials are important because they can potentially help to rule out certain interpretations of the data. For instance, participants might simply have chosen appearance movies on EVERY trial, regardless of whether the voiceover sentence they heard was an exemplar of the novel construction, or an intransitive. Good performance on intransitives would, however, count against this alternative interpretation, and strengthen the argument that participants matched appearance movies to instances of the novel construction based on their understanding of the construction's meaning.

Indeed, we found that participants performed well on intransitive trials regardless of group, or the novelty of the distracter movie. Table 2 summarizes the best-fit logit mixed model of the intransitive data. The model specifies group and novelty as fixed effects, and allows for random participant intercepts. The addition of other random effects did not significantly improve model fit.

The model explicitly tests different pairwise contrasts. For example, the two group parameters listed in Table 2 represent comparisons between five-year-olds and adults, and seven-year-olds and adults. As indicated in Table 1, adults performed at ceiling on intransitive trials (i.e. 100% correct). We are thus testing to see whether there is any evidence that either of the

child groups performed significantly worse than 100%. As the very large p -values for the two group parameters in Table 2 indicate, this was not the case: there is no evidence that either child group performed significantly worse than adults on intransitive trials.

The model shows similar null effects of distracter novelty on intransitive trial performance. Recall that in these trials, target movies showing synchronous repetitive intransitive motion were paired with distracter approach movies containing different levels of novelty, and that distracter novelty increased across the three test sub-blocks. The two novelty parameters listed in Table 2 test to see whether intransitive performance differed based on distracter novelty (or alternatively, across sub-blocks). The first novelty parameter represents performance on trials with novel verb distracters relative to no novelty distracters; the second novelty parameter represents performance on trials with all novel distracters relative to no novelty distracters. As the large p -values associated with each parameter indicate, we found no significant effects of distracter novelty.

The intransitive trial findings are important for a number of reasons. First, they help to constrain the interpretation of good appearance trial performance, as noted above. Second, they suggest that even the youngest participants were fully attentive as the experiment wore on: the fact that no group experienced intransitive accuracy declines across sub-blocks suggests that fatigue did not play a significant role in performance. Finally, these data demonstrate that participants in all three groups, as expected, are fully capable of forming abstract linguistic categories. Recall that intransitive trials contained novel verbs. This means that good performance could not have been due to verb-specific knowledge. Instead, participants had to understand that the pattern *NP₁ and NP₂ are Ving* was associated with intransitive actions performed by two entities. That all groups had such knowledge should not be surprising given that even children in the five-year-old group had undoubtedly been exposed to hundreds of thousands of intransitive exemplars prior to taking part in the experiment. As discussed below, it is clear that children at age five are capable of generalizing constructions with sufficient time and input; our focus is on their facility with generalization relative to older learners.

Approach trials

While the intransitive data are useful because they restrict how the overall experiment can be interpreted, our main questions of interest concerned participant performance on trials involving the novel construction. Specifically, do the different age groups pattern equivalently? And if not, can decrements in child performance be attributed to their learning less general representations?

TABLE 3. *Mixed model parameters for approach trials (parameter estimates indicate the likelihood of a correct response)*

Fixed effects				
	Estimate	SE	Z	p
Intercept	3.47	0.60	5.82	<0.0001
Group (Children vs. Adults)	-5.33	2.17	-2.45	0.014
Group (5 vs. 7-year-olds)	-1.83	0.76	-2.42	0.016
Novelty (Novel V vs. No novelty)	-1.04	0.58	-1.80	0.072
Novelty (All novel vs. No novelty)	-1.66	0.58	-2.87	0.0041
Random effects				
SD				
Participant (intercept)	1.75			

Table 3 provides a summary of the best-fit logit mixed model of the approach trial data. The model specifies group and novelty as fixed effects, and allows for random participant intercepts. The addition of other random effects did not significantly improve model fit, nor did the addition of a group-by-novelty interaction term.

As in the intransitive trial model, the different model parameters represent specific pairwise comparisons. The first group parameter refers to a comparison between children (five-year-olds and seven-year-olds grouped together), and adults. As indicated in Table 3, this effect was significant ($p=0.014$). The sign of the parameter estimate gives the direction of the effect—that is, -5.33 signifies that children as a group were significantly LESS likely than adults to respond correctly. It was not the case, however, that both of the child groups behaved uniformly. The second group parameter estimate represents a comparison between five-year-olds and seven-year-olds, and also shows reliable differences ($p=0.016$). Specifically, the estimate of -1.83 indicates that five-year-olds were significantly less likely to respond correctly than seven-year-olds.

The model’s novelty parameters are also given in Table 3. These represent pairwise comparisons between the novel verb and no novelty conditions, and the all novel and no novelty conditions, respectively. The first novelty estimate of -1.04 shows that participants were marginally less likely to respond correctly when a trial contained a verb that had never been witnessed before ($p=0.072$). The second estimate of -1.66 indicates that participants were significantly less likely to respond correctly

on trials in which none of the vocabulary had been witnessed before ($p = 0.0041$).

Thus far, the approach trial results establish that there were indeed significant differences between groups. What is less clear, however, is whether the performance decrements noted in the child groups can be attributed to their having acquired less abstract constructional representations. The presence of a significant group-by-novelty interaction term would certainly be consistent with this interpretation – children with item-specific representations would be expected to perform well on no novelty trials, but worse on trials in which test items had less lexical overlap with exposure. But, as noted above, the addition of an interaction term did not significantly improve model fit. Should we take this to mean that the noted novelty effects were equivalent across groups?

The answer to this question is no. Figure 4 clearly indicates an interaction pattern in which novelty had no effect on adult behavior, but appears to have had increasingly large, age-related effects in children. The reason that this interaction is not a part of the mixed model is because the non-linear transformation of the model results given in Table 3 from logit space to proportions produces the interaction pattern seen in Figure 4. The model is essentially able to represent interaction patterns without the need for an explicit interaction term.

This fact, however, does not obviate the need to provide positive proof of an interaction. As a means of supplying statistical support for the interaction pattern, we performed follow-up tests on the approach trial data using a conditional inference tree (Strobl, Malley & Tutz, 2009). Conditional inference is a non-parametric method developed in machine learning circles that is especially well suited to investigating interaction patterns. It works by recursively partitioning a dataset into increasingly homogeneous subsets. The end result is a tree structure of the different subsets, along with p -values representing the reliability of each of the hypothesized partitions. Figure 5 illustrates the tree that resulted from application of the conditional inference algorithm to the approach trial data.

The tree confirms many of the results from the mixed model analysis. There are, for instance, significant effects of group. In particular, the partition that is posited at Node 1 demonstrates that five-year-olds are significantly less likely to answer correctly than seven-year-olds and adults, and the Node 5 partition shows that seven-year-olds are significantly less likely to answer correctly than adults as well. The tree also finds significant novelty effects, as evidenced by the Node 2 partition of all novel trials from no novelty and novel verb trials. Note, however, that this effect is confined to five-year-olds – there are no novelty-based partitions posited for the seven-year-old or adult data. This outcome indicates a significant interaction of group and novelty: the ostensibly general novelty effects reported in

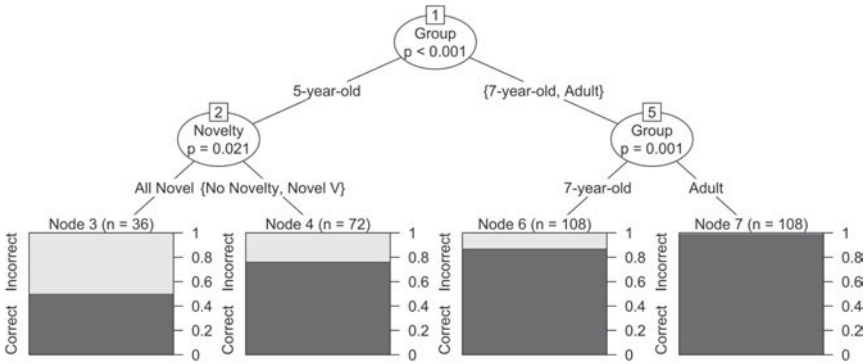


Fig. 5. A conditional inference tree based on the approach trial data. The tree shows significant group and novelty effects, as well as a group-by-novelty interaction in which the novelty effect is primarily driven by five-year-old performance.

the mixed model are actually driven primarily by behavior in the five-year-old group.³ The data thus offer positive evidence in favor of the hypothesis that decrements in five-year-old performance on approach trials are due to their having acquired less than fully abstract representations of the novel construction.

Linking trials

While the good performance of seven-year-olds and adults on approach trials indicates that they were able to learn the general association between NP₁NP₂V forms and APPROACH semantics, it does not demonstrate that they acquired the novel construction’s linking rules. In order to determine whether participants learned that NP₁ links to the construction’s agent role, and NP₂ to its goal role, we fit a logit mixed model to the linking trial data, specifying group as a fixed effect, and allowing for random participant intercepts. Other random effects proved irrelevant insofar as they failed to significantly improve model fit. Table 4 summarizes the model.

As shown in Table 1, adults exhibited near-perfect mastery of the construction’s linking rules. The two parameter estimates for group in Table 4 represent explicit tests of five-year-old and seven-year-old performance measured against the baseline provided by adults. These indicate that five-year-olds were significantly less likely than adults to respond correctly ($p < 0.0001$), but that there were no significant differences between seven-year-olds and adults ($p = 0.65$).

[3] In contrast to approach trials, a conditional inference analysis of the data from intransitive trials showed no interaction of group and novelty.

TABLE 4. *Mixed model parameters for linking trials (parameter estimates indicate the likelihood of a correct response)*

Fixed effects				
	Estimate	SE	Z	p
Intercept	2.35	0.40	5.82	<0.0001
Group (5-year-olds vs. Adults)	-4.39	0.93	-4.71	<0.0001
Group (7-year-olds vs. Adults)	0.46	1.02	0.45	0.65
Random effects				
SD				
Participant (intercept)	1.32			

DISCUSSION

The findings of the present study demonstrate that younger children – exemplified in the present experiment by participants in the five-year-old group – are less likely to generalize than older children and adults when exposed to the same input. That is, children’s failure to generalize a given construction in many previous experiments may not be due simply to a lack of exposure to enough exemplars of that construction, as suggested by the critical mass hypothesis. Instead, we hypothesize that young children fail to generalize because they are less adept at detecting patterns in the input. Item-based behavior is thus the result of younger children’s lack of facility in identifying more abstract generalizations in language, perhaps because of their tendency to be distracted from abstract structural similarity by surface dissimilarity (Gentner & Medina, 1998).

Importantly, we are not claiming that five-year-old children are incapable of generalization. Many studies of considerably younger children have documented abstract knowledge of various constructions (e.g. Pinker, 1989; Tomasello, 2000; Bencini & Valian, 2008). Five-year-olds’ performance on intransitives with novel verbs in the present experiment provides another example of relatively young children’s ability to generalize. It is not the case that five years of age represents a critical threshold for an ability to form abstractions, but rather that younger children have relatively more difficulty in forming abstractions than older children, when the input is held constant.

It is important to keep in mind that exposure to the novel construction in the present study was limited and quite narrow: all of the sixteen exposure scenes involved the same two arguments, and only five novel verbs and corresponding manners of approach were witnessed. Ultimately, in order to

fully generalize, there is clearly an advantage to variability, and we have no doubt that sufficient exposure to a wider variety of tokens would eventually lead young children to form a more abstract representation of the construction. But adults and seven-year-olds show evidence of quick and broad generalization with quite minimal and narrow input. In contrast, five-year-olds' generalization was much more limited.

Alternative interpretations

Before addressing the possible causes of item-based behavior in younger children, we seek to rule out some alternative interpretations of the current data. As noted above, one might argue that good performance on approach trials in the present experiment does not reflect actual learning of the NP₁NP₂V construction. Instead, both children and adults may have adopted a strategy in the first three test sub-blocks whereby they simply favored approach movies, without regard for the meaning of the voiceover sentences that they heard. Recall that two children – whose data were subsequently excluded from analysis – actually chose approach movies on every trial. In order to rule out the possibility that other participants executed a more nuanced version of this strategy wherein they probabilistically (instead of deterministically) favored approach movies, performance on intransitive trials is relevant. Recall that all groups were statistically at ceiling on intransitive trials, that distracter novelty had no effect on intransitive trial performance and that group and novelty did not interact (see Table 2). These findings are inconsistent with the notion that good approach trial performance was due to strategic responding, and instead suggest that participants picked movies – as instructed – based on semantic interpretation of the voiceover sentences.

Another possible strategy involves associating intransitive sentences with intransitive movies, then matching exemplars of the NP₁NP₂V construction to approach movies based on a process of elimination. If, for example, a participant is able to predict that the movie on the left in Figure 2 should be described with an intransitive structure like *The basketball player and the princess are Ving*, then they know that if they hear an NP₁NP₂V structure, it probably refers to the movie on the right. Note that this strategy, like the one outlined above, does not require any actual knowledge of the NP₁NP₂V construction. Previous work that included a no-sound control and a separate control in which only the two nouns were labeled during exposure has ruled out the idea that participants make use this strategy in general (Goldberg *et al.*, 2004; Casenhiser & Goldberg, 2005; Goldberg *et al.*, 2007), and strongly indicate that without actual exposure to a novel construction, participants cannot successfully distinguish it from a known construction that has the same number of participants.

Moreover, the current data also argue against a process-of-elimination strategy because such a strategy incorrectly predicts that young children's performance should be insensitive to the novelty manipulation. Yet our results demonstrate that five-year-olds show significant novelty effects. This outcome is straightforwardly explained by the hypothesis that five-year-olds' knowledge of the NP₁NP₂V construction is, in some respects, item-based. As test items show less lexical overlap with exposure items, they are increasingly unsure about whether to interpret them as descriptions of approach movies or descriptions of intransitive movies.

Note also that both of these strategies are additionally ruled out by seven-year-old and adult performance on linking trials. Always choosing approach movies, or matching the NP₁NP₂V construction to approach movies based on a process of elimination, incorrectly predicts at chance performance when the target and distracter movies both show approach events. Adults and older children, however, were statistically at ceiling on linking trials – an outcome that indicates they have learned the specific linking rules.

Finally, we address the fact that our test item blocks were ordered and that five-year-olds showed numerical declines in accuracy on later trials involving the novel construction. While it is normally legitimate to keep order constant in a between-participants design such as ours, it is possible that five-year-olds had less stamina than seven-year-olds and adults, and that this is responsible for their poorer performance as the experiment went on. However, as noted above, the fact that intransitive trials – which were interspersed with appearance trials in each of the first three test sub-blocks – showed no decrements over time, argues against the notion that the five-year-olds data pattern is due to fatigue. Instead, intransitive performance suggests that all participants remained attentive to the task throughout. Moreover, the entire experiment took 10–12 minutes, of which approximately 5 minutes was devoted to the entire set of test trials. This short time span with children who are kindergarten-aged also argues against an explanation based on increasing fatigue.

When our results are put into a larger context, we find that they are not in fact unexpected. The following section reviews some relevant work about young children's ability to generalize.

Previous indications that younger children tend to miss generalizations

The idea that young children's generalizations may be more tentative or partial, or may need more contextual support than adults', is supported by a great deal of work in non-linguistic category formation (Munakata, McClelland, Johnson & Siegler, 1997; Rovee-Collier, 1997; Munakata, 2001; Fisher & Sloutsky, 2005). Moreover, the finding that young children are more conservative despite having the same input is reminiscent of other

findings in research on children's memory and cognition (Brainerd & Mojaridin, 1998; Brainerd & Reyna, 2004; Brainerd, Reyna & Ceci, 2008). Older children and adults are more likely to 'fill in' gaps in their experience than younger children. In one paradigm, for instance, a list of words is generated by compiling a set of associates for a target word – e.g. the word *doctor* could be used to generate the related words *nurse*, *sick*, *hospital*, *ill*, *patient*, *cure*, *stethoscope* and *surgeon*. After exposure to the list of related words, participants were asked whether the original (non-occurring) word, *doctor*, had appeared in the list. Adults are more likely than children – quite likely, in fact – to falsely recall that they had seen *doctor* in the original list (Brainerd & Reyna, 2004; Fisher & Sloutsky, 2005). This is arguably evidence that adults are responding based on a higher-level category – whether or not the probe word is a member of the semantic field exemplified by the associates of *doctor* – whereas children are responding in an item-based fashion, based on the words that were actually present in the input.

In related work, it has been shown that children have more trouble with systematic reversal shifts than adults do (Brainerd, Reyna & Forrest, 2002). For example, if participants are exposed to instances that exemplify a pattern (e.g. all blue objects are labeled 'winners' and all red objects are labeled 'losers'), adults show an advantage when they need to reverse the pattern (making all blue objects losers, and all red objects winners), as compared with when they have to learn an entirely new rule (e.g. all square objects are winners, and all circular objects are losers). The advantage occurs because adults presumably learn abstract, color-based categories for the labels, and a reversal shift requires these categories to be relabeled, which is significantly easier than learning entirely new shape-based categories. In contrast, children show no advantage for reversal shifts (Tighe, Tighe & Schechter, 1975; Brainerd & Reyna, 2004) – a finding that suggests that they are performing the task in an item-based fashion, by associating each individual object with a label, rather than categorizing objects.

Overall then, there is a significant body of results from linguistic and non-linguistic tasks that demonstrate that children often fail to learn or use abstract categories. Moreover, the results just discussed are crucially similar to our experimental findings in that they demonstrate that children's responses have an item-based character, even when children are exposed to input that is sufficient for category formation in adults. This suggests that the differences between child and adult behavior is not solely due to differences in the amount of input that each group receives.

Children are sometimes ready generalizers

The results reported here at first blush appear to be at odds with recent work by Newport and colleagues that has suggested that in certain types of

situations, children are MORE likely to generalize than adults (Singleton & Newport, 2004, Hudson Kam & Newport, 2005, 2009). For example, a deaf child, referred to as Simon, who was exposed to obligatory motion classifiers in American Sign Language only 70% of the time by his parents, generalized the use of the classifiers to 90% of appropriate contexts, resulting in use that was indistinguishable from native signers (Singleton & Newport, 2004). Similarly, in an experimental study, hearing children who were taught an artificial language that used a determiner only 60% of the time showed a tendency to simplify the pattern either by always producing the determiner, or by omitting the determiner. Adults, on the other hand, were more likely to match the probabilities witnessed in the input, producing the determiner approximately 60% of the time (Hudson Kam & Newport, 2005).

How can children be at once more conservative than adults AND more likely to generalize than adults? There are likely several factors at play in determining whether generalization occurs. The one that is most relevant in the present context is the idea that A PATTERN MUST BE IMPLICITLY RECOGNIZED IN ORDER TO BE GENERALIZED. Clearly, learners have to detect a generalization in order to take advantage of it. Children arguably generalize in Newport and colleague's experimental studies because the pattern of determiner use was easy to identify, since the morpheme was consistent phonologically. Other studies that have investigated children's knowledge of concrete morphological patterns have also found quite early generalization (Dąbrowska & Szczerbiński, 2006; Dąbrowska & Tomasello, 2008).

The patterns required for phrasal construction learning are less easy to detect, however, in that they require children to identify commonalities based on word order (or grammatical relations), and propositional meaning. In order to recognize English argument structure constructions, for example, children have to overlook concrete differences in the various verbs and arguments involved; they need to recognize the common word order and extremely abstract function. Younger children may not recognize a fully abstract pattern as readily; without the shared commonality (in arguments, in the present study) new instances are not easily assimilated to the exemplars witnessed during exposure. Indeed, as reviewed above, studies of children's use of more abstract constructions have consistently shown a conservative bias.

There are other factors – beyond the ability to recognize the patterns involved – that likely lead children and adults to respond as they do in Newport and colleagues' work. We believe the following additional premises are required:

- (1) Learners attempt to learn the input veridically; if it contains probabilistic variation, they will attempt to produce probabilistic variation.

- (2) Unconditioned variation is harder to retain than conditioned variation (since conditioning factors serve to make variation predictable).
- (3) Producing one form (generalization) is easier than producing many forms (variation).

In order to explain why children more readily generalize in Newport *et al.*'s studies than adults do, we need to focus on what it is that participants are trying to do. Most likely, all participants are attempting to reproduce the input veridically. If the input is probabilistic, the target response is probabilistic (see Premise 1). Participants have in fact been shown to match the probabilities in the input, even when it works against their economic self-interest (Myers, 1976).

In Hudson Kam and Newport's studies, the variation in the input was unpredictable; that is, the variation was UNCONDITIONED or inconsistent: the determiner (or classifier) appeared or did not appear, without being conditioned by a differing semantic interpretation or by different accompanying nouns. Premise 2 predicts that such variation should be more difficult to retain. Thus, children's tendency to generalize (or to omit the recalcitrant determiner altogether) may have stemmed from being unable to predict when the determiner should appear. In support of this idea are some of the findings from Hudson Kam and Newport (2009). In their Experiment 1, unconditioned variation was made quite complex – one determiner appeared 60% of the time and sixteen other determiners appeared 2.5% of the time each – and adults also tended to generalize, boosting the probability of the high-frequency determiner. But when the same very complex variation was made predictable in Experiment 2 – i.e. was conditioned such that each of the seventeen determiners now appeared only with certain nouns – the variation WAS learned by adults.

Thus, it seems that when learners are unable to predict variation (because it is unconditioned), they are likely to use some default strategy (e.g. "I'm going to stick with the one determiner that I can remember" or "Just forget those crazy little words, I have no idea what they're doing"). Understandably, with their better short-term memory and meta-cognitive skills, adults have an easier time keeping track of unconditioned probabilistic variation and reproducing it veridically. But when the language gets too complex, even adults rely on a simpler generalization.

The third premise is that generalization is easier than variation, which explains why the default strategy is to generalize WHEN A PATTERN IS RECOGNIZED. This is in some sense trivial: all other things being equal, repeating a pattern is less effortful than producing several unrelated patterns. Repeated actions, whether linguistic or non-linguistic, become routinized and therefore require less effort. Evidence of this comes both from studies indicating that repetition or familiarity leads to increased

perceptual fluency (Bornstein & D'Agostino, 1994), and from studies that demonstrate structural and lexical priming and a tendency to perseverate more generally (Ramage, Bayles, Helm-Estabrooks & Cruz, 1999). The reason that learners do not ultimately generalize or simplify a language completely, using only a single construction, is explained by Premise 1 (they aim to reproduce the target language veridically) in conjunction with the fact that ultimately, of course, multiple constructions exist because communication is improved by being able to utilize various constructions with differing functions. That is, constructions collectively provide the expressive power of a language.

Extant findings are consistent with the idea that both children and adults are subject to a tendency to generalize, because it is easier, and a tendency to reproduce the patterns in the input, because they aim to be 'correct' – i.e. they want to do what others do. The present study does not pit generalization against matching of the input; indeed, the only way to match the pattern in the input on the new items was to generalize. But it is only POSSIBLE to generalize a pattern if there is some (implicit) recognition of the pattern. Older children and adults clearly have better working memory and better meta-cognitive skills, so they are more likely to implicitly or explicitly recognize a pattern when one exists. Conservatism, or initial hugging of the data, results from cases in which children fail to implicitly recognize a pattern.

CONCLUSION

We have presented clear evidence that younger children are more conservative than older learners, even when the younger learners have had just as much exposure to a novel construction as older learners. The results indicate that young children do not simply perform more poorly overall (though they do), but that they face a disproportionately greater challenge when presented with more novel items. That is, young children tend to rely quite closely on the input without taking liberties with it.

We infer that younger children's lack of broad generalization is due to the lack of implicit recognition of the pattern exemplified by the novel construction. Findings from other researchers indicate that children can be ready generalizers when the generalizations are more obvious. The present results are not unexpected when viewed in the context of the non-linguistic memory literature, which has also found that children often do not generalize as fully as older children or adults, but rather show evidence of hugging the data more closely.

Age differences always raise the issue of possible relevance to well-known sensitive period effects. Clearly we can only speculate on this point. It is possible that it is advantageous for young children to be initially

conservative. Much recent work in linguistics has observed that far from being a system of a few simple, elegant rules, linguistic competence requires detailed, partially idiosyncratic, often probabilistic knowledge of a multitude of grammatical patterns (e.g. Lakoff, 1970; Williams, 1994; Culicover, 1999; Wray, 2002; Goldberg, 2006). There are thousands of collocations, idioms and minor constructions that often buck the trends of a language in unexpected ways. A mature learner that leapt too quickly to generalizations might fail to learn the nuanced subregularities that exist.

At the same time, we cannot say on the basis of the present experiment that adults and seven-year-olds FAILED to learn the specifics of the input, because they performed well across the board. It is quite possible that seven-year-olds and adults retained representations of the familiar items, but such lexically specific representations were not required for them to successfully identify new instances of the construction.

If in fact generalization does not eclipse exemplar-based knowledge, the present results would not shed light on children's ultimate advantage in language learning vis-à-vis adults, but instead it may help illuminate adults' and older children's initial outpacing of younger children in language learning tasks (Snow & Hoefnagel-Höhle, 1978). It is clear that while language is rife with semi-idiosyncrasy, it is also replete with regularities. The ability to generalize is required for the interpretation and production of novel utterances.

The present study thus does not determine whether children's failure to generalize represents a net gain or a net loss. But we can safely conclude that their failure is not SIMPLY due to a lack of sufficient exposure to a given construction.

REFERENCES

- Abbot-Smith, K., Lieven, E. & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development* **23**, 48–66.
- Akhtar, N. & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology* **33**, 952–65.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry* **10**, 533–81.
- Bates, E. & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (eds), *Language acquisition: The state of the art*, 175–218. Cambridge: Cambridge University Press.
- Bencini, G. M. L. & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language* **59**, 97–113.
- Bornstein, R. F. & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attribution model of the mere exposure effect. *Social Cognition* **12**, 103–28.

- Boyd, J. K., Gottschalk, E. A. & Goldberg, A. E. (2009). Linking rule acquisition in novel phrasal constructions. *Language Learning* **59**, 64–89.
- Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development* **41**, 1–104.
- Brainerd, C. J. & Mojardin, A. (1998). Children's and adults' spontaneous false memories: Long-term persistence and mere testing effects. *Child Development* **69**, 1361–77.
- Brainerd, C. J. & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review* **24**, 396–439.
- Brainerd, C. J., Reyna, V. F. & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin* **134**, 343–82.
- Brainerd, C. J., Reyna, V. & Forrest, T. (2002). Are young children susceptible to the false-memory illusion? *Child Development* **73**, 1363–77.
- Brooks, P. J. & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language* **75**, 720–38.
- Casasola, M. (2005). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development* **76**, 279–90.
- Casenhiser, D. & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science* **8**, 500–508.
- Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory and language acquisition* reviewed by John R. Taylor, *Cognitive Linguistics* **10**, 251–61.
- Dąbrowska, E. & Szczerbiński, M. (2006). Polish children's productivity with case marking: The role of regularity, type frequency and phonological diversity. *Journal of Child Language* **33**, 559–97.
- Dąbrowska, E. & Tomasello, M. (2008). Rapid learning of an abstract language-specific category: Polish children's acquisition of the instrumental construction. *Journal of Child Language* **35**, 533–58.
- Dittmar, M., Abbot-Smith, K., Lieven, E. & Tomasello, M. (2008). Young German children's early syntactic competence: A preferential looking study. *Developmental Science* **11**, 575–82.
- Fernandes, K. J., Marcus, G. F., Di Nubila, J. A. & Vouloumanos, A. (2006). From semantics to syntax and back again: Argument structure in the third year of life. *Cognition* **100**, B10–B20.
- Fisher, A. V. & Sloutsky, V. M. (2005). When induction meets memory: Evidence for gradual transition from similarity-based to category-based induction. *Child Development* **76**, 583–97.
- Gentner, D. & Medina, J. (1998). Similarity and the development of rules. *Cognition* **65**, 263–97.
- Gertner, Y., Fisher, C. & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science* **17**, 684–91.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, A. E., Casenhiser, D. M. & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics* **15**, 289–316.
- Goldberg, A. E., Casenhiser, D. M. & White, T. (2007). Constructions as categories of language. *New Ideas in Psychology* **25**, 70–86.
- Hudson Kam, C. L. & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* **1**, 151–95.
- Hudson Kam, C. L. & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* **59**, 30–66.
- Ingram, D. & Thompson, W. (1996). Early syntactic acquisition in German: Evidence for the Modal Hypothesis. *Language* **72**, 97–120.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* **59**, 434–46.

- Lakoff, George. (1970). *Irregularity in syntax*. New York: Holt Rinehart and Winston.
- Lieven, E. V. M., Pine, J. M. & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language* **24**, 187–219.
- MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (ed.), *Language development, vol. 1: Syntax and semantics*. Hillsdale, NJ: Lawrence Erlbaum.
- Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M. & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science* **11**, 628–34.
- Marchman, V. A. & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language* **21**, 339–66.
- Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences* **5**, 309–315.
- Munakata, Y., McClelland, J. L., Johnson, M. H. & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review* **104**, 686–713.
- Myers, J. L. (1976). Probability learning and sequence learning. In W. K. Estes (ed.), *Handbook of learning and cognitive processes*, 171–205. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press/Bradford Books.
- Quené, H. & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* **59**, 413–25.
- R Development Core Team (2010). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ramage, A., Bayles, K., Helm-Estabrooks, N. & Cruz, R. (1999). Frequency of perseveration in normal subjects. *Brain and Language* **66**, 329–40.
- Rovee-Collier, C. (1997). Dissociations in infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review* **104**, 467–98.
- 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**, 557–67.
- Schlesinger, I. M. (1982). *Steps to language: Toward a theory of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum.
- Singleton, J. L. & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology* **49**, 370–407.
- Snow, C. E. & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development* **49**, 1114–28.
- Strobl, C., Malley, J. & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* **14**, 323–48.
- Tighe, T. J., Tighe, L. S. & Schechter, J. (1975). Memory for instances and categories in children and adults. *Journal of Experimental Child Psychology* **20**, 22–37.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* **74**, 209–53.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Williams, Edwin (1994). Remarks on lexical knowledge. *Lingua* **92**, 7–34.
- Wray, Alison (2002). *The transition to language*. New York: Oxford University Press.