The cup on the table is green: Children’s Comprehension of Embedded PPs

Erin Hall and Ana T. Pérez-Leroux

1. Introduction

How do children learn to integrate and interpret structurally complex noun phrases? While most of the literature on the acquisition of embedding deals with clausal subordination (de Villiers & Roeper 2016), there are some studies that target a variety of complex NP constructions beyond relative clauses, including pseudo-partitives, recursive possessives, and prepositional modifiers. These studies provide evidence that children experience difficulties with the comprehension of noun phrases (NPs) embedded inside other NPs. Foushee and colleagues, for instance, claim that “despite its systematicity, children do not automatically appreciate the mapping between syntax and semantics” (Foushee, Falkou & Li 2017:1). We argue that this claim is too strong, and that children are sensitive the core semantic consequences of structure. The goal of the present study is to find the simplest possible test of the syntax-semantic mapping of NP-embedded NPs. We developed a novel experiment assessing the comprehension of prepositional phrase (PP) modifiers, as in (1). Concentrated on the stage in which children’s use of such complex NPs emerges but it is not yet productive, our study examines whether children understand that NPs embedded under other NPs are no longer referentially accessible to predication, unlike coordinated NPs, as shown by the contrast between (1) and (2).

(1) The cup on the table is green.

(2) The cup and the table are green.

We first consider the semantics of PP embedding, then review previous studies testing comprehension of complex NPs. Then we present a study of three- and four-year-old English-speaking children. We close by discussing some

*Erin Hall, University of Toronto, e.hall@mail.utoronto.ca. Ana T. Pérez-Leroux, University of Toronto, at.perez.leroux@utoronto.ca. We gratefully acknowledge the support of SSHRC Insight Grant 435-2014-2000 Development of NP complexity in children to A.T. Pérez-Leroux and Y. Roberge, as well as the helpful discussion provided by Susana Bejar, Diane Massam, Yves Roberge, and other members of the Complexity & Recursion Project. Special thanks to Natalia Docteur and Mathusha Sanjeevan for support with testing and recruitment.

implications of children’s understanding of PP modification for general questions about the co-development of syntax and semantics.

2. The semantic consequences of structural configurations: The case of NP embedding

Current linguistic theory posits highly abstract links between form and meaning, assuming a lot of structure is implicit. This makes it challenging to evaluate strong proposals about children’s innate understanding of the mapping between sentence form and interpretation, particularly since the markers of structure are lexical elements whose semantic valuation must by necessity be learned. The general mapping of form and meaning constitutes a locus of learning, under both the classic Whorfian view, where language is considered to have a deterministic influence on thought, and the Cartesian view, which holds that language and thought are independent, and language is but an expressive vehicle for our universal capacity for thought. The last two decades have seen the emergence of a modern fusion, interestingly represented in both functionalism and generativism: One is the Slobin/Jakobson proposal of ‘thinking for speaking’, a neo-Whorfian perspective where it is thought that language differences influence thought only insofar as a given language might require obligatory encoding of certain properties of events. The second take on the modern convergence is perhaps most clearly represented by the neo-universalist ideas of de Villiers (de Villiers 2005, 2007) and Hinzen and Sheehan (2013), that grammar itself (the universal aspects of language, not the variations that characterize the grammars of specific languages) is essential for the format of thought. In their words: “Grammar imposes a novel form of knowledge of the world that could not exist lexically, involving reference and predication” (Hinzen & Sheehan 2013: xvi). This neo-universalism puts language in general and grammar in particular as a key trigger of cognitive development. In a series of empirical studies about the developmental interactions between language and theory of mind, Jill de Villiers proposed a directional view that embedding of sentential complements provides the representational format for false belief understanding (de Villiers 2005, 2007). Under this proposal, true embedding emerges only after children properly represent the point of view of the clause as distinct from the speaker’s point of view. Presumably, children’s initial representation of complement clauses is possibly paratactic, as proposed in Lebeaux’s (2000) account of various points of evidence that suggest high attachment of clausal complements and relative clauses. The key claims are:

i. Acquisition of complex (i.e., embedded) structures proceeds by synthesis (Givón 2009).
ii. Once embedding is acquired, the semantic consequences of embedding must be accessible to children.
For embedding of sentential complements, the consequence is that the subordinated proposition is not part of the speaker’s assertion. As pointed out by Arsenijevic and Hinzen (2012), a sentence complement as in (3a) is opaque (the speaker does not assert its truth). However, a paraphrased version of the same utterance (3b), where the content of the assertion is coordinated rather than embedded, is simultaneously asserted by both the subject of the verb say and by the speaker:

(3) a. Bill said that John left

b. John left, and Bill said so.

Arsenijevic and Hinzen (2012) attribute such effects to the cyclical nature of syntactic computations, arguing that the embedding and the embedded XP exhibit different referential behavior at the semantic interface (423). They observe that embedding inside NPs renders referents inaccessible to predication. Here again we note a contrast between an embedded and a coordinated nominal. A predicate (is green) is construed only with respect to the highest head noun in (4a), whereas in (4b) it applies to both NPs in the coordinate structure:

(4) a. The cup on the table is green

b. The cup and the table are green

If children have mastery of the mapping between syntax and semantics of complex NPs, they should show awareness that referential expressions within a complex NP can combine in three distinct modes of composition, as in (5):

(5) a. The cup and the table
   Coordination: a cumulative addition of referents

   b. My brother, the biologist
      Nonrestrictive modification: additional descriptions of a pre-established referent

   c. The cup on the table
      Restrictive modification: narrowing the set of potential referents

In the present study we aim to assess whether preschool aged children understand the consequences of the contrast between (5a) and (5c).

3. Previous work on children’s interpretation of complex NPs

An extensive current literature examines relative clause comprehension in a variety of languages. This research employs relative clauses as a test of children’s
abilities to generate and comprehend complex dependencies (Friedmann, Rizzi & Belletti 2009, 2017). This work does not directly address the question of how children understand the referential consequences of embedding per se. However, older work on relative clause comprehension had uncovered a pattern where children seemed to construe a relative clause adjacent to a direct object as describing the action of the subject of the clause, as in (6) (Tavakolian 1981). This pattern was attributed to children’s preference to assign a coordinate-like interpretation to embedded clauses.

(6) The lion kissed the duck that hit the pig.
‘the lion kissed the duck and hit the pig’

Lebeaux (2000) proposed that such interpretations originated from a stage in grammatical development where children attach embedded clauses higher, before fully integrating them with the matrix clause. Further evidence of coordinated interpretations of NP-internal phrases comes from work on recursive possessives. Limbach and Adone (2010) tested children and adult L2 speakers’ comprehension of recursive possessives such as (7):

(7) Mary’s father’s bike (Limbach & Adone 2010)

Children showed two main error patterns. Rather than the target referent (‘the father’s bike’), children often chose non-target responses that resulted from reduced (‘Mary’s bike’) or coordinated interpretations (‘the bike jointly owned by Mary and her father’). Roeper (2011) proposed that complex sequences initially emerge in children’s grammar with iterated conjunctive readings.

Other studies tested children’s comprehension of complex NPs modified by adjectives. Ramos (2000) found that children sometimes allow modifying adjectives to have scope over both nouns in a possessive construction, as in (8):

(8) Miss Piggy knocked over the yellow horse’s sign. (Ramos 2000)
‘the horse and the sign are both yellow’

In another study, Stickney (2009) found that children were able to differentiate between pseudo-partitive and prepositional constructions, as in (9), in that they frequently allowed the adjectival modifier to apply to the second noun in the pseudo-partitive construction in (9a) but disallowed this interpretation with the prepositional construction in (9b):

(9) a. The seal wanted a broken plate of cookies. (Stickney 2009)
‘the cookies are broken’

b. The seal wanted a broken plate with cookies.
‘the plate is broken’
However, Stickney (2009) also noted that three-year-olds had higher error rates than older children on the prepositional *with* sentences.

Subsequent work with quantity expressions (Syrett 2013; Foushee et al. 2017) investigated whether children could distinguish the pseudo-partitive (10a) from the attributive construction (10b). In the first construction *three pounds* measures the set; in the second, *three-pound* refers to each individual strawberry.

\[(10)\] a. Three pounds of strawberries (Syrett 2013)

\[b. Three-pound strawberries\]

Syrett (2013) found that if you took away some of the strawberries and then asked children whether the set had been modified (*Do I still have three pounds of strawberries/three-pound-strawberries?*), children accepted the attributive description slightly more often than they did the pseudo-partitive description. Syrett (2013) interpreted her results as evidence that at age four, children are beginning to appreciate the syntax-semantic mappings of these measure phrases. Foushee et al. (2017) conducted a modified replication of the Syrett study and failed to replicate the finding of discrimination between the two types of structures. Foushee and colleagues concluded that children do not automatically appreciate the mapping between syntax and semantics in their comprehension of quantity expressions within complex nominals.

Our goal is to examine the question of the mapping between syntax and semantics by concentrating on the simplest possible property. We sought to test whether children know that nominal expressions embedded as PP modifiers to other nouns are inaccessible to predication from outside of the complex NPs. NP-internal PP modifiers are acquired somewhat late in the preschool years. Five-year olds use modifying adjectives consistently, but PP and relative clause modifiers are infrequent in spontaneous speech until well into the school years (Eisenberg, Ukrainetz, Hsu, Kaderavek, Justice & Gillam 2008). In elicited tasks, two-thirds of three- and four-year olds produce only single NPs (Pérez-Leroux, Castilla-Earls, Bejar & Massam 2012). Our goal is to test whether children at this age nevertheless understand that the domain of embedded PP modifiers is opaque to predication. For that purpose, we designed a study to compare comprehension of comitative (*with*) and locative (*in/on*) PP modifiers with coordinate NPs.

### 4. Methods

#### 4.1. Participants

Twenty-three monolingual English-speaking children were recruited from the Toronto area. Ages ranged between 3;00 (36 months) and 5;02 (62 months), with a mean age of 4;01 (SD = 7.3 months). All of the children were identified by parents or teachers as typically developing. Ten monolingual English-speaking adults from Toronto served as control participants.
4.2. Materials and procedures

The experiment included a novel sentence comprehension task and a short referential elicitation task for PP modifiers (not discussed further here, for reasons of space). Both tasks were created and administered on an iPad using the Educreations (2016) application. The comprehension task relied on a coloring activity following the Coloring Book method proposed by Zuckerman, Pinto, Koutamanis, and van Spijk (2016) for testing the comprehension of pronouns and passives. At the outset, each participant was asked to identify the five colours used in the Educreations (2016) app. Children then heard that the experimenter would read a sentence and the child should repeat the sentence and then color the picture to match it. To facilitate coding, children were asked to use small dots which were practiced on a blank screen.

Three training items ensured children’s familiarity with the procedure. The remainder of the task consisted of 23 items: 12 test items and 11 control items. To mitigate potential effects of order of presentation we used two different sets, with picture and sentence types counterbalanced across stimuli sets. The ordering of items within each list was pseudo-randomized. Table 1 summarizes the test sentence types with examples of each. The test items were divided between coordinated NPs and the two types of modifying PPs, namely comitative (with) and locative (in/on).

<table>
<thead>
<tr>
<th>Sentence type</th>
<th># of pictures</th>
<th>Example</th>
<th># of nouns</th>
<th># of colours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinated NP</td>
<td>4</td>
<td>The book and the apple are yellow</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Comitative PP</td>
<td>2 unique, 2 contrastive</td>
<td>The dog with the bone is blue</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Locative PP</td>
<td>2 unique, 2 contrastive</td>
<td>The pillow on the table is red</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

The embedded PP condition was further divided in terms of the picture used: in the unique referent conditions (Figure 1a), the matrix NP (e.g., the pillow) was the only such object in the picture, while the contrastive pictures (Figure 1b) included a competing object. We hypothesized that the presence of a competing referent might help children interpret embedding correctly because it renders the use of a PP modifier felicitous. Use of a modifier in the absence of contrast, as in Figure 1a, is over-informative: since there is only one pillow in the picture, it is not essential to identify it as the pillow on the table. Davies and Katsos (2010) showed that, given a scalar task, five-year-olds are capable of detecting the over-informativeness of an adjective used in a single-referent context, and consider it
less felicitous than adjectives used contrastively. In Figure 1b, the presence of the second dog justifies elaborating the referential description.

![Figure 1b](image)

**Figure 1. Sample illustrations for the unique and contrastive target referents**

The control stimuli were designed to balance the test sentences in terms of the number of nouns and colors named, as well as the number of items expected to be colored by the participants, as shown in Table 2. The two-color transitive controls included one token where the nouns in subject and object position were the same but had different colors (i.e., two trucks, one yellow, one black), and one token where the subject and object represented different entities.

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th># of pictures</th>
<th>Example</th>
<th># of nouns</th>
<th># of colors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single NP</td>
<td>4</td>
<td>The flower is black</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Single NP intransitive</td>
<td>2</td>
<td>A green mouse is coming out</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1-color transitive</td>
<td>1</td>
<td>A red bus hits a car</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2-color transitive</td>
<td>1 (=NPs)</td>
<td>A yellow truck pulls a black truck</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2-color transitive</td>
<td>1 (≠NPs)</td>
<td>A green bunny chases a blue frog</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3-NPs Coordinated</td>
<td>2</td>
<td>The window, the flower, and the table are red</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

**4.3. Coding**

The Educreations (2016) app recorded the participants’ actions in a video format. After testing, the experimenter watched each video and manually coded the responses in terms of the particular item(s) colored in each picture. Participants’ responses were compared to the expected response pattern. Table 3 summarizes the response patterns for the test sentences. Nouns were identified by
the order of appearance in the stimuli; if the participant colored items not mentioned in the sentence these were coded as "Other N". A few actions were registered as unclear.

Table 3: Expected responses for coordinated NPs and embedded PPs

<table>
<thead>
<tr>
<th>Nouns colored</th>
<th>N1 + N2</th>
<th>N1</th>
<th>N2</th>
<th>Other N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinated NP</td>
<td>√</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PP embedding</td>
<td>X</td>
<td>√</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

5. Results

Adults were 100% accurate across both control and test conditions, and therefore their results are not explored quantitatively. Children’s accuracy varied across sentence types. As shown in Figure 2, accuracy was high for coordinates and controls. Performance with locative PPs was lower than controls by about 10%. However, performance with comitative PPs was very poor, and not different from chance level (set at 33%).

![Figure 2. Percentage of accurate responses to the control and test conditions by children](image)

The data on accuracy was entered into a generalized mixed effects (logit) model in R (R Core Team 2013), using the lme4 (Bates, Maechler, Bolker & Walker 2015) package. Participant was entered as a random effect, and the fixed effects were sentence type as a categorical predictor, and age as a continuous predictor. The model indicated that the effect of condition was highly significant, with both coordinates and locatives significantly differing from comitatives (Type-Coordinate ($\beta=3.676$, $p<.001$); Type-Locative ($\beta=2.727$, $p<.001$)). The model also indicated a significant overall effect of age ($\beta=0.082$, $p=0.02$).
To explore the developmental patterns across conditions, we tested the correlation between age and accuracy for each condition separately using the Kendall method. Considered independently, age correlations for the test conditions were all moderate and not significant. Only the control condition showed a significant effect of age ($z=1.9297, p=0.05$). Figure 3 shows that coordinated NPs and locatives both follow similar patterns as the controls, with performance improving early on and approaching ceiling by age 3;9. For comitatives, the pattern was different, with a quick increase from incorrect responses to an average of 25% correct responses, followed by an extended plateau and a slight improvement to 50% around age 5;0.

Figure 3. Individual participants plotted by number of target responses as a function of age for all four conditions. The line indicates the locally-fitted mean, and the shaded area the 95% confidence interval for the local mean.

We next consider what kinds of errors children make. Table 4 shows the children’s responses to each sentence type; shaded cells represent incorrect responses. For comitative NPs, the double (N1N2) response is the most frequent, making up just over half of the children’s responses. Children interpret with NPs in the same way as coordinates about half of the time. For locative NPs, N1N2 responses were still the most common error, but these were not very frequent. The distribution of response types differs significantly across the three conditions ($\chi^2=159.27, df=6, p<0.001$).
Table 4. Children’s responses by type of complex NP

<table>
<thead>
<tr>
<th>Complex NP type</th>
<th>Response Type</th>
<th>N1</th>
<th>N2</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinated 2-NPs</td>
<td>N1N2</td>
<td>89.1%</td>
<td>1.1%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Embedded-Comitative</td>
<td>N1N2</td>
<td>47.8%</td>
<td>28.3%</td>
<td>17.4%</td>
</tr>
<tr>
<td>Embedded-Locative</td>
<td>N1N2</td>
<td>12.0%</td>
<td>78.3%</td>
<td>5.4%</td>
</tr>
</tbody>
</table>

We then considered individual performance. Figure 4 shows the numbers of individual participants classified according to how many dual responses (i.e., N1N2) they provided. In the coordinated and locative PP conditions, where children showed high accuracy, the response patterns are nearly complementary: most children correctly gave N1N2 responses to at least 3 out of 4 coordinated NPs. For locatives, where N1N2 responses are incorrect, the majority of children gave zero or one of these dual responses. For comitatives, however, we observe that children concentrate in the extremes: 6 children exclusively selected the N1N2 response, and another 7 children never responded this way, while the remaining 10 children were spread fairly evenly across one to three N1N2 responses. This apparent U-shaped distribution suggests that children are sensitive to the lexical semantics of prepositions but are uncertain about the interpretation of *with*, with some children clearly treating it as a preposition, while others clearly prefer a coordinate reading.

![Figure 4](image)

**Figure 4. Individual participants classified according to number of N1N2 responses per condition**

Our final analysis considers the effect of informativeness on children’s performance. We had hypothesized that the presence of a competing referent might help children interpret embedding correctly. Only in this type of scene is
the use of a PP modifier felicitous; using a modifier in the absence of a contrasting referent is considered over-informative. However, the results do not support this pragmatic hypothesis. The presence of a contrastive referent failed to improve overall accuracy. In fact, having more objects present in the visual context resulted in a slight reduction in accuracy (72% vs. 85% for locatives; 26% vs. 30% for comitatives). This difference was not statistically significant.

6. Discussion

Our results can be summarized as follows:

i. Children are very accurate with control sentences (88% on average), and their performance shows a clear effect of age;
ii. Children are also very accurate with coordinate structures;
iii. They are almost as accurate with locatives (78%), with only two of the children making more than one error;
iv. Children had very low accuracy with comitatives; most children preferred giving a consistent target interpretation, or a consistently coordinate interpretation;
v. There is no evidence that the effects are the result of pragmatic limitations. Children’s performance did not improve when the use of the modifier was adequately informative.

Because children are highly accurate with locatives from the outset, we conclude that children are able to understand the difference in the interpretation of coordinates and embedding as early as one might plausibly expect. We find no evidence of an early paratactic or coordinated stage.

It can be argued that our study gave the children specific cues to the difference between coordinate structures and PP modifiers, namely agreement. In (1) and (2), the forms of the plural and singular copula (is/are) provide a clear marker of the distinction between the two structures. However, studies testing whether children use verbal agreement in comprehension suggest discrimination of plural and singular sentences on the basis of verb form is rather poor until the age of five (de Villiers & Johnson 2007). Furthermore, under this interpretation of the data, we would have no way of explaining why agreement helped children with locative PPs, but not with comitative PPs.

The simplest interpretation of our findings is that they provide evidence of sensitivity to structure and of an effect of lexical semantics, since children fail to block predication into comitative modifiers.

Like the youngest children in Stickney’s (2009) study, the children we tested show difficulty with comitative with. From a developmental point of view, it is not easy to account for why this should be the case. Locatives in/on usually emerge earlier than other spatial prepositions, and comitative with also appears early (Tomasello 1987; Kidd & Cameron-Faulkner 2008), usually in contexts of spatial proximity such as those depicted in our materials, where with denotes
accompaniment or attribute. Other meanings of *with* such as instrumental or manner enter later in children’s productions, but they are not relevant within our study. A cross-linguistic comparison of *with* and its analogues in other languages provides a potential clue. Across languages, *with* has a closer semantic relationship to *and* than locatives do. Both *with* and *and* can denote accompaniment, and some languages use a single form to express this meaning, such as Korean (g)wa and hago and Japanese to.

Clearly, the actual preposition matters, and not just across languages, and not just for children. Adults in English will find various of structures to be ambiguous between an interpretation where the higher noun represents a measure and the lower noun heads the structure, or where the higher noun is the head and the lower noun a restrictor. External predication functions accordingly, applying either to the lower or the higher noun. Consider the contrast between the two sentences in (11). Both interpretations are formally available in either case, but as speakers we interpret sentences according to our priorities and world knowledge: we care whether the coffee is hot, but less so the cup, whereas cups are more likely than coffee to be red.

(11)  
   a. That cup of coffee is hot.  
   b. That cup of coffee is red.

*Of* is highly polysemous and semantically ‘light’, helping encode a wide range of relationships. In fact, it is not much beyond a case marker/NP linker; hence the structural ambiguity associated with it. While we do not expect children to know the full range of uses of a relational term, we do expect that once its meaning is learned, children would know the semantic consequences of using such a particle to embed other NPs. Meaningful prepositions make a semantic predicate of the NPs that serve as their complements (Heim & Kratzer 1998; Pérez-Leroux et al. to appear). Once a PP is inserted in NP-internal position, it is treated as a phase, and the domain under it is treated as referentially opaque. Given a pre-NP particle, all that the children need to know is that it is not semantically empty, and the right semantic consequences follow. This minimal threshold of lexical knowledge is met early on for locatives in/on, but not for *with*.

We can conclude that the present study found no support for the existence of a stage where children are indifferent to or unaware of the semantic consequences of hierarchical structure, and treat embedded PPs as accessible. However, lexical meanings of prepositions are clearly required for full mastery of syntax-semantics. Embedding does not happen in the void: lexical knowledge of the embedding particles determines the referential opacity or transparency of the phrase governed by it. Previous studies considered more abstract notions such as quantity expressions, or tested the scope of modifying adjectives with regards to other constituents inside complex NPs. By examining whether predication can reach into a simple PP modifier, our study targets the simplest possible test. Our data shows that young children pass this test, as would be predicted by the neo-
universalist, grammar-centric view of the relationship between language and thought.

References


