COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 36 (2012) 1268–1288

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2012.01249.x

Semantics of the Transitive Construction: Prototype Effects and Developmental Comparisons

Paul Ibbotson,[a] Anna L. Theakston,[b] Elena V. M. Lieven,[c] Michael Tomasello[d]

[a]*Max Planck Child Study Centre, University of Manchester*
[b]*School of Psychological Sciences, University of Manchester*
[c]*Max Planck Institute for Evolutionary Anthropology, Leipzig and School of Psychological Sciences, University of Manchester*
[d]*Max Planck Institute for Evolutionary Anthropology, Leipzig*

Received 20 December 2010; received in revised form 4 November 2011; accepted 12 December 2011

Abstract

This paper investigates whether an abstract linguistic construction shows the kind of prototype effects characteristic of non-linguistic categories, in both adults and young children. Adapting the prototype-plus-distortion methodology of Franks and Bransford (1971), we found that whereas adults were lured toward false-positive recognition of sentences with prototypical transitive semantics, young children showed no such effect. We examined two main implications of the results. First, it adds a novel data point to a growing body of research in cognitive linguistics and construction grammar that shows abstract linguistic categories can behave in similar ways to non-linguistic categories, for example, by showing graded membership of a category. Thus, the findings lend psychological validity to the existing cross-linguistic evidence for prototypical transitive semantics. Second, we discuss a possible explanation for the fact that prototypical sentences were processed differently in adults and children, namely, that children's transitive semantic network is not as interconnected or cognitively coherent as adults'.

Keywords: Prototype; Categorization; Grammar; Semantics

In cognitive linguistics and construction grammar, categories are formed by domain general and well-known processes of categorization such as distributional analysis and analogy (e.g., Goldberg, 2006; Langacker, 1987; Tomasello, 2000). In these approaches, constructions such as the transitive or ditransitive are themselves complex categories or

Correspondence should be sent to Paul Ibbotson, Max Planck Child Study Centre, School of Psychological Sciences, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK. E-mail: paul.ibbotson@manchester.ac.uk

schemas. As such, these constructions should display many of the characteristics of lexical categories, for example, prototype effects.

Since the introduction of the notion of a prototype into the categorization literature by Rosch and her colleagues (e.g., Mervis & Rosch, 1981; Rosch, 1983), the basic idea has been applied to a wide range of linguistic contexts, including lexical semantics (Lakoff, 1987); tense-aspect marking (Andersen & Shirai, 1996; Shirai & Andersen, 1995); relative clauses (Diessel & Tomasello, 2005); questions with long-distance dependencies (Dąbrowska, Rowland, & Theakston, 2009); subject auxiliary inversion (Goldberg, 2006; see also Lakoff & Brugman, 1987; Lambrecht, 1994); and the lexical reorganization that leads to semantic overgeneralization and recovery from overgeneralization, as modeled by an unsupervised neural network (Schyns, 1991). There is also cross-linguistic evidence demonstrating the role of prototypicality in young children's acquisition of linguistic constructions (see review by Ibbotson & Tomasello, 2009) as well as in non-linguistic categories (Boswell & Green, 1982; Ford, 2003; Lasky, 1974). Here we use prototype theory in a novel way: to investigate the semantics of a linguistic construction at two different points in development.

In the classic study of Franks and Bransford (1971) they argued in general that the prototype comprises a maximal number of features common to the category, often ''averaged'' across exemplars. They constructed stimuli by combining geometric forms such as circles, stars, and triangles into structured groups of various kinds. Some of these were then shown to participants who were then later asked whether they recognized these and other shapes they had not seen previously. Importantly, one of the exemplars shown at test contained all of the geometric forms together, an exemplar that had actually never been shown previously (but could be considered the prototype if all of the experienced exemplars were averaged). The participants not only thought that they had seen this prototype, but they were actually more confident that they had seen it than the other previously seen exemplars (or distracter items which they had not seen). Note that these effects were established for an ad hoc non-linguistic category. The first aim of the present study is to see whether this pattern of findings extends to the transitive argument-structure construction, a fundamental building block present in one form or another in all of the world's languages (Hopper & Thompson, 1980; Næss, 2007).

The traditional explanation for ''misremembering'' items is that memory is schematic or constructive (Bartlett, 1932), implying that it is difficult to discriminate what has been experienced from what has been constructed on the basis of inference or knowledge. One developmental prediction of this view is that young children should be less likely to misrecognize true instances as having been presented, because with weaker, less flexible, and less integrated semantic networks they are less likely to make those inferences in the first place. The evidence for this prediction is mixed. Some studies show no developmental difference in false recognition (Liben & Posnansky, 1977; Paris & Carter, 1973); others, an increase in false recognition with age (Johnson & Scholnick, 1979; Prawatt & Cancelli, 1976); still others, a developmental decrease in spontaneous false memory for inferences (Ackerman, 1992, 1994; Reyna & Kiernan, 1994, 1995) and, finally, some studies show different trends in different conditions (Brown, Smiley, Day, Townsend, & Lawton, 1977; Paris & Mahoney, 1974). However, no one to our knowledge has tested the developmental predictions of this constructive view of memory (Bartlett, 1932) on the semantics of a linguistic construction. Just how children's memory affects categorization processes is of fundamental importance. Regardless of one's theoretical position on the innateness of linguistic representations, children have to learn to categorize the particular sound meaning mappings of any one of the 6,000 plus languages they are born into. If the development of children's linguistic semantic networks is characterized as less interconnected to begin with but becomes increasingly organized, coherent, and complex over time (as it is in almost all theories), then it seems reasonable to suggest that developmental differences in this category structure could affect how prototype sentences are processed. Thus, the second aim of using the prototype-plus-distortion methodology is to provide new evidence relating to the structure of the transitive category in adults and children.

If we are going to test the domain-general account of category learning, then we require a working definition of a linguistic prototype. A basic construction in almost all the world's languages is the transitive construction, as in *He kicked the ball*. Building on Hopper and Thompson's (1980, 1984) classic investigation of transitivity across the world's languages, Næss (2007) proposes that the prototypical transitive sentence is characterized by the maximally distinct argument hypothesis: A prototypical transitive clause is one where the two participants are maximally semantically distinct in terms of their roles in the event described by the clause. Thus, this ''motion event'' (Talmy, 1985: 85) is prototypically realized as an agent intentionally instigating an action that directly results in the patient being affected. In line with the gradable nature of concepts advocated by prototype theory, there should be ''better-or-worse'' examples of transitivity. For example, the sentence *John made a vase* semantically overlaps with all the prototypical features described by Næss, while *John dropped the vase accidentally*, *John received the vase*, and *John saw the vase*, are all ''distortions'' from the prototype along the dimensions of agent intentionality, instigation, and affectedness of the patient.

Adapting the prototype-plus-distortion methodology of Franks and Bransford (1971), the current study thus investigates two questions. First, we want to know if the transitive construction will show prototype effects. If it does, this lends novel experimental support for the psychological validity of prototypical transitive semantics, whereas previous accounts have mainly been concerned with how individual languages instantiate aspects of prototypicality, for example, Finnish grammaticalizes the extent to which a patient is affected by a transitive action (Hopper & Thompson, 1980; Næss, 2007). Second, we want to know whether prototype effects with the transitive construction are similar for adults and children. If adults and children process previously unseen prototype sentences differently, it would be suggestive evidence of developmental differences in the way the transitive category structure is represented.

1. Methods

1.1. Participants

1.1.1. Adults

Thirty-nine monolingual English-speaking undergraduates from the University of Manchester took part in the study (29 women, 10 men: age range 18–43 years old, Mean = 18.8).

1.1.2. Children

Thirty monolingual English-speaking children (20 boys; 10 girls: age range 3.9–4.9 years old, Mean = 4.2) took part in this study. A further four children were tested but were excluded from the analysis due to not using the scale during the Recognition phase (1), saying that they remembered all of the test sentences (2) or none of them (1). Testing took place in a quiet room in a nursery in Manchester, UK.

1.2. Test sentences

If we take Næss's prototypical transitive theory with the key features of ''intentionality,'' ''instigation,'' and ''affectedness,'' then sentences can be located according to their semantics in a three-dimensional space (Fig. 1). Intentionality is defined as the extent to which an action is purposefully brought about by the agent (as opposed to being accidental). Instigation is the extent to which the agent is the direct cause of the event, and affectedness is the extent to which the patient is changed in some way by the action. This experiment focuses on a subset of these possible semantic combinations, namely: prototype, instrument, force or involuntary agent (FIA), and neutral.[1]

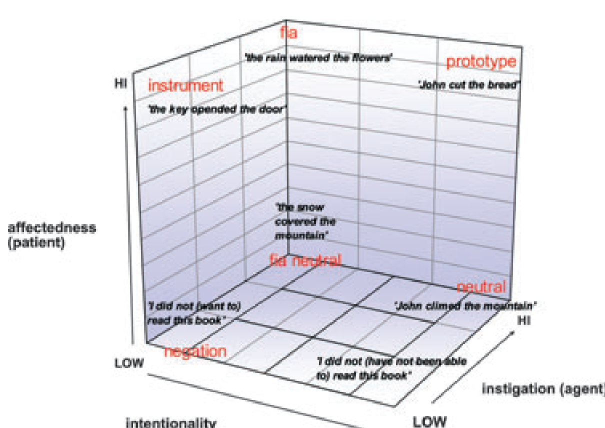

Fig. 1. Semantic dimensions of the transitive construction.

1.2.1. Prototype

According to the maximally distinct argument hypothesis, a transitive clause is prototypically realized as an agent intentionally instigating an action that affects a patient. For example, *John cuts the bread*, *Mary kicks the ball*. Anything that reduces the distinguishability of participants reduces its transitivity; thus, we can think of ''transitivity reducing transformations'' as distortions away from the prototypical semantics of the transitive. Prototype sentences are located in the top right of Fig. 1 because agents score high on intentionality and instigation and patients score high on affectedness.

1.2.1.1. Distortion 1. Instrument: An instrument is an entity being manipulated by another entity in order to achieve an effect on a third entity. For example, *the hammer broke the window*. This is a distortion from the prototype in the sense that the hammer neither has intentionality nor did it instigate the action. Instrument sentences are located in the top left of Fig. 1 because agents score low on intentionality and instigation and patients score high on affectedness.

1.2.1.2. Distortion 2. Force or Involuntary Agent (FIA): This category covers self-driven natural forces with no capacity for volitional action, that is, they bring about events by virtue of their own inherent power. It also includes human or animate actors who are capable of volitional action but whose involvement in the particular action in question is non-volitional. Example: *the wind closed the door* or *John (accidentally) broke the plate*. FIA sentences are located in the top centre of Fig. 1 because agents score low on intentionality and high on instigation and patients score high on affectedness.

1.2.1.3. Distortion 3. Neutral: In this category, the object is not directly involved with the event either in terms of participating in its instigation or in registering its effect. This is different from the above as 'force' or 'non-volitional agents' are not in full control of the event but nevertheless contribute to it being brought about. In 'neutral' the object may be presented as relatively unaffected compared to other, highly affected objects. Example: *Peter climbed the mountain*. In this example, *the mountain* is not affected by the action of climbing to the same extent as *the bread* is affected by the action of cutting in *John cut the bread*. Neutral sentences are located in the bottom right of Fig. 1, because agents score high on intentionality and instigation and patients score low on affectedness.

1.2.1.4. Test sentences for adults: There were two phases to the experiment. In the Acquisition phase, there were six examples of each of the three semantic categories that represent distortions from the prototype (instrument, force/involuntary action, neutral). There were also four ditransitive foils; an argument-structure construction with a different semantic profile to the transitive. In the Recognition phase, there were two examples of each of the distortion categories that appeared in the Acquisition phase which are labeled as ''old'' and two examples of each of the distortion categories that were not in the Acquisition phase labeled as ''new.'' In this sentence type, the lexical items were changed but the underlying semantics is the same with respect to the semantic category they belong to (instrument, force/involuntary agent, or neutral). In the Recognition phase there were also four examples of the prototype which importantly were not in the Acquisition phase. Finally, there were also four new foils in the Recognition phase, included to provide an additional benchmark to judge false-positive recall. The sentences that were used for adults in the study are listed in Appendix A in the order they appeared in the experiment.[2]

Clearly, the more lexical items a sentence in the Recognition phase shares with items in the Acquisition phase, the more confident participants would be that they have heard that sentence before, regardless of semantics. Several investigators (Liben & Posnansky, 1977; Paris & Mahoney, 1974; Small & Butterworth, 1981) have pointed out that in most studies using verbal materials in the prototype-plus-distortion methodology, simple familiarity with the test sentences from previous presentations, rather than consistency with meaning or inference, is sufficient to account for false recognition of true sentences. To control for this methodological confound, it was important that the degree of lexical overlap that *new* and *prototype* (the critical comparison) sentences share with Acquisition items should be comparable. To determine this, we calculated the ratio between Recognition items and Acquisition items, for both their grammatical position and overall frequency (Table 1). For example, for the sentence *Lucy likes cheese,* the frequencies of *Lucy, likes,* and *cheese* were divided by (i) the total number of items in that grammatical slot (Subject, Verb, or Object) in the Acquisition phase and (ii) all items in the Acquisition phase. As Table 1 shows, if there is any difference between new and prototype sentences, it is biased in favor of new items as they have a slightly higher lexical overlap with the sentences during the Acquisition phase than do prototype sentences.

Finally, we needed to be sure that sentences with the prototypical semantics did not have a disproportionally higher frequency than ''deviation'' sentences in English that might lead to higher recognition rates. Table 2 shows the average hits per sentence type from a Google search of the test sentences, ranked from lowest to highest.[3]

Table 1
Ratio of recall items to Acquisition items for adults

	Mean Percent of Lexical Item in the Same Position			Mean Percent of Lexical Item in Any Position			
	Subject	Verb	Object	Subject	Verb	Object	Overall Mean Overlap
New	9.09	4.54	13.62	9.09	4.54	13.62	9.08
Prototype	4.54	0	13.62	4.54	0	13.62	6.05

Table 2
Mean Google hits for sentence types

Sentence Type	Mean Google Hits
Force/involuntary agent	250.665
Instrument	433.583
Prototype	941.75
Foil	1513.125
Neutral	40519.5

All other things being equal, if higher overall frequency leads to higher false-positive recognition, then there should be some relationship between responses and this rank order. As Table 2 demonstrates, the prototype is the median and below the average frequency of all conditions.

1.2.1.5. Test sentences for children: The semantic categories of the sentences were exactly the same for the children as they were for the adults; the only significant difference was that the number of items in the Acquisition phase was reduced from six items per semantic category to four items in order to make the test length more appropriate for the age group. In the Recognition phase, children heard the same number of sentences in each sentence category (i.e., the three distortion categories, foils, and prototype sentences) as adults. A list of test sentences for children appears in Appendix B.

In exactly the same way as we did for the adults, the ratio between Recognition items and Acquisition items was calculated for both their grammatical position and overall frequency (Table 3) and again prototype sentences had a lower mean overlap than new sentences.

1.3. Procedure

For both age groups we matched the procedure as closely as possible; thus, both adults and children were read the test sentences and asked to verbally repeat them. Where the procedure differed was in the motivation for the sentence repetition task and the way in which the recognition scale was modeled, discussed below.

1.3. 1. Adults

1.3.1.1. Acquisition phase: Adults were instructed that they should verbally repeat sentences aloud as read by the experimenter (Acquisition phase, Appendix A). In order to prevent adults anticipating a memory test, participants were told that ''we are interested in recording samples of people reading these sentences aloud for a language processing study.'' This was irrelevant to the aims of the study but was designed to encourage them to pay attention to each sentence as well as to stop them from mentally rehearsing the test items.

1.3.1.2. Recognition phase: After participants had repeated all the sentences they were given a 5-point Likert-style scale: | definitely not heard before | probably not heard

Table 3
Ratio of recall items to Acquisition items for children

	Mean Percent of Lexical Item in the Same Position			Mean Percent of Lexical Item in Any Position			
	Subject	Verb	Object	Subject	Verb	Object	Overall Mean Overlap
New	12.5	6.25	12.5	12.5	6.25	12.5	10.42
Prototype	12.5	0	12.5	12.5	0	18.75	9.37

before | unsure | probably heard before | definitely heard before |. Participants were told by the experimenter ‘‘now I’m going to read some more sentences and I want you to tell me whether you’ve heard them before from the list you have just repeated. I want you to tell me how confident you are that you’ve heard the sentence before. I’d like you place a mark in the box that most closely matches your level of confidence that you have heard the sentence before.’’ The experimenter then read the sentences aloud. After the rating sheets were collected, participants were debriefed as to the true purpose of the task.

1.3.2. Children

1.3.2.1. Acquisition phase: In the first half of the experiment, the child played a sentence repetition game in which s/he was asked to repeat sentences that the experimenter produced (see Appendix B, Acquisition phase). Every time s/he copied a sentence s/he was given a sticker to maintain interest. The sentences that the children were asked to repeat were on average 4.62 words long. None of the 34 children originally tested were excluded on the basis that they failed to repeat the sentences and to the experimenter they all appeared to complete the task with ease.[4] To ensure that the children understood how to use the response scale in the Recognition phase, the experimenter played a brief ‘‘monster game’’ between the Acquisition and Recognition phase to model the scale. The monster game involved the experimenter, with the child, looking at pictures of monsters printed on a card (intended to be analogous to the Acquisition phase). A second, different set of monsters was then used and the experimenter had to remember which ones they had seen from the previous card in a surprise memory game (analogous to the Recognition phase). Some monsters on the second card appeared on the first card and some did not; others had various degrees of resemblance to monsters on the first card. All points of the scale were used during this demonstration. The experimenter then said, ‘‘Now you are going to play the game with words instead of pictures.’’ The Recognition phase then began. It is important to stress that the monster game took place *after* the children heard the Acquisition phase sentences. Thus, the focus on specific differences between monsters in this game could not have led them to focus their attention on the specific lexical items in the Acquisition phase sentences in a way unavailable to the adults as these sentences had already been encoded. This feature of the procedure resulted in a longer gap between the Acquisition phase and the Recognition phase for the children than for the adults (with an ‘‘irrelevant’’ task between), which, if anything, might impair item-specific memory for the children.

1.3.2.2. Recognition phase: In the recognition game the experimenter read out more sentences, some of which were from the previous list and some of which were sentences the child had not heard before (see Appendix B, Recognition phase). The child’s task was then to judge whether s/he thought that s/he had heard the sentence before. Each child did this by placing a counter on a 5-point scale that represented how sure s/he was that s/he had heard the sentence before (developed in Ambridge, Pine, Rowland, & Young, 2008). This was intended to be comparable to the adult 5-point Likert-style scale.[5]

If children thought they had heard the sentence from the previous list, they placed the counter somewhere on the right-hand side depending on how confident they were, if they thought it was new they placed it somewhere on the left-hand side, and if they didn't know they placed it in the middle.

It is worth re-capping that the key comparisons are between (i) new and old sentences; this gives us a baseline accuracy recognition rate and (ii) the more theoretically interesting comparison between prototype sentences and new sentences as these are methodologically identical, in that both are only presented in the Recognition phase.

2. Results

The outcome measure was the confidence rating obtained in the Recognition phase of the experiment. The matrix bar charts in Figs. 2 and 3 show the percentage response for each point on the rating scale across prototype, old, new, and foil conditions. Recall that the old

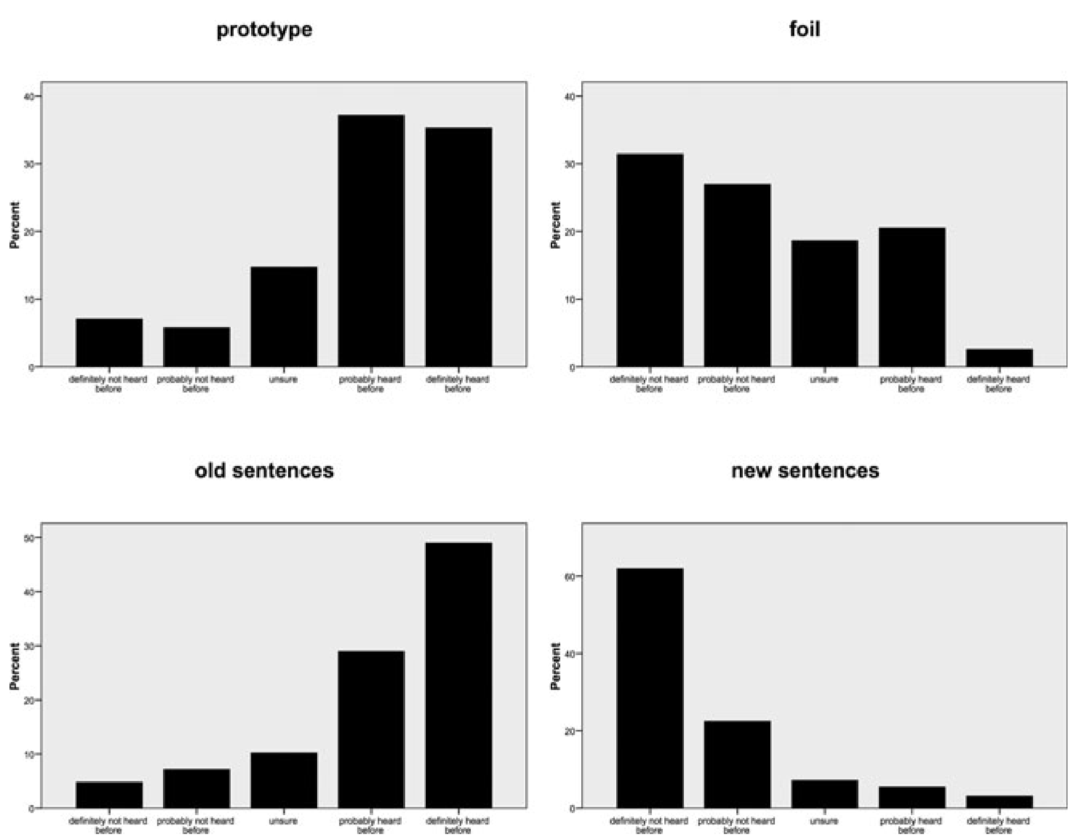

Fig. 2. Percentage confidence rating across conditions for adults.

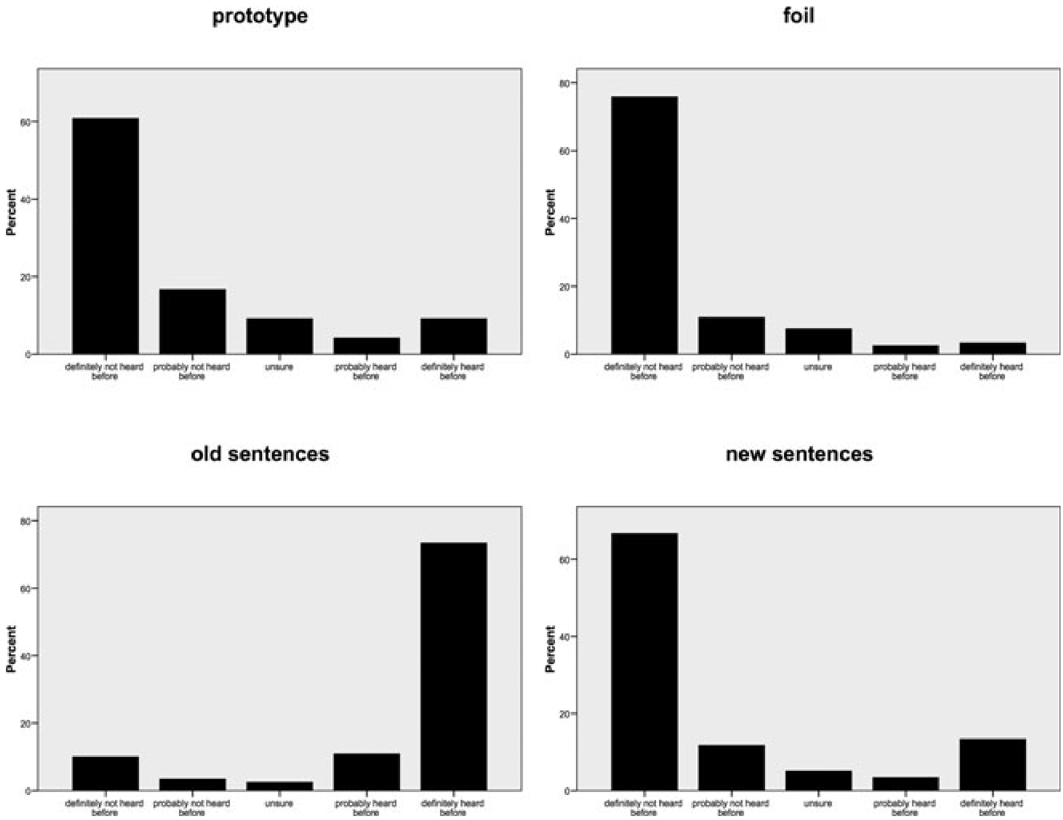

Fig. 3. Percentage confidence rating across conditions for children.

and new conditions comprised sentences from the three distortion categories combined (instrument, force/involuntary action, neutral).

Figs. 2 and 3 show that both adults and children were able to discriminate between new sentences (sentences that were not in the Acquisition phase) and old sentences (sentences that were in the Acquisition phase). The interesting and critical differences lie in the way adults and children treated the prototype sentences, which were only presented in the Recognition phase and so were methodologically identical to new sentences. In summary, adults' responses on prototype items are skewed toward recognizing these sentences as old, whereas children's responses indicate that these are perceived as new.

Figs. 4 and 5 summarize the distributional frequencies of results by taking the median[6] response for each of the rating points across conditions. Fig. 4 clearly shows that adults are treating prototype sentences more similarly to old sentences than they are to new sentences. Conversely Fig. 5 shows that children are on average able to discriminate between old and prototype items.

A Kolmogorov–Smirnov test of normality showed prototype, old, new, and foil conditions to be all significantly non-normally distributed ($p < 0.001$). Wilcoxon Sign Ranked

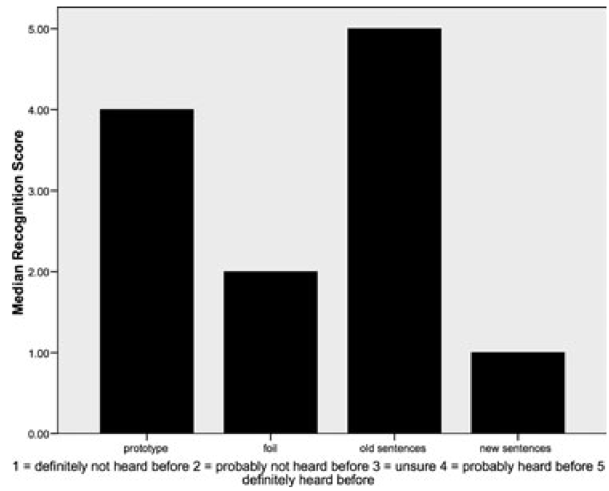

Fig. 4. Median confidence rating across conditions for adults.

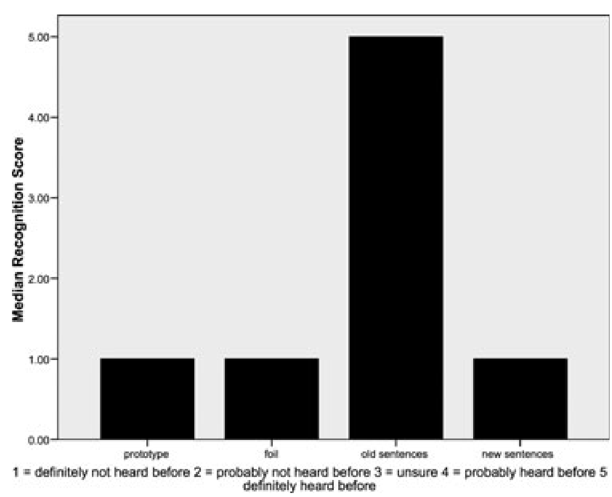

Fig. 5. Median confidence rating across conditions for children.

Tests were therefore used to assess significant differences in the way participants rated the sentences across conditions.

Table 4 shows the comparisons between sentence types. For the adults, all pairs are significantly different, including, crucially, prototype-new, indicating a significantly greater ‘‘recognition’’ of prototype sentences than new sentences. For children, they too were able to discriminate between old and new items, so they were paying attention to the stimuli and knew what they had and had not heard. Unlike adults, there was no significant difference

Table 4
Comparisons between prototype, old, new, and foil conditions

		Prototype-Old	Prototype-New	Prototype-Foil	Old-New
Adults	Median pair	4–5	4–1	4–2	5–1
	Z	−2.45	−9.73	−7.94	−13.90
	Significance (2-tailed)	$p = 0.014$**	$p < 0.001$**	$p < 0.001$**	$p < 0.001$**
	Effect size[a]	0.20	0.78	0.64	0.81
Adults (child matched)	Median pair	3–4	3–1	3–2	4–1
	Z	−5.19	−6.56	−3.14	−9.91
	Significance (2-tailed)	$p < 0.001$**	$p < 0.001$**	$p = 0.002$*	$p < 0.001$**
	Effect size	−0.98	−1.24	−0.59	−1.87
Children	Median pair	1–5	1–1	1–1	5–1
	Z	−8.45	−0.005	−2.43	−8.20
	Significance (2-tailed)	$p < 0.001$**	$p = 0.996$	$p = 0.015$**	$p < 0.001$**
	Effect size	0.77	0.00	0.22	0.75

[a]For the non-parametric statistic used here, estimated as $r = z/\sqrt{N}$.

between prototype and new item median recognition and thus no effect of the prototype eliciting false-positive recall in the children.

Given the comparatively small number of prototype and distortion sentences, we analyzed each sentence to establish if any effect that holds for the group also holds for particular sentences. When we look at the results on a sentence-by-sentence basis (Supplementary Material, Figs. S1 and S2 and Tables S1 and S2), for adults, all prototype sentences are significantly different from new sentences (compare this with the children's result below). The reverse is also true; all new sentences are significantly different from prototypes. However, in comparison to the group analysis, the prototype sentences are individually non-significantly different from old sentences. This is stronger evidence that adults treated prototype sentences more like sentences they had actually heard than methodologically identical sentences they hadn't heard (new sentences). The important point is that, prototype sentences are recognized at a different level to new sentences in both item and group analyses. All old sentences were different from new sentences, suggesting it was not poor memory across the board that was driving the results—false-positive recognition was unique to prototype sentences.

For each prototype sentence type, children showed the opposite pattern to the adults. Prototype sentences were all significantly different from old sentences and non-significantly different from new sentences, meaning that children correctly identified prototype sentences as sentences they had not heard before. All old sentences were significantly different from new sentences, so children, across the board, correctly distinguished between sentences they had and had not heard.

To rule out the possibility that the sequence of the sentences could have driven the results, we analyzed the recognition ratings in terms of the order in which the sentences were presented. Analysis showed the absence of any classic u-shaped primacy and recency curves in either the order of old items presented in the Acquisition phase (Supplementary material Figs. S3 and S4) or in the order of items in the Recognition phase (Supplementary material Figs. S5 and S6) and therefore it is unlikely that that the order of items had a significant effect on the results.

To make the test length appropriate for the 4-year-olds, there were four items per distortion category in the Acquisition phase compared with six for the adults. Obviously recognition performance will be affected by the overall number of items in the Acquisition phase. For that reason, we also tested 28 (different) adults with the same Acquisition stimuli as the children. In this condition there was also a significant difference between the recognition rates of new and prototype items ($Z = -6.56$, $p < 0.001$) albeit the overall effect was weaker, such that the median recognition rates for new items was ''definitely not heard before'' versus ''unsure'' for prototype items. Clearly, as the overall number of items to remember is reduced, this improves performance in the Recognition phase. However, even in this condition the prototype items are being treated differently from other new items by adults in a way that was not the case with the 4-year-olds with the same stimuli.

3. Discussion

We found that adults were lured into false-positive recognition of sentences with prototypical transitive semantics significantly more than other sentences they had also not heard before and thus also had a chance to misremember. This effect was not seen in the 4-year-olds, who were able to discriminate between prototype and new items (recall that neither new nor prototype items were in the Acquisition phase). This effect persists when overall frequency (Table 2) and possible lexical memory effects are accounted for (Tables 1 and 3), thus controlling for the familiarity confound that has been a feature of other studies (for criticisms of this, see Liben & Posnansky, 1977; Paris & Mahoney, 1974; Small & Butterworth, 1981).

The first implication of these results is that they show how a grammatical construction (and one which reflects a fundamental pattern of human experience) can behave in similar ways to non-linguistic categories, for example, by showing graded membership to a category. This is important as the status of the ''language faculty'' has been a controversial issue in linguistics. Generative linguists following Chomsky have claimed linguistic knowledge itself is a separate cognitive faculty, informationally encapsulated and structured according to its own specific principles. Alternatively, the general thrust of the cognitive linguistics enterprise is to render accounts of syntax, morphology, phonology, and word meaning ''consonant with aspects of cognition which are either well-documented, self-evident, or at least highly plausible, and which may well be manifested in non-linguistic activities.'' (Taylor, 2002: 9). The results of the current study would appear to sit most naturally within this second framework.

Regardless of the theoretical approach, one still needs to explain why adults gave false-positive recognition on the items that they did—the sentences with prototypical semantics according to Naess—rather than other sentences in the Recognition phase that they were given equal opportunity to misremember. That this effect was seen in the prototype sentences—and not other new sentences—gives the maximally distinct argument hypothesis

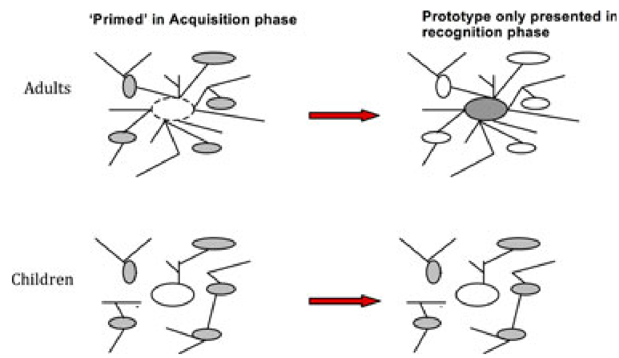

Fig. 6. Schematic representation of radial prototype structure in the Acquisition and Recognition phase.

some experimental support and psychological validity, in addition to the typological evidence with which previous accounts have mainly been concerned (Næss, 2007).

To be clear, the median recognition rate in the prototype condition was not as high as for old items, that is, the adults were still more confident they had heard old items than prototype items so in this sense the prototype effect is not as strong as in Franks and Bransford's original study where prototype items were actually recognized more confidently than all other stimuli, including old items. That said, adults in this study are clearly treating prototype sentences differently from other new sentences and more like sentences they have heard in the Acquisition phase.

The second implication is that adults and children are processing the prototype sentences differently. The results cannot be straightforwardly accounted for by adults' poorer memory for lexical items across the board. Adults and children perform comparably well with new and old items, the crucial difference is that children out-perform adults in recognition accuracy of prototype items.

As discussed in the introduction, one possibility is that the transitive semantic network for the child is not as interconnected as it is in the adult. The top of Fig. 6 schematically shows how the prototype-plus-distortion methodology is proposed to work for a radial prototype category such as the transitive (cf. Goldberg, 2006; Section 8 for discussion of radial categories in the context of Subject-Auxiliary Inversion).

Two relevant features of these radial prototype categories are that there is a strong association between features (having one feature often implies having another) and there is a clustering around the mean (there are more examples ''closer'' to the prototype than there are at the extremes (Collins & Quillian, 1972; Posner & Keel, 1968; Rosch, 1973, 1975). The ovals in Fig. 6 represent clusters of semantic features of the test sentences and the lines show connectivity between shared features. Ovals at the periphery are distortions away from the central prototype, which can be defined as the most representative of the experienced distortions. Both adults' and children's radial categories have a central prototype in Fig. 6 because there is independent evidence that children of a similar age to those tested in this study (and younger) have prototypical knowledge of this construction (e.g., Pyykkönen, Matthews, & Järvikivi, 2009) and more generally show prototype effects in different linguistic constructions and across different languages (Ibbotson & Tomasello, 2009). Activation theories (e.g., Anderson, Budiu, & Reder, 2001) generally account for false word recognition by assuming activation of semantic/encyclopedic neighborhoods (or associatively related links) in a mental network. A similar cognitive process at work in this study would lead to sentences with prototypical semantics being easier to retrieve in the Recognition phase when related but different (distortion) sentences have been activated or primed in the Acquisition phase. A reason for the developmental difference in prototype recognition would therefore be a difference in the structure of the radial category between adults and children. One plausible difference is that children's transitive semantic network is less interconnected to begin with but becomes increasingly organized and interrelated over development. If there is less cognitive coherence to a category for the children, then the transitive semantics will not prime the prototype to the same extent, the children would be less able to ''connect the conceptual dots,'' and thus less susceptible to false-positive recognition. Using a similar prototype-plus-distortion methodology, Boswell and Green (1982) found that when children were asked to remember which figures they had and had not seen—analogous to the Recognition phase in the current study—they successfully avoided false recognition of prototype items, whereas adults did not—an analogous result to the current study. It is worth noting that the same pattern of results was obtained using very different non-linguistic stimuli; the participants had to say whether each shape belonged to one of two ''extraterrestrial space families,'' underlining the domain generality of these effects. Although other studies using the prototype-plus-distortion methodology have found an increase in false recognition with age (Johnson & Scholnick, 1979; Prawatt & Cancelli, 1976), exactly why the adults were not able to recruit exemplar information to the same extent as children remains unclear. One possibility, as suggested by Boswell and Green, is that attention to the ''gist'' or commonalities among figures may function to effectively limit processing of specific item identification.

Although our account of the mechanism underlying the difference between adults and children performance is speculative, it is relatively uncontroversial to characterize semantic/encyclopedic networks as less interconnected to begin with, becoming increasingly organized, coherent, and complex over time. This development obviously has an effect on the way in which sentences are processed. If one adds to this developmental picture the effects of implicit priming (a well established phenomena that this methodology relies on), together they provide one plausible account of why adults and children treated the prototype condition differently. Further studies documenting the same effect are needed to establish whether this account is true.

Given their fundamental and pervasive nature, categorization processes are obvious psychological candidates for cognitive processes that operate over both linguistic and non-linguistic domains. The fact that children showed better verbatim memory than the adults when distinguishing between prototype and new items has implications for usage-based accounts of language acquisition. Indeed, it is a theoretical prerequisite for any usage-based model of language that learners retain significant detailed memories for language over which generalizations and abstractions may be formed (see Goldberg, 2006: 5). We suggest the results are most naturally accommodated in a theoretical framework that views language function, structure, and organization as deeply integrated with processes that are common to the rest of cognition (Goldberg, 2006; Langacker, 1987; Tomasello, 2003).

Notes

1. We have concentrated on the meaning of the transitive construction in this study; prototypicality was defined primarily as a semantic/pragmatic concept, but one with obvious structural implications, of which, the clearest are the traditional grammatical relations of subject and object, which may themselves be generalizations from the core notions of actor and undergoer (Næss, 2007).
2. The sentences we used in this experiment were based on example sentences discussed in *Prototypical Transitivity* (Næss, 2007).
3. An analysis of Child Directed Speech in the Manchester corpus (Theakston et al., 2001), containing 627,645 utterances (1,979,221 words) revealed no matches for any of the test sentences. The prototype is defined with respect to the underlying semantics, not the specific lexical items. The CDS analysis established that none of the test sentences was likely to be remembered more than others due to children's prior linguistic experience with those particular sentences.
4. In a test of phonological working memory, Willis and Gathercole (2001) found that 4-and-a-half-year-olds showed 80% accuracy on repeating six-word-long sentences and 87.5% comprehension accuracy when later asked what the sentences meant (nearly two words longer than the average for this experiment). We do not know for sure the level at which children were processing these sentences, but this evidence is at least suggestive that the sentence length is within the bounds of their processing capacity, and that they would also have ''spare'' capacity to attend to the underlying meaning.
5. Analysis of the responses showed that Adults and Children used the scale in a comparable way. 50% of children used all 5 items of the scale, 33% used 4, and 17% used 3. No child used fewer than 3 points on the scale. The respective figures for Adults were 65%, 24%, and 12%. Again, no adult used fewer than 3 points on the scale.
6. Median is the unit of measurement displayed on the *y*-axis of Figs. 4 and 5 as it is also the measure of central tendency, which the Wilcoxon tests use to establish difference between conditions.

Acknowledgments

The authors wish to thank the staff and the children at Freshfield Nursery School, Manchester UK.

References

Ackerman, B. P. (1992). The sources of children's source errors in judging causal inferences. *Journal of Experimental Child Psychology*, *54*, 90–119.

Ackerman, B. P. (1994). Children's source errors in referential communication. *Journal of Experimental Child Psychology*, *58*, 432–464.

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition*, *106*, 87–129.

Anderson, J. R., Budiu, R., & Reder, L. M. (2001). A theory of sentence memory as part of a general theory of memory. *Journal of Memory and Language*, *45*, 337–367.

Andersen, R., & Shirai, Y. (1996). The primacy of aspect in first and second language acquisition: The Pidgin-Creole connection. In William Ritchie & Tej Bhatia (Eds.), *Handbook of second language acquisition* (pp. 527–570). San Diego, CA: Academic Press.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Boswell, D. A., & Green, H. F. (1982). The abstraction and recognition of prototypes by children and adults. *Child Development*, *53*, 1028–1037.

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. (1977). Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development*, *48*, 1454–1466.

Collins, A. M., & Quillian, M. R. (1972). Experiments on semantic memory and language comprehension. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 117–138). New York: Wiley.

Dąbrowska, E., Rowland, C., & Theakston, A. L. (2009). Children's acquisition of questions with long distance dependencies. *Cognitive Linguistics*, *20*, 571–597.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, *81*, 882–906.

Ford, R. M. (2003). Task variations and attention shifts in young children's category learning. *International Journal of Behavioral Development*, *27* (6), 495–504.

Franks, J. J., & Bransford, J. D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, *90*, 65–74.

Goldberg, A. E. (2006). *Constructions at work*. Oxford, England: Oxford University Press.

Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, *56*, 281–299.

Hopper, P. J., & Thompson, S. A. (1984). The discourse basis for lexical categories in universal grammar. *Language*, *60*, 703–752.

Ibbotson, P., & Tomasello, M. (2009). Prototype constructions in early language acquisition. *Language and Cognition*, *1–1*, 59–85.

Johnson, J. W., & Scholnick, E. K. (1979). Does cognitive development predict semantic integration? *Child Development*, *50*, 73–78.

Lakoff, G. (1987). Cognitive models and prototype theory. In Ulric Neisser (Ed.), *Concepts and conceptual development* (pp. 63–100). Cambridge, England: Cambridge University Press.

Lakoff, G., & Brugman, C. (1987). The semantics of aux-inversion and anaphora constraints. Paper presented at *Linguistic Society of America* meeting, San Francisco, CA.

Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge, England: Cambridge University Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Lasky, R. E. (1974). The ability of six-year-olds, eight-year olds, and adults to abstract visual patterns. *Child Development*, *45*, 626–632.

Liben, L. S., & Posnansky, C. J. (1977). Inferences on inference: Effects of age, transitive ability, memory load, and lexical factors. *Child Development*, *48*, 1490–1497.

Mervis, C., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*, 89–115.

Næss, Å. (2007). *Prototypical Transitivity*. Typological Studies in Language, 72. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Paris, S. G., & Carter, A. Y. (1973). Semantic and constructive aspects of sentence memory in children. *Developmental Psychology*, *9*, 109–113.

Paris, S. G., & Mahoney, G. J. (1974). Cognitive integration in children's memory for sentences and pictures. *Child Development*, *45*, 633–643.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Prawatt, R. S., & Cancelli, A. (1976). Constructive memory in conserving and nonconserving first graders. *Developmental Psychology*, *12*, 47–50.

Pyykkönen, P., Matthews, D., & Järvikivi, J. (2009). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes*, *25*, 115–129. *in press.*

Reyna, V. F., & Kiernan, B. (1994). The development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, *30*, 178–191.

Reyna, V. F., & Kiernan, B. (1995). Children's memory and interpretation of psychological metaphors. *Metaphor and Symbolic Activity*, *10*, 309–331.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.

Rosch, E. (1983). Prototype classification and logical classification: The two systems. In Ellin Scholnick (Eds.), *New trends in conceptual representation: Challenges to Piaget's theory?* (73–86). Hillsdale, NJ: Erlbaum.

Schyns, P. (1991). A modular neural network model of concept acquisition. *Cognitive Science*, *15*, 461–508.

Shirai, Y., & Andersen, R. (1995). The acquisition of tense-aspect morphology: A prototype account. *Language*, *71* (4), 743–762.

Small, M. Y., & Butterworth, J. (1981). Semantic integration and the development of memory for logical inferences. *Child Development*, *52*, 732–735.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (Ed.), *Language typology and syntactic description, vol. III: Grammatical categories and the lexicon* (pp. 57–149). Cambridge, England: Cambridge University Press.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford, England: Oxford University Press.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, *28*, 127–152.

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

Willis, C. S., & Gathercole, S. E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory*, *9*, 349–364.

Supporting Information

Additional Supporting Information may be found in the online version of this article on Wiley Online Library:

Fig. S1. Median recognition responses for adults. Horizontal bars = old sentences; vertical bars = new sentences; red = neutral; orange = instrument; yellow = FIA; blue = prototype; green = foil.

Fig. S2. Median recognition responses for children. Horizontal bars = old sentences; vertical bars = new sentences; red = neutral; orange = instrument; yellow = FIA; blue = prototype; green = foil.

Fig. S3. Adult mean recognition levels on old sentences. 1= definitely not heard before; 2 = probably not heard before; 3 = unsure; 4= probably heard before; 5 = definitely heard before. Red = neutral; orange = instrument; yellow = force/involuntary agent.

Fig. S4. Child median recognition levels on old sentences. 1= definitely not heard before; 2 = probably not heard before; 3 = unsure; 4 = probably heard before; 5 = definitely heard before; red = neutral; orange = instrument; yellow = force/involuntary agent.

Fig. S5. Recognition phase order analysis for adults.

Fig. S6. Recognition phase order analysis for children.

Table S1. Item analysis for adults.

Table S2. Item analysis for children.

Appendix A

List of test sentences for adults

Acquisition Phase	Type
The key unlocked the door	*instrument*
John sees Sophie	*neutral*
The knife cut the bread	*instrument*
He passed Peter the ball	*foil*
The sun warmed the flowers	*force/involuntary agent*
She entered the room	*neutral*
The hammer broke the window	*instrument*
The wind closed the gate	*force/involuntary agent*
John broke the plate accidentally	*force/involuntary agent*
Sophie likes cake	*neutral*
John baked her a cake	*foil*

Appendix A (continued)

Acquisition Phase	Type
Peter crashed the car	*force/involuntary agent*
The oven baked the cake	*instrument*
Sophie read Mary a story	*foil*
The rain watered the flowers	*force/involuntary agent*
The bat hit the ball	*instrument*
Sophie spilt the milk	*force/involuntary agent*
Peter gave the cheese to Mary	*foil*
Peter climbed the mountain	*neutral*
The toothbrush cleaned the teeth	*instrument*
John won the race	*neutral*
Mary heard the song	*neutral*

Recognition Phase	Type
John broke the plate accidentally	*force/involuntary agent (OLD)*
He sang a song	*neutral (NEW)*
James opened the door	*PROTOTYPE*
Lucy loves cheese	*neutral (NEW)*
She made a cake	*PROTOTYPE*
The sponge cleaned the car	*instrument(NEW)*
The wind closed the gate	*force/involuntary agent (OLD)*
She asked Sophie a question	*foil*
The rain filled the bucket	*force/involuntary agent (NEW)*
The hammer broke the window	*instrument(OLD)*
Lisa ate the sandwich	*PROTOTYPE*
The axe chopped the wood	*instrument(NEW)*
She burnt the toast accidentally	*force/involuntary agent (NEW)*
John showed him the room	*foil*
John sees Sophie	*neutral (OLD)*
He sliced the bread	*PROTOTYPE*
She entered the room	*neutral (OLD)*
He gave her the flowers	*foil*
The knife cut the bread	*instrument (OLD)*
Mary sent her a letter	*foil*

Appendix B

List of test sentences for children

Acquisition Phase	Type
The key unlocked the door	*instrument*
John sees Sophie	*neutral*
The knife cut the bread	*instrument*

Appendix B (continued)

Acquisition Phase	Type
He passed Peter the ball	*foil*
The sun warmed the flowers	*force/involuntary agent*
She entered the room	*neutral*
The hammer broke the window	*instrument*
The wind closed the gate	*force/involuntary agent*
Sophie likes cake	*neutral*
John sent her a letter	*foil*
Peter crashed the car	*force/involuntary agent*
Sophie read Mary a story	*foil*
The rain watered the flowers	*force/involuntary agent*
Peter gave the cheese to Mary	*foil*
Peter climbed the mountain	*neutral*
The toothbrush cleaned the teeth	*instrument*

Recognition Phase	Type
The rain watered the flowers	*force/involuntary agent (OLD)*
He sang a song	*neutral (NEW)*
James opened the door	*PROTOTYPE*
Lucy likes cheese	*neutral (NEW)*
She made a cake	*PROTOTYPE*
The sponge cleaned the car	*instrument(NEW)*
The wind closed the gate	*force/involuntary agent (OLD)*
She asked Sophie a question	*foil*
The rain filled the bucket	*force/involuntary agent (NEW)*
The hammer broke the window	*instrument(OLD)*
Lisa ate the sandwich	*PROTOTYPE*
The axe chopped the wood	*instrument(NEW)*
The snow covered the window	*force/involuntary agent (NEW)*
John showed him the room	*foil*
John sees Sophie	*neutral (OLD)*
He sliced the bread	*PROTOTYPE*
She entered the room	*neutral (OLD)*
He gave her the flowers	*foil*
The knife cut the bread	*instrument (OLD)*
Mary sent her a letter	*foil*