The role of incremental parsing in syntactically conditioned word learning

Jeffrey Lidz, Aaron Steven White & Rebecca Baier

jlidz@umd.edu

aswhite@umd.edu

rbaier@umd.edu

Department of Linguistics

University of Maryland, College Park

1401 Marie Mount Hall

University of Maryland

College Park, MD 20742-7505

Abstract

In a series of five experiments we use children's noun learning as a probe into their knowledge of syntax and their ability to deploy this knowledge. We investigate 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 U-shaped pattern of development 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. In the third experiment, we show that this analysis is consistent with the syntactic distributions in child-directed speech. In the final two experiments, we show that the U-shape arises from predictions based on specific verbs' subcategorization frame frequencies and not expectations about the class of verbs as a whole.

Keywords: Syntactic Bootstrapping, Parsing, Prediction, Word learning

1 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 & Crain, 1984; Spelke & 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. 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.

The case of successful acquisition of the grammatical rules in the absence of a robust deployment system can lead to children failing to accurately interpret 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 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 & Trueswell, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999). For example, Trueswell et al. (1999) showed 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.

Similarly, Leddon & Lidz (2006) found that 4-year-old children understand the reflexive pronoun in (2) as only being able to refer to the matrix subject, Janie, whereas adults are able to treat that pronoun as bound by either the matrix or embedded subject, Miss Cruella. However, in cases where the matrix subject is unavailable as an antecedent (3), children are able to link the pronoun to the embedded subject.

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

These authors argued that children's bias for the matrix antecedent in (2) derives from the ballistic nature of the parser, which links the reflexive 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 initially treat the matrix subject as antecedent in (2), but, unlike children, are able to revise that initial commitment when necessary.

Related evidence comes from ambiguous sentences like (4), where we find that 5-year-old children are heavily biased towards interpreting sentences like (4) as meaning (4)a but not (4)b (Musolino, Crain, & Thornton, 2000; Musolino & Lidz, 2003, 2006).

(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, Lidz, & Musolino (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, Fults, Musolino, & Lidz, 2008; Lidz & 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 & Trueswell, 2010; Omaki, 2010; Snedeker & 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, Kaufman, Hafri, & Lidz, 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, Bunger, Leddon, Baier, & Waxman (2010) examined 24-month-old children's ability to learn a novel intransitive verb as a function of the character of the subject NP. We found that children were able to learn a novel verb when its subject was a pronoun (*it is blicking*) but not when it was a lexical NP (e.g., *the truck is blicking*) (see also Childers & Tomasello, 2001). We argued that this asymmetry derived 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 was further supported by the observation that facilitating lexical access for the full NP subject caused the asymmetry to go away. By using the lexical noun in several sentences prior to the verb learning trial, lexical access for that word was facilitated in the sentence containing the novel verb. In turn, easier lexical access of the subject NP made verb-learning easier.

Similarly, Marchman & Fernald (2008) showed 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.

2 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. 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 & Lidz, 2006 for a review). For example, Waxman & 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 & 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, Lidz, Millotte, & Christophe, 2007; Waxman, Lidz, Braun, & Lavin, 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 & 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, Gertner, Scott, & Yuan (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.

3 Experiments

3.1 Experiment 1

Experiment 1 examines how infants use a Noun Phrase's syntactic context to make inferences about its thematic relation. Using a word-learning task in the intermodal preferential looking paradigm (IPLP; Hirsh-Pasek & 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. 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**.

3.1.1 Participants

72 16-month-olds (36 females) with a median age of 16;19 (range: 15;24 to 17;3), 48 19-month-olds (24 females) with a median age of 19;19 (range: 18;29 to 20;5), and 48 28-month-olds (24 females) with a median age of 28;16 (range: 27;29 to 29;3) were included in the final sample. 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 MacArthur-Bates Communicative Development Inventory (CDI) checklist (Fenson, 2007). By this index, 16-month-olds' median verb production vocabulary was 1.5 (range: 0 to 29), 19-month-olds’ median verb production vocabulary was 3 (range: 0 to 83), and 28-month-olds' median verb production vocabulary was 73 (range: 0 to 164). Fourteen additional 16-month-olds, five additional 19-month-olds, and six additional 28-month-olds were tested but were excluded from the final sample for fussiness or inability to complete the experiment.

3.1.2 Method

3.1.2.1 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 infants 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). For this and subsequent experiments 25% of these videos were recoded by a second research assistant. Intercoder reliability was greater than .9 (Kendall's $\tau > .9$).

3.1.2.2 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.

3.1.2.2.1 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 VERBing the NOVEL NOUN" (*V NP* condition), "she's VERBing with the NOVEL NOUN" (*V with NP* condition), or "it's a NOVEL NOUN" (*Control* condition) was associated with each scene as a level in a between-subjects factor. All linguistic stimuli were recorded by an adult female. VERB and NOVEL NOUN in these frames were replaced with a known verb and a novel noun. The linguistic stimulus was presented three times as the scene progressed with different lead-in words---e.g. "Look!".

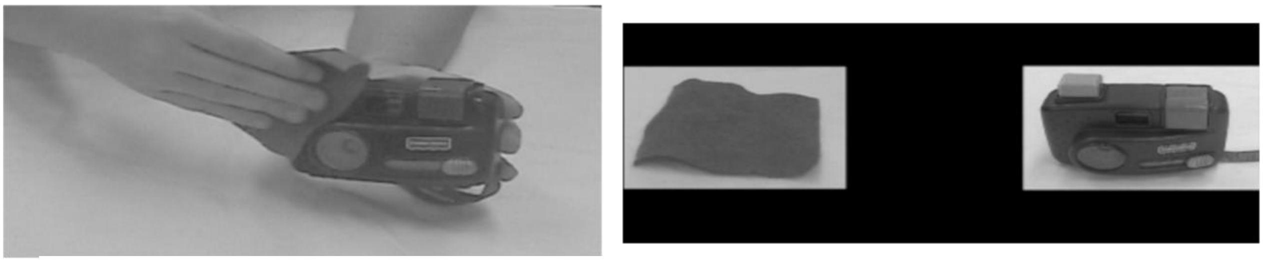

Figure 1: Example of familiarization phase (left; still of video) and test phase (right)

3.1.2.2.2 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.

3.1.2.2.3 Between-subjects factors and participant distribution

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. Order was thus treated as a two-level between-subjects factor (*Order 1* and *Order 2*), which resulted in a six cell design when crossed with our three linguistic stimulus levels (*V NP*, *V with NP*, and *Control*). Participants within each age group and sex were distributed evenly across the cells: 12 16-month-olds (6 females), 8 19-month-olds (4 females), and 8 28-month-olds (4 females) per cell.

3.1.2.3 Materials

Eight verbs contained in the MacArthur-Bates CDI (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. In the *V with NP* condition, children heard *with* during the familiarization, while those in the

V NP condition did not, represented in the table by the parentheses. In the *Control* condition, these sentences were replaced with “it’s a NOVEL NOUN.”

Phase	Length	Video	Audio
	2 seconds	Blank screen	
	5 seconds	Smiling baby	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**!
	2 seconds	Blank screen	Where’s the **tig**?
Test	2 seconds	Split screen: camera and cloth	
	3 seconds		Which one’s the **tig**?

Table 1: Sample script. The smiling baby/laughter attention-getter occurred only after every two trials

Table 2 shows each pairing of verb and novel noun, which were substituted for the italicized and bolded words in Table 1, respectively. The concomitant instrument and patient objects are also shown as well as the number of milliseconds after the beginning of the test phase the novel noun onset began.

Action/Verb	Instrument	Patient	Novel Word	Novel Word Onset (ms)
wipe	cloth	camera	tig	2767
throw	cup	ball	frap	2933
hit	ruler	cone	tam	2733
push	bulldozer	block	gop	2666
touch	pipe cleaner	pumpkin	pint	2767
wash	sponge	toy car	pud	2733
tickle	feather	mouse puppet	seb	2733
pull	fishing pole	train	wug	2866

Table 2: Pairings of action, instruments, and patient in each scene along with the novel noun used with that scene. The novel word offset gives the point in milliseconds during the five second scenes at which the novel word begins.

3.1.3 Results

3.1.3.1 Windowing

Because our data are sampled over time, it is possible that different subsets of the data---that is, data from different time intervals---show different response profiles. One way of dealing with this is to explicitly represent time in our model, e.g. using polynomial or spline regression. But since our hypotheses are about overall trends effected by condition, these sorts of models often contain (i) more fine-grained information than is necessary and (ii) non-zero parameters whose correct interpretation is unclear. We would still like to allow our model to capture some difference in temporal structure of the data, though. One way of doing this is to collapse over windows in the data defined by reasonable time points defined by the experimental structure, as is standard in IPLP designs (Hirsh-Pasek & Golinkoff, 1999).

Our experimental design naturally defines two such windows: 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 onset of the novel word in the second test question until the end of the trial (onset-5000ms; see Table 2 above for time of onset for each novel noun). Both windows are shifted forward 300ms to correct for delays in infant saccades (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998).

3.1.3.2 Statistical tests and software

Since our response variable is binary in nature, we submit them to mixed effects logistic regressions (Jaeger, 2008). All regressions include the maximal random effects structure[1] (Barr, Levy, Scheepers, & Tily, 2013). In our case, this means (i) by-subject random intercepts and

[1] We have not included time dependent random effects---for instance, by-subject or by-item random slopes for some (orthogonally-based) function of time---because it is unclear what function of time is appropriate in this case. And due to the possible problem of performing model comparison between models that differ only in their random effects (Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013), discovering this through model comparison is perhaps unwise.

(ii) correlated verb-by-subject random slopes, as well as (iii) random intercepts for trial nested under verb. The first two are standard practice in many repeated measures designs (verb corresponding to item in standard parlance). They aim to filter out subject-general variability and variability in subjects' responses to a specific verb. The third aims to take into account for both verb-general variability as well as possible order effects by encoding the fact that verbs' variability could differ based on the trials they show up in. This variability in responses to a verb on a specific trial is further controlled by a higher-order random parameter for each verb---hence, the nesting. All subsequent experiments---except for Experiment 3, which was a corpus study---used this random effects structure.

All analyses were run in the statistical computing language R (2.15; R Core Team, 2013). All models were fit using the R package lme4 (0.999999-0; see Bates, Maechler, Bolker, & Walker, 2013 for most recent version). The multcomp package (1.2-15; Hothorn, Bretz, & Westfall, 2008) was used to run comparisons and the ez package (4.1-1; see Lawrence, 2013 for most recent version) was used to bootstrap resample parameter estimates and standard errors for the purpose of plotting. All plots were generated using the ggplot2 package (0.9.3; Wickham, 2009).

3.1.3.3 16-month-olds

Table 3 provides the parameter estimates for the fixed effects of a mixed effects logistic regression fit with the random effects structure above[2]. Note that these estimates are not on the response scale (proportions) but rather a log-odds scale. This is because a logistic regression is in essence a linear regression through log-odds space. Figure 2 shows predictions for each cell on the response scale.

The fixed effects included in the model were condition (*Control* v. *V with NP* v. *V NP*), window (*First Window* v. *Second Window*), and verb production (*Verb Non-Producer* [verb production = 0] v. *Verb Producer* [verb production > 0]) along with all possible two-way and three-way interactions. The reference variables in the treatment coding were *Control* (condition), *First Window* (window) and *Verb Non-Producer* (verb production). A binarization of verb production was used instead of the raw verb scores because we were interested in

[2] Standard errors for the estimates are not presented because they are trivial to calculate from the estimate and the Z-value.

whether having verbs in the production vocabulary at all might affect children's novel noun learning at the cusp of the vocabulary spurt (Anglin, Miller, & Wakefield, 1993; Bloom, 2002)---and by the logic we lay out above, the way they assign thematic roles. Binarizing this variable has an additional benefit: the number of children with no verbs produced is highly inflated with respect to other counts; the number of 0-producers is 27 (making it the mode), the next highest count being 10 1-producers[3].

Term	estimate	Z-value	p-value
(Intercept)	**-0.32**	**2.08**	**<.05**
V with NP	-0.20	0.96	0.34
V NP	-0.14	0.76	0.45
Second Window	**0.15**	**3.40**	**<.05**
Verb Producer	-0.11	0.63	0.53
V with NP : Second Window	**0.47**	**6.33**	**<.05**
V NP : Second Window	-0.10	1.40	0.16
V with NP : Verb Producer	0.27	1.05	0.3
V NP : Verb Producer	0.09	0.36	0.72
Second Window : Verb Producer	-0.10	1.60	0.11
V with NP : Second Window : Verb Producer	**-0.34**	**3.66**	**<.05**
V NP : Second Window : Verb Producer	**0.28**	**2.98**	**<.05**

Table 3: Fixed effects parameter estimates for known verb experiment (16-month-olds).

First, we note a reliably negative intercept in log-odds space (<.5 in probability space). This means that infants in the reference cell (*Control*, *First Window*, *Verb Non-Producer*) are reliably biased to look at the patient (Z=2.08; p<.05). Further, this bias extends to all cells in the first window; all are significantly negative ($\chi^2(6)=16.12$; p<.05). This shows

[3] Though in traditional Analysis of Variance imbalanced designs are problematic, mixed effects models have no such issue (Baayen et al., 2008). Therefore, the fact that there are more verb producers than non-verb producers is not worrying.

an initial underlying bias to look at the patient across all conditions, which is perhaps unsurprising, since the patients generally have more complex visual features than the instruments.

Second, we see that the interaction term *V with NP : Second Window* is significantly greater than zero (Z=6.33; p<.05). A comparison between the *V with NP : Second Window* and *V NP : Second Window* finds that the *V with NP* interaction is reliably greater than the *V NP* interaction (Z=6.8; p<.05). This suggests that infants in the *V with NP* condition are more likely to switch to the instrument from the first window to the second than infants in either the *Control* condition or the *V NP* condition.

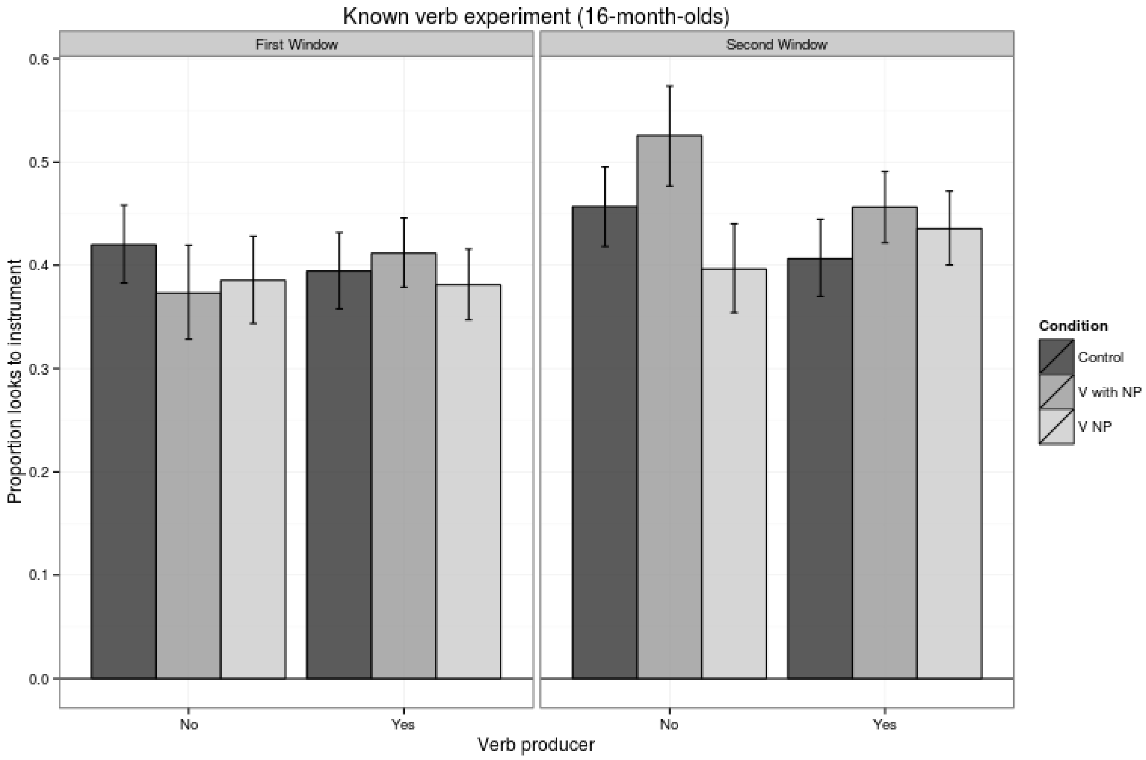

Figure 2: Proportion looks to instrument by window and verb production for 16-month-olds in the known verb experiment. Error bars show one standard error from the predicted proportion.

Third, we note two significant three-way interactions: the negative interaction *V with NP* : *Second Window* : *Verb Producer* and the positive interaction *V NP* : *Second Window* : *Verb Producer*. In comparison to the first window, verb producers are more likely to look to the instrument in the second window in both the *V with NP* (Z=5.03; p<.05) and *V NP* (Z=5.87; p<.05) conditions than in the *Control* condition. But we note that these differences are not reliably different from each other (Z=.85; p=.658). This suggests that, whereas verb non-producers differ in their willingness to switch to the instrument within the two test conditions---switching more to the instrument in the *V with NP* condition than in the *V NP* condition---this is not true of verb producers; infants in both the *V with NP* condition and the *V NP* condition switch to the instrument to the same extent.

3.1.3.4 19-month-olds

Table 4 provides the parameter estimates for the fixed effects of a mixed effects logistic regression with the same random effects structure used for 16-month-olds. Since we had no hypotheses about how vocab would affect 19-month-olds data, we did not introduce a term for verb vocabulary, including only the Condition and Window factors[4].

Term	estimate	Z-value	p-value
(Intercept)	**-0.21**	**2.05**	**<.05**
V with NP	0.00	0.05	0.96
V NP	-0.04	0.49	0.63
Second Window	**-0.18**	**5.07**	**<.05**
V with NP : Second Window	**0.11**	**2.07**	**<.05**
V NP : Second Window	**0.35**	**6.84**	**<.05**

Table 4: Fixed effects parameter estimates for known verb experiment (19-month-olds).

We find that, like the 16-month-olds, the intercept term is reliably negative (Z=2.05; p<.05). Unlike the 16-month-olds, not all conditions in the first window show looking different from chance ($\chi^2(3)=6.03$; p=.11). This suggests that though there looks to be an underlying patient bias, since the *Control* condition in the first window is reliably below 0 (.5 on the data scale), this bias does not surface in the test conditions. Taken together, we cannot conclude that all conditions reflect this in the first window.

[4] Since the 19-month-old sample did not show the same zero-inflation in verb production vocabulary, the binarization used for the 16-month-olds would yield many fewer children in the non-verb-producer group. And using a median split in this case (at 3 verbs produced) is not clearly as theoretically interesting as the zero vs. non-zero distinction. This does not preclude the possibility of controlling for verb production in 19-month-olds, however. Indeed, we will have cause to include such a control for 19-month-olds in Experiment 5, where we treat verbs produced as a continuous variable.

Also unlike either verb-producing or verb-non-producing 16-month-olds, we find 19-month-olds in both the *Control* (Z=5.07; p<.05) and *V with NP* (Z=2.06; p<.05) conditions reliably switching to the patient instead of the instrument in the second window. But like verb-producing 16-month-olds, we see 19-month-olds in the *V NP* condition reliably switching to the instrument (Z=4.6; p<.05). This rate of switching to the instrument is greater than infants in the *V with NP* condition (Z=4.7; p<.05). This is the exact opposite pattern we found for verb-non-producing 16-month-olds, though it shares similarities with that found for the verb-producers in that infants in the *V NP* condition switch more to the instrument.

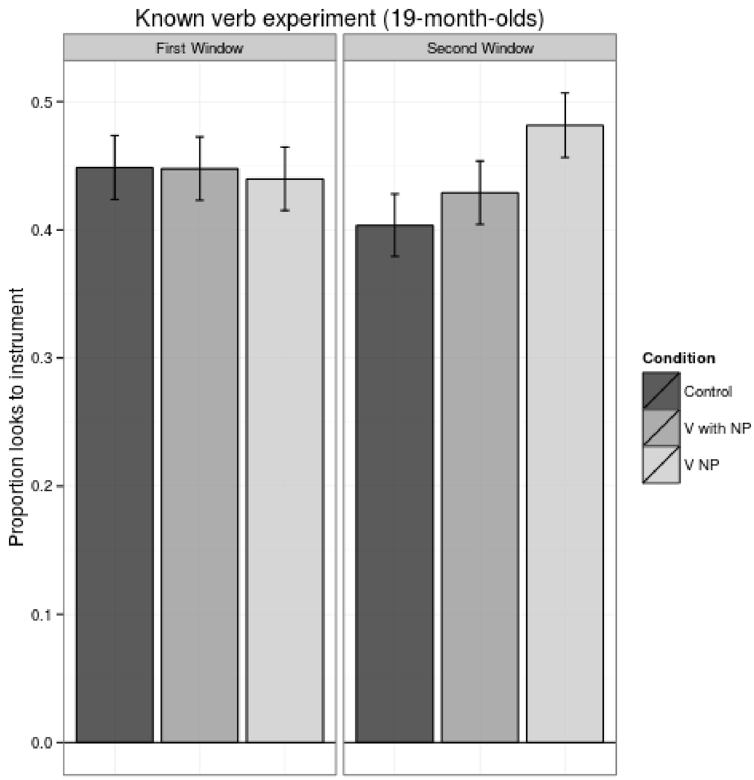

Figure 3: Proportion looks to instrument by window for 19-month-olds in the known verb experiment. Error bars show one standard error of the predicted proportion.

3.1.3.5 28-month-olds

Table 5 provides the parameter estimates for the fixed effects of a mixed effects logistic regression fit with the same random effects structure used for both the 16-month-olds and the 19-month-olds.

Term	estimate	Z-value	p-value
(Intercept)	-0.02	0.21	0.84
V with NP	0.01	0.1	0.92
V NP	-0.07	0.67	0.51
Second Window	**-0.17**	**4.65**	**<.05**
V with NP : Second Window	-0.06	1.09	0.28
V NP : Second Window	**-0.23**	**4.46**	**<.05**

Table 5:Fixed effects parameter estimates for known verb experiment (28-month-olds).

As with the 19-month-olds, we find a reliable negative effect for *Second Window* (Z=4.65; p<.05). Unlike with the 19-month-olds, we find a reliable negative interaction term *V NP* : *Second Window* (Z=-4.46; p<.05). Further, this interaction is reliably different from the *V with NP : Second Window* interaction (Z=3.38; p<.05). This suggests that infants in the *V NP* condition were more likely to switch from the instrument to the patient between the first window and the second window compared with both *Control* and *V with NP*.

Though the *V with NP* switches to instrument are not different than *Control* (Z=-1.09; p=.28), we note that this may have to do with the preset windows we chose. Due to the coarse-grained nature of collapsing across an entire window, there may be subtle effects missed within the window that get averaged out over the window.

To investigate this further, we inspected a model fit with more fine-grained time course information. To build the model, we held constant the random effects structure and introduced successively higher-order polynomial terms along with their interactions with condition. We tested each *n*-order polynomial model against the *n-1*-order polynomial model using a likelihood ratio test. This allowed us to ascertain the smallest polynomial justified by its ability to predict the data. We found that the 12^{th}-order polynomial model meets this criterion.

Figure 5 shows the predictions over time from the estimated parameters as well as predictions from 100 parameter values resampled from the fixed effects variance-covariance matrix. We found a reliable negative estimate (Z=2.28; p<.05) for the *V NP* effect in this model, suggesting that infants in the *V NP* look reliably more at the patient than in the *Control* condition. This estimate was also significantly less than that for the *V with NP* condition (Z=2.28; p<.05). This corroborates the pattern found in the above model collapsing over windows.

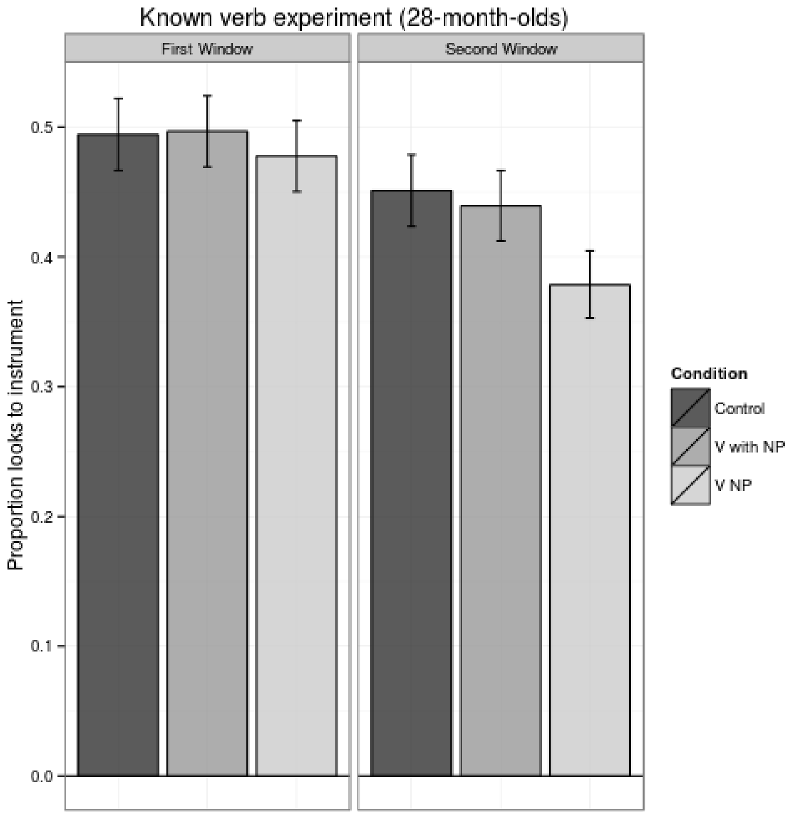

Figure 4: Proportion looks to instrument by window for 28-month-olds in the known verb experiment. Error bars show one standard error of the predicted proportion.

We further note two clear spikes above *Control* in the *V with NP* condition. These spikes seem to be counteracted by looks in the last 500 ms of the trials in both conditions, thus resulting in a lower collapsed estimate. This suggests that, in collapsing over the two windows, we missed important aspects of the temporal dynamics of the trial and that infants in the *V with NP* condition do in fact show a preference for the instrument.

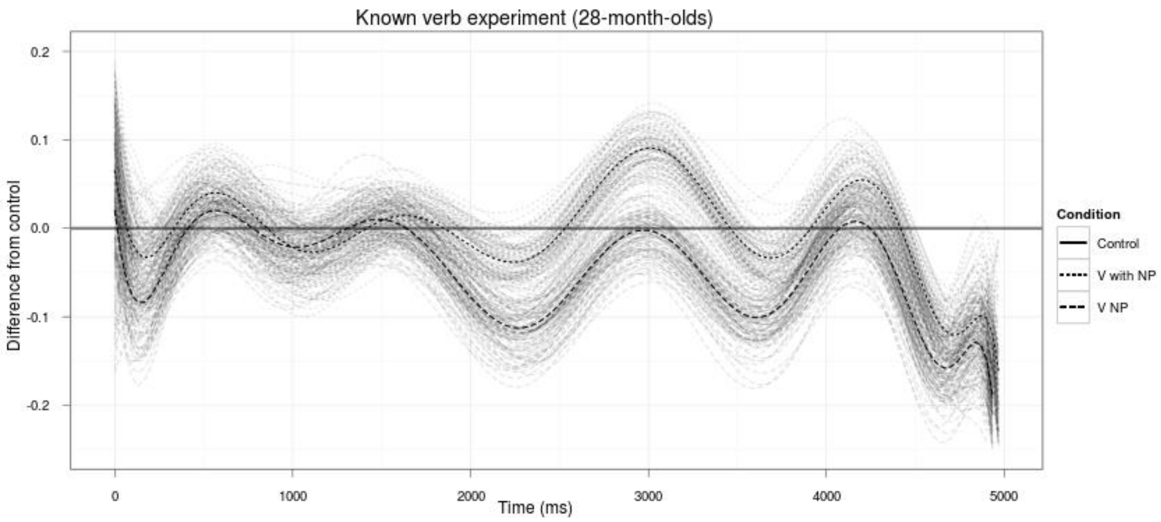

Figure 5: Difference in proportion from Control over time for 28-month-olds. Dark lines represent predictions from the model's fixed effects estimates; light lines represent predictions from multivariate normal samples from the fixed effects estimates and covariance matrix.

3.1.3.6 Discussion

The data in Experiment 1 support a U-shaped pattern of performance on this task, with the syntactic frame influencing performance only for the 28-month-olds and for the 16-month-olds without a productive verb vocabulary, but not for the two intermediate groups. The two groups who are influenced by the syntactic context may be computing an inference by which the syntactic position of a Noun Phrase 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 U-shaped pattern of facts raises 3 questions. First, how do 16-month-olds with no 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 for the change in behavior that appears to be associated with the onset of a verb vocabulary? Third, what is responsible for the reappearance of syntactic sensitivity in the 28-month-olds. We address these questions in turn.

3.2 Experiment 2: More Prepositions

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) He's pushing <u>on</u> **the tiv**.

(11) He's pushing <u>gub</u> **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.

Now, 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.

3.2.1 Participants

32 16-month-olds (16 females) with a median age of 16;19 (range: 16;3 to 17;5) were tested on our two new preposition conditions. This sample was merged with the 24 16-month-olds' that participated in the control condition from Experiment 1 to yield a final sample of 56 16-month-olds with a median age of 16;18 (range: 15;24 to 17;5). 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 verb production vocabulary was 1 (range: 0 to 83). Ten additional infants were tested exclusively for this experiment but were excluded from the final sample for fussiness or inability to complete the experiment.

3.2.2 Method

3.2.2.1 Apparatus and Procedure

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

3.2.2.2 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 VERBing the NOVEL NOUN" (*V NP* condition) or "she's VERBing <u>with</u> the NOVEL NOUN" (*V with NP* condition), infants heard either "she's VERBing <u>on</u> the NOVEL NOUN" (*V on NP* condition) or "she's VERBing <u>gub</u> the NOVEL NOUN" (*V gub NP* condition).

3.2.2.3 Materials

All materials were identical to those from Experiment 1, described in section 3.1.2.3. As noted above, the only change made was to the linguistic stimuli during the training phase.

3.2.3 Results

Table 6 provides the linear parameter estimates for the fixed effects of a mixed effects logistic regression fit with the same random effects structure used for all ages in Experiment 1. The fixed effects were also the same as those used in the 16-month-old model from Experiment 1. The reference values reported for that model are the same ones reported here.

First, we note that, as in the model for the 16-month-olds in Experiment 1, *Second Window* has a reliable positive effect (Z=3.4; p<.05). This suggests that non-verb-producing infants have a baseline preference to switch from the patient to instrument in the second window. Verb-producing 16-month-olds do not appear to have such a reliable switching preference (Z=1.17; p=.12). Both of these are expected, since the *Control* data is the same as that found in Experiment 1.

Second, we note the lack of a reliable interaction between *Second Window* and either *V gub NP* (Z=-1.22; p=.22) or *V on NP* (Z=-1.79; p=.07)[5]. Since we saw a positive interaction *V with NP : Second Window* and a negative interaction *V NP* : *Second Window* in Experiment 1, this pattern suggests that verb-non-producers are not treating the *V gub NP* or the *V on NP* frame like either the *V with NP* or *V NP* frames.

Term	estimate	Z-value	p-value
(Intercept)	-0.38	2.69	<.05
V gub NP	0.05	0.32	0.75
V on NP	0.03	0.13	0.9
Second Window	**0.15**	**3.41**	**<.05**
Verb Producer	0.00	0.03	0.97
V gub NP : Second Window	-0.08	1.22	0.22
V on NP : Second Window	-0.15	1.79	0.07

[5] The *V on NP : Second Window* term in fact trends in the direction we might expect if verb-non-producers knew the meaning of *on*, but since it's p-value (.07) does not pass the .05 threshold, no inference is licensed.

V gub NP : Verb Producer	-0.07	0.31	0.76
V on NP : Verb Producer	-0.34	1.33	0.18
Second Window : Verb Producer	-0.1	1.62	0.11
V gub NP : Second Window : Verb Producer	**0.29**	**2.98**	**<.05**
V on NP : Second Window : Verb Producer	**0.45**	**4.23**	**<.05**

Table 6: Fixed effects parameter estimates for novel preposition experiment (16-month-olds). Estimates represent linear model through log-odds space.

Third, we note positive three-way interactions between *V gub NP*, *Second Window*, and *Verb Producer* as well as *V on NP*, *Second Window*, and *Verb Producer.* These interactions parallel what was seen for verb-producers in the *V NP* condition from Experiment 1 and are exactly opposite of those found for the *V with NP* condition, bolstering the claim that 16-month-olds are not treating *gub* or *on* like *with*.

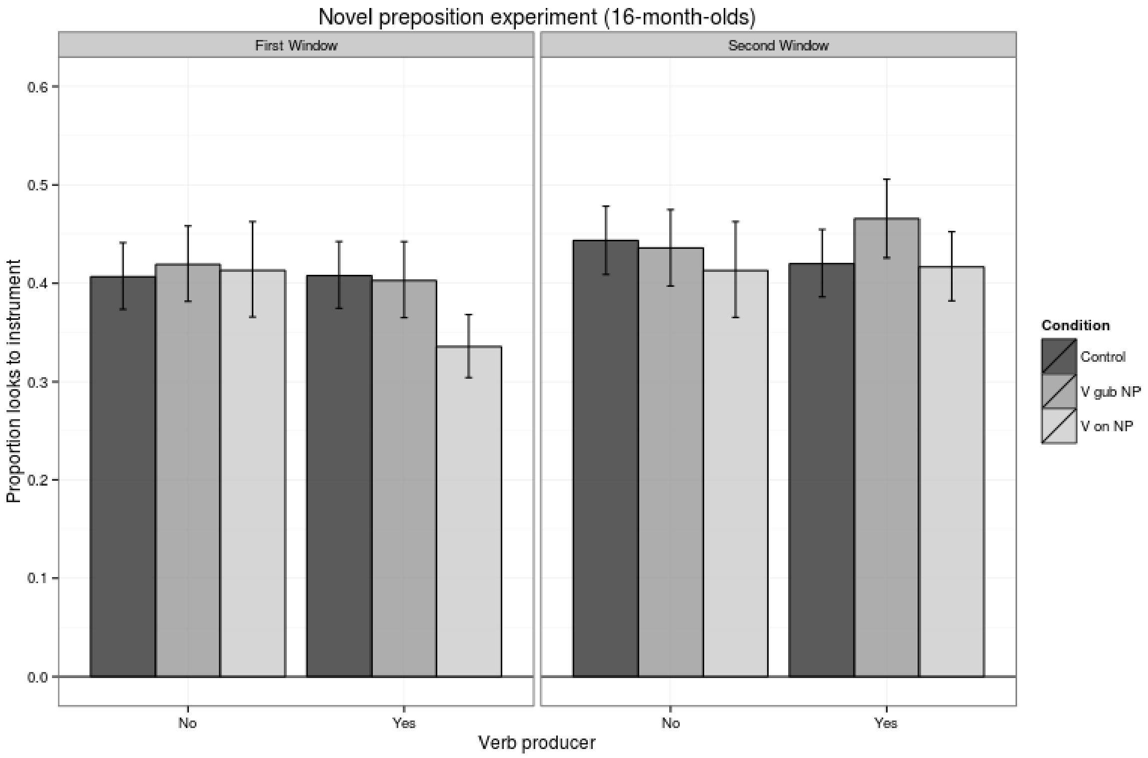

Figure 6: Proportion looks to instrument by window for 16-month-olds in the novel preposition experiment. Error bars show one standard error of the predicted proportion.

3.2.4 Discussion

Experiment 2 suggests that children do know the content of the preposition *with*. 16-month-olds in the *V with NP* condition looked at the instrument significantly more than chance and significantly more than children in the *on* and *gub* conditions. Moreover, children in the *V on NP* condition, looked at the instrument significantly less than chance; and, children in the *V gub NP* condition did not differ from chance.

All 16-month-olds (independent of verb vocabulary) appear to be interpreting the object of *on* as a patient. And, performance in the *V on NP* condition appears to be no different from chance. This data argues against the view that early success is based on a parsing heuristic. Instead, this pattern of data seems to indicate that by 16-months of age, children are aware of the semantic contributions of the prepositions *with* and *on* in establishing a thematic relation between their objects and the verb in the same clause.

This data also provides 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 (Dowty, 1991; 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 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 & Kamide, 1999; Gordon & Chafetz, 1990; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Kello, 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 sentence (8) from Experiment 1 (reproduced below). 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 & Kamide, 1999; Omaki, 2010; Sussman & 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.

(8) She's pushing with the tiv.

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.

3.3 Experiment 3: Corpus Analysis

Experiment 3 tests the prediction that the verbs used in Experiment 1 are more likely to be used as transitives than intransitives with a PP. In order to test this prediction, we examined the distribution of complement types for each of the 8 verbs used in Experiment 1 in Pearl & Sprouse's (2013) parsed version of the Brown corpus obtained from CHILDES (MacWhinney, 2000). We asked what proportion of the instances of each verb occurred in a transitive clause not also containing a PP ($[_{VP}$ __ NP]), an transitive clause also containing a PP ($[_{VP}$ __ NP PP]), an intransitive clause containing a PP ($[_{VP}$ _ PP]), or an intransitive clause containing a PP headed by *with* ($[_{VP}$ _ $PP_{[with]}$]). The results of this search are given in Table 7.

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*[6]. 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.

Verb	Total	Count $[_{VP}$ __ NP]	Count $[_{VP}$ __ NP PP]	Count $[_{VP}$ _ PP]	Count $[_{VP}$ _ $PP_{[with]}$]
hit	218	136 (.62)	53 (.24)	5 (.02)	0 (.00)
pull	291	203 (.70)	28 (.10)	10 (.03)	0 (.00)
push	302	206 (.68)	30 (.10)	8 (.03)	2 (.01)
throw	310	196 (.63)	82 (.26)	9 (.03)	1 (.003)
tickle	39	22 (.56)	0 (.00)	0 (.00)	0 (.00)
touch	184	151 (.82)	2 (.01)	2 (.01)	0 (.00)

[6] The reader will note that this higher percentage probably has to do with the presence of various other meanings of *with* in the corpus and not necessarily the instrument-marker meaning---e.g. g*oing with*.

wash	186	144 (.77)	13 (.07)	3 (.02)	0 (.00)
wipe	87	68 (.78)	13 (.15)	0 (.00)	0 (.00)
Mean (proportion of total)	**202**	**140.8 (.70)**	**27.6 (.14)**	**4.6 (.02)**	**0.4 (.002)**

Table 7: Counts from obtained from Pearl & Sprouse (2013) corpus.

3.4 Experiment 4: Satisfying the verb prediction

Experiment 4 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 thematic relation while also putting the verb in its preferred syntactic context, then this knowledge should reemerge.

Consider (12) and (13).

(12) He's pushing that thing <u>with</u> **the tiv**,

(13) He's pushing **the tiv** <u>with</u> 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.

3.4.1 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. This sample was merged with the 16 19-month-olds' that participated in the control condition from Experiment 1 to yield a final sample of 48 16-month-olds with a median age of 19;10 (range: 18;27 to 19;29). As in Experiments 1 and 2, 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 verb production vocabulary was 3.5 (range: 0 to 96). Four additional infants were tested exclusively for this experiment but were excluded from the final sample for excessive fussiness.

3.4.2 Method

3.4.2.1 Apparatus and Procedure

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

3.4.2.2 Design

All elements of the design were identical to that found in Experiment 1 except for (1) the form of linguistic stimuli used in the training phase; and (2) the number of trials. Instead of hearing “she's VERBing the NOVEL NOUN” (*V NP* condition) or “she's VERBing with the NOVEL NOUN” (*V with NP* condition), infants heard either “she's VERBing the NOVEL NOUN with that thing” (*V NP with that thing* condition) or “she's VERBing that thing with the NOVEL NOUN” (*V that thing with NP* condition). Instead of seeing eight trials, children saw six.

3.4.2.3 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.

Action/Verb	Instrument	Patient	Novel Word	Novel Word Onset (ms)
tap	pipe cleaner	train	pint	2800
brush	brush	mouse	seb	2733
stop	block	ball	frap	2767
hit	ruler	cone	tam	2733
wipe	cloth	camera	tig	2767
push	bulldozer	block	gop	2666

Table 8: Pairings of action, instruments, and patient in each scene along with the novel noun used with that scene. The novel word offset gives the point in milliseconds during the five second scenes at which the novel word begins.

3.4.3 Results

Table 9 provides the parameter estimates for the fixed effects of a mixed effects logistic regression fit with the same random effects structure used for all ages in Experiments 1 and 2. The fixed effects were the same as those used in the 19-month-old model from Experiment 1. The reference values reported for that model are the same ones reported here.

Term	estimate	Z-value	p-value
(Intercept)	**-0.15**	**2.6**	**<.05**
V that thing with NP	-0.18	1.53	0.13
V NP with that thing	-0.16	1.43	0.15
Second Window	**-0.18**	**5.02**	**<.05**
V that thing with NP : Second Window	**0.46**	**8.06**	**<.05**
V NP with that thing : Second Window	-0.09	1.55	0.12

Table 9: Fixed effects parameter estimates for two argument experiment (19-month-olds).

We note first a reliable bias to look to the patient over all conditions in the first window ($\chi^2(3)=23.23$; $p<.05$). This pattern contrasts with 19-month-olds in Experiment 1, who show no such reliable bias, but it is similar to both of the other age groups, who do show such a bias.

Second, we note a reliable negative effect of *Second Window*, this is consonant with our findings in both 19-month-olds and 28-month-olds in Experiment 1. Further, it is unsurprising since the *Control* condition is the same used in Experiment 1.

Third, we note the reliable positive interaction *V that thing with NP* : *Second Window* ($Z=8.6$; $p<.05$). This suggests that infants in the *V that thing with NP* are more likely to switch to the instrument than infants in the *Control* condition. Further, this interaction term is reliably greater than *V NP with that thing : Second Window.* This suggests that infants in the *V that thing with NP* are more likely to switch to the instrument than infants in the *V NP with that thing* condition.

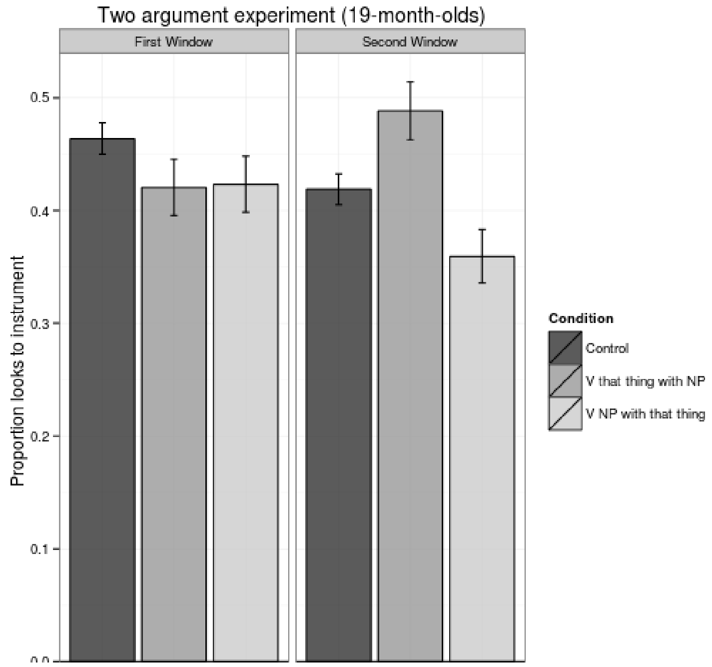

Figure 7: Proportion looks to instrument by window for 19-month-olds in the two-argument experiment. Error bars show one standard error of the predicted proportion.

3.4.4 Discussion

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

An additional question raised by the combination of results from Experiments 1 through 4 concerns the character of children's distributional expectations. Is the expectation for transitive clauses in these experiments driven by knowledge of specific verbs or is it a more general expectation that clauses will be transitive? Experiment 5 addresses this question.

3.5 Experiment 5: Eliminating the verb prediction

The hypothesis currently under consideration is that 19-month-old children's failure to use syntactic context as a cue to meaning in Experiment 1 derived from three factors: (i) their use of lexical subcategorization frequencies to predict structure, (ii) the mismatch between this prediction and the actual structure, and (iii) children's difficulty revising their initial commitment leading them to rely on predictive cues over bottom-up cues to structure. This hypothesis further predicts that for novel verbs, children will make no syntactic prediction and so should be able to rely on the bottom-up cue to structure (i.e., the preposition). Thus, with novel verbs, as in (14) and (15), the 19-month-olds should be able to interpret the novel noun differently as a function of its syntactic context, just like the non-verb-knowing 16-month-olds. Moreover, if non-verb-knowing 16-month-olds' success in Experiment 1 reflects the lack of a verb-specific prediction, then we expect these children to perform similarly in this experiment.

(14) He's meeking **the tiv**.

(15) He's meeking with **the tiv**.

These predictions, however, are valid only to the extent that the children's expectations about clause structure are based on lexical subcategorization frequencies of particular verbs and not a general expectation that all clauses will be transitive. If children at this age simply expect all clauses to be transitive, then this expectation should exert its influence in this experiment as it did in Experiment 1.

Experiment 5 tests two groups of 19-month-olds in each of these syntactic environments in (14) and (15). This design allows us to determine whether, in the absence of knowledge of lexical subcategorization frequencies for a given verb, infants can use the syntactic context as a cue to thematic relation.

3.5.1 Participants

32 19-month-olds (16 females) with a median age of 19;15 (range: 19;0 to 20;0) were tested on our two-argument task. This sample was merged with the 16 19-month-olds' that participated in the control condition from Experiment 1 to yield a final sample of 48 19-month-olds with a median age of 19;14 (range: 18;29 to 20;0). As in Experiments 1, 2, and 4, 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 MacArthur-Bates Communicative Development Inventory (MCDI) checklist. By this index, the median verb production vocabulary was 5 (range: 0 to 71). Six additional infants were tested exclusively for this experiment but were not included in the final sample due to excessive fussiness or inability to complete the experiment.

3.5.2 Method

3.5.2.1 Apparatus and Procedure

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

3.5.2.2 Design

All elements of the design were identical to that found in Experiment 1 except for the linguistic stimuli used in the training phase. As in Experiment 1, infants heard “she's VERBing the NOVEL NOUN” (*V NP* condition) or “she's VERBing <u>with</u> the NOVEL NOUN” (*V with NP* condition); but unlike Experiment 1, VERB was replaced by a novel verb.

3.5.2.3 Materials

All materials were identical to those from Experiment 1, described in section 3.1.2.3. As noted above, the only change made was to the verbs used in the linguistic stimuli during the training phase. The known verbs from the materials in Experiment 1 were replaced with the novel verbs in Table 10.

Action	Novel Verb	Instrument	Patient	Novel Noun	Novel Noun Onset (ms)
wipe	meek	cloth	camera	tig	2767
throw	doadge	cup	ball	frap	2933
hit	lonk	ruler	cone	tam	2733
push	tiz	bulldozer	block	gop	2666
touch	rem	pipe cleaner	pumpkin	pint	2767
wash	sloob	sponge	toy car	pud	2733
tickle	chiff	feather	mouse puppet	seb	2733
pull	stip	fishing pole	train	wug	2866

Table 10: Pairings of action, instruments, and patient in each scene along with the novel verb and novel noun used with that scene. The novel noun offset gives the point in milliseconds during the five second scenes at which the novel word begins.

3.5.3 Results

Table 11 provides the linear parameter estimates for the fixed effects of a mixed effects logistic regression fit with the same random effects structure used for all ages in Experiments 1, 2, and 4. The fixed effects were the same as those used in the 19-month-old model from Experiment 1 with the addition of a continuous predictor *Verbs*, which corresponds to the number of verbs produced.

This last variable was included to control for a possible effect of lexical knowledge on 19-month-olds' predictions about upcoming structure. The idea behind including this control is that the more verb types a learner has access to, the better that learner may be in their second-order generalizations about the distributional properties of verbs in the verb lexicon as a whole. An example of this might be a learner who knows that there are verbs that sometimes take PPs absent a bare NP---e.g. *go with*---and thus is less likely to erroneously predict an upcoming NP for new verbs. The fewer verb types a learner has access to, the more likely that learner may be to make second-order generalizations about new verbs that are based on a biased subset. An example of this might be a learner who has noticed that verbs overwhelming show up with NP complements and thus that new verbs probably will too.

The reference values reported for the discrete predictors are the same ones reported for 19-month-olds in Experiments 1 and 4.

Term	estimate	Z-value	p-value
(Intercept)	**-0.15**	**1.26**	**0.21**
V with NP	0.05	0.39	0.69
V NP	**-0.44**	**3.86**	**<.05**
Second Window	**-0.22**	**4.89**	**<.05**
Verbs	-0.01	1.02	0.31
V with NP : Second Window	-0.07	0.97	0.33
V NP : Second Window	**0.48**	**7.08**	**<.05**
V with NP : Verbs	0.00	0.09	0.93
V NP : Verbs	0.01	1.83	0.07
Second Window : Verbs	0.00	1.25	0.21
V with NP : Second Window : Verbs	**0.01**	**3.52**	**<.05**
V NP : Second Window : Verbs	**-0.01**	**2.05**	**<.05**

Table 11: Fixed effects parameter estimates for novel verb experiment (19-month-olds). Estimates represent the linear model through log-odds space.

We first note the standard first-window patient bias in all conditions ($\chi^2(6)=30.4$; p<.05). We also find an interesting difference between the *V NP* condition in the first window on the one hand and the *Control* and *V with NP* conditions on the other hand. First, the intercept (0 verbs produced) for the *V NP* condition is significantly lower than *Control* (Z=-3.86; p<.05) as well as *V with NP* (Z=-4.04; p<.05). The interaction term *V NP : Verbs* is not reliably greater than *Control* (Z=1.83; p=.07), but it is reliably greater than *V with NP : Verbs* (Z=2.45; p<.05). *V with NP : Verbs*, however, is not reliably different from *Control* (Z=.09; p=.93). This suggests that infants with smaller verb-production vocabularies in the *V with NP* condition are more likely to look at the instrument in the first window than infants with smaller verb-production vocabularies in the *V NP* condition, though those in the *V with NP* condition are not more likely to look at the instrument than those in the *Control* condition.

Second, we note that the *V with NP : Second Window* interaction term is reliably positive, meaning the intercept of the *V NP* model (0 verbs produced) is reliably greater than that in the first window (Z=7.08; p<.05). This suggests that infants with smaller verb-production vocabularies in the *V NP* condition are switching to the instrument from the first to the second window. This is not true of infants with smaller verb-production vocabularies in the *V with NP* condition (Z=.97; p=.33).

Finally, we note a reliable positive interaction *V with NP : Second Window : Verbs* (Z=3.52; p<.05) and a reliable negative interaction *V with NP : Second Window : Verbs* (Z=2.05; p<.05). These interactions are also significantly different from each other (Z=7.07; p<.05). This is reminiscent of the three-way interactions found in the model of 16-month-olds

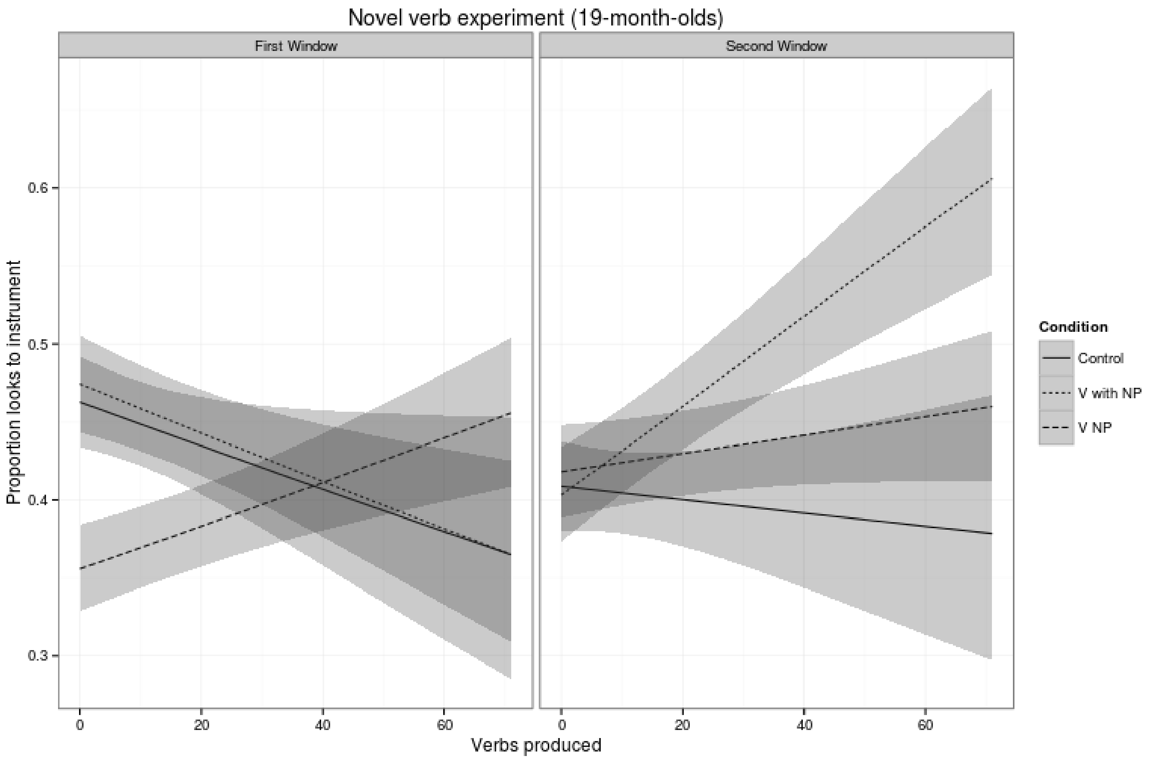

Figure 8: Proportion looks to instrument by window for 19-month-olds in the novel verb experiment. Error ribbons show one standard error of the predicted proportion.

in Experiment 1. The difference is that the signs are reversed: higher verb production in the *V with NP* condition is associated with more looks to instrument while higher verb production in the *V NP* condition is not. That is, infants with larger verb-production vocabularies are more likely to switch to the instrument in the *V with NP* condition than in the *V NP* condition from the first to the second window.

3.5.4 Discussion

Here we found that infants with larger verb production vocabularies interpreted the transitive object in sentences like (14) as the patient, devoting more of their looking time to the patient than the instrument. And critically, they interpreted the prepositional object in sentences like (15) as the instrument, devoting more of their looking time to the instrument than the patient. Infants with smaller verb production vocabularies also show some evidence of this trend, though they do so in a region of the time course in which we have not previously seen effects. Further, their looking is not different from control in these cases.

These results suggest that, at the very least, 19-month-old infants with larger verb production vocabularies make adult-like interpretive decisions when given a novel verb. This contrasts with their behavior when presented with known verbs in the same context. 19-month-old infants with smaller verb production vocabularies show a trend for such interpretations, but further research is necessary to understand the complexities of the time course information.

In sum, this suggests that 19-month-old infants use their knowledge of a verb's subcategorization frequency to predict upcoming structure, but in the absence of such a prediction (i.e., with a verb whose distributional profile is unknown), they are able to rely on bottom up cues to structure. Thus, errors in the assignment of thematic relation found in Experiment 1 can be attributed to children building erroneous syntactic structures in real time, but making the correct thematic inferences from those structures. When the factors that give rise to these erroneous structures are removed, these infants build the correct structure and so are able to make the appropriate thematic inferences.

4 General discussion

In a series of five experiments we have uncovered the following effects. First, we see a U-shaped pattern of development 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 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 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. The impact of expectations appears to be overcome by the time children are 28-months old. Second, children's ability to use a preposition as a cue to the meaning of a novel noun is in place by 16-months. At this age, children distinguish *with* from *on* semantically, and they distinguish both of these from a novel preposition. Third, 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. Finally, this difficulty can also be attenuated by using a novel verb, which blocks the child from generating expectations about the distribution of the verb and hence allows them to make inferences from syntactic position to 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 (Gagliardi & Lidz, in press) or as a function of their developing information processing mechanisms.

References

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

Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary development: A morphological analysis. *Monographs of the society for research in child development*, i–186.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390–412.

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

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4*.

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

Bloom, P. (2002). *How children learn the meaning of words*. The MIT Press.

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

Choi, Y., & 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*. University of Maryland.

Conroy, A. M., Fults, S., Musolino, J., & Lidz, J. (2008). Surface Scope as a Default: The Effect of Time in Resolving Quantifier Scope Ambiguity. Presented at the 21 st CUNY Sentence Processing Conference, University of North Carolina, Chapel Hill.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 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*. Baltimore: Paul H. Brookes Publishing Company.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & 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., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(2), 143–149.

Gagliardi, A., & Lidz, J. (in press). Statistical insensitivity in the acquisition of Tsez noun classes. *Language*.

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

Gordon, P., & 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., & Crain, S. (1984). Acquisition of cognitive compiling. *Cognition*, *17*(2), 85–136.
Hirsh-Pasek, K., & Golinkoff, R. M. (1999). *The origins of grammar: Evidence from early language comprehension*. The MIT Press.
Hollich, G. (2005). *Supercoder: A program for coding preferential looking*.
Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363.
Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434–446.
Lawrence, M. A. (2013). *ez: easy analysis and visualization of factorial experiments*.
Leddon, E., & Lidz, J. (2006). Reconstruction Effects in Child Language. Presented at the 30th Boston University Conference on Language Development, Cambridge: Cascadilla Press.
Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. R. (2010). When one cue is better than two: lexical vs. syntactic cues to verb learning. *Unpublished manuscript*.
Lidz, J., & Conroy, A. M. (2007). The Role of Working Memory in Explaining Isomorphism. Presented at the Generative Approaches to Language Acquisition, Barcelona, Spain.
MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review*, *101*(4), 676.
MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. transcription format and programs* (Third.). Mahwah, NJ: Lawrence Erlbaum Associates.
Marchman, V. A., & 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., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, *38*(1), 1–32.
Musolino, J., & Lidz, J. (2003). The scope of isomorphism: Turning adults into children. *Language Acquisition*, *11*(4), 277–291.
Musolino, J., & 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*(02), 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*. University of Maryland.
Pearl, L., & 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*. The MIT press.

R Core Team. (2013). *R: A language and environment for statistical computing* (Manual). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

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

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., & 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., & Lidz, J. (2012). Development of parsing abilities interacts with grammar learning: Evidence from Tagalog and Kannada. Presented at the 36th Boston University Conference on Language Development, Cascadilla Press.

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

Trueswell, J. C., Tanenhaus, M. K., & 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., & Musolino, J. (2010). Priming of abstract logical representations in 4-year-olds. *Language Acquisition*, *17*(1-2), 26–50.

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

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

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67–95.

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

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer Publishing Company, Incorporated.

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