Beyond Associations:

Sensitivity to structure in pre-schoolers' linguistic predictions

Chiara Gambi

Martin J. Pickering

Hugh Rabagliati

Edinburgh University

Address for correspondence:

Chiara Gambi

Developmental Lab

Department of Psychology

7, George Square

University of Edinburgh

EH8 9JZ Edinburgh, U.K.

cgambi@exseed.ed.ac.uk

Phone: +44 (0)131 650 33 43

Abstract

One influential view of language acquisition is that children master structural generalizations by making and learning from structure-informed predictions. Previous work has shown that from 3 years of age children can use semantic associations to generate predictions. However, it is unknown whether they can generate predictions by combining these associations with knowledge of linguistic structure. We recorded the eye movements of pre-schoolers while they listened to sentences such as *Pingu will ride the horse*. Upon hearing *ride*, children predictively looked at a horse (a strongly associated and plausible patient of *ride*), and mostly ignored a cowboy (equally strongly associated, but an implausible patient). In a separate experiment, children did not rapidly look at the horse when they heard *You can show Pingu ... "riding"*, showing that they do not quickly activate strongly associated patients when there are no structural constraints. Our findings demonstrate that young children's predictions are sensitive to structure, providing support for predictive-learning models of language acquisition.

Keywords: prediction; association; linguistic structure; visual-world.

Beyond Associations:

Sensitivity to structure in pre-schoolers' linguistic predictions

Introduction

A growing consensus in cognitive science is that our expertise in a variety of domains, from low-level action and perception to high-level cognition, is underlain by prediction (Clark, 2013). For example, the ability to generate expectations about others' actions, thoughts and words may underlie smooth turn-taking in social interaction (Magyari, Bastiaansen, de Ruiter, & Levinson, 2014), and could contribute to expert (i.e., adult) language processing (Pickering & Garrod, 2013). But is prediction just a tool deployed by expert systems, or rather the driving force behind the development of such systems? A number of computational models have proposed that prediction is critically important for acquiring language in the first place. For example, the connectionist models described in Elman (1990) and Chang, Dell, and Bock (2006) not only use prediction to process sentences, but also to master structural (i.e., syntactic and semantic) generalizations. Prediction, then, might serve as the unifying principle for processing and learning (Chang, Kidd, & Rowland, 2013; Dell & Chang, 2014).

If prediction drives language acquisition, then children must be able to generate the right kinds of predictions from early on. But while there is strong evidence that adults generate sophisticated predictions, the evidence that children make (and learn from) equally sophisticated predictions is much weaker (Rabagliati, Gambi, & Pickering, 2015). As one example, in order to learn structural generalizations, children need to be able to make predictions using their knowledge of linguistic structure, rather than solely relying on more basic knowledge such as semantic associations. Semantic associations comprise both world knowledge (e.g., that the event of "arresting" typically involves both policemen and robbers) and word co-occurrences (e.g., that *policeman* and *robber* are often mentioned close to the word *arrest*), and they play an important role in the language processing of both adults (e.g., Ferretti, McRae, & Hatherell, 2001) and children (Arias-Trejo & Plunkett, 2009, 2013; Mani, Johnson, McQueen, & Huettig, 2013). This includes an important role in prediction, as highly-associated words are often highly predictable. However, associations alone (even sophisticated ones) can be fallible guides to prediction. For example, the verb *arrest* has semantic associations to both *policeman* (a likely agent) and *robber* (a likely patient), but only the latter is structurally predictable in an active sentence, such as *Toby arrests the…* (Kukona, Fang, Aicher, Chen, & Magnuson, 2011). That is to say, semantic associations are poor guides to prediction unless they can be combined with knowledge of linguistic structure.

To illustrate why structure-based predictions are so important for learning structural generalizations, consider the example of a child who has already learned the active transitive construction, and is now acquiring the passive. This child could, in principle, use their knowledge of the active voice to predict, on hearing the verb *arrests*, that a potential patient (e.g., a robber) will be mentioned next. If so, then their prediction will be dramatically disconfirmed when they hear a passive, which could gradually cause them to learn that agents (e.g., *policeman*) can also follow the verb. By contrast to this, if the child only predicted on the basis of associations, then upon hearing *arrests* they would expect to hear either *policeman* or *robber* or both, and would therefore not learn any useful structural generalization from encountering *policeman* after the verb in a passive sentence.

In this study, we test whether young children are able to combine knowledge of both semantic associations and linguistic structure in order to generate predictions that can be learned from. Previous work has shown that adults' predictions make use of linguistic structure in this way. Kukona and colleagues (2011) demonstrated that, after hearing *Toby arrests the…*, adults quickly direct their attention to a picture of a robber, but after hearing *Toby was arrested by the…*, they look at a policeman. Similarly, in earlier studies by Kamide and colleagues (Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003) adults' predictive looks were driven by the meanings of words in combination with the words' case marking, which signalled their structural role in the sentence. Therefore, there is clear evidence that adults make use of structural knowledge when predicting upcoming words.

But this does not mean that semantic associations have no role in adults' predictions: In Kukona et al.'s (2011) study, after hearing *Toby arrests the…*, adults looked more at the associated but structurally unpredictable *policeman* than at the completely unrelated *surfer*. Similarly, in Kamide, Altmann, and Haywood (2003), participants who heard *The man will ride…* looked at a motorbike (which is strongly associated to both *man* and *ride*) the most, and those who heard *The girl will ride* looked at the motorbike more than those who heard *The girl will taste*. Thus, looks to the motorbike increased with the number of words associated with it in the preceding sentence. In sum, there is clear evidence that adults make use of associations as well as structure when predicting upcoming words. Importantly, they are able to combine their knowledge of associations with their knowledge of structure, so that when associations support multiple alternatives to an equal extent, they usually entertain structurally unpredictable alternatives to a lesser extent than structurally predictable ones (Kukona et al., 2011).

Whether preschool-aged children can generate predictions based on linguistic structure is less clear. Visual-world studies have shown that children generate predictions about upcoming words by 2 years of age (Borovsky & Creel, 2014; Borovsky, Elman, & Fernald, 2012; Borovsky, Sweeney, Elman, & Fernald, 2014; Mani & Huettig, 2012; Fernald, 2004, as reviewed in Fernald, Zangl, Portillo, & Marchman, 2008), but the mechanisms underlying those predictions have not been well established. In fact, work by Borovsky and colleagues suggests that children's predictive eye movements may be based on semantic associations, rather than structural knowledge. For example, on hearing *The pirate chases the…* children as young as three tended to look towards a depicted ship, which is associated with both *pirate* and *chases* (Borovsky et al., 2012), and is a plausible patient of *chases*. However, they also looked to treasure (associated with *pirate*) and to a cat (associated with *chases*) more than to unrelated distractors (e.g., a bone), even though these were not plausible patients. That is to say, their predictive looks could be explained as the result of a simple summation of the associations between the pictures on the screen and the words heard so far.

Other work suggests that these associations may be more complex than simple word-to-picture associations. For example, on hearing *I want to hold the…* spoken by a character who previously introduced himself as a pirate, children as young as three look towards a depicted sword, suggesting that they can generate predictions based on a speaker's identity. However, these predictions still appeared to be driven by associations of some form: The children also looked towards a ship (associated with the character but not holdable), and a wand (associated with *hold* and not with a pirate) more than to unrelated distractors (Borovsky & Creel, 2014). That is to say, the children in this study did not appear to be ruling out associated but unpredictable continuations.

In sharp contrast with the extensive evidence for association-based predictions, there is only more limited evidence for structural predictions in young children. Older children, such as 5- to-6-year-olds, appear to process active and passive constructions (Arai & Mazuka, 2014; Huang, Zheng, Meng, & Snedeker, 2013) by predicting upcoming arguments based on their structural knowledge of these constructions. Most interestingly, a recent study (Lukyanenko & Fisher, 2016) found that 3-year-olds will predictively look to a plural subject when they hear *Where are the ...*[1]. This shows that they can use the number feature of the verb (a syntactic feature) to predict the number of an upcoming subject noun, and therefore suggests that they use a syntactic relation (i.e., agreement) to guide their predictions (see also Melançon & Shi, 2015). However, this study was not set up to examine whether young children are able to combine association-based with structure-based predictions. Rather, structure-based predictions were the only type of predictions afforded by the sentence preambles used in this study, because none of the words preceding the structurally predictable subjects were semantically associated to these subjects.

Here, we pit structure against associations directly. We ask whether young (3-to-5 year olds), language-learning children are able to combine their knowledge of associations and linguistic structure to generate predictions in the same way as adults do. For example, are they able to predict that the verb *arrests* in an active sentence is more likely to be followed by *robber* than by *policeman*? From previous studies (e.g., Borovsky et al., 2012) we know that children aged 3 and older have acquired knowledge about the typical participants in common events, and are able to deploy such knowledge predictively. However, these studies have only tested whether children predict strongly or weakly associated *patients*, and have shown that they predict proportionally to the strength of the association (see also Mani, Daum, & Huettig, in press). But because in these studies the most associated patient was also the most associated word *tout court*, it remains unclear whether children were simply predicting on the

[1] Lukyanenko and Fisher also found that 2.5-year-olds were faster to orient to a plural noun when it was heard in an informative context, a result that could also potentially be driven by prediction. However it is also explicable by facilitated integration (see also Lew-Williams & Fernald, 2007). Unambiguously predictive effects (i.e., registered before or at noun onset) were not fully reliable in 2.5-year-olds.

basis of the strongest association, or were combining associations and linguistic structure to predict the most strongly associated patient.

Experiment 1

In order to test if young children predict using a combination of linguistic structure and associations, Experiment 1 used a task inspired by the visual-world study of Kukona et al. (2011): A large sample of preschool-aged children, and adults, listened to sentences such as *Pingu will ride/pull the horse*, while looking at the subject of the sentence (Pingu), an associated patient (e.g., horse), an associated agent (e.g., cowboy), and a distractor. We compared children's predictive looks to patients when they were associated with the verb (*ride*) and when they were unrelated (*pull*); similarly, we also tested whether children's predictive looks to agents were affected by the presence of an associative link between these and the verb. Crucially, while both agents and patients were associated, only patients were structurally predictable. Since children's predictions lag behind adults' (Borovsky et al., 2012), we included both short and long sentences (e.g., *Pingu will ride/pull the very tired horse*) to give children more time to generate predictions. Listeners whose predictions are solely driven by associations should launch predictive eye-movements towards patients and agents alike when they are associated with the verb. But listeners who make use of linguistic structure to generate predictions should predominantly look at patients.

Method

Participants. We assumed an effect size slightly lower than in Mani and Huettig (2012), and planned to recruit 80 children to achieve 80% power. Due to the ending of the school year, we recruited seventy-seven English-speaking children from nurseries in and around

Edinburgh. Five children's data were discarded for not following instructions (2), language impairment (2) or bilingualism (1), leaving 72 children in the final sample (mean age: 49.3 months, range [34,66] months, 33 males). We also tested twenty-four English-speaking students from the University of Edinburgh (mean age: 21.8 yrs, range [19, 33], 8 males); sample size was set based on previous studies in this case (e.g., Kukona et al., 2011).

Materials. Transitive sentences containing predictive or non-predictive verbs were paired with sets of four toys: Pingu (a well-known British penguin), an associated agent of the predictive verb, an associated patient, and a distractor (see Tables 1 and S1 online). Sentences varied in the distance between verb and direct object noun; long sentences contained pre-nominal modifiers (4-5 syllables) that were absent in short sentences. Different pre-nominal modifiers were used for each item (i.e., each target noun), but the same modifiers were used across predictive and non-predictive versions of each sentence as shown in Table 1. Verb Type (non-predictive vs. predictive) and Length (short vs. long) were fully crossed in a within-items, within-subjects design. Items were assigned to four lists using a Latin Square, with two random orders per list.

Table 1.

Example materials; bracketed words were used only in long sentences. The critical verb is highlighted in bold.

Verb Type		Patient	Agent	Distractor
Predictive	In this one, Pingu will **ride** the (very tired) horse.	Horse	Cowboy	Nurse
	Now, Pingu will **milk** the (incredibly fast) cow.	Cow	Farmer	Pony
Non-predictive	In this one, Pingu will **pull** the (very tired) horse.	Horse	Cowboy	Nurse
	Now, Pingu will **listen to** the (incredibly fast) cow.	Cow	Farmer	Pony

Importantly, each predictive verb was strongly associated with both an agent and a patient (e.g., *ride* had Agent: cowboy*,* Patient: horse). The association strength from verb to agent was matched to the association strength from verb to patient. In addition, the agents were highly plausible as agents but implausible as patients, and vice versa, while the plausibility of agents as agents was equal to the plausibility of patients as patients. Each predictive verb was yoked to a non-predictive verb (e.g., *pull* had Agent: cowboy, Patient: horse), which had no strong association to either the agent or the patient, and for which both objects were equally plausible as agent or patient.

To develop these stimuli we conducted two norming studies. First, following Kukona et al. (2011), adults rated whether characters were plausible agents or patients of the verbs. Then, critically, we asked a separate group of children to select two pictured characters from a set of eight (the agent, the patient, and two distractors, each represented by two easily distinguishable exemplars) and use them to act out the meaning of each verb in front of a puppet. We calculated the proportion of children who selected each character as agent (agent-hood rating) or patient (patient-hood rating). Eight pictures were used to ensure the association between agent and verb could be measured independently of the association between patient and verb (i.e., participants could potentially choose the same character as both agent and patient). After norming, we selected 12 sets of materials, whose characteristic agent-hood and patient-hood ratings and association scores can be seen in Table 2. Distractors were unrelated to both predictive and non-predictive verbs. Further details and statistical analyses can be found in the Supplemental material online.

Table 2.

Latent Semantic Analysis (LSA) association scores, agent-hood, and patient-hood ratings for the agents and patients used in this study; means over 12 items (standard deviations in brackets).

		Association strength	Children norming study[b]		Adult norming study[c]	
Verb Type	Entity	LSA score [a]	Agent-hood	Patient-hood	Agent-hood	Patient-hood
Predictive	Agent	.156 (.147)	.70(.20)	.07(.12)	6.42 (0.37)	3.53 (1.30)
	Patient	.176 (.149)	.056(.11)	.72(.32)	3.50 (1.29)	6.54 (0.52)
Non-predictive	Agent	.084(.065)	.20(.15)	.24(.20)	6.00 (0.88)	5.28 (1.11)
	Patient	.093(.083)	.23(.24)	.22(.17)	5.31 (1.12)	5.20 (1.77)

[a] Based on the following corpus: general reading up to 3rd grade (http://lsa.colorado.edu/).

[b] Proportion of children (N=15, 7 males; M=52.7 months, range=[38;66]) who selected the entity as agent or patient (respectively) when asked to act out the verb.

[c] Average rating assigned by adults (N=31) on a 7-point Likert scale. Higher values indicate higher plausibility.

Sentences were spoken in child-directed Scottish English by a female speaker. Verb duration was similar across the four versions of each sentence (predictive: short 734 ms, long 706 ms; non-predictive: short 671 ms, long 670 ms; Length $F(1,11) = 2.09$, $p = .176$, $r = 0.40$; Verb Type $F(1,11) = 2.16$, $p = .160$, $r = 0.41$; Length:Verb Type $F(1,11) = 0.16$, $p > .250$, $r = 0.12$). The direct object noun's onset was on average 1.7 seconds after the verb's offset in short sentences and 3.7 seconds after the verb's offset in long sentences.

Procedure. We followed Snedeker and Trueswell (2004): Participants sat in front of an inclined wooden stage containing four shelves. A camera housed in the center of the stage recorded participant's eye-movements at 25 frames per second. Children's actions were recorded by a second camera behind their shoulder. Sentences were played through loudspeakers. Participants were told they would act out short stories about Pingu using the toys, and completed one practice trial. Before each trial, the experimenter laid out and named the toys. The toys' positions on the stage were counterbalanced across items. Adults were tested in the lab, children at their nursery in 10-to-20 minute sessions. Children's productive vocabulary was assessed using the Expressive Vocabulary (EV) sub-test of the Clinical Evaluation of Language Fundamentals (CELF-Preschool-2, UK Edition; Wiig, Secord, & Semel, 2006).

Coding. Trials (non-predictive verbs: 7.87% in short, 8.33% in long sentences; predictive verbs: 9.72% in short, 13.43% in long sentences) were excluded because of experimenter error, or because the child was distracted or performed the wrong action (adults' actions were always correct). The first author and three trained research assistants determined the participant's direction of gaze for every frame from sentence onset to either the onset of an action or 2 seconds after sentence offset, whichever was earlier. Gaze was coded as being directed at one of the four shelves, at the center, off-stage, or missing (blinks, track loss). The first author independently recoded 25% of participants coded by each of the other coders. Inter-coder agreement was high and similar across coders, based both on the percentage of agreed-upon total frames and on the percentage of agreed-upon shift frames (the latter is reported between square brackets): 92%[96%] (Coder 1), 94%[97%] (Coder 2) for adult data and 91%[92%] (Coder 1), 90%[90%] (Coder 3) for child data.

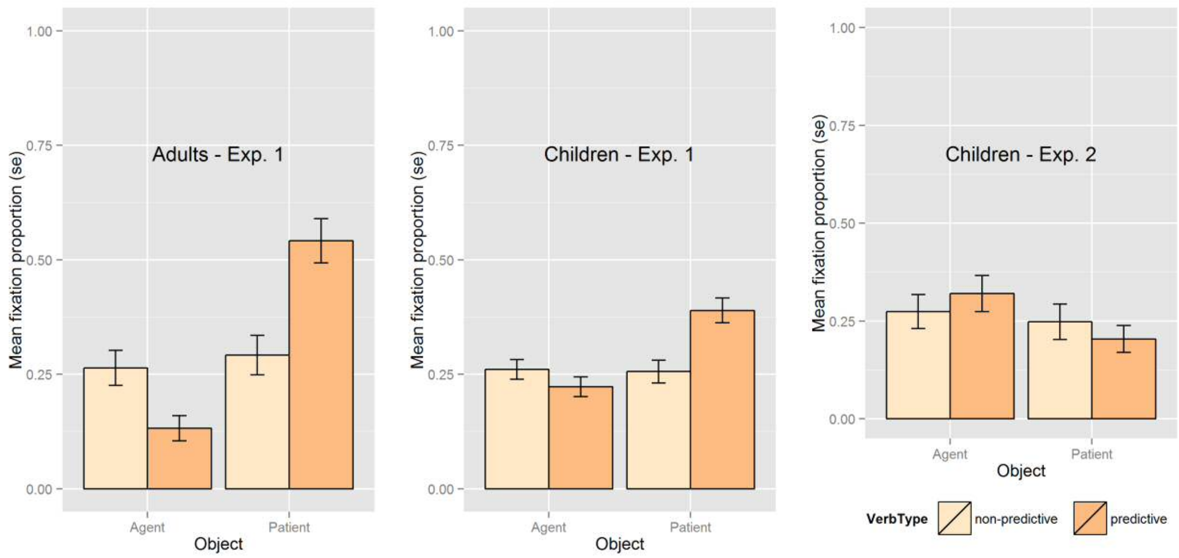
Results

We analysed whether the likelihood of participants looking to the agent and patient varied depending on the predictive power of the verb (Verb Type) and the amount of time available for prediction before the onset of the noun (Length). We did this in two ways. Our first analysis (Figure 1) provided a snapshot of participants' predictions just before the onset of the noun, during a 300ms window ending 100ms after noun onset (to account for delays in launching saccades; Trueswell, 2008); in separate mixed-effects logistic regressions we tested how Verb Type, Length, and their interaction affected the likelihood of looks to the patient and the likelihood of looks to the agent. We chose a short time window defined with respect to target noun onset for these analyses because they included the factor Length, and Long and Short sentences differed up until the target noun. Our second analysis, following Kukona et al. (2011), used growth curve modelling (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008) to provide an exploratory assessment of how looks to each character changed over time during a 2200ms window beginning 500ms before the offset of the critical verb and ending

1700ms after (Figure 2). Separate mixed-effects linear regressions tested how Verb Type affected the change in proportion of looks to the agent and the patient over time; data were averaged over items to obtain more robust estimates of the curves. Since this analysis was time-locked to the verb rather than the noun, Length was not included in these models. All analyses used the lme4 package (Bates, Maechler, & Dai, 2014) in R (R, Version 3.1.3). Fixed effects were contrast coded and centered. Random effects structure was maximal (Barr, Levy, Scheepers, & Tily, 2013), but correlations between random effects were sometimes set to zero to aid convergence (Bates, Kliegl, Vasishth, & Baayen, 2015). All *p* values are from log-likelihood ratio tests; 95% confidence intervals for model estimates are from the *confint* function (method="Wald").

Figure 1.

Snapshot analysis. Mean proportion of predictive looks to the patient and the agent after predictive and non-predictive verbs. See text for details of the time window used in this analysis. Error bars represent ± 1 SEM.



Adults. Our snapshot analysis (Table 4, top) confirmed that adults' predictions use structure, and are not just driven by associations. Average fixations proportions to the patient and agent in the four conditions are reported in Table 3. Figure 1 (left-most panel) shows the same data in graphic form, collapsing over short and long sentences. Adults were much more likely to predictively look at the patient upon hearing a predictive than a non-predictive verb (Table 3; log-odds Beta= 1.44, SE= 0.35, *CI*= [0.74,2.13], z= 4.07; $\chi2(1)$= 11.5, p< .001), and this effect did not vary with Length (log-odds Beta= -0.97, SE= 0.78, *CI*= [-2.51,0.56], z= -1.24; $\chi2(1)$= 1.49, p=.222). By contrast, participants did not generate more predictive looks to the agent after a predictive than a non-predictive verb; in fact there was a marginal tendency to generate fewer looks (log-odds Beta= -0.94, SE= 0.56, *CI*= [-2.04,0.16], z= -1.67; $\chi2(1)$= 3.50, p= .061), an effect that did not depend on Length (log-odds Beta= 0.60, SE= 1.04, *CI*= [-1.44,2.63], z= 0.58; $\chi2(1)$= 0.33, p>.250). In fact, Length did not affect looks to either the patient or agent (see Table 4, top).

Table 3.

Proportion of looks to the patient and agent in the snapshot analysis (Adults). Means over subjects (SE).

Verb Type	Length	Patient	Agent
Non-predictive	Long	.40 (.07)	.22 (.06)
Predictive	Long	.57 (.07)	.13 (.03)
Non-predictive	Short	.18 (.04)	.31 (.05)
Predictive	Short	.51 (.07)	.14 (.04)

The growth curve analysis confirmed these results (Table 4, bottom). In lme4 syntax, we used the following structure: 1 + Verb Type + Time + Time^2 + Verb Type:Time + Verb

Type:Time^2, plus random effects. The intercept term represents the mean proportion of looks over the entire window. The first order effect of Verb Type captures variation in the intercept term. The interaction between Verb Type and the linear time term captures variation in how rapidly looks to a character rise over time, while the interaction with the quadratic time term captures variation in the curvature of the line representing looks to each character. As in the snapshot analysis, adults looked to patients more after predictive than non-predictive verbs (Verb Type, Beta= 0.14, SE= 0.04, *CI*= [0.06,0.22], *t*= 3.55; $\chi 2(1)$= 10.14, *p*= .001) and, in addition, they looked *faster* to the patient after predictive than non-predictive verbs, as shown by a significant interaction between Verb Type and the linear time term (Beta= 0.58, SE= 0.14, *CI*= [0.30, 0.86], *t*= 4.10; $\chi 2(1)$= 12.72, *p*< .001). Verb Type did not affect the quadratic time term (Beta= -0.08, SE= 0.14, CI= [-0.37,0.20], t= -0.59; $\chi 2(1)$= 0.34, p>.250). By contrast, there was no overall effect of Verb Type on looks to the agent (Beta= -0.02, SE= 0.02, *CI*= [-0.06,0.03], *t*= -0.72; $\chi 2(1)$= 0.51, *p*>.250), and instead participants were *slower* to gaze at the agent after predictive than non-predictive verbs (Verb Type: Time, Beta= -0.35, SE= 0.14, *CI*= [-0.62, -0.07], *t*= -2.46; $\chi 2(1)$= 5.71, *p*= .017). Again, Verb Type did not affect the quadratic time term (Beta= -0.09, SE= 0.09, *CI*= [-0.26,0.08], *t*= -1.08; $\chi 2(1)$= 1.15, *p*>.250). See Figure 2.

Figure 2.

Growth curve analysis (Experiment 1). Proportion of looks to the patient (bottom panels) and agent (top panels) over time in the non-predictive (solid line) and predictive (dashed line) conditions; 0 is at verb offset. Error bars represent 95% confidence intervals computed over 1000 bootstrapped samples.

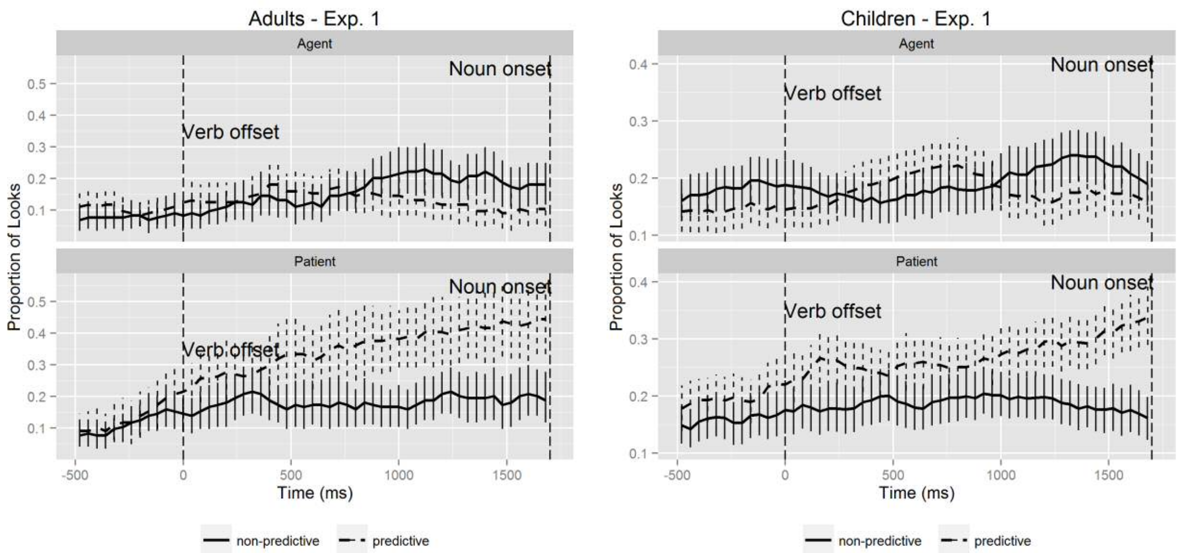


Table 4. Snapshot (top) and growth curve models (bottom) for adults in Exp. 1.

Snapshot analyses

Predictor	Object	Estimate (SE)	*z*	CI	χ2 and p value
Verb Type	Patient	1.44 (0.35)	4.07	[0.74,2.13]	χ2(1)= 11.5, *p*< .001
	Agent	-0.94 (0.56)	-1.67	[-2.04,0.16]	χ2(1)= 3.50, *p*= .061
Length	Patient	0.76 (0.39)	1.95	[-1.14,0.09]	χ2(1)= 2.47, *p*= .116
	Agent	-0.46 (0.59)	-0.79	[-1.62,0.69]	χ2(1)= 1.09, *p*>.250
Verb Type: Length	Patient	-0.97 (0.78)	1.24	[-2.51,0.56]	χ2(1)= 1.49, *p*=.222
	Agent	0.60 (1.04)	0.58	[-1.44,2.63]	χ2(1)= 0.33, *p*>.250

Growth Curve Analyses

Predictor	Object	Estimate (SE)	*t*	CI	χ2 and p value
Verb Type	Patient	0.14 (0.04)	3.55	[0.06,0.22]	χ2(1)= 10.14, *p*= .001
	Agent	-0.02 (0.02)	-0.72	[-0.06,0.03]	χ2(1)= 0.51, *p*>.250
Verb Type: Time	Patient	0.58 (0.14)	4.10	[0.30, 0.86]	χ2(1)= 12.72, *p*< .001
	Agent	-0.35 (0.14)	-2.46	[-0.62, -0.07]	χ2(1)= 5.71, *p*= .017
Verb Type: Time^2	Patient	-0.08 (0.14)	-0.59	[-0.37,0.20]	χ2(1)= 0.34, p>.250
	Agent	-0.09 (0.09)	-1.08	[-0.26,0.08]	χ2(1)= 1.15, *p*>.250

Children. As with adults, our snapshot analysis (Table 6, top) indicated that children's predictions are driven by linguistic structure, and not just associations. Average fixations proportions to the patient and agent in the four conditions are reported in Table 5. Figure 1 (middle panel) shows the same data in graphic form, collapsing over short and long sentences. Children were more likely to predictively look at the patient upon hearing a predictive than a non-predictive verb (log-odds Beta= 0.78, SE= 0.25, *CI*= [0.30,1.26], *z*= 3.19; $\chi 2(1)= 8.59$, *p*= .003), and this did not vary with Length (log-odds Beta= -0.22, SE= 0.52, *CI*= [-1.25,0.81], *z*= -0.42; $\chi 2(1)= 0.18$, *p*>.250). In contrast, hearing predictive verbs did not cause more predictive looks to the agent compared to hearing non-predictive verbs (log-odds Beta= -0.22, SE= 0.21, *CI*= [-0.64,0.19], *z*= -1.06; $\chi 2(1)= 1.33$, *p*>.250, and again this effect of Verb Type did not vary with Length (log-odds Beta= -0.10, SE= 0.48, *CI*= [-1.04,0.85], *z*= -0.20; $\chi 2(1)= 0.04$, *p*>.250). As with adults, Length did not affect looks to patient or agent (see Table 6, top). Unlike adults, however, children did not show a tendency to look at agents *less* after hearing predictive than non-predictive verbs[2].

[2] In additional snapshot analyses, we checked for potential order effects, which might have occurred if adults and children were able to identify likely agents and patients, and learn that patients would always be mentioned. There was no evidence for this: Order did not affect the likelihood of looking at the patient or agent, nor the magnitude of the Verb Type effect (all $|z|$'s < 1.45).

Table 5.

Proportion of looks to the patient and agent in our snapshot analysis (Children). Means over subjects (SE in brackets).

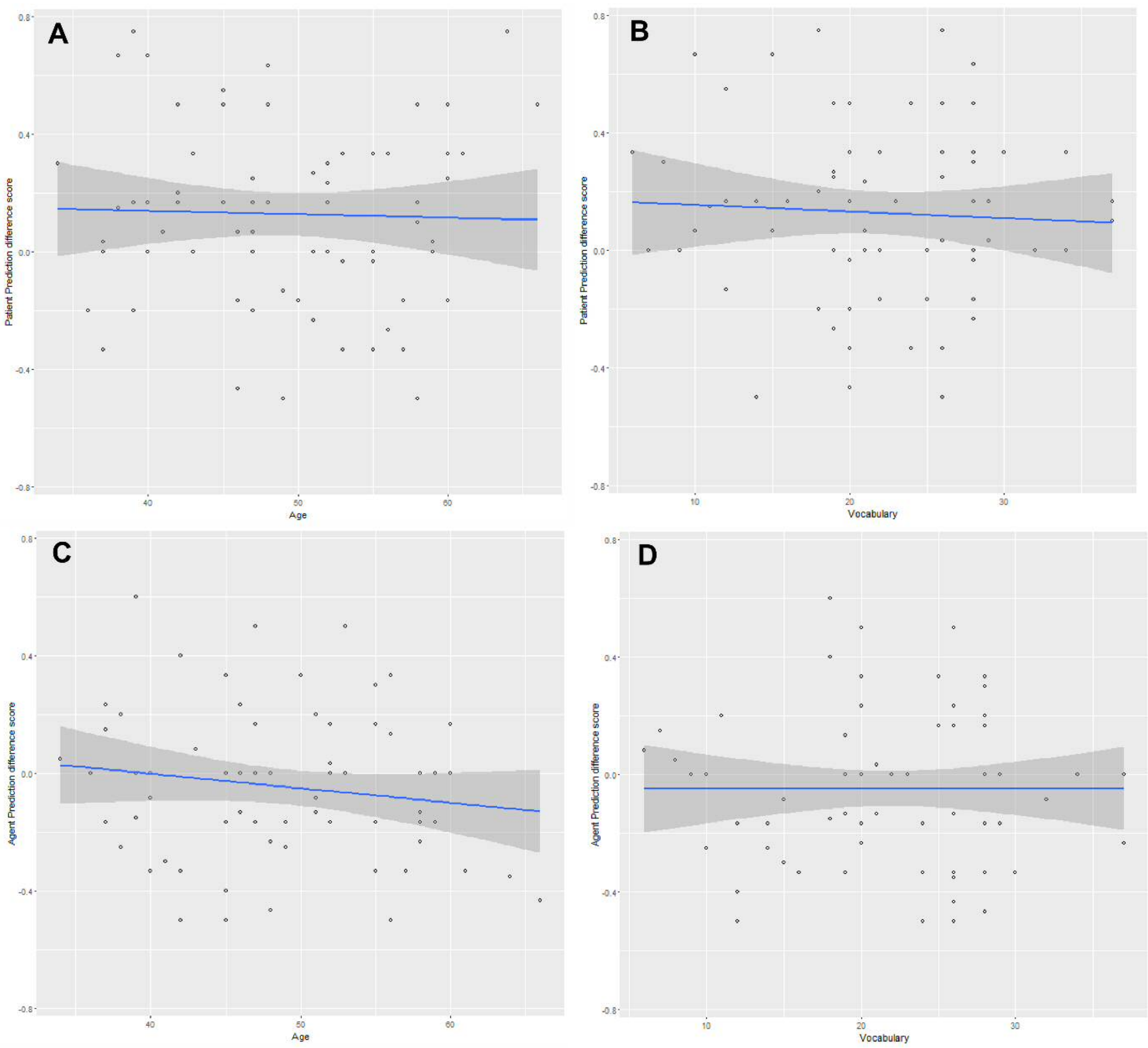
Verb Type	Length	Patient	Agent
Non-predictive	Long	.28 (.03)	.28 (.03)
Predictive	Long	.43 (.04)	.23 (.03)
Non-predictive	Short	.24 (.04)	.24 (.03)
Predictive	Short	.35 (.04)	.22 (.03)

Next we asked if these effects varied with age or linguistic knowledge. In fact, there was no evidence that children in this study processed the sentences differently depending on their age or vocabulary. When expressive vocabulary (centered raw scores) or age (centered age in months) were entered into separate regression analyses, neither factor interacted with either Verb Type or Length (all p's >.05).[3] The absence of age differences is also evident in the two panels of Figure 3, which show prediction summary scores for each child plotted against their age or vocabulary. These summary scores were computed with reference to the same time window used in the snapshot analyses: for each child, the proportion of fixations to the patient (top panel) or agent (bottom panel) after a non-predictive verb was subtracted from the proportion of fixations to the patient or agent after a predictive verb. The sizes of these prediction effects did not vary with age: The slopes of the regression lines do not differ from zero (Patient: t=0.03, *CI*=[-0.01, 0.01], p>.250; Agent: t=-1.50, *CI*=[-0.01,0], p=.138).

[3] Age and productive vocabulary were entered into separate regressions as they were strongly correlated (r(70) = 0.64, p<.001).

They also did not vary with vocabulary size (Patient: t=-0.44, *CI*=[-0.01, 0.01], p>.250; Agent: t=-0.04, *CI*=[-0.01, 0.01], p>.250).

Figure 3. (top panels) Patient Prediction difference scores (gaze in predictive minus non-prediction conditions) plotted against age in months (A) and productive vocabulary (B); (bottom panels) Agent Prediction difference scores plotted against age in months (C) and productive vocabulary (D).



The growth curve analysis (Table 6, bottom) confirmed the importance of linguistic structure in children's predictions. Like adults, children were overall more likely to look at the patient after a predictive verb (Verb Type, Beta= 0.07, SE= 0.02, *CI*= [0.03,0.10], *t*= 3.65; $\chi 2(1)$= 12.24, *p*< .001), and in addition looked faster to the patient upon hearing a predictive than a non-predictive verb (Verb Type: Time, Beta= 0.24, SE= 0.09, *CI*= [0.07,0.42], *t*= 2.68; $\chi 2(1)$= 7.01, *p*= .008). Verb Type did not interact with the quadratic time term (Beta= 0.08, SE= 0.08, *CI*= [-0.08,0.24], *t*= 1.03; $\chi 2(1)$= 1.05, *p*>.250). Also like adults, there was no overall effect of Verb Type on looks to the agent (Beta= -0.02, SE= 0.02, *CI*= [-0.05,0.02], *t*= -0.98; $\chi 2(1)$= 0.96, *p*>.250), confirming the snapshot analysis. Verb Type did not affect the speed with which children looked at the agent (Verb Type: Time, Beta= -0.05, SE= 0.09, *CI*= [-0.23,0.12], *t*= -0.58; $\chi 2(1)$= 0.34, *p*>.250). There was an effect of Verb Type on the quadratic time term (Verb Type:Time^2, Beta= -0.18, SE= 0.07, *CI*= [0.32,-0.03], *t*= -2.44; $\chi 2(1)$= 5.73, *p*= .017); an examination of the fitted curves (see Figure S1 online) suggests that this was driven by a graded tendency to look *away* from the agent more quickly after a predictive than a non-predictive verb. Again, these effects did not seem to vary as a function of age or expressive vocabulary, and neither factor interacted with Verb Type (all p's >.05). Figure S2 in the online Supplemental Material shows that the patterns depicted in Figure 2, right panel, were highly comparable in younger (<48 months, according to a median split of age) and older children.

Table 6. Snapshot (top) and growth curve models (bottom) for children in Exp. 1

Snapshot analyses

Predictor	Object	Estimate (SE)	*z*	CI	χ2 and p value
Verb Type	Patient	0.78 (0.25)	3.19	[0.30,1.26]	χ2(1)= 8.59, *p*= .003
	Agent	-0.22 (0.21)	-1.06	[-0.64,0.19]	χ2(1)= 1.33, *p*>.250
Length	Patient	0.42 (0.23)	1.81	[-0.04,0.88]	χ2(1)= 3.35, *p*= .067
	Agent	0.12 (0.22)	0.57	[-0.30,0.55]	χ2(1)= 0.29, *p*>.250
Verb Type: Length	Patient	-0.22 (0.52)	-0.42	[-1.25,0.81]	χ2(1)= 0.18, *p*>.250
	Agent	-0.10 (0.48)	-0.20	[-1.04,0.85]	χ2(1)= 0.04, *p*>.250

Growth Curve Analyses

Predictor	Object	Estimate (SE)	*t*	CI	χ2 and p value
Verb Type	Patient	0.07 (0.02)	3.65	[0.03,0.10]	χ2(1)= 12.24, *p*< .001
	Agent	-0.02 (0.02)	-0.98	[-0.05,0.02]	χ2(1)= 0.96, *p*>.250
Verb Type: Time	Patient	0.24 (0.09)	2.68	[0.07,0.42]	χ2(1)= 7.01, *p*= .008
	Agent	-0.05 (0.09)	-0.58	[-0.23,0.12]	χ2(1)= 0.34, *p*>.250
Verb Type: Time^2	Patient	0.08 (0.08)	1.03	[-0.08,0.24]	χ2(1)= 1.05, *p*>.250
	Agent	-0.18 (0.07)	-2.44	[0.32,-0.03]	χ2(1)= 5.73, *p*= .017

Comparison between children and adults. Finally, we pooled the child and adult data and compared the two groups using growth curve analysis. Overall, children looked at agents more than adults did (Beta= -0.05, SE= 0.02, *CI*= [-0.09,-0.01], *t*= -2.23; $\chi2(1)$= 4.72, *p*= .030), and they looked to patients less quickly than adults (Age Group: Time, Beta= 0.31, SE= 0.10, *CI*= [0.13,0.50], *t*= 3.22; $\chi2(1)$= 10.87, *p*< .001), but neither effect varied with Verb Type (all *p*'s >0.5). That is to say, children's predictive eye movements were both qualitatively and quantitatively similar to the adults' eye movements.

Discussion

Experiment 1 found that pre-school children are savvy predictors. Like adults, they looked more to associated and structurally predictable patients after hearing predictive than non-predictive verbs, and they also looked at these patients more quickly in the former than the latter case. In contrast, both children and adults failed to pay more attention to strongly associated but structurally implausible agents. This suggests that children use what they already know about linguistic structure to guide their predictions. Surprisingly, the magnitude and time course of prediction effects did not differ between children and adults, nor did they vary with the children's age or expressive vocabulary.

Experiment 2

We have argued that Experiment 1 shows language-learning children use structural information to inform their predictions. However, this conclusion rests on the assumption that, upon hearing predictive verbs, children rapidly activate both strongly associated agents and strongly associated patients, but disregard agents because they do not fit with the sentence structurally. Another possibility, though, is that, for children, verbs are differentially associated with their agents and patients, either through different types of association, or through different strengths of association (despite our best efforts in the pre-test).

For example, children might represent agent-verb associations in semantic memory (as other forms of world knowledge) but represent patient-verb associations as part of a verbs' meaning, and so would be slower to retrieve the agent information than to retrieve the patient information. Priming studies have shown that adults immediately activate associated agents when they hear a verb (e.g., Ferretti et al., 2001), but there is no comparable evidence for children. Alternatively, children might have a general bias towards gazing at associated patients more than towards associated agents, because they have learned associations that are ordered. For example, children may have learned an association that when they hear the verb *arrest,* then they tend to hear *robber* soon after, and this temporally ordered association could drive their predictive looks to the patient; the ordered association between *arrest* and *policeman* would instead be much weaker. Crucially, both of these alternative explanations predict that children should launch rapid predictive looks towards associated patients regardless of which structural cues are present in the sentence.

We tested these alternative explanations in Experiment 2. Children listened to structurally neutral instructions (e.g., children heard *Now, you can show Pingu ... riding/pulling*) while viewing the same visual displays used in Experiment 1. If children activate patients more strongly than agents regardless of structure, then we should again see rapid looks to patients but not to agents after hearing predictive verbs like *arrest*, just as in Experiment 1. But if children's predictive looks to patients in Experiment 1 were instead due to their use of structure to constrain prediction, then we would expect much reduced looks to patients when cues to structure are removed, along with, perhaps, more looks to agents.

Method

Participants. We recruited twenty-five additional English-speaking children from nurseries and a database of families in the Edinburgh area. We discarded the data from one child who did not follow instructions, leaving 24 children (mean age: 50.3 months, range [39, 68] months, 12 males).

Materials. The same verbs from Experiment 1 were spoken by a different female speaker using child-directed British English in structurally neutral sentences, such as *Now, you can show Pingu ... riding/pulling.* Verb duration was similar between predictive (1078 ms) and non-predictive verbs (1121 ms; $F(1,11) = 0.28$, $p > .250$, $r = 0.16$). Items were assigned to one of two lists in a Latin Square, with two random orders per list.

Procedure. Children were asked to demonstrate a word to Pingu using two toys of their choice; if they did not spontaneously do so, the experimenter prompted them to act out the word. After the task, children received the same vocabulary test used in Experiment 1. Sessions lasted 20 minutes, and took place at nurseries or the Developmental Lab at the University of Edinburgh.

Coding. Trials (non-predictive: 5.56%, predictive: 11.11%) were excluded and eye-movements coded (by the first author and a trained assistant) as in Experiment 1, except that gaze was only coded up to 1 second after sentence offset (or the onset of an action, if earlier). Inter-coder agreement was 92% (94% based on shift frames only). For details of performance in the act-out task, see the Supplemental material online.

Results and Discussion

Results. Eye-movement data were analysed as in Experiment 1, except that because there was no noun following the verb, the window used in the snapshot analysis began 200ms before verb offset; to avoid overlap with actions, the growth curve analysis used a 700ms time window starting 500ms before verb offset. Children's raw vocabulary scores ranged from 15 to 36, and correlated with their age (r(22) = 0.59, *p*=.002).

The snapshot analysis (Figure 1, right-most panel and Table 7, top) showed that children's looks to the agent were unaffected by the predictive power of the verb, and the same was true of their looks to the patient (Agent: predictive, M= .32, SE= .05, non-predictive, M= .27, SE= .04, log-odds Beta= 0.14, SE= 0.56, *CI*= [-0.96,1.24], *z*= 0.24; $\chi 2(1)$= 0.06, *p*>.250; Patient: predictive, M= .20, SE= .03, non-predictive, M= .25, SE= .05, log-odds Beta= -0.09, SE= 0.43, *CI*= [-0.93,0.76], *z*= -0.20; $\chi 2(1)$= 0.03, *p*>.250). Confirming the snapshot analysis, the growth curve analysis (Table 7, bottom) found that children did not look more to the agent overall (Verb Type, Beta= 0.02, SE= 0.05, *CI*= [-0.08,0.13], *t*= 0.47; $\chi 2(1)$= 0.22, *p*>.250) after a predictive verb than a non-predictive verb. However, the growth curve analysis also revealed that children rapidly associate agents to verbs (Figure 4): Children's looks to the agent rose faster (Verb Type: Time, Beta= 0.23, SE= 0.07, *CI*= [0.09,0.37], *t*= 3.21; $\chi 2(1)$= 8.64, *p*= .003) after a predictive than a non-predictive verb. In addition, the curvature of the line representing looks to the agent tended to be more pronounced after a predictive verb (Verb Type: Time^2, Beta= -0.08, SE= 0.04, *CI*= [-0.16,-0.003], *t*= -2.04; $\chi 2(1)$= 3.84, *p*= .050), but this effect was driven by children with larger vocabularies (Verb Type: Time^2: Vocabulary, Beta= -0.02, SE= 0.006, *CI*= [-0.03,-0.01], *t*= -3.61; $\chi 2(1)$= 10.39, *p*= .001). There were no effects for patients (see Table 7, bottom), nor other effects of vocabulary or age (all *p*'s> .05).

Figure 4.

Growth curve analysis (Experiment 2). Proportion of looks to the patient (bottom panel) and agent (top panel) over time in the non-predictive (solid line) and predictive (dashed line) conditions; 0 is at verb offset. Error bars represent 95% confidence intervals computed over 1000 bootstrapped samples.

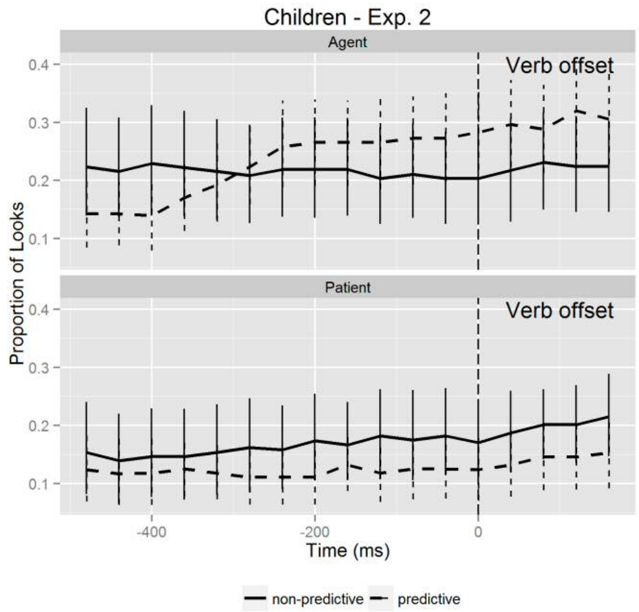


Table 7. Snapshot (top) and growth curve models (bottom) for children in Exp. 2

Snapshot analyses

Predictor	Object	Estimate (SE)	*z*	CI	χ2 and p value
Verb Type	Patient	-0.09 (0.43)	-0.20	[-0.93,0.76]	χ2(1)= 0.03, *p*>.250
	Agent	0.14 (0.56)	0.24	[-0.96,1.24]	χ2(1)= 0.06, *p*>.250

Growth Curve Analyses

Predictor	Object	Estimate (SE)	*t*	CI	χ2 and p value
Verb Type	Patient	-0.05 (0.03)	-1.57	[-0.10,0.01]	χ2(1)= 2.34, *p*= .126
	Agent	0.02 (0.05)	0.47	[-0.08,0.13]	χ2(1)= 0.22, *p*>.250
Verb Type: Time	Patient	-0.04 (0.07)	-0.61	[-0.19,0.10]	χ2(1)= 0.37, *p*>.250
	Agent	0.23 (0.07)	3.21	[0.09,0.37]	χ2(1)= 8.64, *p*= .003
Verb Type: Time^2	Patient	0.01 (0.03)	0.52	[-0.04,0.07]	χ2(1)= 0.27, *p*>.250
	Agent	-0.08 (0.04)	-2.04	[-0.16,-0.003]	χ2(1)= 3.84, *p*= .050

Discussion. Experiment 2 found no evidence that, in the absence of structural constraints, children look more, or more quickly, at strongly associated patients after hearing predictive than non-predictive verbs. In addition, although children did not show an overall preference for strongly associated agents, we did find that their looks to agents rose more quickly after hearing a predictive than a non-predictive verb, which suggests that children can indeed activate associated agents on hearing verbs, albeit weakly[4]. Importantly, these findings fail to support the possibility that children's predictive looks in Experiment 1 were driven by knowledge of simple, ordered associations between verbs and nouns. For that to be the case, we should have uncovered strong evidence that children gaze to the patient upon hearing a predictive verb, but we did not.[5] Instead, the findings of Experiment 2 are most consistent with the hypothesis that children generate predictions based on their knowledge of linguistic structure.

General Discussion

Influential models of the acquisition of grammar, such as Chang et al. (2006), propose that children compare their predictions about upcoming words to the words they actually hear, and use the discrepancy (prediction error) to learn linguistic generalizations. But for this

[4] It is possible that presenting verbs outside of a structural frame and in an unusual sentence final position is responsible for the weakness of the effects observed in Experiment 2.

[5] Bayes factor calculations also suggested that, in Experiment 2, hearing a predictive verb did not cause participants to gaze to the patient. We assessed whether the relevant regression terms in our analyses were more consistent with a null effect, or a positive effect. Following Dienes (2014), we compared the null with a range of potential positive effects, between 0 and twice the effect sizes observed in Experiment 1. The resulting Bayes factors were consistently less than 0.33, indicating strong evidence in favor of the null hypothesis. In the snapshot analysis, the Bayes Factor for the effect of Verb Type was 0.25; in the growth curve analysis the Bayes Factors for the effect of Verb Type, and for the interaction of Verb Type with the linear time term were both less than 0.1.

to be possible, children's predictions must incorporate information at the linguistic level to which the generalization pertains. So, for example, to learn a structural generalization, such as the passive construction, children must be able to make sophisticated structure-based predictions, such as "the next word will be a patient".

Here we have demonstrated that pre-schoolers predictively direct their attention towards strongly associated and structurally predictable patients, while they largely ignore equally strongly associated but structurally unpredictable agents (Experiment 1). Importantly, when they hear sentences that provide no cues to structure, they do not look at those same patients, and instead rapidly orient their attention towards the agents (Experiment 2). Although our findings do not show that children learn structural generalizations by making structure-informed predictions, they demonstrate that language-learning children make predictions that are critical for learning such generalizations. Quite strikingly, children's sensitivity to structure was not reduced compared to adults', and did not depend on their age (or vocabulary). This indicates that even the youngest children (3 year olds) can make correct predictions informed by structure while processing active sentences. If they make the same kind of predictions while processing other constructions, such as passive sentences, then they could use them to compute suitable prediction errors, which they could in turn use to learn the relevant structural generalization.

Importantly, while our data shows a critical role for structure in children's predictions, we do not claim that associations play no role. Previous studies have shown that children's predictive looks increase with associative strength (Borovsky et al., 2012; Mani et al., in press), and, in our own study (Experiment 2), we found some indication that children launched rapid looks to associated agents in structurally neutral contexts. Moreover, associations can be useful for prediction and for learning. In fact, words that co-occur with a larger number of words in parental input (and have more associative links in adult semantic networks) tend to be acquired earlier by children (Hills, Maouene, Riordan, & Smith, 2010). This, combined with the evidence that children use associations to generate predictions from early on, is strong evidence for a role of associations in learning, alongside structure.

Prediction and learning: age and vocabulary effects.

One of our more striking findings was that children's predictions (as indexed by the difference in the speed of looks to patients after predictive versus non-predictive verbs) did not differ from adults' in the degree to which they relied on structure, nor did children's predictions vary as a function of their age or vocabulary knowledge. This contrasts with previous findings showing that 2 year olds with larger production vocabularies are more likely to predictively look at a cake upon hearing *The boy will eat the…* (Mani & Huettig, 2012), and that 3-10 year olds direct their attention to a ship more quickly upon hearing *The pirate chases…* the larger their comprehension vocabularies (Borovsky et al., 2012).

Interestingly, in these studies children could make predictions on the basis of associations alone. This suggests that the previously-found developmental changes in prediction ability may be driven by changes in lexical associations. As they learn more and more words, children's lexicons change dramatically, and newly learnt words might change the strength of the associations between words already in the lexicon (Hills et al., 2010). For example, learning the verb *pet* might strengthen the existing association between *stroke* and, say, *cat*, because *pet* links to *stroke* (of which it is a synonym) and to *cat* (with which it often co-occurs). In addition, vocabulary size or age might simply be good proxies for children's world knowledge: The greater the number and type of events they experience, the more likely children are to know many words, but also the more likely they are to associate events with their typical participants.

In contrast to this, the ability to make structure-based predictions might vary less gradually with vocabulary or age. For example, once a child has acquired knowledge of the active transitive construction, and has begun to use it predictively, he or she may do so quite consistently across verbs. Incidentally, we chose to test this construction (and not, for example, the passive construction) precisely because we expected all children in our target age range would have consolidated their knowledge of it. Nonetheless, it is possible that children who are in the early stages of acquiring a new construction would make structure-based predictions only when they encounter familiar verbs. If this is the case, then one should observe a relationship between vocabulary size and structure-based prediction abilities. Future longitudinal studies might be able to uncover such a relationship by tracking, for example, a child's developing knowledge of the passive construction (e.g., in off-line interpretation tasks), and their predictive looks while they listen to passive sentences.

How are structure and associations combined in prediction?

Children's and adults' dual sensitivity to structure and associations raises the question: How are associations and structure combined in real-time to predict the most likely upcoming word? This question is particularly important because of some discrepancies between our work and previous studies. Most notably, Kukona et al. (2011) found that adults' predictive looks were sometimes directed to strongly associated agents that were structurally unpredictable. Instead, we found that, when associations favour two words equally, structural knowledge determines which one adults predict.

The discrepancy between our results and Kukona et al.'s (2011) could be explained by differences in the experimental set-up of the two studies. First, all our sentence materials had very similar structure. In addition, in order to make the task comparable for adults and children, we had adults listen to sentences spoken at a rate much slower than the one used by

Kukona et al, perhaps allowing more time for structure-based prediction. Interestingly, even in that study, participants were strongly influenced by associations only in Experiment 1, which used active sentences; looks to associated but structurally unpredictable agents were much weaker in Experiment 2, which used passive sentences instead. As the authors discuss, the presence of clear cues to structure and the additional time available for prediction during the beginning of the post-verbal *by*-phrase might have enhanced the role of structure in their participants' predictions. Similarly, it is possible that the high rate of structural repetition and the slow speech rate made the role played by structure more prominent in our study compared to theirs.

On the basis of their findings, Kukona et al. (2011) argued in favour of models of sentence interpretation that consider structure and semantics as parallel, separate, but constantly interacting processing streams (Kuperberg, 2007; MacDonald, Pearlmutter, & Seidenberg, 1994; Snedeker & Trueswell, 2004; Trueswell & Gleitman, 2004). Also consistent with this idea, Chang et al.'s (2006) model implements a dual architecture, in which one processing system learns sequences of thematic roles (e.g., agent - patient), and is largely independent of a second system, which learns relations between concepts (e.g., cowboy, ride and horse).

The existence of separate semantic and structural streams is also supported by recent evidence that, under some conditions, adults might compute predictions mostly or solely on the basis of associations. Chow, Smith, Lau, and Phillips (2015) showed that readers have difficulty processing verbs that are atypical given the participants mentioned in the sentence (e.g., *The superintendent overheard which realtor the landlord had evicted…*, compared to *The superintendent overheard which tenant the landlord had evicted…*), but not verbs that are atypical because the participants' roles have been reversed (e.g., *The superintendent overheard which landlord the tenant had evicted* compared to *The superintendent overheard* *which tenant the landlord had evicted*). While these findings do not demonstrate that readers predicted typical verbs (as difficulty was measured at the encountered verb *evicted*), they suggest that associations might sometimes trump structural relations during on-line interpretation (though see Kim, Oines, & Sikos, 2015). This could be especially likely in cases where structural relations are complex (such as in object-extracted questions), causing structure-building to be slow.

In contrast to this, our findings — an effect of structure but no effect of association — suggest that models in which structure and semantics are independent contributors to interpretation might not be fully adequate. Instead, we propose they are most compatible with the idea that undirected spreading activation at the semantic level generates a wide range of candidates for prediction, while a structure-based mechanism funnels processing resources and attention towards the more focussed set of candidates that fit with the unfolding structure (i.e., semantics proposes, structure disposes; cf. Crain & Steedman, 1985).

This account is inspired by the idea that prediction during language comprehension can make use of the production system (Pickering & Garrod, 2007, 2013). If this is true, then prediction during language comprehension should sometimes follow the stages involved in production, and there is consensus on the fact that semantics largely precedes syntax in production (Bock & Levelt, 1994; Dell, 1986; Levelt, Roelofs, & Meyer, 1999). Such an account suggests an architecture that allows for interactions between structural analysis and semantic interpretation, but assumes an ordered set of processes, with semantic predictions being computed before structural predictions. In this regard, it also differs from the proposal that structural (thematic) knowledge is directly encoded in the lexico-semantic network, which amounts to a blurring of the distinction between semantics and structure (McRae, Ferretti, & Amyote, 1997). Our account is compatible with findings that semantics can have immediate effects on the structural analysis of sentences (e.g., Taraban & McClelland, 1988), and can sometimes cause syntactically congruent sentences to be processed as syntactically anomalous (e.g., Kim & Osterhout, 2005).

Note that according to a production-based account, predictions must be compatible with the unfolding semantic interpretation of the sentence and will (additionally) be compatible with its unfolding structural interpretation if the comprehender has enough time to compute structural relations. Because structural computations will mostly be completed after semantics, though, there will be situations in which predictions will only be compatible with the unfolding semantic interpretation but not with the structure of the sentence (such as in Kukona et al., 2011, Experiment 1).

Structure-based predictions will instead be more likely when the comprehender is given more time to predict (and the time needed may be longer for children than adults). As mentioned above, the rate at which sentences were presented in our study was much slower than in Kukona et al. (2011), which fits well with the fact that structure was more prominent in our adult participants' predictions. However, accounts that posit separate but interacting processing streams can also accommodate variations in the degree to which one stream guides interpretation or prediction over the other. Such accounts could therefore accommodate the discrepancies between our findings and Kukona et al.'s as well.

In sum, our findings are clearly incompatible with the idea that language comprehenders, whether adults or young children, merely predict on the basis of associations. They show that language-learning children and adult expert language users are able to use their knowledge of structure in real time to constrain association-based predictions. One possibility, which is compatible with several existing accounts, is that semantic associations and structural relations are computed roughly at the same time and jointly influence the level of activation of candidates for prediction. Another possibility is that semantic associations are computed before structural relations in a way that resembles the ordered stages of production. Either way, our findings suggest that prediction in language-learning children and adults is supported by a strikingly similar architecture, one in which different sources of knowledge are combined in real time. Determining the precise details of this architecture is an open question for future research.

Before concluding, we note that, if prediction is at the heart of language learning, the way in which semantics and structure are combined in young children's predictions has important implications for how they can learn. For example, in case of a wrong prediction, it will determine at what linguistic level (or levels) the learning triggered by the resulting prediction error will occur. If learning occurs at more than one level, encountering *policeman* after *arrests* (when *robber* was expected) might strengthen the associative link between *policeman* and *arrests* at the same time as it weakens the expectation that patients should follow verbs, thus potentially hindering the learning of a new structural generalization. But if learning only occurs at the structural level (because it is computed last), then more focussed learning may be possible. Thus, questions about processing and prediction might bear on the issue of how quickly children can learn.

Conclusion

We have shown that adults and pre-schoolers are able to combine their knowledge of structure and of semantic associations to predict only structurally plausible continuations among those that are strongly associated. Therefore, our study demonstrates that children can take advantage of what they already know about linguistic structure to make structure-informed predictions, which are the kinds of predictions that they could use to learn more sophisticated structural generalizations. Our findings thus provide support for a key assumption behind models of language learning that assume a central role for prediction (Dell & Chang, 2014).

Supplementary Material

The data are available at https://github.com/chiara-gambi/structpred

Acknowledgements

This project was supported by a Leverhulme Trust Research Project Grant RPG-2014-253 to Hugh Rabagliati and Martin Pickering. We thank all the participating children, families and nurseries, and Sarah Hampton, Clara Smidt-Nielsen, Gintare Siugzdinyte, Alexandra Nemetschke, Cara Connachan, and Ellie Drake.

References

Arai, M., & Mazuka, R. (2014). The development of Japanese passive syntax as indexed by structural priming in comprehension. *The Quarterly Journal of Experimental Psychology, 67*(1), 60-78.

Arias-Trejo, N., & Plunkett, K. (2009). Lexical–semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3633-3647.

Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition, 128*(2), 214-227.

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

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. Retrieved from http://arxiv.org/pdf/1506.04967.pdf

Bates, D., Maechler, M., & Dai, B. (2014). Lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7). Retrieved from http://lme4.r-forge.r-project.org/

Bock, K., & Levelt, W. J. M. (1994). Language Production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Borovsky, A., & Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. *Developmental psychology, 50*(5), 1600-1613.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417-436.

Borovsky, A., Sweeney, K., Elman, J. L., & Fernald, A. (2014). Real-time interpretation of novel events across childhood. *Journal of Memory and Language, 73*, 1-14.

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

Chang, F., Kidd, E., & Rowland, C. F. (2013). Prediction in processing is a by-product of language learning. *Behavioral and Brain Sciences, 36*(4), 350-351.

Chow, W.-Y., Smith, C., Lau, E., & Phillips, C. (2015). A "bag-of-arguments" mechanism for initial verb predictions. *Language, Cognition and Neuroscience*, 1-20. doi:10.1080/23273798.2015.1066832

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(2), 181-204.

Dell, G. S. (1986). A spreading-activaction theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321. doi: http://dx.doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634), 20120394.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, http://dx.doi.org/10.3389/fpsyg.2014.00781

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*(2), 179-211.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*(4), 516-547.

Fernald, A. (2004). The search for the object begins at the verb. Paper presented at the 29th Annual Boston University Conference on Language Development, Boston, Nov 4-7, 2004.

Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina, I.A., Fernandez, E.M., & Clahsen, H. (Eds.) *Developmental Psycholinguistics: On-line methods in children's language processing* (pp. 113-132). John Benjamins: Amsterdam/Philadelphia.

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language, 63*(3), 259-273.

Huang, Y. T., Zheng, X., Meng, X., & 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. doi:http://dx.doi.org/10.1016/j.jml.2013.08.002

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156.

Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*(1), 37-55.

Kim, A. E., Oines, L. D., & Sikos, L. (2015). Prediction during sentence comprehension is more than a sum of lexical associations: the role of event knowledge. *Language, Cognition and Neuroscience*, 1-5. doi:10.1080/23273798.2015.1102950

Kim, A. E., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*(2), 205-225.

Kukona, A., Fang, S.-Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition, 119*(1), 23-42.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23-49.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-75.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193-198.

Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition, 146*, 349-370. doi:http://dx.doi.org/10.1016/j.cognition.2015.10.012

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676-703.

Magyari, L., Bastiaansen, M. C., de Ruiter, J. P., & Levinson, S. C. (2014). Early anticipation lies behind the speed of response in conversation. *Journal of Cognitive Neuroscience, 26*(11), 2530-2539.

Mani, N., Daum, M. M., & Huettig, F. (in press). "Pro-active" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. *Quarterly Journal of Experimental Psychology*. Retrieved from http://pubman.mpdl.mpg.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:2181001

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake - but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843-847.

Mani, N., Johnson, E., McQueen, J. M., & Huettig, F. (2013). How yellow is your banana? Toddlers' language-mediated visual search in referent-present tasks. *Developmental psychology, 49*(6), 1036-1044.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic Roles as Verb-specific Concepts. *Language and Cognitive Processes, 12*(2-3), 137-176. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0031511049&partnerID=40&md5=fe1f4e94f2c9afda75dfa6c05f90c9c5

Melançon, A., & Shi, R. (2015). Representations of abstract grammatical feature agreement in young children. *Journal of child language, 42*(6), 1379-1393.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: CRC Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475-494.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105-110.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-392. doi:http://dx.doi.org/10.1017/S0140525X12001495

R. (Version 3.1.3) [Computer Software]. Vienna, Austria: R Development Core Team. Retrieved from http://www.R-project.org

Rabagliati, H., Gambi, C., & Pickering, M. J. (2015). Learning to predict or predicting to learn? *Language, Cognition, and Neuroscience, 31*(1), 94-105. doi:10.1080/23273798.2015.1077979

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.

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language, 27*(6), 597-632.

Trueswell, J. C. (2008). Using eye-movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Language processing in children* (pp. 73-96). Amsterdam: John Benjamin.

Trueswell, J. C., & Gleitman, L. (2004). Children's eye movements during listening: Developmental evidence for a constraint-based theory of sentence processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 319-346). New York: Psychology Press.

Wiig, E. H., Secord, W. A., & Semel, E. (2006). *Clinical Evaluation of Language Fundamentals - Preschool, Second UK Edition*. London: Pearson.

Supplemental Material

Materials and Norming

Table S1.

Full list of materials used in Experiment 1. Bracketed words were only used in long sentences.

Verb Type		Patient	Agent	Distractor
Predictive	(1) In this one, Pingu will **ride** the (very tired) horse.	Horse	Cowboy	Nurse
	(2) Now, Pingu will **milk** the (incredibly fast) cow.	Cow	Farmer	Pony
	(3) This time, Pingu will **wash** the (really dirty) baby.	Baby	Mum	Princess
	(4) This time, Pingu will **walk** the (incredibly fat) dog.	Dog	Grandpa	Mechanic
	(5) In this one, Pingu will **save** the (incredibly tall) girl.	Girl	Fireman	Donkey
	(6) Now, Pingu will **rock** the (really happy) baby.	Baby	Mum	Sheep
	(7) Now, Pingu will **bite** the (really small) child.	Child	Dog	Queen
	(8) In this story, Pingu will **feed** the (very hungry) pig.	Pig	Farmer	Builder
	(9) In this story, Pingu will **catch** the (incredibly big) fish.	Fish	Fisherman	Old woman
	(10) In this story, Pingu will **arrest** the (noisy and fun) robber.	Robber	Policeman	Girl

	(11) In this one, Pingu will **scare** the (sweet and nice) child.	Child	Witch	Man
	(12) This time, Pingu will **stroke** the (sleepy and quiet) kitty.	Kitty	Grandma	Bull
Non-predictive	(1) In this one, Pingu will **pull** the (very tired) horse.	Horse	Cowboy	Nurse
	(2) Now, Pingu will **listen to** the (incredibly fast) cow.	Cow	Farmer	Pony
	(3) This time, Pingu will **see** the (really dirty) baby.	Baby	Mum	Princess
	(4) This time, Pingu will **watch** the (incredibly fat) dog.	Dog	Grandpa	Mechanic
	(5) In this one, Pingu will **point at** the (incredibly tall) girl.	Girl	Fireman	Donkey
	(6) Now, Pingu will **think of** the (really happy) baby.	Baby	Mum	Sheep
	(7) Now, Pingu will **find** the (really small) child.	Child	Dog	Queen
	(8) In this story, Pingu will **meet** the (very hungry) pig.	Pig	Farmer	Builder
	(9) In this story, Pingu will **hear** the (incredibly big) fish.	Fish	Fisherman	Old woman
	(10) In this story, Pingu will **touch** the (noisy and fun) robber.	Robber	Policeman	Girl
	(11) In this one, Pingu will **speak to** the (sweet and nice) child.	Child	Witch	Man

(12) This time, Pingu will **push** the (sleepy and quiet) kitty.	Kitty	Grandma	Bull

Adult pre-test. Adult participants generated agent-hood or patient-hood ratings, following the same procedure used by Kukona et al. (2011). Each participant rated either the predictive or the non-predictive verb in a pair, in combination with 7 different nouns: the associated agent and patient, three nouns that were relatively plausible agents/patients for the predictive verb, and two nouns that were implausible agents/patients for this verb. One group of participants was asked to produce agent-hood ratings and answered the question: "How common is it for a NOUN to VERB somebody/something"? Another group of participants produced patient-hood ratings and answered the question: "How common is it for a NOUN to be VERB-ed by somebody/something?". Ratings were given on a 7-point Likert scale. For half the lists, 7 corresponded to "extremely common" and 1 to "extremely uncommon"; for the other half, the scale was reversed (averages reported in the article, Table 2, were computed after recoding all data in such a way that higher scores correspond to higher agent-hood/patient-hood ratings). Participants completed the questionnaire online. We report statistical analyses for the 12 verb pairs that were included in the experiment (see Table 2 in the main article). With predictive verbs, the associated agents were rated as better agents than associated patients (agents, $M = 6.42$, $SD = 0.37$; patients, $M = 3.50$, $SD = 1.29$; $t(11)=7.15$, $p<.0001$), and the associated patients were rated as better patients than associated agents (patients, $M = 6.54$, $SD = 0.52$; agents, $M = 3.53$, $SD = 1.30$; $t(11)=6.26$, $p <.0001$). Importantly, the difference between the agent-hood scores of agents and the patient-hood scores of patients was similar across non-predictive and predictive verbs (non-predictive, $M = 0.80$, $SD = 1.84$; predictive, $M = -0.12$, $SD = 0.60$; $t(21)=1.64$, $p =.116$), and the average difference score for predictive verbs did not differ significantly from zero ($t(11)=-0.71$, p

=.492). This shows that predictive verbs did not elicit a stronger bias towards their associated patients than towards their associated agents. Finally, for non-predictive verbs, the agent-hood of agents did not differ from the agent-hood of patients (agents, $M= 6.00$, $SD = 0.88$; patients, $M = 5.31$, $SD = 1.12$; $t(10)=1.70$, $p = .119$), and the patient-hood of patients did not differ from the patient-hood of agents (patients, $M= 5.20$, $SD = 1.77$; agents, $M = 5.28$, $SD = 1.11$; $t(10)=0.16$, $p = .879$).

Children pre-test. To obtain agent-hood and patient-hood ratings from children, we developed a new act-out game. Children sat at a table containing a cardboard stage, as did the experimenter and a puppet. In the game, children acted out the meanings of verbs for the puppet on the stage, using pictures. On each trial, the experimenter displayed and named eight pictures for the child: these depicted toy characters or animals. Then, the experimenter said "Now, we have to show [Puppet name] "VERB-ing"!", and waited for the child to choose two pictures and demonstrate the action to the puppet. If the child did not pick any pictures, or did not use the pictures to act out the action, the experimenter encouraged the child by asking "Can you show [Puppet name] "VERB-ing?". If the child's demonstration of the action was unclear, the experimenter asked "Can you tell [Puppet name] what's happening?" to elicit a verbal description. If needed, the experimenter followed this up with a more specific question (e.g., "Who's VERB-ing?").

The proportion of children who selected the associated agent as agent (patient) gave the agent-hood (patient-hood) score for the agent, and similar scores were computed for the associated patient. Unlike in the adult pre-test, we used every trial in the computation of both agent-hood and patient-hood scores. To ensure independence of these two sets of scores, each of the eight pictures shown to the child depicted one of only four different characters or animals (the associated agent, the associated patient, and two others); each entity was thus depicted twice. We took care that the two depictions were easily distinguishable from one another (for example, one picture for *dog* depicted a brown puppy, while the other depicted a black and white puppy of a different breed). In this way, it was possible for children to pick the same entity as both agent and patient, which they often did (on 30.32% of codable correct trials; see below).

Children were tested at their nursery in a quiet room or inside the Developmental Lab at the Department of Psychology, University of Edinburgh. First, the experimenter played the game on one practice trial, and then children played it with 16 verbs (half predictive, half non-predictive). The order in which pictures were displayed on the table was randomised separately for each trial and child. There were 2 lists, so that each child was tested either on the predictive or the non-predictive verb in a pair, and we created two different presentation orders for each list. Children's actions were recorded on camera for off-line scoring. Before computing the scores, the first author discarded all trials on which the child did not act out any verb meaning, acted out a different verb meaning than the one intended, or produced an ambiguous action whose meaning could not be determined (39.63% of trials in total). In addition, she discarded trials on which the child picked one or more pictures before the experimenter mentioned the verb (a further 4.91% of trials). Finally, she also excluded a small number of cases in which the agent or patient were missing because the child interpreted the verb as intransitive or demonstrated the action on himself/herself or the puppet instead of a second picture.

After excluding such cases, the first author coded which of the two pictures selected by the child was the intended agent (the other picture was taken to be the patient). The following criteria were used to identify agents: (a) If the child verbalized the event using a transitive sentence, the agent of this sentence was coded as the agent. (b) If in the child's demonstration only one picture was moving while the other remained static, then the moving picture was coded as the agent. (c) If the child moved both pictures, the picture that moved first was coded as the agent. (d) If the child moved both pictures at the same time, the picture that occupied the left-most position in the direction implied by the action was coded as the agent. If none of the above conditions was satisfied, the trial was treated as non-codable.

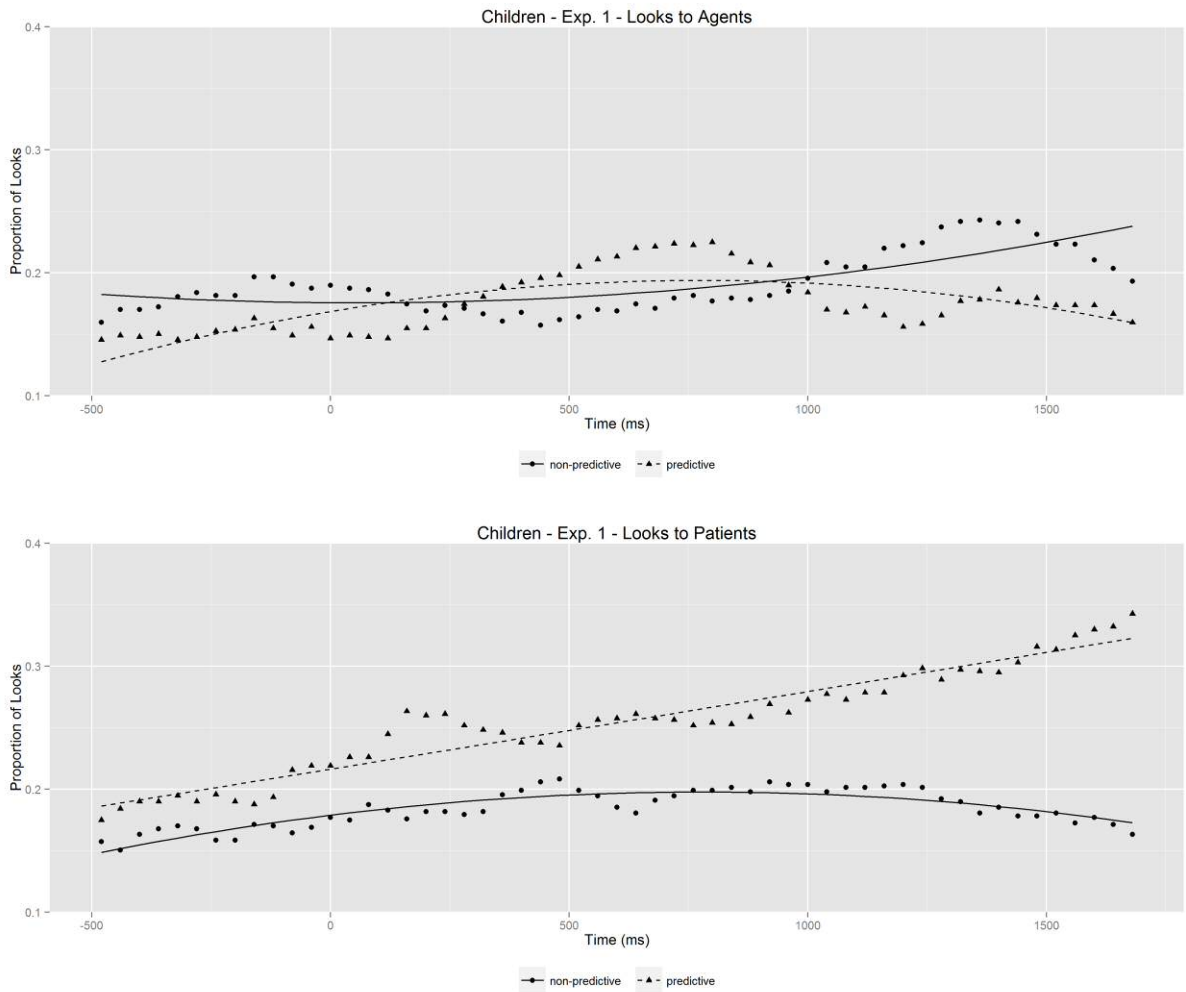
On the basis of this pre-test, we discarded 4 predictive verbs, either because they did not have a clear associated agent/patient (*hug, chase, marry*), or because most children did not understand them (*cure*). For the remaining 12 predictive verbs (see Table 2), associated agents were more often selected as agents than associated patients were (agents, $M = 0.70$, $SD = 0.06$; patients, $M = 0.20$, $SD = 0.11$; $t(11)=8.24$, $p<.0001$), and conversely associated patients were more often selected as patients than associated agents (patients, $M = 0.72$, $SD = 0.32$; agents, $M = 0.07$, $SD = 0.12$; $t(11)=5.45$, $p <.0005$). One of the non-predictive verbs (*look for*) had to be replaced (with *touch*), because it behaved like a predictive verb with the agent and patient we had selected. Therefore, scores are available for only eleven of the twelve non-predictive verbs used in the experiment. Importantly, the difference between the agent-hood of agents and the patient-hood of patients was similar between predictive and non-predictive verbs (non-predictive, $M = -0.03$, $SD = 0.24$; predictive, $M = -0.02$, $SD = 0.37$; $t(21)=0.05$, $p =.955$), and the average difference score for predictive verbs did not differ significantly from zero ($t(11)=-0.19$, $p =.854$). In sum, we replicated the outcome of the adult pre-test with children, confirming that predictive verbs did not elicit a stronger bias towards their associated patients that towards their associated agents. Finally, for non-predictive verbs, associated agents were no more likely to be selected as agents than associated patients (agents, $M= 0.20$, $SD = 0.15$; patients, $M = 0.23$, $SD = 0.24$; $t(10)=0.34$, $p = .741$), and similarly associated patients were no more likely to be selected as patients than associated agents (patients, $M= 0.22$, $SD = 0.17$; agents, $M = 0.24$, $SD = 0.20$; $t(10)=0.20$, $p = .847$).

Experiment 2 Act-out Task

The act-out task used in Experiment 2 yielded additional data on children's preferences about agents and patients associated with predictive verbs, which further confirm the outcome of our norming study. Note that agent-hood and patient-hood scores were now not independent of one another, as children saw only one exemplar of each entity, unlike in the pre-test norming study. Children's actions were analysed using the same criteria as in the child pre-test norming study just described. Once again for predictive verbs the agent-hood of agents was higher than the agent-hood of patients (agents, $M = 0.85$, $SD = 0.12$; patients, $M = 0.05$, $SD = 0.07$; $t(11)=17.01$, $p<.0001$), and the patient-hood of patients was higher than the patient-hood of agents (patients, $M = 0.85$, $SD = 0.17$; agents, $M = 0.02$, $SD = 0.04$; $t(11)=15.06$, $p <.0001$). Importantly, the difference between the agent-hood of agents and the patient-hood of patients was similar for predictive and non-predictive verbs (non-predictive, $M = 0.02$, $SD = 0.18$; predictive, $M = 0.01$, $SD = 0.15$; $t(22)=-0.22$, $p =.828$), and the average difference score for predictive verbs did not differ significantly from zero ($t(11)=-0.19$, $p =.854$). Finally, for non-predictive verbs the agent-hood of agents did not differ from the agent-hood of patients (agents, $M= 0.32$, $SD = 0.20$; patients, $M = 0.26$, $SD = 0.23$; $t(11)=0.65$, $p = .530$), and the patient-hood of patients did not differ from the patient-hood of agents (patients, $M= 0.30$, $SD = 0.26$; agents, $M = 0.27$, $SD = 0.20$; $t(11)=0.26$, $p = .796$).

Experiment 1 Growth Curve Analysis of Children Data

Figure S1. Growth curve analysis (Experiment 1, Children), fitted curves.



This figure plots the same data as shown in the right panel of Figure 1 in the main article: The proportion of looks children directed to the patient (bottom panel) and agent (top panel) over time is shown for the non-predictive and predictive conditions, in a time window ranging from 500 ms before to 1700 ms after verb offset. Note that the observed data are now plotted as filled circles (non-predictive condition) or triangles (predictive conditions). Fitted curves derived from our models including linear and quadratic time terms are superimposed on the observed data as solid (non-predictive) or dotted (predictive) lines. Note how, according to the fitted model, children's looks to the agent (top panel) follow an inverted U-shape pattern after predictive verbs, suggesting they have a tendency to gradually look away from the agent more quickly when they hear a predictive than a non-predictive verb.

Figure S2. Growth curve analysis (Experiment 1), separately for older (>48 months, top panel) and younger children (bottom panel).