Passives in first language acquisition: What causes the delay?

Susannah Kirby
University of British Columbia, suki@ibiblio.org
Passives in first language acquisition: What causes the delay?

Abstract
The passive construction is a late acquisition in child language. In this paper I evaluate the claim that difficulty with noncanonical semantics, rather than non-mature subject-object A-chains, underlie young children's poor performance on the passive. In a series of truth-value judgment tasks, 4- and 5-year-old English-speaking children were tested on their comprehension of matrix passives and passives embedded under raising-to-object (RO: want, need) and object control (OC: ask, tell) verbs. RO passives (Olivia wanted/needed Scott to be kissed by Misha) entail object-subject-object A-chains, but allow semantic patients to surface as syntactic objects; OC passives (Olivia asked/told Scott to be kissed by Misha) and matrix passives (Scott was kissed by Misha) involve similar A-chains but do not result in this syntactic-semantic configuration. I found that although 4-year-olds failed to comprehend matrix passives and passives embedded under OC verbs, they correctly interpreted passives under RO verbs. (5-year-olds performed above chance in all tasks.) I propose a “semantic scaffolding” account of children's comprehension of the passive that rests on the prototypicality of subjects being agents and objects being patients, arguing against the view that difficulty with passives results from an inability to form A-chains.
Passives in first language acquisition: What causes the delay?

Susannah Kirby

1 Passive Acquisition: The data and the debate

Lexical items and structures that are late to arise in language acquisition are naturally of interest to researchers, as they constitute the fundamental evidence behind the developmental problem of language acquisition (Hyams, 1988). If language is guided, as many believe, by some internal structure or module such as Universal Grammar, then why should there be any delay?

The passive construction is a prime example of just such a structure. Naturalistic data from a number of languages, including English (Horgan, 1978), French (Sinclair, Sinclair, and Marcellue, 1971; cited in Suzman, 1985), German (Mills, 1985), and Hebrew (Berman, 1985), has indicated that spontaneous full passives, which have an agent by-phrase, as in (1), are quite rare in child language until sometime between the ages of 4;0 (English, German) and 8;0 (Hebrew).

(1) Scott was kissed by Misha

The passive construction is a marked one in many adult languages (Keenan, 1985); it appears to be less common (or less “basic”) in adult speech, and thus constitutes a smaller proportion of the input to children than does the active. However, while input frequency often contributes to order of acquisition (e.g., Newport et al., 1977), frequency is not completely deterministic in language learning (e.g., Tanz, 1980). The questions then remains: what causes the acquisitional delay?

Two prominent analyses have been put forth to answer this question. The linguistic maturation hypothesis (initially proposed by Borer and Wexler, 1987, 1992, and later used in the analyses of data presented by Horgan, 1978; Mills, 1985, and Pierce, 1992, among others) asserts that certain linguistic constructions made available by UG (including those involving A-chains, such as the passive) are not immediately accessible to the child, but must rather mature over time, just like secondary sex characteristics. Before maturation, which for the passive should occur around age 4;0, the A-Chain Deficit Hypothesis (Babyonyshev et al., 2001) suggests that “non-trivial,” or subject-object, A-chains are ungrammatical for the child, and predicts that utterances involving them will not appear in spontaneous speech.

In contrast, a second analysis of the passive’s late appearance in child speech builds on the notion that semantics play a significant foundational role in language acquisition. Proponents of the semantic bootstrapping hypothesis claim that children have innate biases about semantics which aid in the acquisition of syntax (Bowerman, 1973; Grimshaw, 1981; Pinker, 1984). While the details of the particular account vary slightly from researcher to researcher, there are two major components generally hypothesized to be guided by bootstrapping: first, the categorization of words into syntactic classes, and second, the identification of syntactic functions within an utterance. Both are thought to be connected inextricably to the semantics (both linguistic and cognitive) of an utterance.

Semantic bootstrapping comes into play for the passive question with its assumption that children rely on canonical notions of “subject” and “object” in their interpretation of all sentences. Specifically, children are biased to associate particular syntactic positions with particular thematic roles in a consistent way: “subjects” should be agents, and “objects” should be patients/themes. This assumption, along with a linguistic description of an observable event, should aid children in determining the word order of the ambient language. For instance, upon hearing The girl is kissing the boy in the context of a kissing event, the child will be able to determine that the word order of English is SVO, since the agent, mentioned first, maps to the subject, and the patient, mentioned last, maps to the object.

(2) OBSERVABLE EVENT: girl kisses boy

The-girl is-kissing the-boy
AGENT(subj) ACTION(verb) THEME(obj)

CONCLUSION: Language is SVO
In this task children will first need to rely on utterances that display canonical semantics (e.g. active-voice clauses with agentive-transitive verbs) as they build up a basic level of knowledge; only after building this foundation will it be possible to extrapolate, thereby identifying subjects and objects that do not conform to this canonical scheme.\(^1\)

In general, though, the assumption of such a thematic association is a good strategy, since syntactic and semantic roles often align in adult language. As a result, agent-subjects and patient/theme-objects are a crosslinguistic norm (Fillmore, 1968; Keenan, 1976; Dowty, 1991; Baker, 1997). But the passive construction results in a mismatch between syntax and semantics, such that themes, and not agents, now surface in subject position. And in fact, just as the semantic bootstrapping approach would predict, children who have not yet acquired the relevant syntax interpret “reversible” passives, which have two animate arguments as if they were active sentences, construing the first NP as the agent of the action (Bever, 1970; Maratsos, 1974; Mills, 1985).

(3) a. **Utterance:** Scott was kissed by Misha  
   b. **Child’s interpretation:** Scott kissed Misha

With such evidence as support, Hyams et al. (2006) have argued that children’s trouble with the passive stems precisely from this marked arrangement of theta roles. They propose the *Canonical Alignment Hypothesis* (CAH), which states that in the earliest grammar, any external argument (which is typically an agent) must map onto subject position (Spec, IP). In short, it is not A-chains *per se* that give children trouble, but instead only those A-chains whose configuration violates the CAH.

The two accounts detailed here set up a neat dichotomy, in terms of their predictions: if children’s failure on passive comprehension is simply due to problems overcoming a syntax-semantics mismatch, rather than trouble forming or interpreting A-chains, children should perform better on passives if the prototypical alignment of theta roles were to obtain. Fortuitously enough, such passives do in fact exist: passive clauses embedded under raising-to-object verbs (RO-EPs) circumvent CAH violations, thereby providing an utterance type with which to test this prediction. In RO-EPs, semantic objects surface as syntactic objects, for instance, in (4), Scott is the semantic patient of *kiss* and the syntactic object of *want*.

(4) Olivia wanted/needed Scott, \([t_t \text{ to be kissed} t_t \text{ by Misha}]\)

In contrast, passives under object control verbs (OC-EPs) function syntactically just like matrix passives (MPs); the embedded clause, with its PRO subject, is essentially a MP with regards to its syntax-semantics alignment, and mismatch.

(5) Olivia asked/told Scott, \([\text{PRO}_j \text{ to be kissed} t_j \text{ by Misha}]\)

As a result, the accounts described above can be assessed by comparing children’s comprehension of EPs with their comprehension of MPs. If trouble with the passive is due to CAH violations, children’s performance on RO-EPs should be better than their performance on MPs, which in turn should be roughly equal to (or better than) their performance on OC-EPs (depending on whether the biclausality of the latter results in a higher processing load, and thus lower performance). On the other hand, if children’s trouble with passives stems from non-trivial A-chains, their performance on RO-EPs should be equal to (or poorer than) their performance on OC-EPs, which should be be equal to (or poorer than) their performance on MPs (subject to the same caveat on biclausality).

(6) Predictions:  
   a. **Trouble due to CAH violation**:  
      RO-EPs > MPs ≥ OC-EPs

\(^1\)The same caveat applies to syntactic categorization. Children will first need to rely on lexical items that respect the “canonical structural representation” (Grimshaw, 1981): nouns that map to objects, verbs that map to actions, and so forth. A store of such nouns and verbs will ultimately allow children to identify and categorize words which do not label tangible or observable elements, on the basis of their morphosyntactic privileges of occurrence.
2 Method

The experiments described in the following sections tested the working hypothesis that children’s difficulty with the passive construction is a result not of a pre-mature grammar which is unable to parse non-trivial A-chains, but rather the consequence of the thematic mismatch posed by this construction. I tested this hypothesis with English-speaking children (ages 4–5) using two truth-value judgment tasks.

2.1 Experiment 1: Matrix Passives

In Experiment 1, children ages 4–5 were tested on their ability to interpret matrix (non-embedded) passives. 32 monolingual English-speaking children (ages 4;1.15–5;11.15) were recruited from the region around Chapel Hill, North Carolina to take part in the study. The 4-year-old group contained 8 boys and 8 girls and had a mean age of 4;5 (range: 4;1.15–4;11.12); the 5-year-old group contained 7 boys and 9 girls, and had a mean age of 5;5 (range: 5;0.18–5;11.15). Participants had no known linguistic (speech or hearing) impairments, or other cognitive or developmental delays. Children received a small token gift for their participation.

Experiment 1 comprised a truth-value judgment task (TVJ; Crain and McKee 1985; Gordon 1996), in which each child heard stories and saw them acted out with small figurines. After each story, the child listened to a puppet make a comment about the story. The child was asked to reward the puppet for his correct comments by “feeding” him a plastic orange, and to punish him (i.e., provide him with a less attractive reward) for his incorrect comments by “feeding” him a plastic piece of lettuce. In the latter case, children were also asked to explain why the puppet was wrong. Children received test items separated by filler items, which were included to check for answer biases and/or inattention to the task.

The test items in Experiment 1 were all full reversible passives (i.e., they included an agent by-phrase, and had 2 animate arguments). All arguments were full lexical NPs, rather than pronouns. An example story appears in (7), and test items appear with their target answers in (8). Those test items with target “false” answers had their arguments swapped, from what would have been a correct description of the event.

(7) Example story: Passive
This woman wanted to draw a picture of someone. The nurse was busy, and she had already drawn a picture of farmer, so she decided to draw the policeman.

Experiment: What happened?

Puppet: The woman was drawn by the policeman (F)

(8) Passive test items
a. The pig was sent by the farmer (T)
   b. The woman was drawn by the policeman (F)
   c. The farmer was picked by the sheep (F)

The outcome of Experiment 1 was that only 5-year-olds behaved in an adultlike manner in their interpretations of matrix passive utterances. That is, the 5-year-olds, but not the 4-year-olds, appeared to have mastered matrix passives. The results of Experiment 1 are given in Table 1.

The data was analyzed by age group (4, 5). A series of logistic regressions (with the standard error adjusted for multiple observations within subjects) were performed to compare the number of...
Table 1: Performance on Matrix Passives

<table>
<thead>
<tr>
<th>Age</th>
<th>Percent Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>64.6</td>
</tr>
<tr>
<td>5</td>
<td>79.2*</td>
</tr>
</tbody>
</table>

*p < .01

The test of the hypothesis that children performed at a chance level on matrix passive constructions was rejected for 5-year-olds (z = 3.08, p = .0020), but not for 4-year-olds, who did not perform above chance (z = 1.83, p = .0667).

In short, children performed as would be predicted, given their age and native language. However, given the details of the semantic bootstrapping hypothesis, it is possible that children who failed on the task in Experiment 1 could still succeed on interpreting embedded passives (RO-EPs), in which syntactic and semantic roles align. Testing children who fail to comprehend matrix passives on their comprehension of RO-EPs will serve as a test for the hypothesis that it is specifically the non-trivial (object-subject) A-chain in the passive construction that causes children trouble. The A-chains in RO-EPs (e.g. \textit{Olivia wanted Scott, \{t to be kissed \} by Misha\}) are exceedingly non-trivial, in that they constitute object-subject-object A-chains. As a result, if the maturation account is right, RO-EPs should prove just as problematic (if not more so) for “pre-mature” children as do matrix passives. On the other hand, if the mismatch account is correct, children should have less trouble with embedded passives than they do with matrix passives.

2.2 Experiment 2: Embedded Passives

In Experiment 2, children were tested on their ability to interpret passives embedded under raising-to-object and object control verbs. The participants and methods in Experiment 2 were the same as those in Experiment 1.

The test items in Experiment 2 each had a pronoun matrix subject (to reduce processing load) and a lexical NP embedded subject and oblique object. Each child received either want/ask items, or need/tell items (i.e., items for one RO and one OC verb). Example stories appear in (9), and test items appear with their target answers in (10).

(9) Example stories: Embedded passives

RO: Winnie the Pooh said to Tigger, “Somebody should call Elmo and invite him over to play with us. Do you have his telephone number? Can you call him up?” Tigger said, “Yes, I can call Elmo,” and he went to call him and invite him over.

Experimenter: What did Winnie the Pooh do?

Puppet: He needed Tigger to be called by Elmo (F)

OC: Elmo said to Cookie Monster, “Dora really loves hugs. Can you go and hug her? She’ll give you a big hug back.” Cookie Monster agreed, and went and gave Dora a big hug.

Experimenter: What did Elmo do?

Puppet: He asked Dora to be hugged by Cookie Monster (F)

(10) Embedded passive test items

W: She \textit{wanted} the tiger to be tickled by the bear (T)
She \textit{wanted} Clifford to be drawn by Patrick (T)
He \textit{wanted} the horse to be seen by the farmer (F)

N: He \textit{needed} Clifford to be fed by Shrek (T)
He \textit{needed} Tigger to be called by Elmo (F)
She needed Cookie Monster to be photographed by Elmo (T)

A: He asked Dora to be hugged by Cookie Monster (F)
Shes asked the cat to be licked by the dog (F)
He asked the goat to be washed by the cow (T)

T: He told the horse to be ridden by the bear (F)
She told the policeman to be sniffed by the dog (T)
He told the cat to be scratched by the dog (F)

Note that if children interpreted the EPs as actives, they would fail on the task (He needed/told [Clifford to be fed by Shrek] ≠ He needed/told [Clifford to feed Shrek]).

The outcome of Experiment 2 was that only 5-year-olds performed above chance in their comprehension of passives embedded under OC verbs. However, both 4- and 5-year-olds performed significantly above chance in their interpretations of passives embedded under RO verbs. The results of Experiment 2 are given in Table 2.

Table 2: Performance on Embedded Passives

<table>
<thead>
<tr>
<th>Age</th>
<th>Percent Correct</th>
<th>RO</th>
<th>OC</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>75*</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>79.2*</td>
<td>64.6*</td>
<td></td>
</tr>
</tbody>
</table>

As in Experiment 1, the data was analyzed by age group, with logistic regressions performed to compare the number of correct responses per age group to a chance level of performance. For the OC-EP items, the test of the hypothesis that children performed at a chance level was rejected for the 5-year-olds, but not the 4-year-olds (4: z = 0.00, p = 1.0000; 5: z = 1.99, p < .0461), but a similar hypothesis was rejected for both age groups for performance on RO-EP items: here, both 4- and 5-year-olds’ performance was significantly above chance (4: z = 3.30, p = .0010; 5: z = 4.41, p < .001).

Thus, contrary to the maturation account, but as predicted by the mismatch account, both groups performed in an adult-like manner on interpreting RO-EPs. Note that this was even true for 4-year-olds, who had failed to comprehend matrix passives at above-chance levels! The data from Experiment 2 thus dovetails with independent evidence suggesting that children have no trouble with A-chain formation (Köppe, 1994; Fox and Grodzinsky, 1998; Guasti, 2002), as well as with adult aphasia studies indicating that processing load increases with non-canonical theta role ordering (Caplan and Hildebrandt, 1988).

Table 3 presents the data from Experiments 1 and 2 in comparative form, for easy reference.

Table 3: Comparative Performance on Passives (Percent Correct)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>75*</td>
<td>64.6</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>79.2*</td>
<td>79.2*</td>
<td>64.6*</td>
</tr>
</tbody>
</table>

It is crucial to note that success on comprehension of RO-EPs was not significantly correlated with success on matrix passives ($r^2 = .076, p = .1273$); thus, it was not simply the case that the only children who succeeded on RO-EP items in Experiment 2 were those who succeeded in their comprehension of MPs in Experiment 1.
In summary, children’s comprehension of RO-EPs is equal to or better than their comprehension of MPs, which in turn is better than their comprehension of OC-EPs. These results support the mismatch account over the maturation account.

3 Conclusions

In the experiments presented here, English-speaking 4-year-olds failed to comprehend matrix passives or passives embedded under object control verbs in an adultlike way, instead interpreting them as active-voice utterances (to at least some extent). But even these younger children performed significantly above chance in interpreting passives embedded under raising-to-object verbs. These results are especially noteworthy when considered in light of the facts that RO-EPs are both longer and more syntactically “complex” than MPs, and that these two features are generally assumed to be correlated with a higher processing load.

Children’s pattern of performance in these two experiments are evidence of a cluster of strategies that I call semantic scaffolding (Kirby, 2009). The semantic scaffolding hypothesis submits that children rely strongly on the semantics of the utterances in a number of ways, to aid in interpretation until syntactic representations and processing power reach adult levels. The semantic scaffolding strategy subsumes the bias towards canonical syntax-semantics matches, as claimed by the CAH, and also makes claims about what children expect the general shape and character of basic clauses to be.

The patterns seen in children’s performance on comprehension of the three types of passives tested here provide strong support for the semantic scaffolding hypothesis, and especially for the notion that children’s interpretation of the passive does not rest on the (in)ability to form A-chains, but instead hinges on the match between syntactic and thematic roles. Furthermore, the results of Experiment 2 seem to indicate that the CAH acts as a kind of filter, only looking at the head and the tail of the chain, and not at any intermediate traces, which would result in a violation of canonical alignment in RO-EPs. These utterances involve D-structure objects which surface as S-structure objects, even though they pass through an external argument position between the two levels of representation.

Apart from its ability to account for the data presented above, the semantic scaffolding approach has several other benefits. First, the effects of semantic scaffolding can be seen elsewhere. For instance, in a study conducted by Chapman and Kohn (1977), children were better able to interpret sentences with agentive subjects and inanimate theme objects (i.e., Dowty-style proto-agent subjects and proto-theme objects) than sentences with the opposite configuration. Similarly, Marantz (1982) found that children have less difficulty learning novel verbs which take preverbal agents and postverbal patients than those with the opposite alignment. And in a study by Lempert (1989), children ages 2;6–5;3 produced more novel passive utterances if they had been trained on stimuli with animate patients. Since patients surface as subjects in passives, the findings from all three studies are consistent with the proposal that children have a concept/category of subject which includes agency/animacy among its chief characteristics. Finally, it may be fruitful to interpret the so-called “perspective maintenance” effects seen in the processing of subject- and object-extracted relative clauses (i.e., SS>OO,OS>SO clauses; MacWhinney, 1982) as being “mismatches” of case which violate alignment.

A second benefit to the semantic scaffolding approach is that it may explain the get-passive conundrum. Researchers have been interested in the discrepancy in children’s comprehension of get-passives and be-passives; specifically, children who fail on tests of full be-MPs sometimes succeed on get-MPs (Crain et al., 1987; Crain and Fodor, 1989; Crain, 1991; Fox and Grodzinsky, 1998). While the syntax of these two constructions appears to be identical (Fox and Grodzinsky, 1998), there is a subtle semantic distinction between the two, with the result that the patient-subjects in get-passives are more “active” in their semantic role. Compare the be-passive in (11a) with the get-passive in (11b).

(11)  a. Misha was kissed  
b. Misha got kissed
When we question what the patient-subject Misha did in each case, we find that the be-passive is congruent with a situation in which all Misha did was to sit still during the kiss. In contrast, in the get-passive, Misha may have played a much more active role in bringing the kissing event to fruition: for instance, she may have flirted conspicuously, used her feminine wiles, or even just brazenly asked someone to kiss her (cf. Misha got herself kissed).

Abstracting away from the subject’s role in the event, there is also an aspectual difference in the auxiliary verbs, since be is a stative verb and get is eventive. As a result, get-passives more closely match children’s default expectation that the verb in a NVN sequence represents an action (a proposal made by members of the semantic bootstrapping camp; again, see Bowerman, 1973; Grimshaw, 1981; Pinker, 1984). Combined with observations about the subject’s involvement in the event, the result is that get-passives are simply less “passive-y,” and it may be precisely this semantic distinction, rather than any syntactic or pragmatic one, that allows children to master them before they master the corresponding be-form.

The final benefit which I will discuss here relates to cognitive cost. Semantic scaffolding makes use of the concept of “processing load” without abandoning a grammatical account of acquisition, and it also provides a more refined definition for this notion than is traditionally found in the literature. The passive construction clearly carries with it some cognitive cost: in another experiment not detailed here (Kirby, 2009), the same children were tested on their comprehension of active-voice RO and OC utterances (e.g. He wanted the hippo to rub the lion). Children performed above chance on these items (4: 87.5% correct, $z = 4.26, p = 0.0001; 5: 91.2%$ correct, $z = 3.93, p < .0001$) and significantly better on actives than EPs (RO->RO-EP: $\chi^2 = 9.94, d.f = 1, p = .0016; OC->OC-EP$: $\chi^2 = 19.02, d.f = 1, p = .0001$). And length (biclausality) also appears to play a role; recall that children’s performance on MPs was better than their performance on OC-EPs, although the syntax of the passive clauses in these utterances is essentially identical, and that 5-year-olds’ performance on RO-EPs was equivalent to their performance on MPs (see Table 3).

However, the picture that emerges here is that cognitive/processing load is not monotonically correlated with either sentence length or transformation complexity. If it were, children should not have been able to comprehend RO-EPs at the rates that they did.

In contrast with the imprecise (and apparently overly simplistic) notion that cognitive load simply results from length or transformations, the semantic scaffolding hypothesis equates processing load with “novelty of construction,” including the noncanonical assignment/appearance of thematic roles, and the basic shape of utterances as monoclausal. It also makes the prediction that utterances which do not violate children’s default assumptions about the syntax-semantics interface will serve a protective function (i.e., will not derail children) in the face of high-load transformations, including the passive. Moreover, it explains how UG may constrain and guide non-adultlike performance, even in the face of processing limitations. Children’s “errors” should tend in the direction of their initial biases about the semantics of utterances; in the case presented here, this involves non-adultlike interpretation of passive-voice utterances as actives.

Child data is often helpful in how it contributes to debates in adult syntax, and studies on raising and control are certainly no exception (Kirby et al., Submitted). A current proposal in the adult literature is that both raising and control are instances of movement, and that the only distinction between the two classes rests with the number of thematic roles they assign (Hornstein, 1999, and many others). RO verbs (12) embed a subject which receives a single theta role, from the lower predicate, while OC verbs (13) themselves assign the embedded subject a second thematic role.

(12) RO: Olivia wanted/needed Scott, to be kissed Scott, by Misha

(13) OC: Olivia asked/told Scott, to be kissed Scott, by Misha

In both utterances above, Scott is hypothesized to be generated as the semantic object of kiss, and to promote (for Case reasons) to the subject of the embedded infinitive. The NP moves again to become the object complement of matrix verb, and everything in the double-chain except the overt NP is deleted at PF.

But conflating RO and OC would seem to predict equivalent performance by children on both constructions, and this prediction is not borne out in the experiments presented above. Moreover, in further experiments not detailed here, children show asymmetries on sentence judgments of active...
RO and OC utterances (Kirby, 2009); this is true for judgments on the semantic felicity of embedded inanimate subjects (*Misha wanted/*asked the hug to comfort Olivia*), as well as for judgments on the grammaticality of embedded expletive subjects (*Misha needed/*told there to be enough yogurt for Olivia*).

Unfortunately, the results of these experiments are not conclusive in evaluating the movement theory of control, precisely because RO and OC are still assumed to differ with respect to the number and configuration of theta roles involved. The phenomenon of a single NP bearing multiple distinct theta roles may be less canonical than a situation in which there is a one-to-one correspondence between NPs and semantic roles, and the former may therefore require greater processing resources than the latter. However, because children perform at ceiling on active RO and OC utterances (Kirby, 2009), the sheer number of thematic roles alone is unlikely to be the deciding factor.

As a result, it would appear that the current data can neither support or refute this proposal. Future research should continue to explore children’s comprehension and processing by carefully controlling for all variables except these two theta role configurations.

**References**


Kirby, Susannah, William D. Davies, and Stanley Dubinsky. Submitted. Up to d[eb]ate on raising and control.


