The role of incremental parsing in syntactically conditioned word learning

Jeffrey Lidz
University of Maryland

Aaron Steven White
Johns Hopkins University

Rebecca Baier

In a series of three experiments, we use children's noun learning as a probe into their syntactic knowledge as well as their ability to deploy this knowledge, investigating how the predictions children make about upcoming syntactic structure change as their knowledge changes. In the first two experiments, we show that children display a developmental discontinuity in their ability to use a noun's syntactic environment as a cue to its meaning. We argue that this pattern arises from children's reliance on their knowledge of verbs' subcategorization frame frequencies to guide parsing, coupled with an inability to revise incremental parsing decisions. We show that this analysis is consistent with the syntactic distributions in child-directed speech. In the third experiment, we show that the discontinuity arises from predictions based on verbs' subcategorization frame frequencies.

Keywords: language acquisition, parsing, prediction, thematic roles

Introduction

In language acquisition, and indeed in many areas of child development, researchers often find themselves struggling with questions of competence and performance. Do children fail at some task because they lack the relevant knowledge or because that knowledge is masked behind the performance systems used to deploy that knowledge (Hamburger and Crain, 1984; Spelke and Newport, 1998)? Rarely, however, do we face the question of how children's developing performance systems constrain the generalizations that they ultimately make and how errors of interpretation feed forward for subsequent learning (Elman, 1990; Newport, 1990). In this paper, we take up this issue in the domain of syntactic development and word learning. In particular we ask how children's immature parsers lead to the assignment of erroneous grammatical structures and how such errors contribute to the acquisition of unknown words in those structures. This paper thus contributes to discussions of syntactic development, the role of syntax in word learning, and the role of parsing in syntactic development.

In understanding the interaction between parsing and learning, it is important to consider ways that parsing impacts understanding. We can consider two situations. First, the child may have acquired the grammatical rules for some construction without being able to deploy this knowledge consistently and robustly in real time (Hamburger and Crain, 1984; Huang et al., 2013; Omaki et al., 2014; Trueswell et al., 1999). Second, the child may not have acquired a given grammatical construction but nonetheless succeeds in interpreting sentences exhibiting it due to heuristics that promote understanding without relying on precise grammatical knowledge (Gagliardi et al., 2016; Yuan et al., 2012).

The case of successful acquisition of the grammatical rules in the absence of a robust deployment system can lead children to fail at accurately interpreting sentences for which they have appropriate grammatical knowledge. This could happen because the construction places high demands on component processes of understanding, such as lexical access, structure building, temporary ambiguity resolution, or retrieval from working memory, making the child's success with the construction dependent on the ease with which these subprocesses can be completed. For example, if a sentence uses low frequency words that are difficult to access from the lexicon, or if it contains a temporary ambiguity, then demands on the parser could be amplified in a way that hinders understanding, despite the child having an appropriate grammar for that construction.

In older children, there is mounting evidence that parsing dynamics shape understanding in a way that gives rise to children behaving in non-adult-like ways (Snedeker and Trueswell, 2004; Trueswell et al., 1999). For example, Trueswell et al. (1999) show that in certain discourse contexts, both adults and 5-year-old children initially interpret the first PP (*on the napkin*) in (1) as if it were the locative argument of the verb. Whereas adults can recover from this initial misinterpretation upon encountering the second PP (*in the box*), children have difficulty doing so.

(1) Put the frog on the napkin in the box.

Similar behavior has been found in at least four other domains: pronoun resolution, WH question interpretation, argument structure construction, and quantifier scope computation.

In the domain of pronoun resolution, Leddon and Lidz (2006) find that 4-year-old children only resolve reflexive pronouns to the closest syntactically licit antecedent, even in the presence of other licit antecedents. For instance, 4-year-olds resolve *herself* in (2) to Janie but not to Miss Cruella.

(2) Janie knew which picture of herself Miss Cruella put on the wall.

They argue that this bias derives from the ballistic nature of the parser, which links the reflexive pronoun to an antecedent as quickly as possible (Sturt, 2003), coupled with children's inability to revise their initial interpretive commitments (Trueswell et al., 1999). Indeed, recent eye-tracking work (Omaki, 2010) shows that adults also initially resolve *herself* in (2) to Janie, but unlike children, are able to revise that initial commitment when necessary.

In the domain of WH question interpretation, Omaki et al. (2014) find that adults and 5-year-old children prefer to associate adjunct WH words like *where* in (3) to the closest verb in terms of linear order (*say*).

(3) Where did Lizzie say that she was going to catch butterflies?

This finding is cross-linguistically robust. In Japanese, which is a head-final language, the order of *say* and *catch* is reversed. Omaki et al. found that the biases displayed by Japanese-speaking adults and 5-year-old Japanese-learning children were concomitantly flipped: both adults and children prefer to associate *where* with *catch* in Japanese.

In the domain of argument structure construction, Huang et al. (2013) find that, upon hearing a subject that is a plausible agent, 5-year-old Mandarin-learning children begin to construct an active interpretation for the sentence, and if they receive information that the sentence is actually passive, they have trouble recovering from this initial misparse. Huang and Arnold (2016) find a similar pattern in a word-learning task with 5-year-old English-learning children.

Finally, in the domain of quantifier scope computation, Musolino et al. (2000); Musolino and Lidz (2003,0) find that 5-year-old children are heavily biased towards interpreting sentences like (4) as meaning (4-a) but not (4-b).

(4) Every horse didn't jump over the fence.
  a. All of the horses failed to jump over the fence.
  b. Not every horse jumped over the fence.

Conroy (2008) and Viau et al. (2010) argue that children's bias results from the interpretation in (4-a) being the first interpretation constructed, paired with children's difficulty to revise their initial parsing commitments. Support for this view comes from several adult on-line parsing studies demonstrating that children's only interpretation corresponds to adults' initial interpretation (Conroy et al., 2008; Lidz and Conroy, 2007).

Given these and other findings showing that preschool aged children's parsers are more brittle than adults' and less able to integrate across multiple information sources (Choi and Trueswell, 2010; Omaki, 2010; Snedeker and Trueswell, 2004), it stands to reason that younger children will be at least as susceptible to failures of understanding due to parsing difficulty as older children are. Moreover, to the degree that parsing derails understanding, we expect that any process of language development that depends on information that would have been gleaned from a successful parse of a given sentence will succeed only to the degree that the parse can be accomplished (Trueswell et al., 2012).

In the domain of word learning, there are several recent findings suggesting that successful parsing of the initial part of a sentence is a prerequisite for learning words downstream. For example, Lidz et al. (2010) examine 24-month-old children's ability to learn a novel intransitive verb as a function of the character of the subject NP. They find that children are able to learn a novel verb when its subject is a pronoun (*it is blicking*) but not when it is a lexical NP (e.g., *the truck is blicking*) (see also Childers and Tomasello, 2001). They argue that this asymmetry derives from the fact that pronouns are both more frequent and have less complex semantic representations than lexical nouns, making lexical access easier for pronouns.

This hypothesis is further supported by the observation that facilitating lexical access for the full NP subject causes the asymmetry to go away. By using the lexical noun in several sentences prior to the verb learning trial, lexical access for that word is facilitated in the sentence containing the novel verb. In turn, easier lexical access of the subject NP makes verb-learning easier.

Similarly, Marchman and Fernald (2008) show that infants who are faster in interpreting familiar words in continuous speech are also more successful in learning novel words downstream, suggesting that early fluency in parsing and understanding has cascading consequences for word learning. Such findings are all the more important in light of recent evidence that socioeconomic status correlates with real-time processing ability, both at the level of lexical access (Fernald et al., 2013) and at the level of argument structure construction (Huang et al., 2016).

Parsing, syntactic inference and argument structure

Consider sentences (5) and (6).

(5) She's pushing the tiv.

(6) She's pushing with the tiv.

Even without a referential context, one can conclude that in (5) *the tiv* refers to the patient of the pushing event (i.e., the pushee) whereas in (6) *the tiv* refers to the instrument of pushing. This conclusion derives from a link between syntactic position and thematic relations (i.e., the role that an individual plays in an event): direct objects are generally interpreted as patients and the object of the preposition *with* is generally interpreted as an instrument. If we add a referential context to these sentences—for example, a scene in which a woman pushes a block with a truck—then we can use the conclusion about the thematic relation between the novel NP and the verb, to determine the referent of the phrase containing the novel noun, and hence the meaning of that word. *The tiv* refers to the block in (5) but the truck in (6).

Because these inferences depend on syntactic structure, they provide an ideal window into the development of real-time mechanisms for constructing phrase structure representations (Altmann and Kamide, 1999; Gordon and Chafetz, 1990; MacDonald et al., 1994; Snedeker and Trueswell, 2004; Trueswell et al., 1993). Moreover, because lexical differences in subcategorization frequency have been shown to play a critical role in guiding initial parsing decisions in adults and older children, probing the emergence of these effects in the earliest learners can also help us to determine the degree to which links between thematic relations and syntactic positions are acquired via a generalization over individual verbs (Dowty, 1991; Tomasello, 2000) or from more abstract principles of structure mapping (Gleitman, 1990; Pinker, 1989).

A long line of research beginning with Brown (1957, 1973) examines the role of syntax in driving inferences about word meaning. One stream of this research has shown that children can use the syntactic category of a novel word to make inferences about its meaning (see Waxman and Lidz 2006 for a review). For example, Waxman and Markow (1995) show that by 12-months of age English learning infants expect a novel word presented as a noun to refer to a category of objects, but have less specific expectations for a novel word presented as an adjective. By 14-months, the expectations for adjectives become more specific, with infants concluding that a novel adjective refers to an object property (e.g., color/texture) and not to an object category (Waxman and Booth, 2001). At least by 24-months, infants expect a novel verb to refer to a category of events and not to a category of objects (Bernal et al., 2007; Waxman et al., 2009).

Beyond inferences from syntactic category to meaning, a number of studies have shown that infants and toddlers can use the number and type of arguments in a sentence to make inferences about the meaning of a novel verb in that sentence (Gleitman, 1990; Yuan and Fisher, 2009). For example, Naigles (1990) showed that 25-month-old infants use the transitivity of a clause as a cue to whether a novel verb in that clause refers to a pair of events related by causation (for transitives) or temporal synchrony (for intransitives)—see Fisher et al. 2010 for a review.

While this kind of work compellingly demonstrates children's ability to use syntactic information to draw inferences about word meaning, the range of syntactic environments that has been examined to date is relatively narrow. In addition, the kinds of inferences that learners must make from syntactic distribution to verb meaning are somewhat indirect. The syntactic environment provides some evidence about the thematic relations to the verb and which NPs bear those thematic relations. The thematic relations provide some evidence about which event is being referred to, which in turn provides evidence about the meaning of the verb. In what follows, we expand the range of syntactic environments that trigger semantic inferences.

In addition, we focus in on the first piece of this process: what do children know about the link between the syntactic context of an NP and its interpretation. By directly exploring this link, we can understand children's knowledge of argument structure without relying on their ability to make complex inferences from argument structure to verb meaning.

Experiment 1

Experiment 1 examines how infants use a syntactic context of a noun phrase (NP) to make inferences about its thematic relation. Using a word-learning task in the intermodal preferential looking paradigm (Hirsh-Pasek and Golinkoff, 1999; Spelke, 1976), we tested children's abilities to assign a meaning to a novel noun contained in a direct object NP as compared to a prepositional object NP and a syntactically uninformative control. In adult English, the NP containing the novel word is interpreted as a patient in (7) but as an instrument in (8). In (9), there is no syntactic cue to the meaning of the novel word.

(7) She's pushing **the tiv**.

(8) She's pushing *with* **the tiv**.

(9) It's **a tiv**.

If children are able to use this thematic role information to learn the meaning of a novel noun, in (7), we expect them to be able to link *the tiv* to the object being pushed, or in (8), to the object used to do the pushing.

Their ability to do this may well depend on their knowledge of verbs specifically or words more generally. For instance, under accounts such as the Verb Island Hypothesis (Tomasello, 2000), wherein thematic roles are constructed from individual verbs, a lack of verb knowledge would make this task impossible. As such, we also collect information

Table 1
An example of a single test trial.

Phase	Length	Video	Audio
Pre-trial	2 seconds 5 seconds	Blank screen Smiling baby	*Silence* Baby giggle
Familiarization	15 seconds	Camera being wiped by a cloth	Hey, look at that! She's wiping (with) the tig! Wow, do you see her wiping (with) the tig? Yay, she's wiping (with) the tig!
Test	2 seconds 2 seconds 3 seconds	Blank screen Split screen: camera and cloth	Where's the tig? *Silence* Which one's the tig?

about children's productive verb vocabulary to use as a predictor in our analysis.

Method

Apparatus and procedure. Each infant arrived with his/her parent and was entertained by a researcher with toys while another researcher explained the experiment to the parent and obtained informed consent. The infant and parent were then escorted into a sound proof room, where the infant was either seated on the parent's lap or in a high chair, centered six feet from a 51" television, where the stimuli were presented at the infant's eye-level. If the infants were on the parents' laps, the parents wore visors to keep them from seeing what was on the screen. Each experiment lasted approximately 5 minutes, and the infants were given a break if they were too restless or started crying. In the case that the infant did not complete the experiment or were extremely fussy over the entire course, this was noted for later exclusion from the sample.

The infant was recorded during the entire experiment using a digital camcorder centered over the screen. A researcher watched the entire trial with the audio off on a monitor in an adjacent room and was able to control the camcorder's pan and zoom in order to keep the infant's face in focus throughout the trial. Videos were then coded offline frame-by-frame for direction of look by a research assistant blind to the experimental condition and without audio using the SuperCoder program (Hollich, 2005).

Design. Participants were presented with eight trials, each involving a different verb and concomitant scene. Each of these trials was separated into two phases: the familiarization phase and the test phase. These phases are described below and Table 1 gives a sample script.

Familiarization Phase. During the familiarization phase, children were shown videos of 15 second dynamic scenes involving three objects: a human hand, an instrument manipulated by the hand, and a patient causally affected via the instrument. A recorded linguistic stimulus of the form either *she's* VERB*ing the* NOVEL NOUN (*V NP*), *she's* VERB*ing with the* NOVEL NOUN (*V with NP*), or *it's a* NOVEL NOUN (*It's a NP*) was associated with each scene. Each of these pairing constitute a level in the between-subjects STRUCTURE factor. VERB and NOVEL NOUN in these frames were replaced with a known verb and a novel noun. All linguistic stimuli were recorded by the same adult female. The linguistic stimulus was presented three times as the scene progressed with different lead-in words—e.g. *Look!*.

Test Phase. A blank screen was then shown for two seconds after each scene, during which the question *where's the* NOVEL NOUN*?* was asked once. The test video began at the offset of the novel noun in the first of these questions, when a screen with separate static images of both the instrument and the patient from the previous dynamic scene was displayed. One of these images took up approximately one third both by-width and by-height of the left portion of the screen and the other took up approximately one third by-width and by-height of the right portion, with an approximately one-third by-width separation in the middle of the screen. The side on which the instrument appeared was counterbalanced and pseudorandomized such that the instrument did not show up on the same side more than twice in a row.

Two seconds after the two images were presented, the question—*which one's the* NOVEL NOUN*?*—was played. The split screen was presented for five seconds total, after which the screen went blank. After a two second blank screen, either the next learning phase started or an attention-getting phase involving a picture of an infant and laughter was presented.

For the purposes of analysis, we section the test phase into two windows based on the placement of the two questions. The first window includes the interval from the beginning of the video—concurrent with the offset of the novel word in the first test question—until the beginning of the second test question at the two-second mark (0ms-2000ms); the second window includes the interval from the offset of the novel word in the second test question until the end of the trial (~3000ms-5000ms). The first window is shifted for-

Table 2
The verbs and novel nouns used in the linguistic stimuli and the objects used in the visual stimuli for Exps. 1 and 2.

Verb	Noun	Instrument	Patient
wipe	*tig*	cloth	camera
throw	*frap*	cup	ball
hit	*tam*	ruler	cone
push	*gop*	bulldozer	block
touch	*pint*	pipe cleaner	pumpkin
wash	*pud*	sponge	toy car
tickle	*seb*	feather	mouse puppet
pull	*wug*	fishing pole	train

ward 300ms to correct for delays in infant saccades (Fernald et al., 1998), thus yielding two windows of approximately two seconds each (*Window 1* 300ms-2300ms and *Window 2* ~3000ms-5000ms). These windows are treated as fixed effects (WINDOW) in our analysis.

Materials. Eight verbs contained in the MCDI checklist were chosen with the criterion that their associated event concept must support the use of an instrument. Eight novel nouns were constructed and one associated with each verb. Table 1 gives a sample script summarizing the above description. In the *V with NP* conditions, children heard *with* during the familiarization, while those in the *V NP* conditions did not, represented in the table by the parentheses. In the *control* condition, these sentences were replaced with *it's a* NOVEL NOUN.

Table 2 shows each tuple of verb, novel noun, instrument object, and patient object. To control for possible order effects, we created two presentation orders for the trials by first building one pseudorandomized order according to the above sequencing criterion, then inverting it to create the second order. When crossed with the three linguistic structure levels (STRUCTURE: *V NP*, *V with NP*, *It's a NP*), this yielded six stimulus sets.

Participants

We recruited 48 16-month-olds (24 females) with a median age of 16;19 (mean: 16;17, range: 15;25 to 17;3) and 48 19-month-olds (24 females) with a median age of 19;19 (mean: 19;17, range: 18;29 to 20;5). 14 additional 16-month-olds and five additional 19-month-olds were tested but were excluded from the final sample for fussiness or inability to complete the experiment.

Participants were recruited from the greater College Park, MD area and were acquiring English as a native language. All participants heard English at least 80% of the time. Participants within each age group and sex were distributed evenly across the six stimulus sets.

Parents completed the MacArthur-Bates Communicative Development Inventory (MCDI) checklist (Fenson, 2007). By this index, the 16-month-olds' median productive verb vocabulary was 1 verb (mean: 3.1 verbs, IQR: 0–4 verbs), and their median productive total vocabulary was 29 words (mean: 41.2 words, IQR: 16–51.5 words); the 19-month-olds' median productive verb vocabulary was 3 verbs (mean: 9.7 verbs, IQR: 0.8–9 verbs), and their median productive total vocabulary was 59.5 words (mean: 90.7 words, IQR: 25.8–107 words). The parent of one 16-month-old participant in the *V NP* condition did not submit an MCDI checklist, and for the purposes of analysis, that participant's verb vocabulary value was set to the mean across 16-month-olds.

Results

We analyze the results of this experiment using mixed effects logistic regression with LOOKS TO INSTRUMENT as the dependent variable and random intercepts for PARTICIPANT and VERB.[1] Our predictions are that, if children both know the relation between syntactic position—in the current case, presence or absence of a preposition—and thematic relation—in the current case, instrument v. patient—and if they can furthermore deploy that knowledge, they will map the NP in *V NP* to the patient, and thus look more to the patient when they receive the *V NP* structure, and they will map the NP in *V with NP* to the instrument and thus look more to the instrument when they receive the *V with NP* structure.

Figure 1 plots the proportion of looks to instrument by STRUCTURE and AGE over the five-second trial. Grey areas show the two-second windows described above. As a reminder, the first window starts at the beginning of the trial, which corresponds to the end of the first question (*Where's the* NOVEL NOUN*?*), and ends at the beginning of the second question (*Which one's the* NOVEL NOUN*?*). It is offset 300 seconds to account for saccade delay. *Window 2* starts at the end of the second question and ends at the end of the trial.

We begin by fitting a model with fixed effects for STRUCTURE, AGE, and WINDOW as well as all possible two-way, and three-way interactions. We report these effects using a reference coding, wherein *It's a NP* is the reference level for STRUCTURE, *16 months* is the reference level for AGE, and *window 1* is the reference level for WINDOW.

We first test whether the three-way interaction between STRUCTURE, AGE, and WINDOW is warranted using a log-likelihood ratio test. This test suggests that there is a significant three-way interaction between these three variables ($\chi^2(2)$ =42.51, $p < 0.001$). Therefore, we retain the model with all interaction terms. Table 3 shows the fixed effects estimates for this mixed effects model, standard errors for those estimates, and the estimate of the variance due to participants and items.

[1]Note that the structure of the experiment does not support estimation of by-participant random slopes (Barr et al., 2013), since each verb was presented to each participant only once.

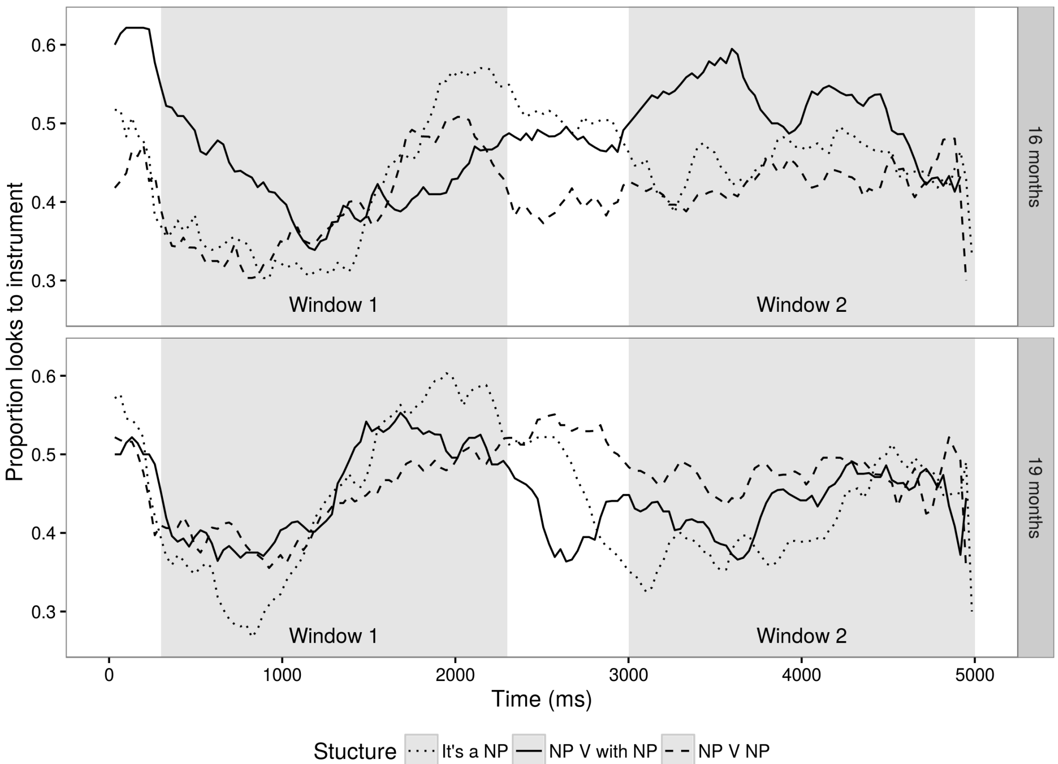

Figure 1. Mean proportion looks to instrument by STRUCTURE and AGE plotted over the five second trial. Grey areas show two second windows. *Window 1* starts at the beginning of the trial (the end of the first question) and ends at the beginning of the second question. It is offset 300 seconds to account for saccade delay. *Window 2* starts at the end of the second question.

We see five significant fixed effects here. The first is the intercept term, which is reliably negative. Given the reference coding described above, this means that 16-month-olds in the control condition prefer to look at the patient in the first window. The same preference, which we henceforth refer to as the *patient bias*, appears to manifest itself in the 19-month-olds as well. This can be seen in Table 3, first, by the fact that the simple effect of AGE is much smaller than the intercept term estimate and, second, by the fact that the 19 months olds in the control condition show a similar looking pattern in the first window, with an initial look to the patient, followed by a look to the instrument. This pattern of first looking to the patient and then to the instrument may suggest that, in the first part of a trial both 16-month-olds and 19-month-olds are getting their bearing on where particular objects are in the scene.[2]

The second significant effect we see is a positive effect of WINDOW. This means that 16-month-olds in the control condition look more to the instrument in the second window than in the first. In contrast, 19-month-olds in the control condition tend to look more to the patient than the 16-month-olds, corroborated by the third significant effect: a negative interaction between AGE and WINDOW. This can be seen in Figure 1 by the fact that 19-month-olds in the control condition go back to looking at the patient after having looked to the instrument in the first window, where switches back to the patient for 16-month-olds in the control condition are more attenuated.

The fourth significant effect we see is a positive interaction between the *V with NP* structure and WINDOW. This means that 16-month-olds in the *V with NP* condition switch more to the instrument after the first window, over and above the switching behavior we see for 16-month-olds in the control condition. This is our first piece of evidence that infants of any age can do the task of mapping an NP marked by *with* to the instrument. We do not find a reliable interaction between the *V NP* structure and WINDOW, though the estimate

[2]If we set the reference level for AGE to *19 months* (instead of *16 months*), this first window patient bias is not reliable for 19-month-olds, though it remains reliable for 16-month-olds. Referencing Figure 1, this appears to be due to the fact that 19 months olds switch to the instrument earlier in the trial than the 16-month-olds, meaning they look longer to the instrument during the first window than in the second. The important point to take away here is that both 16-month-olds and 19-month-olds show the same behavior of first looking to the patient, then to the instrument.

Table 3
Fixed effects and random effects estimates for mixed effects logistic regression of LOOKS TO INSTRUMENT *in Experiment 1*

Term	Estimate	Std. Err.	
Fixed effects			
(Intercept)	−0.39	(0.14)	**
STRUCTURE:*V with NP*	0.07	(0.10)	
STRUCTURE:*V NP*	−0.05	(0.10)	
AGE:*19 months*	0.16	(0.10)	
WINDOW:*window 2*	0.19	(0.04)	***
AGE:*19 months* × WINDOW:*window 2*	−0.33	(0.05)	***
STRUCTURE:*V with NP* × AGE:*19 months*	−0.03	(0.14)	
STRUCTURE:*V NP* × AGE:*19 months*	0.01	(0.14)	
STRUCTURE:*V with NP* × WINDOW:*window 2*	0.19	(0.05)	***
STRUCTURE:*V NP* × WINDOW:*window 2*	−0.05	(0.05)	
STRUCTURE:*V with NP* × AGE:*19 months* × WINDOW:*window 2*	−0.13	(0.07)	
STRUCTURE:*V NP* × AGE:*19 months* × WINDOW:*window 2*	0.33	(0.07)	***
Random effects			
PARTICIPANT	0.07		
ITEM	0.11		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

for this interaction goes in the right direction (negative). This is not particularly worrying, since infants in the control condition show a patient bias, and thus it is perhaps not surprising that infants in the *V NP* condition do not differ from the *control* condition in this respect.

The final significant effect we see is a reliable positive interaction between the *V NP* structure, WINDOW, and AGE. This effect is quite surprising, since it suggests that 19-month-olds in the *V NP* condition are switching more to the instrument than we would expect from the behavior of the 16-month-olds and the 19-month-olds in the other conditions. This can be seen in Figure 1 by the fact that the condition with the most looking to the instrument, among 19 months olds, is the *V NP* condition. This is doubly surprising given the fact that, though it is not reliable, the interaction between the *V NP* structure, WINDOW, and AGE is negative. This suggests that the switching behavior we see with the 16-month-olds is attenuated in the 19-month-olds. Indeed, if we set the reference level for AGE to *19 months* (instead of *16 months*), the interaction between the *V NP* structure and WINDOW no longer remains (though it is positive). This suggests that, unlike 16-month-olds, 19-month-olds do not appear to be reliably mapping NPs marked by *with* to the instrument and unmarked NPs to the patient.

Why should 19-month-olds fail where 16-month-olds succeed? Setting aside the 16-month-olds' behavior for the moment, one potential reason we do not see 19 months olds reliably constructing the correct mappings could be that infants have to know at least some verbs to be able to even start using the syntax to infer which thematic relation is assigned to an NP (which is in turn a prerequisite for using the thematic relation to infer the noun's meaning). Then, the failure of infants that don't know enough verbs might be obscuring the success of infants that do know enough verbs. This would make sense under, e.g., Tomasello's (2000) Verb Island Hypothesis, since under that hypothesis all thematic information is verb-specific (at least at this stage in development), and so children should not be able to construct mappings from syntactic structure to thematic structure for verbs they do not know. It may also be consistent with our data, since as noted above, the interaction between the *V NP* structure and WINDOW, though not reliable, goes in the right direction. Of course, this hypothesis does not address why 16-month-olds as a whole appear to be succeeding, but if we find that, among 16-month-olds, those who succeed are those with larger verb vocabularies, we may have at least preliminary evidence for this verb-based hypothesis.

To test the possibility that the behavior of infants with little verb knowledge is obscuring the success of those infants with greater verb knowledge, we ask whether there is a correlation between verb knowledge and success at our task. We again use a mixed effects logistic regression with LOOKS TO INSTRUMENT as the dependent variable and random intercepts for PARTICIPANT and VERB, but instead of including fixed effects for STRUCTURE, AGE, and WINDOW, we replace AGE with LOG VERB VOCAB, investigating its interaction with STRUCTURE and WINDOW. (We take the log of VERB VOCAB since this variable has heavy right skew.) The

Table 4
Fixed effects and random effects estimates for mixed effects logistic regression of LOOKS TO INSTRUMENT *in Experiment 1*

Term	Estimate	Std. Err.	
Fixed effects			
(Intercept)	−0.31	(0.14)	
LOG VERB VOCAB	0.00	(0.04)	
STRUCTURE:*V with NP*	−0.08	(0.11)	
STRUCTURE:*V NP*	0.01	(0.11)	
WINDOW:*window 2*	0.04	(0.04)	
LOG VERB VOCAB × STRUCTURE:*V with NP*	0.11	(0.06)	
LOG VERB VOCAB × STRUCTURE:*V NP*	−0.04	(0.06)	
LOG VERB VOCAB × WINDOW:*window 2*	−0.01	(0.02)	
STRUCTURE:*V with NP* × WINDOW:*window 2*	0.38	(0.05)	***
STRUCTURE:*V NP* × WINDOW:*window 2*	−0.05	(0.05)	
LOG VERB VOCAB × STRUCTURE:*V with NP* × WINDOW:*window 2*	−0.19	(0.03)	***
LOG VERB VOCAB × STRUCTURE:*V NP* × WINDOW:*window 2*	0.13	(0.03)	***
Random effects			
PARTICIPANT	0.07		
ITEM	0.11		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

idea here is that, if more verb knowledge helps infants map NPs to the correct referent in this task, we should see verb knowledge predicting better performance.

As with the AGE-based models we fit a model with all simple effects of LOG VERB VOCAB, STRUCTURE, and WINDOW as well as all two-way and three-way interactions. We then compare this model to one without the three-way interactions but with all simple effects and two-way interactions. A log-likelihood ratio test suggests that the three-way interactions are significant ($\chi^2(2)$ =97.06, $p < 0.001$), and so we retain them for our analysis. Table 4 shows the fixed effects estimates for this mixed effects model, the standard errors for those estimates, and the estimate of the variance due to participants and items. Significance annotations are based on Bonferroni-corrected p-values. The only effect whose *p*-value prior to corrections is below 0.05 but which is not reliable after corrections is the intercept term.

What we see here is that, *contra* the hypothesis that more verb knowledge improves infants' ability to learn a noun based on the syntactic structure, more verb knowledge appears to be a hindrance. This can be seen also in Figure 2. This figure plots the difference in proportion looks to instrument (window 2 - window 1) by VERB VOCAB. (We plot a difference score here, since it more directly visualizes what the interactions in Tables 3 and 4 correspond to.) We see here that only infants with a very small verb vocabulary map NPs marked by *with* to an instrument and unmarked NPs to a patient. This ability rapidly drops off as verb knowledge grows.

This finding is quite surprising, since it is unclear why lower verb knowledge should make infants better able to learn novel nouns. But insofar as 16-month-olds know fewer verbs, it is at least consonant with the finding that 16-month-olds but not 19-month-olds are able to do this task. That is, if verb vocabulary is the real predictor of ability to correctly map NPs in our task to the correct referent—at least in this age range—then it makes sense that 16-month-olds, who tend to know fewer verbs, would do better on this task in aggregate. In other work, we demonstrate that 28-month-olds successfully map the NP to the correct referent based on syntactic position.

Discussion

The data in Experiment 1 support a developmental discontinuity on this task, with the syntactic frame influencing performance only for infants who know few verbs but not for those with greater verb knowledge. The infants who are influenced by the syntactic context may be computing an inference by which the syntactic position of an NP containing a novel noun determines its thematic relation, which in turn allows the learner to identify the referent of the novel NP and hence to determine the meaning of the novel noun. The decline in performance raises two questions. First, how do infants with a limited verb vocabulary succeed at this task? Are they succeeding because they know the meaning of *with* or because of a parsing heuristic that leads to correct behavior despite a lack of knowledge? Second, what is responsible

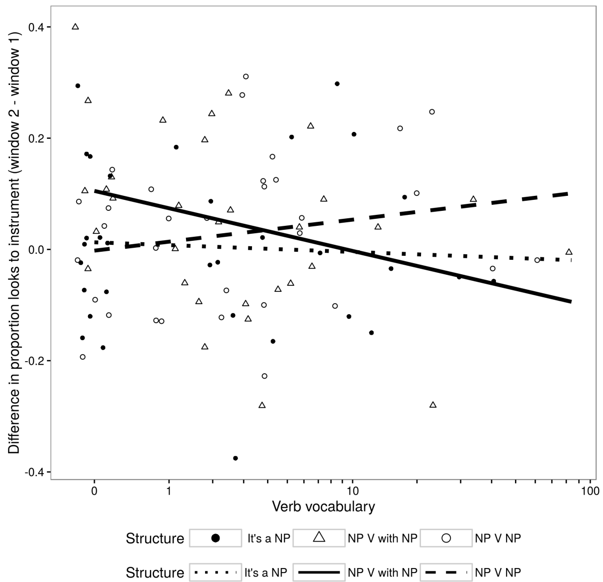

Figure 2. Difference in proportion looks to instrument (*window 2 - window 1*) plotted against VERB VOCAB.

for the change in behavior that appears to be associated with the onset of a verb vocabulary? We address these questions in turn.

Experiment 2

Experiment 2 examines two groups of 16-month-olds in order to identify the source of their early success in Experiment 1. We consider two hypotheses. First, these children may be exhibiting a heuristic whereby any NP that is directly adjacent to the verb is interpreted as a patient and any NP that is not directly adjacent to the verb is interpreted as bearing some other thematic relation. Given the materials in Experiment 1, the only possible other role would be instrument. Initial success in this task would therefore be explained not by children representing a link between the object of *with* and the instrumental relation, but rather by a link between being a non-direct object and being a non-patient. Alternatively, these children may be succeeding because they understand the link between syntactic position and thematic relation. In particular, they know that direct objects are patients and that objects of *with* are instruments.

In order to tease these possibilities apart, we tested 16-month-olds in 2 conditions:

(10) She's pushing *on* **the tiv**.

(11) She's pushing *gub* **the tiv**.

These conditions have the following properties: (10) uses a different preposition that assigns a patient-like thematic role to its complement and (11) uses a novel preposition.

If the success of 16-month-olds in Experiment 1 derives from children not knowing the meaning of the prepositions, but using a parsing heuristic whereby any NP that is not adjacent to the verb is interpreted as a non-patient, then we would expect a similar pattern of behavior in Experiment 2 as in Experiment 1. In both conditions, the NP is not adjacent to the verb and so in both conditions we would expect the novel NP to be interpreted as an instrument.

However, if the early success derived from the children having knowledge of the content of the preposition *with*, then we would expect a different pattern of results here. If these children already know the content of the prepositions, then we expect them to look more at the patient in the *on* condition. Moreover, because there is no meaning associated with the novel preposition, children should not know what thematic relation to assign to its object and so we expect chance performance in that condition.

Method

Apparatus and Procedure. The apparatus and procedure used for this experiment were identical to that used for Experiment 1.

Design. All elements of the design were identical to that found in Experiment 1 except for the linguistic stimuli used in the training phase. Instead of hearing *she's* VERB*ing the* NOVEL NOUN (*V NP* condition) or *she's* VERB*ing with the* NOVEL NOUN (*V with NP* condition), infants heard either *she's* VERB*ing on the* NOVEL NOUN (*V on NP* condition) or *she's* VERB*ing gub the* NOVEL NOUN (*V gub NP* condition).

Materials. All materials were identical to those from Experiment 1. As noted above, the only change was to the linguistic stimuli during the training phase.

Participants

32 16-month-olds (16 females) with a median age of 16;19 (mean: 16;19, range: 16;3 to 17;5) were tested on our two new preposition conditions. As in Experiment 1, the new participants were recruited from the greater College Park, MD area and were acquiring English as native language. All participants heard English at least 80% of the time. Parents completed the MCDI checklist. By this index, the median productive verb vocabulary was 1 verb (mean: 6.8 verbs, IQR: 0–4 verbs), and there median productive total vocabulary was 25 words (mean: 48.8 verbs, IQR: 11.5–50.3 verbs). Ten additional infants were tested but were excluded from the final sample for fussiness or inability to complete the experiment.

Results

As in the Experiment 1 analyses, we analyze the results of this experiment using mixed effects logistic regression with LOOKS TO INSTRUMENT as the dependent variable and random intercepts for PARTICIPANT and VERB. Figure 3 plots

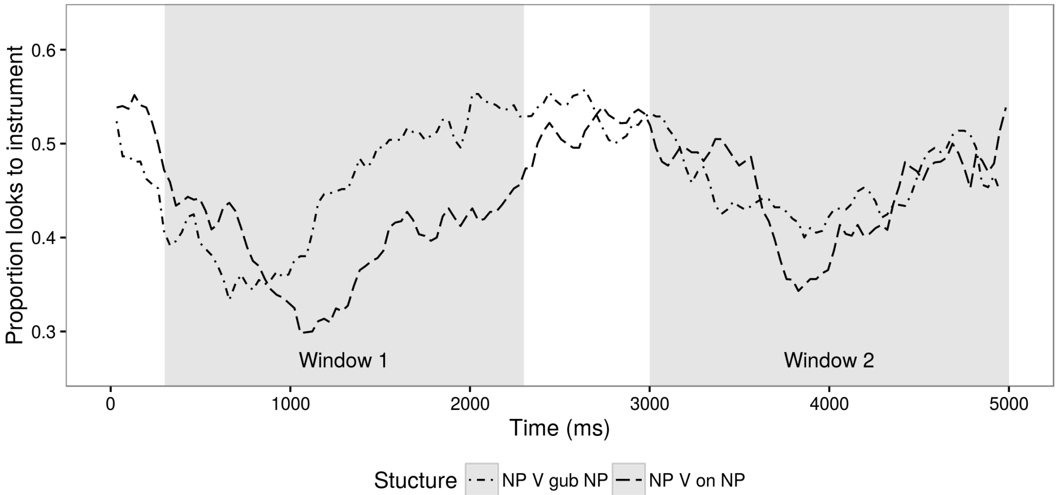

Figure 3. Mean proportion looks to instrument by STRUCTURE and AGE plotted over the five second trial. Grey areas show two second windows. *Window 1* starts at the beginning of the trial (the end of the first question) and ends at the beginning of the second question. It is offset 300 seconds to account for saccade delay. *Window 2* starts at the end of the second question.

the proportion of looks to instrument by STRUCTURE and AGE over the five-second trial. Grey areas show the two-second windows described above. As a reminder, the first window starts at the beginning of the trial, which corresponds to the end of the first question (*Where's the* NOVEL NOUN*?*), and ends at the beginning of the second question (*Which one's the* NOVEL NOUN*?*). It is offset 300 seconds to account for saccade delay. *Window 2* starts at the end of the second question and ends at the end of the trial. Note, as a sanity check, that we again see initial looks to the patient (the patient bias) followed by switches the the instrument, corroborating our interpretation of the Experiment 1 data, under which participants are getting their bearings on the scene in window 1.

We begin by fitting a model with all simple effects of LOG VERB VOCAB, STRUCTURE, and WINDOW as well as all two-way and three-way interactions (analogous to the second model from Experiment 1). We then compare this model to one without the three-way interactions but with all simple effects and two-way interactions. A log-likelihood ratio test suggests that the three-way interactions are not significant ($\chi^2(1) = 2.37$, $p > 0.1$), and so we remove them.

We then test each possible two-way interaction by removing that interaction while retaining the other two-way interactions and comparing the resulting model against one with all two-way interactions. The STRUCTURE × WINDOW interaction ($\chi^2(1) = 11.80$, Bonferroni corrected $p < 0.01$) and the STRUCTURE × LOG VERB VOCAB interaction ($\chi^2(1) = 20.68$, Bonferroni corrected $p < 0.001$), but not the LOG VERB VOCAB × WINDOW interaction ($\chi^2(1) = 4.74$, Bonferroni corrected $p > 0.1$), were significant after Bonferroni correction, and so we drop the LOG VERB VOCAB × WINDOW interaction.

Next, we test the two remaining two-way interactions by removing each in turn and comparing the resulting models against model that retains both. The STRUCTURE × WINDOW interaction ($\chi^2(1) = 13.92$, Bonferroni corrected $p < 0.001$), but not the STRUCTURE × LOG VERB VOCAB interaction ($\chi^2(1) = 5.75$, Bonferroni corrected $p > 0.05$), was significant after Bonferroni correction, and so we drop the STRUCTURE × LOG VERB VOCAB interaction.

Next, we test the simple of effect of LOG VERB VOCAB by removing it and comparing the resulting model against one that retains it. This simple effect is not significant ($\chi^2(1) = 0.01$, Bonferroni corrected $p \approx 1$), and so we remove it. Finally, we test the STRUCTURE × WINDOW interaction by comparing a model with the interaction to one with only simple effects for STRUCTURE and WINDOW. The interaction is significant as per this test, so we retain it ($\chi^2(1) = 14.10$, Bonferroni corrected $p < 0.01$).

Table 5 shows the fixed effects estimates for this mixed effects model, the standard errors for those estimates, and the estimate of the variance due to participants and items. We report these effects using a reference coding, wherein *V gub NP* is the reference level for STRUCTURE and *window 1* is the reference level for WINDOW. We see two significant effects here. The first is a significant negative effect of *V on NP*. This effects suggests that participants looked more to the patient in the first window when they received an NP marked by *on*, which is corroborated by the separation of the solid and dotted lines in Figure 3. This could mean that they used *on* to infer that the NP referred to the patient. Caution is warranted here, however, since in Experiment 1, we did not see similar effects of STRUCTURE in the first window, and

Table 5
Fixed effects and random effects estimates for mixed effects logistic regression of LOOKS TO INSTRUMENT

Term	Estimate	Std. Err.	
Fixed effects			
(Intercept)	−0.21	(0.14)	
STRUCTURE:*V on NP*	−0.26	(0.11)	*
WINDOW:*window 2*	0.00	(0.04)	
STRUCTURE:*V on NP* × WINDOW:*window 2*	0.20	(0.05)	***
Random effects			
PARTICIPANT	0.08		
ITEM	0.12		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

thus this effect may be spurious.

The second significant effect we see is a positive interaction between STRUCTURE and WINDOW. This means that participants switched more to the instrument from the first window to the second with the *V on NP* structure than with the *V gub NP* structure. This is somewhat surprising if those participants know that *on* marks NPs that refer to patients. One potential explanation for this might be that the effect of *V on NP* is arising earlier than the effect of *V with NP* in Experiment 1. This would explain the negative simple effect of *V on NP*, which suggested participants look more to the patient in the first window, but again, it is surprising from the standpoint of Experiment 1, where the effect of STRUCTURE only seems to arise in switching between window 1 and window 2.

Crucially, however, what we do not see is a positive simple effect of *window 2*, which would mean that participants were switching more to the instrument between window 1 and window 2 in the *V gub NP* condition, compared to the *V on NP* condition. Indeed, the fact that the estimate for the simple effect of WINDOW is near 0 may suggest that they are not changing their behavior between window 1 and window 2 much at all. And to the extent that their behavior is changing from window 1 to window 2, the looking trajectory for *V gub NP* is completely unlike that for *V with NP* from Experiment 1. This is important for our purposes, since if 16-month-olds were succeeding in Experiment 1 by linking non-direct objects to non-patients, we would expect such a similar trajectory.

Discussion

Experiment 2 suggests that children do know the content of the preposition *with*. Whereas 16-month-olds in the *V with NP* condition in Experiment 1 switched to the instrument significantly more than in the *control* condition, children in the novel preposition *V gub NP* position do not switch at all.

Experiments 1 and 2 together also provide a novel argument against the view that the links between thematic relation and syntactic position are acquired on a verb-by-verb basis and only generalized after a sizeable verb vocabulary has been acquired (Tomasello, 2000). Because the children who succeed at using syntactic context to determine the thematic relation of the NP in Experiment 1 are reported to have little to no productive verb vocabulary, it cannot be the case that the thematic relations are constructed by a process of generalizing over the distributional and interpretive properties of known verbs. This argument goes through even if the MCDI does not provide a perfect measure of children's verb vocabulary. No version of the exemplar-driven generalization theory predicts that having a larger vocabulary would be detrimental to acquiring the link between syntactic position and thematic relations.

Returning now to the main thread, if we accept the conclusion that the non-verb-knowing 16-month-olds are aware of the relation between syntactic context and thematic relations, then we must determine the source of the dip in performance associated with the onset of a productive verb lexicon. What changes in the child's grammar or parser could cause them to fail to use information that they apparently already have?

We pursue the hypothesis that the dip in performance exhibited by verb-knowing 16-month-olds and 19-month-olds derives from developmental changes in the weighting of predictive vs. bottom-up cues in parsing. As children develop a larger verb vocabulary, they begin to use their knowledge of subcategorization frequencies to anticipate syntactic structure (Altmann and Kamide, 1999; Gordon and Chafetz, 1990; MacDonald et al., 1994; Trueswell et al., 1993). When these predictions conflict with bottom-up information from the sentence itself, they have difficulty resolving this conflict and rely instead on their early commitments. However, relying on early commitments comes at the expense of building a parse that is fully consistent with the bottom-up information.

To make this hypothesis more concrete, consider again

Table 6

Subcategorization frame frequencies for verbs in Experiments 1 and 2 calculated using the Brown corpus from CHILDES (MacWhinney, 2014a,1) parsed by Pearl and Sprouse (2013)

Verb	**[$_{VP}$ _ NP]**	**[$_{VP}$ _ NP PP]**	**[$_{VP}$ _ PP]**	**[$_{VP}$ _ PP$_{\text{with}}$]**	**Total**
hit	136	53	5	0	218
pull	203	28	10	0	291
push	206	30	8	2	302
throw	196	82	9	1	310
tickle	22	0	0	0	39
touch	151	2	2	0	184
wash	144	13	3	0	186
wipe	68	13	0	0	87
Total	1126	221	37	3	1617

sentence (8) from Experiment 1.

(8) She's pushing with the tiv.

Imagine that the child has heard the subject and the verb. At this stage, if the child expects the verb to be used transitively, it is possible to predict that a direct object NP is coming and to build that structure in advance of hearing it (Altmann and Kamide, 1999; Omaki, 2010; Omaki and Lidz, 2015; Sussman and Sedivy, 2003). When the next word turns out to be a preposition and hence is inconsistent with the predicted structure, the parser must revise its initial commitment in order to successfully parse the sentence. However, because this revision is too difficult for children to execute (Trueswell et al., 1999), they treat the object of the preposition as the object of the verb and effectively ignore the preposition for the purposes of parsing and interpretation.

If this hypothesis explains the pattern of data seen in verb-knowing 16-month-olds and 19-month-olds, then we can make several predictions. First, the verbs in our study are predicted to be significantly more likely to be used transitively than intransitively with a PP. This asymmetry is a presupposition of the account based on a differential weighting of predictive vs. bottom-up cues because subcategorization frequency can function as a predictive cue only to the degree that asymmetries in subcategorization frequencies exist. Second, if we could satisfy the verb's subcategorization preference in sentences containing a preposition, then we expect sensitivity to the content of the preposition to re-emerge. Third, if 19-month-olds were given a verb for which they had no subcategorization expectations, sensitivity to the preposition should re-emerge. Fourth, children with no productive verb vocabulary should behave identically with real and novel verbs. Finally, if 19-month-olds were given substantial exposure to a novel verb in one subcategorization frame, then their sensitivity to the preposition should be a function of the degree to which the preposition is consistent with that exposure. We address the first two predictions below, and we address the second two in other work.

To test the first prediction, we examined the distribution of complement types for each of the 8 verbs used in Experiments 1 and 2 in Pearl and Sprouse's (2013) parsed version of the Brown corpus obtained from CHILDES (MacWhinney, 2014a,1). We asked what proportion of the instances of each verb occurred in a transitive clause not also containing a PP ([$_{VP}$ V _ NP]), an transitive clause also containing a PP ([$_{VP}$ V _ NP PP]), an intransitive clause containing a PP ([$_{VP}$ V _ PP]), or an intransitive clause containing a PP headed by *with* ([$_{VP}$ V _ PP$_{\text{with}}$]). The results of this search are given in Table 6.

The verbs that we used occurred on average 70% of the time in a transitive clause and .2% of the time in intransitive clauses with PPs headed by with. In addition, we also asked what proportion of all verbs in the corpus occurred in these 4 environments, finding that 33% occurred in transitive clauses with no PP, 7% occurred in transitive frames containing PPs, 13% occurred in intransitive clauses containing PPs, and 2% occurred in intransitive clauses containing PPs headed by with. These data are consistent with the hypothesis that children who fail to use the syntactic context to determine the thematic relation of the novel NP are doing so because they rely on their knowledge of subcategorization frequencies to guide their parsing decisions.

Experiment 3

Experiment 3 tests the prediction that satisfying the verb's subcategorization expectations in sentences containing a preposition would allow sensitivity to the content of the preposition to re-emerge. We hypothesized that the two groups of children in Experiment 1 who failed to use syntactic context as a cue to meaning failed to do so because they were relying more on their knowledge of the verb's likely subcategorization than on the verb's actual subcategorization in the experiment. Thus, if we could find a way to test their knowledge of the relation between syntactic context and the-

Table 7
The verbs and novel nounds used in the linguistic stimuli and the objects used in the visual stimuli in Exp. 3.

Verb	Noun	Instrument	Patient
tap	*pint*	pipe cleaner	train
brush	*seb*	brush	mouse
stop	*frap*	block	ball
hit	*tam*	ruler	cone
wipe	*tig*	cloth	camera
push	*gop*	bulldozer	block

matic relation while also putting the verb in its preferred syntactic context, then this knowledge should reemerge.

Consider (12) and (13).

(12) She's pushing that thing *with* **the tiv**.

(13) She's pushing **the tiv** with that thing.

Both of these sentences contain two referentially ambiguous expressions (*that thing*, *the tiv*). In (12), the novel word is used as the object of the preposition *with*. In (13) it is used as the direct object of the verb. But, without knowledge of the link between syntactic position and thematic relation, it would be impossible to know what the NP containing the novel word refers to. Hence, to the degree that children can use syntactic context to infer the meaning of the novel word, it follows that they represent the link between syntactic context and thematic structure. Moreover, because these clauses are all transitive, they satisfy the preferred subcategorization frame of the verb, allowing the effect of syntactic context to emerge independent of subcategorization preferences.

We use the same visual stimuli as in experiment 1, with the two audio conditions in (12) and (13). If children are able to use syntactic position as a cue to thematic relation, then they should interpret the novel word as referring to the instrument in (12) but the patient in (13).

We tested 16 19-month-olds. This helps to determine the viability of our hypothesis for their failure in Experiment 1. If the 19-month-olds in experiment 1 failed to use the preposition as a cue to meaning because they do not know the meaning of the preposition or the link between syntactic position and thematic relation, then they should be unable to identify the meaning of the novel word here. However, if they failed because they were relying on the subcategorization frequency of the verb, then they should succeed here. Because the verb occurs in its preferred syntactic environment, then if children have knowledge of the semantic contribution of the preposition, it should emerge here.

Method

Apparatus and Procedure. The apparatus and procedure used for this experiment were identical to that used for Experiment 1.

Design. All elements of the design were identical to that found in experiment 1 except for (i) the form of linguistic stimuli used in the training phase and (ii) the number of trials. Instead of hearing *she's* VERB*ing the* NOVEL NOUN (*V NP* condition) or *she's* VERB*ing with the* NOVEL NOUN (*V with NP* condition), infants heard either *she's* VERB*ing the* NOVEL NOUN *with that thing* (*V NP with that thing* condition) or *she's* VERB*ing that thing with the* NOVEL NOUN" (*V that thing with NP* condition). Instead of seeing eight trials, children saw six.

Materials. Three verbs from the original experiment (*hit*, *push*, *wipe*) were retained along with their associated dynamics scenes and novel nouns. Three new verbs from the MCDI were used (*brush*, *stop*, *tap*) along with associated novel nouns. Two new pseudorandomized orders were created. Table 7 shows each tuple of verb, novel noun, instrument object, and patient object.

Participants

32 19-month-olds (16 females) with a median age of 19;9 (range: 18;27 to 19;29) were tested on our two-argument task. As in Experiments 1 and 2, participants were recruited from the greater College Park, MD area and were acquiring English as native language. All participants heard English at least 80% of the time. Parents completed the MCDI checklist. By this index, the median productive verb vocabulary was 3.5 (range: 0 to 96). Four additional infants were tested but were excluded from the final sample for excessive fussiness.

Results

As in the Experiment 1 and Experiment 2 analyses, we analyze the results of this experiment using mixed effects logistic regression with LOOKS TO INSTRUMENT as the dependent variable and random intercepts for PARTICIPANT and VERB. Figure 4 plots the proportion of looks to instrument by STRUCTURE and AGE over the five-second trial. Grey areas show the two-second windows described above. As a reminder, the first window starts at the beginning of the trial, which corresponds to the end of the first question (*Where's the* NOVEL NOUN*?*), and ends at the beginning of the second question (*Which one's the* NOVEL NOUN*?*). It is offset 300 seconds to account for saccade delay. *Window 2* starts at the end of the second question and ends at the end of the trial. Note, as a sanity check, that we again see initial looks to the patient (the patient bias) followed by switches the the instrument, corroborating our interpretation of the Experiment

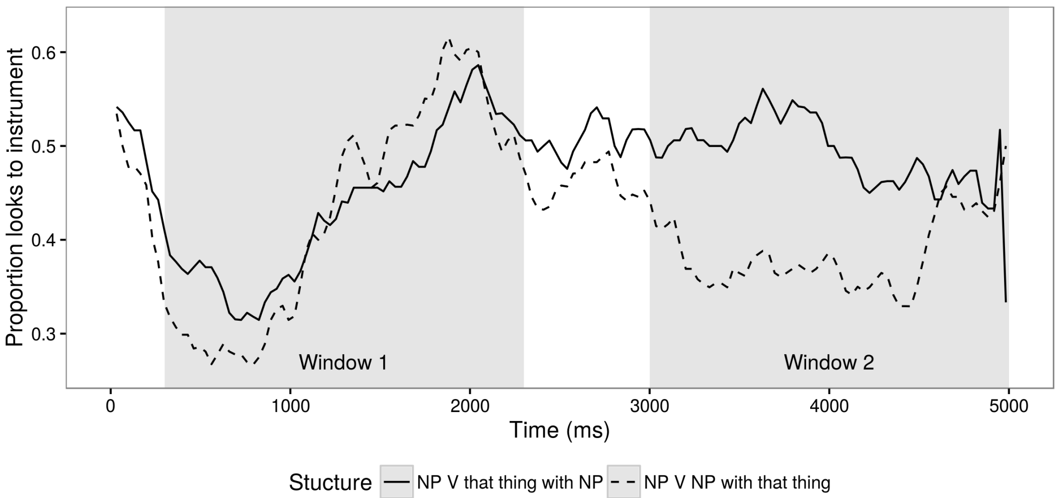

Figure 4. Mean proportion looks to instrument by STRUCTURE, plotted over the five second trial. Grey areas show two second windows. *Window 1* starts at the beginning of the trial (the end of the first question) and ends at the beginning of the second question. It is offset 300 seconds to account for saccade delay. *Window 2* starts at the end of the second question.

1 data and Experiment 2 data, under which participants are getting their bearings on the scene in window 1.

We begin by fitting a model with all simple effects of LOG VERB VOCAB, STRUCTURE, and WINDOW as well as all two-way and three-way interactions. We then compare this model to one without the three-way interactions but with all simple effects and two-way interactions. A log-likelihood ratio test suggests that the three-way interactions are not significant ($\chi^2(1) = 1.64$, $p = 0.20$), and so we remove them.

We then test each possible two-way interaction by removing that interaction while retaining the other two-way interactions and comparing the resulting model against one with all two-way interactions. The STRUCTURE × WINDOW interaction ($\chi^2(1) = 83.35$, Bonferroni corrected $p < 0.001$) and the STRUCTURE × LOG VERB VOCAB interaction ($\chi^2(1) = 63.00$, Bonferroni corrected $p < 0.001$), but not the LOG VERB VOCAB × STRUCTURE interaction ($\chi^2(1) = 0.28$, Bonferroni corrected $p > 0.5$), were significant after Bonferroni correction, and so we drop the LOG VERB VOCAB × STRUCTURE interaction.

Next, we test the two remaining two-way interactions by removing each in turn and comparing the resulting models against model that retains both. Both the STRUCTURE × WINDOW interaction ($\chi^2(1) = 83.31$, Bonferroni corrected $p < 0.001$) and the STRUCTURE × LOG VERB VOCAB interaction ($\chi^2(1) = 20.60$, Bonferroni corrected $p < 0.001$) are significant after Bonferroni correction, and so we retain the model with both interactions.

Table 8 shows the fixed effects estimates for this mixed effects model, the standard errors for those estimates, and the estimate of the variance due to participants and items. We report these effects using a reference coding, wherein *V NP with that thing* is the reference level for STRUCTURE and *window 1* is the reference level for WINDOW. We see four significant effects here. The first is a negative intercept. This suggests that, as in Experiments 1 and 2, the participants in this experiment show a patient bias in window 1.

This patient bias is attenuated to some extent by the second significant effect, which is a positive effect of LOG VERB VOCABULARY. That is, there is a positive correlation between knowing more verbs and having less of a patient bias in the first window. This attenuation is effectively neutralized by the third significant effect: the interaction between LOG VERB VOCAB and WINDOW, which is a positive interaction between STRUCTURE and WINDOW. We see this neutralization in the fact that, if we relevel WINDOW so that *window 2* is the reference level, the simple effect of LOG VERB VOCAB is not significant but the interaction between LOG VERB VOCAB and WINDOW is significantly positive. The take-away from this is that, while verb knowledge apparently predicts something about patient bias in the first window, it is likely irrelevant to the effect of interest, which occurs in the second window (or more precisely, in switching behavior from the first to second window).

The final significant effect is a positive interaction between STRUCTURE and WINDOW. This is the crucial effect that shows, regardless of verb vocabulary, all 19-month-olds in this experiment were able to correctly map an NP marked by *with* to the instrument.

Discussion

These data support the hypothesis that 19-month-old children know the content of the preposition *with* and can use it

Table 8
Fixed effects and random effects estimates for mixed effects logistic regression of LOOKS TO INSTRUMENT

Term	Estimate	Std. Err.	
Fixed effects			
(Intercept)	−0.51	(0.17)	**
LOG VERB VOCAB	0.16	(0.05)	***
STRUCTURE:*V that thing with NP*	−0.03	(0.12)	
WINDOW:*window 2*	0.01	(0.05)	
LOG VERB VOCAB × WINDOW:*window 2*	−0.19	(0.02)	***
STRUCTURE:*V that thing with NP* × WINDOW:*window 2*	0.56	(0.06)	***
Random effects			
PARTICIPANT	0.10		
ITEM	0.11		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

as a cue to the thematic relation borne by its object. Moreover, it supports the view that the 19-month-olds' failure in experiment 1 was caused not by a lack of knowledge, but by interference from the mechanics of parsing. Because these children are better able to use their knowledge of subcategorization frequency to predict upcoming structure, these predictions interfere with children's ability to display their syntactic knowledge. This finding highlights the critical nature of understanding the linguistic input as it is represented by learners. An accurate model of learning must treat the input not as it is intended, but rather as it is represented by immature learners.

General discussion

In a series of three experiments we have uncovered the following effects. First, we see a developmental discontinuity in children's ability to use the syntactic position of a noun phrase headed by a novel noun to learn the meaning of that noun. Prior to acquiring a substantial verb vocabulary, children are able to distinguish the interpretation of a novel NP when it is a direct object as compared to when it is a prepositional object. Upon acquiring a substantial verb vocabulary, children appear to rely more on their expectations about a verb's syntactic distribution than on the actual sentence it occurs in, blocking the inference from syntactic position to thematic relation, and consequently the inference from thematic relation to lexical meaning. Second, by 16-months children have the ability to use a preposition as a cue to the thematic relation of an NP, and hence as a cue to the meaning of a novel noun in that NP. At this age, children distinguish *with* from *on* semantically, and they distinguish both of these from a novel preposition. Finally, 19-month-olds' difficulty in using *with* as a cue to novel noun meaning can be attenuated by placing the *with* PP after a direct object NP. We have argued that the pattern of findings in 19-month-olds reflects the contribution of their immature parsers. When infants are able to make a prediction about upcoming syntactic structure, but that prediction turns out to be wrong, they have difficulty recovering from the misparse and hence make incorrect inferences about meaning.

This pattern of effects highlights the importance of identifying and keeping separate the contributions of syntactic knowledge and those of parsing mechanics. The immaturity of a child's parser can lead them to assign erroneous syntactic structures and to consequently make incorrect inferences about the meanings of novel words. In turn, this conclusion emphasizes the importance of separating children's linguistic input, what they are exposed to, from their linguistic intake, how they represent their input. In characterizing the role of input in shaping language development, we must take care to think of the input not in terms of how it was intended by those who produced it, but rather in terms of the information that children are able to glean from that input. The degree to which an utterance is informative for some learning inference is a function of how that utterance is represented. This representation, as we have seen, can be shaped by properties of the developing parser. An important goal for the future, therefore, is to identify the various ways that children can distort their input as a function of either their current knowledge state (Lidz and Gagliardi, 2015) or their developing information processing mechanisms (Omaki and Lidz, 2015).

Acknowledgments

We are grateful to audiences at the University of Maryland, BUCLD 34, BUCLD 35, LSA 2010, CUNY 2011, and SRCD 2013 for useful comments on this work. This work was supported by NSF BCS-1551629, and White was supported by NSF DGE IGERT-0801465.

References

Altmann, G. and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3):247–264.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3):255–278.

Bernal, S., Lidz, J., Millotte, S., and Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development*, 3(4):325–341.

Brown, R. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology*, 55(1):1.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Childers, J. B. and Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*, 37(6):739.

Choi, Y. and Trueswell, J. C. (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of experimental child psychology*, 106(1):41–61.

Conroy, A. M. (2008). *The role of verification strategies in semantic ambiguity resolution in children and adults*. PhD thesis, University of Maryland.

Conroy, A. M., Fults, S., Musolino, J., and Lidz, J. (2008). Surface Scope as a Default: The Effect of Time in Resolving Quantifier Scope Ambiguity. University of North Carolina, Chapel Hill.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3):547–619.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2):179–211.

Fenson, L. (2007). *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual*. Paul H. Brookes Publishing Company, Baltimore.

Fernald, A., Marchman, V. A., and Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental science*, 16(2):234–248.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., and McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9(3):228–231.

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

Gagliardi, A., Mease, T. M., and Lidz, J. (2016). Discontinuous development in the acquisition of filler-gap dependencies: Evidence from 15-and 20-month-olds. *Language Acquisition*, pages 1–27.

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

Gordon, P. and Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, 36(3):227–254.

Hamburger, H. and Crain, S. (1984). Acquisition of cognitive compiling. *Cognition*, 17(2):85–136.

Hirsh-Pasek, K. and Golinkoff, R. M. (1999). *The origins of grammar: Evidence from early language comprehension*. The MIT Press, Cambridge, MA.

Hollich, G. (2005). Supercoder: A program for coding preferential looking.

Huang, Y. T. and Arnold, A. R. (2016). Word learning in linguistic context: Processing and memory effects. *Cognition*, 156:71–87.

Huang, Y. T., Leech, K., and Rowe, M. L. (2016). Exploring socioeconomic differences in syntactic development through the lens of real-time processing. *Cognition*, page to appear.

Huang, Y. T., Zheng, X., Meng, X., and Snedeker, J. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of memory and language*, 69(4):589–606.

Leddon, E. and Lidz, J. (2006). Reconstruction Effects in Child Language. Cambridge. Cascadilla Press.

Lidz, J., Bunger, A., Leddon, E., Baier, R., and Waxman, S. R. (2010). When one cue is better than two: lexical vs. syntactic cues to verb learning. *Unpublished manuscript*.

Lidz, J. and Conroy, A. M. (2007). The Role of Working Memory in Explaining Isomorphism. Barcelona, Spain.

Lidz, J. and Gagliardi, A. (2015). How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics*, 1(1):333–353.

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review*, 101(4):676.

MacWhinney, B. (2014a). *The CHILDES project: Tools for analyzing talk, Volume I: Transcription format and programs*. Psychology Press.

MacWhinney, B. (2014b). *The CHILDES project: Tools for analyzing talk, Volume II: The database*. Psychology Press.

Marchman, V. A. and Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental science*, 11(3):F9–F16.

Musolino, J., Crain, S., and Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, 38(1):1–32.

Musolino, J. and Lidz, J. (2003). The scope of isomorphism:

Turning adults into children. *Language Acquisition*, 11(4):277–291.

Musolino, J. and Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics*, 44(4):817–852.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of child language*, 17(2):357–374.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, 14(1):11–28.

Omaki, A. (2010). *Commitment and flexibility in the developing parser*. PhD thesis, University of Maryland.

Omaki, A., Davidson White, I., Goro, T., Lidz, J., and Phillips, C. (2014). No fear of commitment: Children's incremental interpretation in English and Japanese wh-questions. *Language Learning and Development*, 10(3):206–233.

Omaki, A. and Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, 22(2):158–192.

Pearl, L. and Sprouse, J. (2013). Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition*, 20(1):23–68.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. MIT Press, Cambridge, MA.

Snedeker, J. and Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49(3):238–299.

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, 8(4):553–560.

Spelke, E. and Newport, E. L. (1998). Nativism, empiricism, and the development of knowledge. In Damon, W. and Lerner, R. M., editors, *Handbook of child psychology: Volume 1: Theoretical models of human development (5th ed.)*, pages 275–340. John Wiley & Sons Inc, Hoboken, NJ, US.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48(3):542–562.

Sussman, R. S. and Sedivy, J. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes*, 18(2):143–163.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3):209–253.

Trueswell, J. C., Kaufman, D., Hafri, A., and Lidz, J. (2012). Development of parsing abilities interacts with grammar learning: Evidence from Tagalog and Kannada. Cascadilla Press.

Trueswell, J. C., Sekerina, I., Hill, N. M., and Logrip, M. L. (1999). The kindergarten-path effect: Studying online sentence processing in young children. *Cognition*, 73(2):89–134.

Trueswell, J. C., Tanenhaus, M. K., and Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(3):528.

Viau, J., Lidz, J., and Musolino, J. (2010). Priming of abstract logical representations in 4-year-olds. *Language Acquisition*, 17(1-2):26–50.

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

Waxman, S. R. and Lidz, J. L. (2006). Early Word Learning. In Kuhn, D., Siegler, R. S., Damon, W., and Lerner, R. M., editors, *Handbook of child psychology: Vol 2, Cognition, perception, and language (6th ed.)*, pages 299–335. John Wiley & Sons Inc, Hoboken, NJ, US.

Waxman, S. R., Lidz, J. L., Braun, I. E., and Lavin, T. (2009). Twenty four-month-old infantsâĂŹ interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, 59(1):67–95.

Waxman, S. R. and Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive psychology*, 29(3):257–302.

Yuan, S. and Fisher, C. (2009). "Really? She Blicked the Baby?" Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. *Psychological Science*, 20(5):619–626.

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