



Hall, Erin & Pérez-Leroux, Ana T. 2022. Children's comprehension of NP embedding. *Glossa: a journal of general linguistics* 7(1), pp. 1–41. DOI: <https://doi.org/10.16995/glossa.5816>



Open Library of Humanities

Children's comprehension of NP embedding

Erin Hall, California State University, San Bernardino, US, erin.hall@csusb.edu

Ana T. Pérez-Leroux, University of Toronto, CA, at.perez.leroux@utoronto.ca

How do children learn to interpret structurally complex noun phrases? NPs embedded inside other NPs are not accessible to predication, so that in a sentence with a subject NP containing a PP modifier such as *the cup on the table is green* or *the dog with the bone is blue*, the adjectival predicate has scope over the highest but not the embedded nominal referent (Arsenijevic & Hinzen 2012). We used a coloring task to examine children's comprehension of sentences containing these complex NPs, comparing PP modifiers (locative and comitatives) to coordinated NPs (*the cup and the table are green*), where both referents are accessible. Three- to five-year-old children were highly accurate with control and coordinate sentences, and performed well with locative PPs, but were not different from chance level for comitative sentences, which many children treated as coordinates. That children differentiate between coordinate and locative sentences provides evidence that children have early access to the syntax-semantics of complex nominals. The contrast between locatives and comitatives suggests that comprehension is not merely guided by subject agreement (since the agreement patterns are the same for both types of PP-modified subjects), and that children still need to learn the lexical semantics of prepositions. Diachronically, languages with comitative modifiers evolve into languages with comitative coordination (Haspelmath 2007). Thus, we propose that these error patterns for comitative prepositions can be explained by the assumption that children's errors align with the direction of systematic language change.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



1 Introduction

The present study investigates children’s comprehension of prepositional phrase modifiers (PPs). Our focus is not the acquisition of their pragmatic restrictive function, which has been investigated elsewhere (e.g., Nadig & Sedivy 2002), nor the developmental changes brought by maturation in this domain, but children’s understanding of the semantic consequences of structural embedding, and the potential interplay between such structural knowledge and lexical learning. Specifically, we investigate what children know about the referential accessibility of the elements that form part of a complex noun phrase (NP) structure. By this we refer to the observation that, if we take a noun phrase (*the book*) and embed a modifier PP within it (*on the table*), any assertion we make about the resulting complex NP (i.e., *the book on the table*) will not apply to the referent introduced by the embedded PP modifier. In other words, in saying *the book on the table is green*, we are saying nothing about the color of the table. Do young children understand modifiers in this precise way?

Our goal is to examine how such knowledge is acquired by children. Broadly speaking, language acquisition integrates distinct cognitive-developmental processes that bring an infant to full status as a speaker of a language in the span of a few years. These developmental processes yield three forms of knowledge and abilities: language knowledge that is learned, language knowledge that is not learned, and yet others, which, whether learned or not, seem to grow over time. Let us consider each briefly.

Learned representations: Children show remarkable capacities for distributional learning. Preverbal infants’ ability to extract statistical information from linear sequences underlies many early language learning processes: formation of phonological categories (Kuhl 2004), word segmentation (Saffran, Aslin & Newport 1996), word learning (Fisher et al. 2006), abstraction of rule-like patterns (Marcus et al. 1999), emergence of word classes (Reeder, Newport & Aslin 2013; Aslin & Newport 2014), and others. With time, these distributional learning capacities support and accelerate word learning, bootstrapping the relational words that then serve as the foundation for complex syntax (Gleitman 1990; Gleitman et al. 2005; Fisher et al. 2020).

Implicit knowledge of hierarchical structure: Children represent the learned sequences as hierarchically structured configurations, mapping these objects with precision into sentence-level semantics. Young children filter the output of statistical learning according to a set of structural principles, which manifest as various learning biases about the structure of the system being learned. Child learning biases have been argued to account for the distribution of typological variants in the world’s languages (Culbertson & Newport 2015; Martin et al. 2020), or cyclic patterns of grammaticalization (Cournane & Pérez-Leroux 2020). Evidence for this is found in the early mastery of grammatical characteristics that are either underdetermined by properties of the surface structure, and/or poorly represented in the input (i.e., what Goldin-Meadow

(1982) refers to as “resilient properties”, which develop robustly despite potential fluctuations in experience). For example, children demonstrate early sensitivity to core structural factors, including agreement with complex NPs (Koulaguina et al. 2019) and the interpretation of *one*-pronominalization (Lidz, Waxman & Freedman 2003), as well as robust and early mastery of aspects of sentence interpretation, in phenomena such as plural conjunction (Tieu et al. 2018), antecedent-contained deletion (Syrett & Lidz 2009), and many others. The conclusion in these studies is that the formal consequences of grammatical structure are not “learned”, in the sense above, but generated by children’s internal biases to build hierarchically-structured combinatorial systems (Goldin-Meadow & Yang 2017).¹

Growth in general abilities: Some language abilities grow over time. These properties may be learned early, but their full capacity is deployed on a timetable that seems similar across typologically different languages. Such developmental shifts may reflect general maturational processes rather than language-specific learning. For example, the abandonment of telegraphic speech in analytic languages coincides with important morphological changes in agglutinative languages in young children.² Similarly, increases in the productivity of complex syntactic structures occur in the preschool and early school years across languages (see de Villiers 2021, for recent investigation into sentence complements, and Pérez-Leroux & Roberge 2018, for comparative work on NP embedding). Such shifts broadly align with important developmental changes in other relevant domains of cognition such as Theory of Mind (Arslan et al. 2017), and with changes in the cortical organization of language that correlate with growing syntactic abilities (Weber-Fox & Neville 1996; Fengler et al. 2015; Enge et al. 2020).

Which aspects of the syntax-semantics of syntactic embedding are learned, which aspects are not, and which aspects merely grow? Which of these categories does comprehension of PP modifiers fall into? Our point of departure is the assumption that grammatical complexity mediates the organization of grammatical meaning. According to Arsenijevic & Hinzen (2012), referential opacity (or, to distinguish from substitutability, what we will call here referential inaccessibility) is a consequence of the formal organization of phrasal structure, be it the embedding of a clause or a nominal. They depart from the more familiar case of clausal (CP)

¹ A *Glossa* reviewer cautions against overinterpreting innateness from early mastery of the consequences of hierarchical structure. Our core aim is to highlight that some aspects of sentence interpretation have a visible developmental profile, whereas others seem stable and target-like from the outset. For instance, consider the developmental differences in the interpretation of co-referent pronouns in quantificational and non-quantificational contexts (Chien & Wexler 1990).

² Franchetto & Santos (2017) describe how in Kuikuro (an agglutinative Carib language of Southern Amazonia), mothers describe the initial (morphologically impossible) words of their children as *egipanetoho* (eg-ipaN-ne-toho, ‘made for learning’). Franchetto & Santos report that *egipanetoho* are fully abandoned by age 2;04. See Allen (2017) for a description of the early stages of development of Inuktitut, as well as Grinstead (2016) for root infinitive stages across languages.

embedding, which is opaque in terms of the assertion expressed by the main clause: in *Bill said that John is a student*, the truth value of the embedded clause *John is a student* does not carry over into the evaluation of the entire sentence. In concrete terms, this complex sentence is true if Bill spoke of such a thing, but Bill can be mistaken, or lying, about it. Whether John is or is not a student is irrelevant to the overall evaluation of the complex sentence. Arsenijevic & Hinzen (2012) argue that parallel semantic consequences in the nominal domain are manifested in terms of predication, as mentioned above. In *the book on the table is green*, only the highest referent (*the book*) is asserted to be green. Arsenijevic & Hinzen (2012) propose that these referential inaccessibility effects are a consequence of the cyclic nature of syntactic derivations, where computations are cyclic, and some syntactic objects (phases) are transferred to the semantic interface. Following Chomsky (2001; 2005), after phasal transfer, the complement of a phase head is not accessible to operations outside the phase. In their analysis, referential inaccessibility is a consequence of syntactic transfer to the semantic interface: once a phrasal category has been transferred at a phase boundary, it is closed off to reference (that is, to predication, for nominal referents, and for the evaluation of truth values, for propositional referents).

Work in the CP domain suggests that children's understanding of the referential inaccessibility of embedded clauses is not automatic (de Villiers & de Villiers 2000, and subsequent work). One could assume that such fundamental aspects of the logic of grammatical organization should be directly derived from structure and thus not an actual target of learning. Nonetheless, although universal, this knowledge is mediated by lexical learning of the language-specific elements. For complements, these elements include the nature of a complement-taking verb (whether factive/non-factive), the expression of finiteness, and the clausal connectors. For nominals, the relevant element is the preposition, conjunction, particle, etc., which determines the type of connection between two noun phrases, whether one is subordinated to the other, or they are coordinated.

Let us consider the referential differences in the various complex NPs in (1)–(8). In the case of two coordinated NPs (1), predication affects both NPs. In complex NPs containing a prepositional phrase (PP) modifier, as in (2)–(3), only the (initial) head noun is accessible to predication, and the embedded nominal remains non-referential. This does not exhaust all possibilities. Linear sequences of noun-preposition-noun produce other underlying configurations, where the lower, not the higher nominal, is the syntactic head. This is the case of quantifier structures, as in (4). Here reference is given by the lower noun, *students*, and the top-most noun expresses quantity. The surface-identical *of* construction is actually ambiguous between a pseudo-partitive configuration³, where the lower noun is the head,

³ Pseudo-partitive constructions contain a DP (referring to measures, containers, atoms, portions or groups) interpreted as measuring an embedded NP (Selkirk 1977).

and a modified noun configuration, where the higher noun is the head. In (5), either noun is accessible, and the predicate applies to either the lower or higher noun, as made evident by the choice of predicative adjectives in (6) and (7). Even then, the semantics of predication is strict: co-predication (Collins 2017) is not felicitous (8).

- (1) The box and the ball are green => both are green
- (2) The box with a ball is green => only the box is green
- (3) The box on the table is green => only the box is green
- (4) The majority of the students are nice => students are nice
- (5) This cup of coffee is very nice => either the coffee or the cup is nice
- (6) This cup of coffee is hot and black => the coffee is hot and black
- (7) This cup of coffee is white => the cup is white
- (8) #The cup of coffee is hot and white

Learning to interpret complex NPs seems less trivial than it might appear at first glance. The potential structural ambiguity of pseudo-partitive constructions does not exist for true modifier structures, which are interpreted with strict headedness. Is this form of structural knowledge robust from the outset, or does it exhibit a pattern of gradual learning in children? If children's understanding of hierarchical structure is constrained or resilient, in the sense of Goldin-Meadow (1982), we should expect that from the point they are able to understand modification, children might demonstrate awareness of the referential inaccessibility of this domain. Nonetheless, we acknowledge that there will always be a lexical learning component: learning about PP modification entails learning the meaning of the relevant prepositional connectors, and disentangling their uses from those of coordinating conjunctions. In some languages, the so-called comitative coordination languages, coordinated subjects such as *John and Mary left* are expressed via comitative modification: *John left with Mary* (Haspelmath 2007). In such languages, coordination appears with a comitative particle to denote accompaniment, while retaining the structural properties of coordination.

In this study, we focus on children's comprehension of sentences like (1), (2), and (3). The article is organized as follows. Section 2 summarizes previous work on children's comprehension of PP modification and associated features, including the lexical learning of prepositions, the acquisition of complex NPs, and sensitivity to agreement as a correlate of NP structure. We then detail our research questions and hypotheses, before outlining the methods and design of our study in Section 3. Section 4 presents our results, which are then discussed in Section 5, and Section 6 concludes.

2 Background

2.1 Lexical learning of prepositions

Studies of the acquisition of prepositions in English and German generally show that locative prepositions emerge earliest in child language, with words like *in*, *on*, and *under* appearing before more complex spatial concepts such as *between* and *through* (e.g., Grimm 1975; Johnston & Slobin 1979; Durkin 1981). In a comparison of English- and French-speaking children ages 1;8 to 2;4, the English speakers showed early use of spatial prepositions, while the French speakers used functional prepositions such as *for* and *of* first; in both languages, however, these prepositions first appeared with pragmatic functions (e.g., making requests or justifying actions), rather than syntactic ones (Morgenstern & Sekali 2009). In Germanic languages, other preposition types such as temporal, instrumental, and dative tend to follow the locatives (Grimm 1975; Tomasello 1987). In Tomasello's (1987) data from a child between the ages of 1;5 and 1;11, all types of prepositional phrases were mainly used as verb modifiers, as in *Daddy help me with this* (p. 87), *Put it in* (p. 85), and *I throw it to you* (p. 88); examples of PPs as noun modifiers were absent or restricted to rote phrases used as nouns (e.g., *Piece-of-ice down here table*, p. 89). In addition to the early use of locative prepositions, the child in Tomasello's (1987) study showed early and productive use of the phrase *with me*, denoting accompaniment, similar to her use of spatial prepositions and in contrast to later-developing instrumental and other uses of *with*. Tomasello (1987) suggests that comitative *with* and spatial prepositions were modelled by the child's parents in similar ways, which made this particular meaning of *with* easier to acquire.

These findings raise the issue of polysemy in prepositions, as spatial forms like *in* and *on* tend to have more limited semantic roles than a preposition like *with*. McKercher (2001) points out that *with* can potentially encode more than 100 distinct meanings, including the main senses of instrument (*I ate the soup with a spoon*), accompaniment (*I ate the soup with my friend*), attribute/modifier (*I ate the soup with salt*), and manner (*I ate the soup with pleasure*). In Rice's (2009) study of 9 different English prepositions, children used *with* initially with the verbs *coming* and *going*, and only later acquired the comitative, instrumental, and attributive uses of the preposition. Kidd & Cameron-Faulkner (2008) examined one child's acquisition of the multiple meanings of *with*. They noted that initial productions from ages 2;0 to 2;4 denoted spatial proximity (accompaniment or attribute/modifier, i.e., comitative *with*), while instrumental uses emerged between 2;5 and 2;8. Extension of the use of *with* to more adult-like constructions and meanings was achieved by age 4;0. These studies suggest that children produce prepositions denoting spatial meanings such as *in*, *on*, and *with* at similarly early stages.

The meaning of the connecting particle is an important factor for children's learning. An interesting issue, which we touched upon briefly in the introduction, is that some languages express coordination with the same form as accompaniment. In Korean, the particles (g)wa 'and' and hago 'with' are interchangeable in sentences like (9) and (10) (N-Y Ryu, p.c.):

- (9) Na-neun nae chingu-hago/wa nonda
 I-SUBJ my friend-with/and play(present plain)
 ‘I play with my friend’
- (10) Piteo-hago/wa meri-neun hakgyo-e gassda
 Peter-with/and Mary-SUBJ school-to go(past)
 ‘Peter and Mary went to school’

In Japanese, the particle *to* is similarly used in both conjunctive ‘and’ and prepositional ‘with’ contexts. Even in English one can argue for some degree of overlapping functions for *with* and *and*: both *Peter and Mary went to school* and *Peter went to school with Mary* express accompaniment. This relationship between coordinates and *with*-phrases does not seem to hold for locative prepositions.

Given these facts (and our discussion of sentences (1) to (8)), it seems worth asking a) whether or not children undergo a stage where they confuse the comitative preposition *with* and the coordinating conjunction *and*, and b) if this is so, whether errors occur randomly or in a specific direction. As mentioned above, children’s early use of *with* as a verbal modifier can involve both the subject NP and the NP in the modifier carrying out the same action (as in *the girl is playing with the boy*). An anonymous *Glossa* reviewer suggests that this might lead to potential confusion about the function of *with* in complex NPs. Here we explore an alternative possibility, based on the proposal that children’s developmental patterns play a role in predictable diachronic processes such as grammaticalization patterns (Cournane 2019). In the historical record of comitatives and related structures, grammaticalization is unidirectional: comitative conjunctions evolve from grammars with comitative prepositional modifiers (Haspelmath 2007). If children’s errors indeed align with this direction of change, they may be expected to interpret the comitative preposition *with* as a conjunction, but not to show the reverse pattern of treating *and* as a preposition.

In sum, children show early use of prepositions such as *in*, *on*, and *with* as verb modifiers. It is possible that the comitative preposition *with* may present additional learning challenges, due to its close relationship to the conjunction *and*. In the next section, we turn to children’s comprehension and use of PP modifiers of nouns.

2.2 Children and noun modifiers

Spontaneous use of noun modifiers is infrequent, both in the input and in children’s use. Koulaguina et al. (2019) found complex NPs to represent 0.02% of the NPs extracted from the data in a corpus of close to 200,000 input utterances spoken to French children. In English, Lorimor et al. (2019) examined child-directed speech and found that complex NPs accounted for 0.07% of utterances in the Pearl-Sprouse corpus. Complex NPs do not appear in early child

speech. In narratives, which tend to contain more elaborate language than dialogue, more than 40% of five-year-olds do not use postnominal modifiers of any type (Eisenberg et al. 2008). It is interesting to note that prenominal adjectives are used more often by children. Using an elicited production task, Pérez-Leroux et al. (2012) compared elicitation of coordinate NPs and complex NPs containing comitative PP modifiers and possessives in children aged 3–5. Even the younger children were able to fluently coordinate three nouns, but NP embedding was much more constrained. Half of the children did not produce complex NPs, only 40% produced complex NPs with PPs, and the remainder produced NPs with possessors but not PP modifiers. Subsequent work eliciting production of recursive structures shows better performance in preschoolers, but still significant differences from adults (Pérez-Leroux & Roberge 2018; Giblin et al. 2019). Studies that target levels of embedding explicitly find differences between recursive NPs and sequences of non-recursive modifiers (Pérez-Leroux et al. 2018; Béjar et al. 2020).

It is possible that children have knowledge of PP modifiers, but underproduce them as part of a general tendency to be under-informative in their referential descriptions (Nadig & Sedivy 2002; Grigoroglou & Papafragou 2019; among others). For example, comprehension studies show that young children understand that contrasting objects in a scene introduce a need for modification (Nadig & Sedivy 2002; Katsos & Bishop 2011). Like adults, children demonstrate sensitivity to under-informative utterances, and give lower ratings to over-informative utterances than to optimal ones, but they are more tolerant of over-informativity than adults overall (Davies & Katsos 2010). These results suggest that the felicity conditions for modification may influence children’s performance on referential tasks. Despite such early sensitivity to context, children rarely provide explicit characterization of contrast in their verbal description. When they do, the most common properties they use to distinguish object referents are size, color, and location.

Only two studies have directly investigated children’s interpretations of complex NPs involving adjectival modification. Ramos (2000) tested the interpretation of adjectives modifying possessors, as in *the yellow horse’s sign*. Adults assign the adjective scope over the possessor (‘the sign has a yellow horse in it’), but younger children (ages 3;8 to 4;5) perform at chance, often giving the adjective scope over the possessum (‘the sign is yellow, and it is about/for horses’). The older children in her study (ages 4;9 to 5;5) performed above chance. Stickney (2009) used a similar design to compare complex NPs containing an ambiguous pseudo-partitive *of* construction in (11) and unambiguous prepositional modifiers (12):

(11) The seal wanted a broken plate of cookies.

(12) The seal wanted a broken plate with cookies.

Of-NPs are ambiguous between the higher or the lower noun interpreted as the head of the structure. Adults allowed adjective construal with the lower noun (i.e., where the cookies were

broken, not the plate) a quarter of the time for stimuli such as (11), whereas such interpretations were negligible for the prepositional modifier (12). Children had a strong preference for the lower reading in the pseudo-partitive condition (40–90% of the time) and clearly distinguished it from the prepositional condition. However, three-year-olds in this study made errors 33% of the time, with this error rate decreasing to 15% by the age of four.

Another series of studies that has shown comprehension difficulties with complex structures comes from the comprehension of recursive structures. Limbach & Adone (2010) tested children and adult L2 speakers' comprehension of two-level possessives such as (13):

(13) Mary's father's bike

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'). In this and subsequent literature (Hollebrandse & Roeper 2014), these findings are interpreted not as general evidence of children's difficulties with hierarchical structure, but pertaining to the specific challenge of recursive configurations, i.e., those containing iterated embeddings of the same type. The argument is because recursive configurations vary across languages (Roeper & Snyder 2004) and need to be learned. Roeper (2011) proposed that complex sequences initially emerge in children's grammar with iterated conjunctive readings.

These previous experiments suggest that children do not reliably attribute hierarchical structure to complex NPs in comprehension. It is important to note that the studies reviewed constitute very complex cases: recursive structures (i.e., with three nested NPs) and the scope of modifying adjectives in two-nominal configurations. The comprehension of simple modification cases has not been investigated. In the following section, we turn to evidence on a potential cue to the structural organization of complex NPs, comprehension of noun-verb agreement.

2.3 Agreement as a test of understanding NP structure

Infant sensitivity to agreement mismatches manifests quite early, albeit in somewhat constrained fashion. Infants under the age of two showed a preference for listening to sentences in which agreement was grammatical over ungrammatical combinations (Soderstrom et al. 2007; Sundara, Demuth & Kuhl 2011). This evidence indicates early distributional learning of the matching between verb form and subject form and reflects sensitivity to sentence form. However, as Soderstrom (2008: 673) points out, learning the distributional properties of dependencies is very easy for infants, "while extracting or attaching meaning is not." In comprehension studies, where speakers can use singular and plural verb agreement to select a referent from visual scenarios depicting single or multiple actors, the evidence shows divergent results across studies, languages, and methods.

Kouider et al. (2006) show robust looking preferences for stimuli combining copula and NP number in English (*look there is a blicket/there are some blickets*) at 24 months. In Legendre et al. (2010), 30-month-old French-speaking children used agreement cues to identify which video display matched the event described by a sentence. In subsequent work, using the same methods and materials, Legendre and colleagues (2014) found no evidence of comprehension in English, and revealed only partial success in Spanish at a comparable age. The authors propose that the saliency of the markers across the relevant agreement systems may explain asymmetries in performance across languages.

Comprehension seems to appear later in the preschool years in tasks where participants have to explicitly demonstrate their understanding by pointing. Results vary according to a number of language-specific morphophonological and syntactic properties (Johnson et al. 2005; Pérez-Leroux 2005; Barrière et al. 2019). English-speaking four- and five-year-olds cannot reliably use verbal agreement to accurately choose a singular or plural scene (Johnson et al. 2005; see also Pérez-Leroux 2005 for Spanish), or to determine whether the stem is a noun compound or a generic verb (de Villiers & Johnson 2007). Variation in age of comprehension has also been found within dialects in English. Barrière et al. (2019) find that performance across conditions of different degrees of difficulty depends on which variety of English is spoken by the child.

One study by Brandt-Kobele & Höhle (2010) found asymmetric results across tasks within the same language: German children between the ages of three and four showed a preference in looking to the target that reflected the correct interpretation of the sentences on the basis of verb agreement. However, when preferential looking was combined with a pointing task, children did not perform above chance level. Some further evidence that comprehension at this age is facilitated by reducing task demands comes from a study of Mexican children aged three to five by Gonzalez-Gomez and colleagues (2017). In their study, children were able to use Spanish verbal agreement for novel objects when using an underspecified noun (*objeto* ‘object’) but not with specific pseudo-nouns assigned to the same objects. These authors conclude that results from some tasks may underestimate underlying competence.

Formal factors also matter, such as copula agreement vs. affixal agreement, or noun morphology. For example, copula comprehension might be established ahead of the verbal affix -s, at least in implicit comprehension (eye tracking) studies: three-year-olds reliably show anticipatory looking to single vs. multiple agent pictures on the basis of the copula (*Are the nice little dogs running?* (Deevy et al. 2017); *Where are the good cookies?* (Lukyanenko & Fisher 2016)). Davies et al. (2020) show that preschoolers prefer the target picture (single or multiple referent) in sentences with novel nouns on the basis of number in the copula, with transparent number

(**Where is the tups?*). Children significantly discriminated, but accuracy was not very high for three- and four-year-olds (about 60–70%). However, when nominal morphology was ambiguous (*Where is/are the gex/gecks?*), these younger children were significantly below chance, which suggests that the nominal ending was a more salient cue than the copula.

Sensitivity to agreement has also been probed with complex NPs. Koulaguina and colleagues (2019) explored the age at which French toddlers show sensitivity to agreement between a left-dislocated conjoined subject and the adjacent subject pronoun, a common configuration in French (14). Their corpus study showed that the most common use of the coordinating conjunction *et* is sentence-initial, as a discourse marker. The structure of interest was quite rare:

- (14) Le N₁ et le N₂ ils... ‘The N₁ and the N₂, they...’
 Papi et mamie, ils vont... ‘Papi and Mamie, they ...’
 (Koulaguina et al. 2019: 162)

Using a head turn preference procedure, they found that 18-month-olds did not differentiate between grammatical (plural) agreement and ungrammatical singular agreement. Looking times for 24- and 30-month-olds were significantly longer for grammatical sentences, but only during the first half of the experiment. The authors concluded that children are able to track agreement with conjoined subjects, and as a consequence, have the ability to integrate two conjoined subjects as a constituent, but this ability comes at a high processing cost.

Another study of French agreement in coordinates simultaneously tested gender and number agreement between subject topic and subject pronoun. Shi, Legrand & Brandenberger (2020) presented young children age 2.5 with sentences as in (15) and (16). They contrasted coordinate structures (which would have default masculine agreement if one noun is masculine) with modified NPs (where agreement should match the first noun):

- (15) La banane et le chapeau... √ils/*elle/*elles (Coordinated NPs)
 The banana and the hat ... they-MASC/*she/*they-FEM
- (16) La banane dans le chapeau... √elle/*ils/*il (Modified NP)
 The banana in the hat... she/*they-MASC/*he

These French toddlers consistently looked longer at the ungrammatical trials, demonstrating that they understood the grammaticality of gender agreement. They showed no preference or advantage for construction (coordinate/modified NPs). The authors argue against this being a demonstration of sequential learning abilities, due to the scarcity of the input, as in Koulaguina et al.’s (2019) corpus analysis. Instead, they conclude that the observed grammaticality effects indicate structure-dependent processing, supporting the assumption that the hierarchical

organization of utterances is present very early in acquisition. A subsequent replication with toddlers ages 17 to 18 months (Shi, Emond & Badri 2020) similarly found a novelty preference for ungrammatical sentences at an even younger age.

A recent production study by Lorimor et al. (2019) asked whether preschoolers could ignore potentially ambiguous cues to number that arise in complex NPs, where a modifier noun might have a number specification that differs from that of the head noun (*the man with the cars*). These contexts, which sometimes give rise to number attraction errors (Bock & Miller 1991), might be particularly challenging for children. Their analysis of modified NP subjects in both input data (see above) and in storybooks established that most complex NPs were headed by quantifiers or indefinite pronouns (*one of them*). Also uninformative were cases where the head noun and the modifier noun match in number. In their elicitation task, Lorimor et al. (2019) found robust sensitivity to number: children made more errors than adults, but both groups made more errors in the mismatch conditions (where the head noun and the modifier noun had different number specifications). They found that when facing feature mismatch in the nouns of a complex NP, many children had a bias for giving one type of response (either singular or plural), regardless of the number of the head noun. This data suggests that children are sensitive to number and headedness, but show difficulties in producing agreement in complex NPs.

In sum, the agreement literature suggests that infants demonstrate early sensitivity to agreement mismatches in simple NPs. Because ability to discriminate fluctuates across conditions, it is reasonable to conclude that children's difficulties reflect performance rather than representation. Therefore, it is possible that agreement might provide a cue to the internal structure of complex NPs. To what extent children could rely on agreement in explicit tasks, however, is not clear. Alongside indications of early agreement comprehension in complex NPs, we see evidence of declining performance with more challenging tasks and structures with high processing costs. Children are sensitive to the grammaticality of agreement mismatches with coordinated and modified NPs. In production, children are most accurate when there is number matching between the components, and much less so in the mismatch environments that induce adults to produce number attraction errors. Given these various sources of evidence combined, we cannot be certain that children's understanding of verb agreement will definitely support their interpretation of the scope of predication. This question remains open for now.

2.4 Research questions and hypotheses

The preceding review shows that children typically learn spatial prepositions like *in*, *on*, and *with* quite early in English. Less is known about their use in complex NPs, or about children's general abilities with complex NP comprehension. The agreement data provide some evidence

of early sensitivity to structure, but agreement (a distributional phenomenon) is different from comprehension (semantics; the structuring of thought). For us, this justifies the need for a simpler test of the syntax-semantics of complex NPs. For this, we look at basic comprehension data on predication. In particular, we ask:

1. Can children correctly interpret sentences with complex NPs, i.e., those that contain multiple NPs, including coordinated NPs (*the cup and the table are green*) and NPs embedded as prepositional modifiers (*the cup on the table is green*)?
2. Are children's interpretations of embedded NPs dependent on the type of preposition involved (locative *in/on* vs. comitative *with*), and what error types are seen in each case?

The first question allows us to examine which type of developmental process underlies our understanding of the referential inaccessibility of modifiers. Two competing hypotheses can thus be contrasted:

Hypothesis 1. Referential inaccessibility is learned: We will see an initial stage of no discrimination between coordinates and PP modifiers, and improved comprehension of both types of complex NPs with age.

Hypothesis 2. Referential inaccessibility is not learned: We will see discrimination between coordinates and modifiers from a young age, with no age-related improvements in comprehension. This result is ambiguous between understanding hierarchical structure and being sensitive to agreement.

A third, ancillary hypothesis, about the directionality of errors, pertains to our second descriptive question.

Hypothesis 3. Grammaticalization hypothesis: If children have substantive numbers of comprehension errors, these patterns will not be random but reflect underlying learning biases. Given that grammars with comitative conjunctions evolve from those with comitative prepositional modifiers, theories linking language change and language acquisition would predict unidirectional biases: children should interpret comitatives as coordinates. This hypothesis predicts two asymmetries in the distribution of errors: one, that comitatives will be treated as coordinates, but not vice versa; and two, that locative and comitative PPs will differ in accuracy, because only comitatives should elicit coordinate-type interpretations.

The remainder of this paper presents the methods, results, and discussion of a novel experimental task designed to test these hypotheses.

3 Methods

3.1 Participants

We recruited 51 English-speaking children from the Toronto area in Ontario, Canada. To focus on the period when children begin to use prepositional modification, we targeted children from three to four years. A few younger five-year-olds were also recruited. Of these participants, three had significant exposure to another language in the home (reported by parents as level 2 or higher on a proficiency scale from 1 to 5), and one had previously received intervention from a speech-language pathologist; the remaining 47 children were identified by parents or teachers as typically-developing monolingual English speakers. Based on preliminary analysis of the control items, a further seven children were excluded from the final sample; these children (six 3-year-olds and one 4-year-old) each scored 67% (6/9 correct) or lower on the intransitive, transitive, and single NP control conditions (see section 4.2 for more details). The ages of the 40 children in the final sample ranged between 3;00 and 5;04 (mean = 48.7 months; SD = 7.8 months). **Table 1** summarizes the participants by age group, both prior to and following performance-based exclusions. In addition to the child participants, ten monolingual English-speaking adults from Toronto served as control participants.

Group	Age range	N before exclusions	N after exclusions
3-year-olds	3;00–3;11	27	21
4-year-olds	4;01–4;11	14	13
5-year-olds	5;00–5;04	6	6
Total		47	40

Table 1: Study participants by age group.

3.2 General procedures

The present study consisted of a sentence comprehension task, and two baseline measures: an elicited production task, targeting PP modifiers, and a standardized general measure of grammatical development, the Recalling Sentences subtest of the Clinical Evaluation of Language Fundamentals Preschool, Second Edition (CELF-P2; Wiig, Secord & Semel 2004). The two novel tasks were created and administered on an iPad using the Educreations (2016) application.

Test sessions were conducted individually with each participant in the daycare or home setting, with all tasks administered by the first author or a research assistant trained in language assessment. Because of the novel nature of the experimental tasks, the first author piloted the methods with 11 children in the target age range, to evaluate their level of engagement and understanding of the iPad-based format. Children were generally eager to start the task, but to maintain the young children's engagement throughout, it was deemed optimal for the

experimenter to read the stimulus sentences aloud.⁴ The experimenters were trained to read the sentences with a consistent speech rate and natural prosody, and to encourage children to repeat each sentence in the comprehension task themselves, to optimize their attention and listening skills.

Each session started with a general introduction to the tasks, and then the production task was administered. This was followed by the sentence recall task, and then by the sentence comprehension task. The Educreations (2016) app recorded all of the participants' actions on the iPad in a video format, accompanied by audio recording of the instructions and comments of both the experimenter and the participants. After testing, the experimenter or another trained research assistant watched each video and transcribed the participants' verbal and manual responses to each task into a spreadsheet for coding (the coding schemes for the experimental tasks are described further in the following sections). Each child's transcriptions and response codes were later verified independently by two additional research assistants who were blind to the study's purpose and hypotheses (see following sections for more details).

3.3 Production task: Procedure, materials, and coding

To assess children's ability to use PP modifiers, we designed a novel elicited production task. A secondary purpose was to determine that this specific set of children were, as suggested by previous work on this topic, at the stage in which they begin to produce complex NPs containing PP modifiers. This task was presented in the Educreations (2016) app. Following common elicitation approaches, the first picture was used to introduce two competing referents (e.g., two different bears, as shown in **Figure 1**). The experimenter provided a verbal description of the picture, where all potential target referents were named. The second picture showed a star on one of the two competing items, and the child was asked a referential question of the form *Which X has the star?* The first training item demonstrated the task by inviting a color adjective description ("the blue books"). Subsequently, children were presented with four test items. Two were designed to elicit NPs containing comitative (*with*) PPs and two targeted locatives (*in* and *on*). For the test items, the experimenter covered her eyes during the prompting picture, to discourage the child from giving pointing responses. We illustrate this task with a comitative test item in **Figure 1**. (See Supplementary Files for the full set of instructions, training, and test items).

⁴ Reading the stimuli also meant that the experimenter did not need to introduce additional equipment to play pre-recorded stimuli (since pre-recording could not be done through the Educreations (2016) app itself). Although pre-recorded stimuli can help reduce potential experimenter bias, direct interactions with children remain the norm in spoken language assessment, and as a speech-language pathologist, the first author carefully trained all research assistants to conduct each task in a standardized manner.

Non-verbal responses were followed up with encouragement to use words, and prompts were repeated up to two additional times. Incomplete, under-informative responses (e.g., “The bear”) were followed up by an additional referential prompt (“But which one?”).

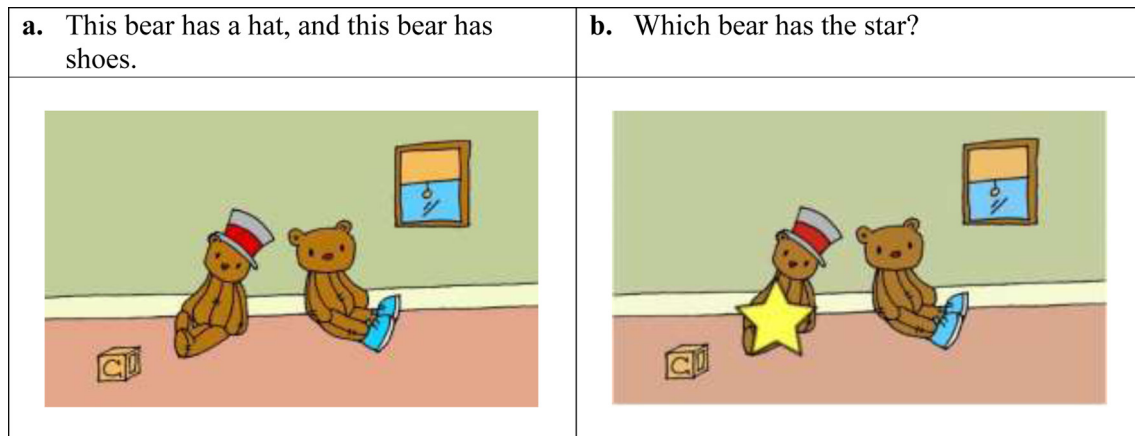


Figure 1: Sample comitative test item in the production task.

Children’s recorded responses were transcribed verbatim into a spreadsheet by the first author or a research assistant, and each transcription was verified independently by two additional research assistants. The elicited responses were then coded in terms of referential completeness, whether or not participants used the target PP embedding, and the actual prepositions used (*in*, *on*, *with*, other). As in similar production studies, children produced various types of approximate but non-target responses, including those given in (17). For the purposes of further analysis, only responses that were both referentially accurate and contained the syntax of PP embedding were classified as target, as in (18).

- (17)
- a. Incomplete responses: Descriptions that did not include sufficient information to identify the referent (e.g., *the bear*, *that one*).
 - b. Alternative responses: Descriptions that identified the correct referent by some means other than the target PPs (e.g., *the hat one*, *the bear that has a hat*).
- (18) Target responses: Descriptions that used the target configuration (embedded PP) to correctly identify the referent (e.g., *the bear with the hat*, *the one with the hat*).

3.4 Sentence recall task

To obtain a general assessment of the children’s language skills, we administered the Recalling Sentences subtest of the CELF-P2 (Wiig et al. 2004) standardized language test. Sentence repetition tasks have been shown to provide a useful measure of individual differences in language ability (e.g., Klem et al. 2015). The task was administered according to the instructions

in the test manual: children were instructed to listen to each sentence and repeat it after the experimenter, and two trial sentences were administered for training purposes. Each child then continued to repeat sentences of gradually increasing length, with the experimenter using the scoresheet to keep track of errors on each sentence. The experimenter assigned each sentence a score based on the number of errors in the repetition, and the task was discontinued after the child received three consecutive zero scores, or after they repeated all 13 sentences in the subtest. After the test sessions, a second research assistant listened to each recording and verified the scoring completed by the experimenter during the session, and the raw score was tallied for each participant.

3.5 Comprehension task: Procedure, materials, and coding

As explained above, the main comprehension task was presented after the other tasks. It was set up as a coloring activity, following the Coloring Book method proposed by Pinto & Zuckerman (2019), which has been found to be an excellent tool for assessing sentence interpretation (Gerard et al. 2018). At the outset, each participant was asked to identify the five colors (black, blue, green, yellow, red) in the palette given by the app. The experimenter then explained that she would read a sentence and the child should repeat the sentence and color the picture on the iPad to match it. The pilot phase revealed that some children enjoyed spreading the digital brush across the image. Thus, to facilitate coding, children were asked to use small dots of color which were practiced on a blank screen. Three training trials were then used to model the procedure for the child, and to reinforce the practice to color only the items mentioned in the sentence. In the first two training trials, the experimenter read a sentence (e.g., *the balloon is yellow*) and then colored the picture to match, intentionally making mistakes and asking the child to help correct them. In the third trial, the child carried out the task themselves, with correction by the experimenter as necessary. After the training items were complete, the experimenter proceeded to administer the test items, after reinstructing the child to repeat each sentence and color the picture to match. Corrective feedback was not provided for the test items, but the experimenter repeated each sentence on request from the child, or if the initial presentation did not elicit a response. Children were prompted to repeat the sentences, but some children frequently chose not to repeat, even with prompting.

The test materials contained 23 items: 12 test items and 11 control items. To mitigate potential ordering effects, the order of presentation was pseudorandomized, and counterbalanced for picture and sentence types across two stimuli lists (participants were randomly assigned to complete List A or List B). **Table 2** summarizes the test sentence types with examples of each (see Supplementary Files for the full sets of instructions and stimuli for the two lists). The test items were evenly divided between coordination, comitative PPs (*with*) and locative PPs (*in/on*). An additional methodological manipulation, contrastiveness, was nested within the PP conditions.

Sentence type	# of pictures	Example	# of nouns mentioned	# of items to be colored
Coordinated NP	4	The book and the apple are yellow	2	2
Comitative PP	2 unique 2 contrastive	The dog with the bone is blue	2	1
Locative PP	2 unique 2 contrastive	The pillow on the table is red	2	1

Table 2: Summary of test sentences.

Contrastiveness was implemented visually: in the unique referent conditions (**Figure 2a**), the matrix NP (e.g., *the pillow*) was the only object of its kind in the picture (which renders the prompt description over-informative), while the contrastive picture (**Figure 2b**) was not over-informative because it included a competing referent. (See section 2.2 for a brief discussion of pragmatic factors in modification). The purpose of this manipulation was to control whether the number of objects in the design affected children’s accuracy. When the scenarios for the coordinated and the PP conditions were visually matched, the use of PP modification was over-informative, and less felicitous. When the use of modification was set up as contrastive, the PP modification trials contained more graphic objects than those of the coordination trials. Given the possibility that either feature (having additional objects, or the uninformative use of modifiers) could influence children’s comprehension accuracy, including contrastiveness as a factor would allow for a more comprehensive interpretation of results.

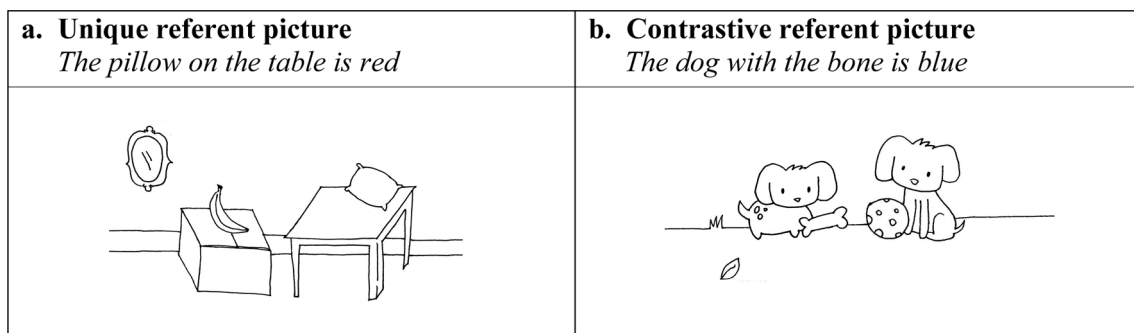


Figure 2: Sample pictures for the unique and contrastive target referents.

Controls were designed to vary in terms of syntactic format, number of nouns mentioned, and nouns to be colored. The simple controls (one noun mentioned, one object colored, $n = 6$) were a baseline for the basic ability to carry out the task. Five additional controls were manipulated to add structural variety to the stimuli, and to increase task complexity by adding

potential referents. These trials had various formats: two nouns mentioned, one colored; two identical nouns mentioned, both colored; two different nouns mentioned, both colored; and three coordinated NPs, all three colored. **Table 3** summarizes the structure of the control sentences, and sample pictures are given in **Figure 3**.

Sentence Type	# of trials	Example	# of nouns mentioned	# of items to be colored
Single NP predicative	4	The flower is black	1	1
Single NP intransitive	2	A green mouse is coming out	1	1
1-color transitive	1	A red bus hits a car	2	1
2-color transitive (same object)	1	A yellow truck pulls a black truck	2	2
2-color transitive (diff objects)	1	A green bunny chases a blue frog	2	2
3-NPs coordinated	2	The window, the flower, and the table are red	3	3

Table 3: Summary of control sentences.

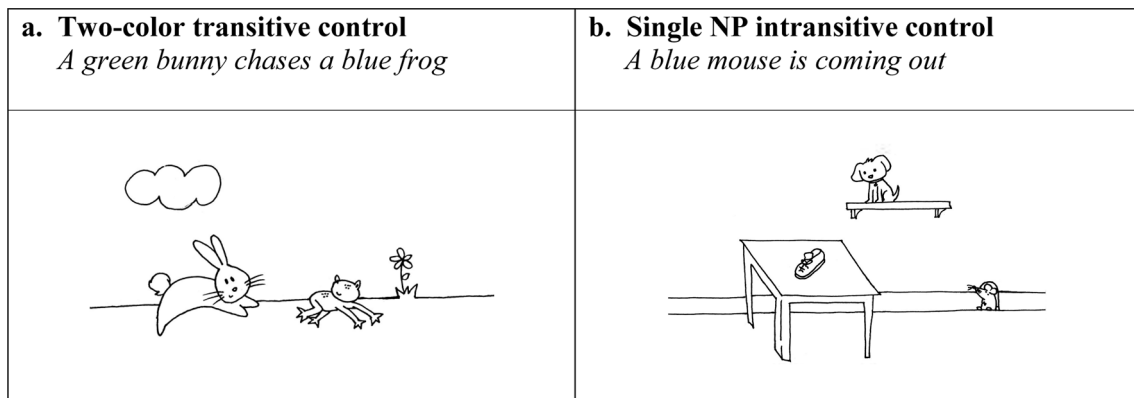


Figure 3: Sample pictures for two types of control items.

The Educreations (2016) app recorded the participants' actions in a video format. After testing, the experimenter or another trained research assistant watched each video and manually coded the participant's responses in terms of the particular item(s) colored in each picture. Children's responses were coded according to the entity colored, which were labelled by the order of appearance in the stimuli sentence (N1, N2, etc.). If the participant colored items not

mentioned in the sentence, these were coded as “Other N”. After the initial coding, two additional research assistants watched each video independently and recoded the data. Only four out of the total of 1311 responses by all participants (before exclusions) showed discrepancies across the three coders, representing 99.7% agreement. These four responses were deemed “unclear” and not codable by the scheme given in **Table 4**, which summarizes the target response patterns for test sentences.⁵

Nouns colored	N1 + N2	N1	N2	Other N
Coordinated NP	√	X	X	X
PP embedding	X	√	X	X

Table 4: Expected responses for coordinated NPs and embedded PPs.

4 Results

The comprehension, production, and sentence recall data for the final sample of 40 children and 10 adult controls were analyzed, and the following sections present the results.

4.1 Baseline measures: Production and Sentence Recall

We start by reporting the two baseline measures of language development, including children’s ability to produce PP modifiers and their performance on a standardized sentence recall task. These measures were intended to allow us to characterize our participants as typical language learners and to serve as the developmental baselines against which our comprehension results would be matched.

Production of PP modifiers was an open-response task, so that while all adult participants produced at least one PP modifier, adults as a group used the target PP strategy 72.5% of the time (29/40 responses) overall. Child participants, in contrast, had much lower production of target responses (40–50%). In **Table 5** we compare accuracy across conditions. The difference in the distribution of target/non-target responses was not significantly different across conditions ($\chi^2 = 1.6123, p = 0.204$).⁶

⁵ To mitigate potential pressure against the coordinated NP response, which involves coloring more objects than the embedded cases, the control items were designed to balance the task-related effort associated with each condition by requiring one to three objects to be colored (see Table 3).

⁶ In terms of non-target response types, the youngest children frequently pointed at the picture instead of providing a verbal response (26/84 responses by 3-year-olds; 29/160 responses by all children). Children of all ages also produced compound responses (*the hat one*; 17/160), full clauses (*the bear has a hat*; 15/160), relative clauses (*the bear who/that has a hat*; 10/160), and responses that were missing the head noun (*the hat* or *with the hat*; 11/160 responses). For the adults, the non-PP responses consisted of compounds (5/40 responses), relative clauses (4/40), and one response with the head noun omitted.

Age group	Locatives	Comitatives
3-year-olds	15/42 (35.7%)	8/42 (19.0%)
4-year-olds	16/26 (61.5%)	15/26 (57.7%)
5-year-olds	10/12 (83.3%)	9/12 (75.0%)
All children	41/80 (51.3%)	32/80 (40.0%)
Adults	16/20 (80.0%)	13/20 (65.0%)

Table 5: Numbers of target responses (with % in parentheses) produced by condition (locative vs. comitative) across age groups.

Individually, we observed that 31 out of the 40 children were able to produce at least one PP modifier in the production task. Performance improved sharply between ages 3 and 4: among 3-year-olds, 61.9% (13/21) produced one PP modifier, while 92.3% of 4-year-olds (12/13) and 100% of 5-year-olds (6/6) were able to produce at least one PP response. In general, then, the majority of children in the sample showed the ability to produce PP modifiers of the type targeted in the comprehension task (e.g., *the bear with the hat*), with performance improving to adult-like levels by age 5.

The other developmental measure included was the Recalling Sentences subtest of the CELF-P2 (Wiig et al. 2004). Three children declined to complete this task. The remaining children obtained raw scores of between 3 and 36, with a mean score of 20.54 (SD = 8.73 points). This is an expected range for children ages three to five, as a raw score of 3 corresponds to a scaled score of 8 for children ages 3;0 to 3;5, indicating average performance, while participants over age five are often able to repeat most of the sentences correctly (a perfect score is 37). Our developmental analyses rely on raw scores, rather than scaled (or standardized) scores that are age-referenced, in order to allow us to conduct correlation analyses with children's age in months.

4.2 Comprehension task

The 10 adult participants were all 100% accurate in completing the comprehension task, and their results are therefore not discussed in further detail. Below, we first discuss children's performance with control items, and then compare the three experimental conditions.

4.2.1 Accuracy in control conditions

Overall, children were highly accurate in understanding the control items, obtaining an average score of 86.8% on all control items combined. Recall that the 11 control items consisted of five different subtypes (see **Table 3**): four of these (one-color transitive, two-color transitive, intransitive, and single NP) involved simple NP structures, while the other involved a complex

three-noun coordinate NP structure, and three items to be colored rather than one or two. Although overall performance on the control items was high, performance was not uniform across the subtypes. As shown in **Figure 4**, children are at ceiling with the simple controls, with one referent to be colored (predicative sentences, intransitives, and one-color transitives). Performance is lowest with three coordinated NPs, where three different objects had to be colored. Interestingly, for the two-color transitives, children made more coloring errors when the type of referent was the same (32/40) than when two different types of objects were given different colors (38/40).

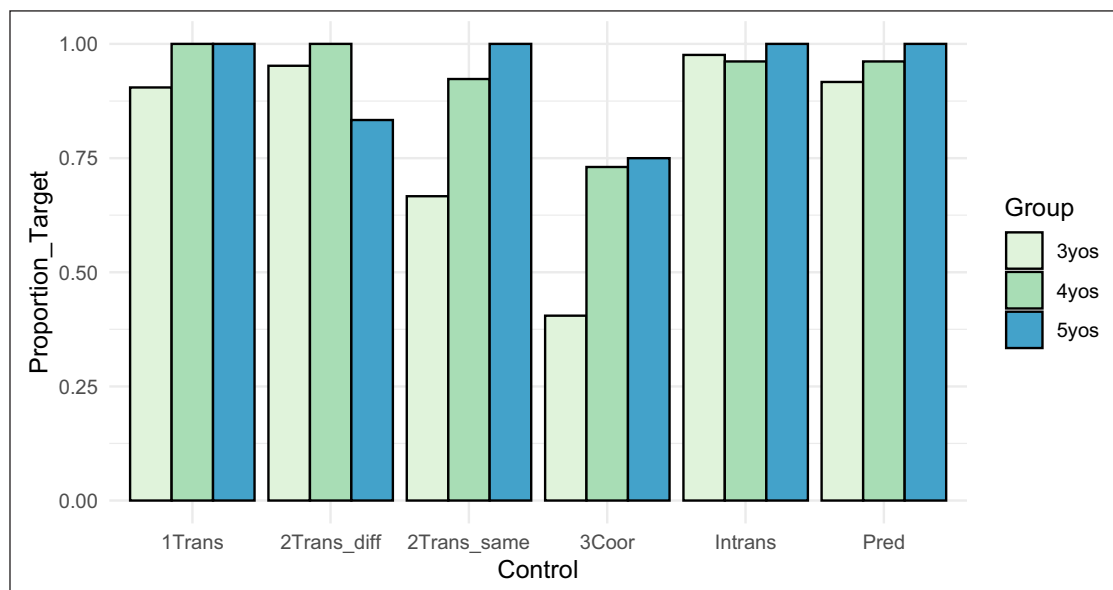


Figure 4: Proportion of target responses to the various control items, by group.

4.2.2 Accuracy in experimental conditions

The main question in this study relates to children's comprehension of coordinated NPs as compared with PP modifiers, specifically locative and comitative PPs. We hypothesized that, if children have some implicit knowledge of hierarchical structure, young children would clearly distinguish between coordinates and modifiers. In addition, we predicted that locative and comitative PPs would show different patterns, with comitatives being interpreted as coordinates more often.

Our results, illustrated in **Figure 5**, largely support these predictions. Children's performance in the coloring task varied by condition: they were very accurate with coordinated sentences (mean = 82.5%, SD = 38.1%), slightly less accurate with locative PP sentences (mean = 73.8%, SD = 44.1%), and much less accurate with comitative PPs (mean = 30.6%, SD = 46.2%).

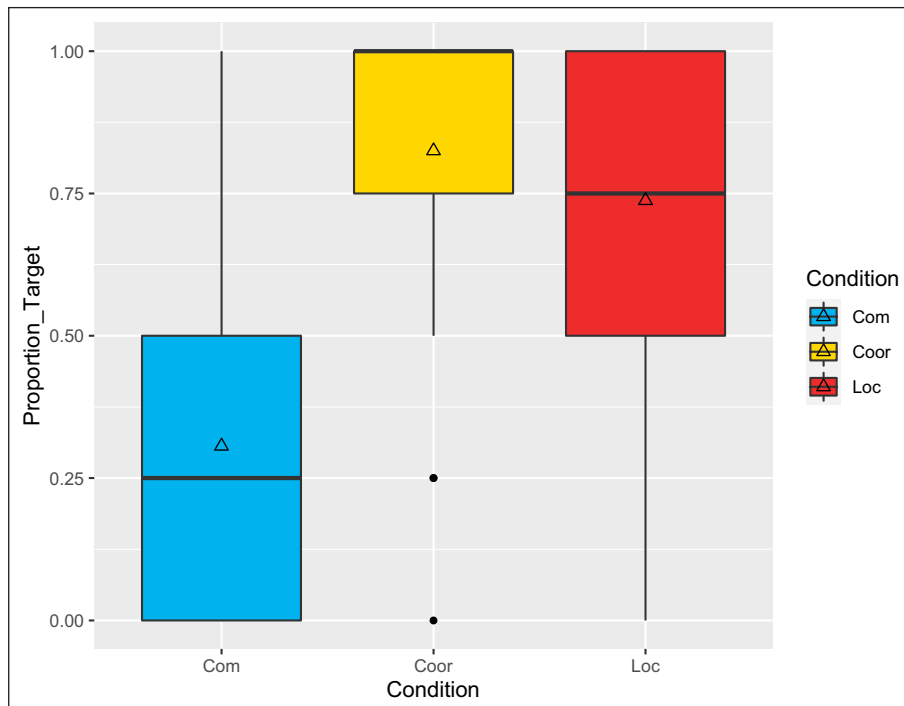


Figure 5: Children’s overall accuracy in the comprehension task by condition (Comitative: *The dog with the bone is blue*; Coordinate: *The book and the apple are yellow*; Locative: *The pillow on the table is red*). Bolded line indicates median score per condition; the small triangle indicates mean score.

We statistically analyzed the results using a generalized mixed effects logistic regression, using the *glmer* function in the *lme4* package (Bates et al. 2015) in R (R Core Team 2020). The data on accuracy (480 observations by 40 participants) were fit to three incrementally built models using maximum-likelihood estimation (Laplace), including random by-participant slopes for Condition.

The simplest model contained a single fixed effect, Age (in months) as a continuous variable, which had a small but highly significant effect ($\beta = 0.078$, $Z = 3.843$, $p < 0.001$) on responses. A second model was augmented with Condition, using treatment coding and coordinates as the reference level. This model had a better fit to the data than the simpler model (AIC = 515.5 vs. 552.7), and this difference was highly significant ($\chi^2(2) = 41.177$, $p < 0.001$). The results for this model also showed a significant effect of age ($\beta = 0.071$, $Z = 3.808$, $p < 0.001$). As for condition, the model revealed that the difference between comitatives and coordinates was highly significant ($\beta = -2.698$, $Z = -6.584$, $p < 0.001$), but the difference between the locative and coordinated conditions was not significant ($\beta = -0.618$, $Z = -1.586$, $p = 0.113$). The comparison between locatives and comitatives was obtained by changing the reference level to locatives, and this difference was highly significant ($\beta = -2.080$, $Z = -6.740$, $p < 0.001$).

The third model was built by adding the interaction of Age and Condition. The interaction did not improve the fit of the data (AIC = 518.0), and the difference in fit from the model without interactions was not significant ($\chi^2(2) = 1.545, p = 0.462$).

A secondary analysis focused on the presence or absence of contrast in the pictures used in the PP conditions. The goal was to evaluate whether these contextual conditions that make the use of a modifier more or less informative influenced children's accuracy in this coloring task. **Figure 6** shows children's accuracy with the two types of PP items for each contrast condition (contrastive = two competing items in the picture; non-contrastive = only one named item in the picture).

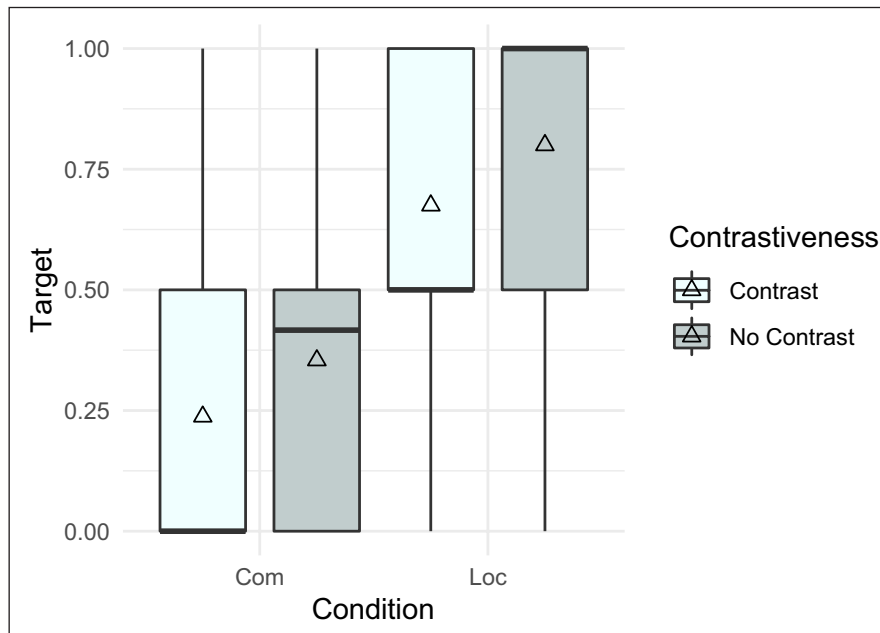


Figure 6: Average proportion of accurate responses on embedded items by contrast in picture. Bolded line indicates median score per condition; the small triangle indicates mean score.

These results show that children's accuracy was around 10% lower for stimuli with contrast compared to non-contrastive stimuli in both PP conditions. In other words, the presence of a competing item in the picture was associated with poorer performance overall. We ran a single model on the PP conditions only in order to evaluate the potential role of the contrast manipulation: Contrast (implemented with treatment coding: contrastive vs. non-contrastive) was entered as a fixed effect, with a random intercept for Participant.⁷ The results of this model

⁷ We also ran a model with random by-participant slopes for contrast, but this did not improve the fit of the data (AIC = 443.7 vs. 440.1).

(320 observations by 40 participants, AIC = 440.1) confirmed that contrast played a significant role: children were more accurate in the non-contrastive trials than in the contrastive trials ($\beta = 0.466$, $Z = 1.976$, $p = 0.048$).

4.2.3 Response patterns in experimental conditions

In addition to overall accuracy, we analyzed children’s responses in terms of the type of response given. This analysis had two goals: to determine main error patterns for each experimental condition, and to assess group performance against chance. A listener guessing the interpretation of a complex NP containing two nouns connected by a marker (N1 marker N2) has three main choices of response: a top head interpretation, in which N1 is colored, a lower head interpretation, in which N2 is colored, and a coordinate interpretation, in which both N1 and N2 are colored. We therefore set the probability of guessing the correct response at 33%. A series of *t*-tests of children’s performance against chance showed that children appear to be guessing at the meaning of comitative sentences ($t = -0.335$, $df = 38$, $p = 0.630$), while their comprehension of coordinated and locative NPs was significantly different from chance (Coordinated: $t = 12.27$, $df = 38$, $p < 0.001$; Locative: $t = 9.223$, $df = 38$, $p < 0.001$). **Table 6** shows the response types by condition for all children combined: the unshaded box represents the correct answer for that condition, while the shaded boxes show the incorrect response types (including an “Other” category for all responses other than N1, N2, and N1N2).

For the coordinated and locative conditions, in which overall accuracy is relatively high, children show a variety of error types. For the comitative condition, by contrast, the incorrect N1N2 response is more common than all of the other response types, including the correct N1 response. In other words, the children frequently interpreted comitative sentences in the same way as coordinates.

	N1N2	N1	N2	Other
Coor	82.5%	5.0%	1.3%	11.3%
Com	45.6%	30.6%	15.6%	8.1%
Loc	11.3%	73.8%	5.6%	9.4%

Table 6: Group performance by response type. Target response types are left unshaded.

4.2.4 Individual analyses of the three main test conditions

The next step in our analysis was to examine individual patterns of responses. Since our primary question asks whether individual children know that PP modifiers, unlike coordinate NPs, are referentially inaccessible to color predication, we classified each child according to the number of coordinate (i.e., N1N2) responses they gave per condition. We then plotted the frequencies of

individual children so classified. In the histograms given in **Figure 7**, the x-axis represents the number of N1N2 responses given (from 0 to 4), and the y-axis, the number of individual children who gave that number of N1N2 responses.

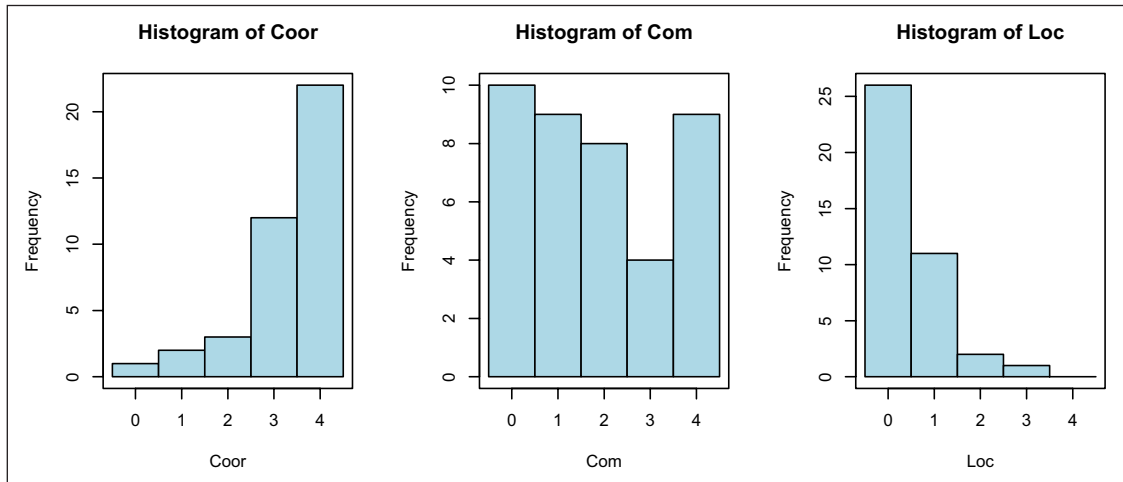


Figure 7: Number of children classified according to how many coordinated (N1N2) responses they gave in each condition (from 0 to 4).

The coordinated and locative condition results resemble mirror images of each other: most children produced three or four N1N2 responses in the coordinated condition, where this is the correct answer. In the locative condition, where this response type is incorrect, very few children fell into this pattern (two children gave two N1N2 responses, and another child gave three). In the comitative condition, performance appears much more mixed, and seemingly bimodal.

As in the previous section, we depart from the assumption that children who are paying attention to the sentence, but do not necessarily understand the structure, would color only the named elements. Since the three most relevant possible responses would be to color the first noun (N1), the second noun (N2), or both nouns (N1N2), the probability of guessing the correct response by chance is set at 33%. Using the binomial distribution, the probability of getting four out of four trials correct is estimated at 1.2%, and the probability of getting three out of four correct is 9.6%. We consider children in these two accuracy groups to be “reliable comprehenders”.

For the coordinate trials, 34/40 children reliably chose N1N2, the correct response for this condition. For the locative trials, 27/40 children reliably chose N1, treating locatives as modifiers; an additional 9 chose it twice out of the four trials. These same children were quite ambivalent about comitative trials: only one-sixth of children (6/40) were reliably assigning those trials the target modifier (N1) interpretation, and an additional 10 children chose the target response

half the time. On the other hand, one-third of children (13/40) reliably assigned comitative PPs a coordinate N1N2 interpretation, with 8 additional children showing a bias in this direction, choosing the coordinate interpretation half the time. In sum, for the comitative condition, 40% of the children were reliable or biased toward the target response, and 53% were reliable or biased towards the coordinate interpretation. The answers given by the remainder of the children ($n = 3$) had no discernible pattern.

4.3 Developmental patterns

Our final analysis was an exploration of the developmental patterns observable in our data. Does comprehension develop with age, along with general language development, as represented by the standardized sentence recall measure, or by the ability to produce complex NPs with modifiers? Domains that are learned or grow along with domain-general capacities during the preschool years should exhibit positive correlations with age. Domains that are learned should also be associated with general language measures.

To examine language development, we calculated correlations between age in months and children's raw scores on the various tasks: sentence recall, production of PPs (including each subtype), and comprehension tasks. **Table 7** presents the descriptive statistics for all tasks, and the Spearman correlations between all variables; the correlations of interest are discussed below.

As shown in **Table 7**, sentence recall scores show positive, moderate correlations with age, as do scores on the production task, both overall and by condition (locative and comitative). This is also the case for the control items of the comprehension task, while the coordinate scores show a small but significant correlation with age. These age results reflect that these measures are developing during the age span investigated. Accuracy with control sentences is also significantly correlated with the sentence recall and overall production scores. Interestingly, there was a significant correlation between accuracy in comprehending coordinates and control sentences. Finally, we note that comprehension of coordinates had a moderate but significant association with PP production, whereas PP comprehension scores were not correlated with PP production scores.

As for the three main test conditions, we calculated Kendall correlations between age in months and response accuracy for each condition separately. Only the coordinated condition was found to be significant ($T_b = 0.302, p = 0.018$). There was no evidence of association between age and PP comprehension, neither for locatives ($T_b = 0.178, p = 0.153$) nor comitatives ($T_b = 0.170, p = 0.169$). The scatterplots in **Figure 8** show these results for age, separated by condition. For coordinated NPs, we observe a sharp increase for the younger children. During the same age range of observation, the PP conditions remain somewhat stable.

Variable	<i>N</i>	<i>range</i>	<i>M</i>	<i>SD</i>	<i>Age</i>	<i>SR</i>	<i>Ptotal</i>	<i>Ploc</i>	<i>Pcom</i>	<i>Cont</i>	<i>Ccoo</i>	<i>Cloc</i>	<i>Ccom</i>
1. Age	40	36–64	48.73	7.76	–								
2. Sentence Recall	37	3–36	20.54	8.73	0.68***	–							
3. Prod Total	40	0–4	1.83	1.43	0.65***	0.35*	–						
4. Prod Loc	40	0–2	1.03	0.83	0.50**	0.29	0.87***	–					
5. Prod Com	40	0–2	0.80	0.82	0.61***	0.32	0.84***	0.49**	–				
6. Controls	40	7–11	9.55	1.38	0.56***	0.45**	0.34*	0.30	0.26	–			
7. Comp Coor	40	0–4	3.30	0.99	0.37*	0.27	0.44**	0.46**	0.28	0.46**	–		
8. Comp Loc	40	0–4	2.95	1.08	0.22	0.34*	–0.04	–0.08	0.00	0.09	0.03	–	
9. Comp Com	40	0–4	1.23	1.19	0.22	0.28	–0.06	–0.01	–0.08	0.09	–0.14	0.18	–

Table 7: Descriptive statistics and Spearman correlations between study measures. Significant correlations are indicated with asterisks ($p \leq 0.05^*$, $p \leq 0.01^{**}$, $p \leq 0.001^{***}$).

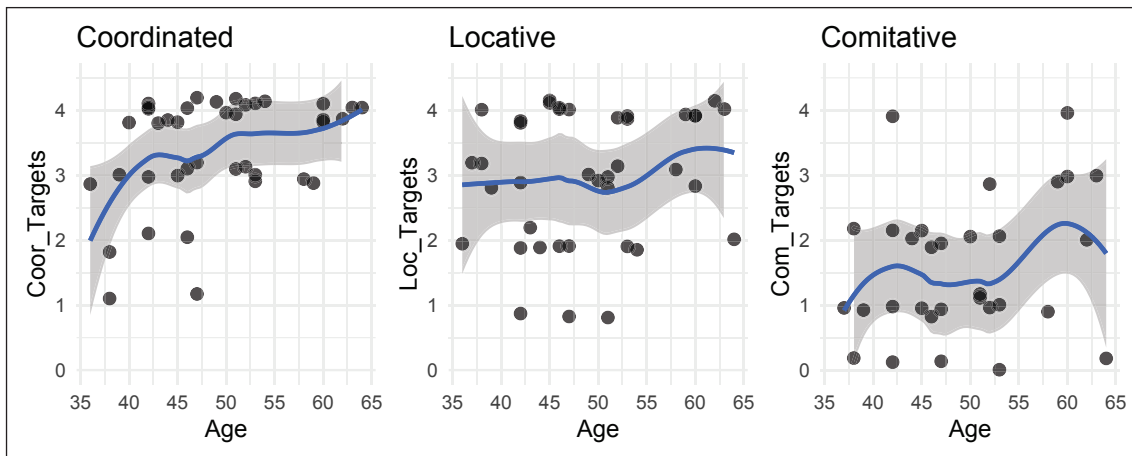


Figure 8: Correlations between age and target responses in each of the experimental conditions.

The age effect for coordinates indicates that better performance with multiple objects grows with age, similar to the results for the control items, which showed less accuracy when more objects were involved. By contrast, the lack of an age effect for PPs suggests stable representations, mediated by the particular preposition involved, as discussed in the previous subsection.⁸

5 Discussion

Our first question asked about children’s ability to interpret complex NPs. Our analysis compared performance with sentences containing PP modifiers and coordinated NPs. Children performed very well with PP modifiers containing *in* and *on*, showing similar accuracy as with coordinate structures. For comitative PPs, on the other hand, the data looks quite different: children performed at chance level for *with* sentences. This characterization is confirmed by the individual analyses, where we see that most children are very accurate in interpreting both coordinate and locative PP sentences. Our second question concerned the lexical type of prepositions. In contrast to locative structures, for comitatives we instead observe what seems like a bimodal distribution: most children treat it as either a coordinate structure or a PP modifier, and about one-quarter of the children are in the middle, choosing equal numbers of N1 and N1N2 responses. The number of other responses given by children is comparable across conditions.

When we explore separately the effect of age, we see that the coloring task shows an overall developmental effect. Errors with control sentences are strongly correlated with age; when we

⁸ Given the prevalence of the coordinate (N1N2) response in the comitative condition (46% of all responses, with 30 out of 40 children giving at least one such response), a *Glossa* reviewer recommends exploring the developmental patterns for this response type. A Kendall correlation between age and number of N1N2 responses in the comitative condition shows no indication of developmental changes for this prominent pattern ($T_b = 0.065, p = 0.589$).

examine performance for the various types of control sentences, errors seem to increase with the number of objects named, suggesting a task performance effect. Accuracy with control sentences is also moderately correlated with the other developmental measures of PP production and sentence recall.

For the experimental conditions, only comprehension of the coordinate NPs is significantly correlated with age. Descriptively, this effect appears driven by the younger children. As in the case of control items, we interpret this finding to mean that children's performance with multiple objects improves with age. On the other hand, the lack of age effects for the PP conditions suggests stable representations. For locatives, children's interpretations are accurate from a young age, whereas for comitatives, the bimodal character of their responses seems to reflect an unresolved decision. In addition, we observed that contrastive scenarios elicited slightly more incorrect responses than non-contrastive (unique item) scenarios. This likely indicates another task performance effect: in the contrastive condition, the additional object in the visual scenario (which rendered the use of modifiers felicitous) made it more difficult for children to identify the correct item.

In general, the present study has examined the comprehension performance of children at the time in which they begin to produce NP modification. Our correlational data confirm that their ability to produce PP-modified NPs, both locative and comitative types, is growing with age, as suggested by previous literature (Eisenberg et al. 2008). However, children's ability to interpret these complex NPs correctly is not uniform across these two lexical classes. We see very good performance with locative PPs, with most children exclusively or near-exclusively identifying the correct referent. On its own, this finding is sufficient to establish that three- to five-year-old children clearly understand an important semantic consequence of the hierarchical structure of complex NPs. As such, this is congruent with the insights of work with younger children, as in studies by Lidz et al. (2003), Koulaguina et al. (2019), and Lorimor et al. (2019).

One potential objection to our interpretation could be raised with regards to the work that shows early sensitivity to the ungrammaticality of agreement mismatches. It is possible that children are using agreement as a cue to the referential status of the complex NPs. Under that account, the presence of the plural copula in coordinate structures is what guides the child to the two-referent target, and to only one referent in the locative case. Note that use of agreement alone does not explain why agreement helps locatives but is not sufficient to guide comitatives, nor how children decide that predication should affect only the highest noun (N1), not the embedded noun.

Our experimental comparison is very similar to the materials in Shi, Legrand & Brandenberger (2020) mentioned in section 2.3. Recall that those authors presented children with complex

subjects (coordinates or locative-modified), testing gender/number agreement between the left dislocated subject and the redundant French subject pronoun. They found clear discrimination of feature grammaticality in 30-month-old French-speaking toddlers, who showed a novelty preference towards the ungrammatical stimuli, that is, they looked longer while listening to the incorrect trials. Subsequent work by Shi, Emond & Badri (2020) finds discrimination in younger children, at 17 months. Both results are interpreted correctly as evidence of sensitivity to agreement as well as to the structural differences between coordinates and PP structures. Our results parallel those findings: children clearly discriminate between coordinates and locative PPs. Our results further suggest that agreement is not by itself a reliable cue.⁹ In striking contrast to the locative condition, in the comitative condition we found chance-level performance overall, and many children preferring the coordinate interpretation of the PP modifier *with*. This pattern of results, specifically the difference between the two PP conditions, suggests that it takes time for children to disambiguate whether the comitative marker functions as a preposition or as a conjunction.

From another perspective, we note that in the comitative condition, children as a group are not simply making random errors, but seem to be ambivalent between a coordinate and a modifier interpretation. The specific pattern of children's errors we observe in the comitative condition is in line with the expected direction of diachronic change. As previously discussed, the typological literature documents unidirectional change: languages with comitative conjunctions evolve from grammars with comitative prepositional modifiers (Haspelmath 2007). Our findings therefore align with an emergent literature suggesting that child acquisition drives language change (Hall 2020), and more specifically, that patterns of cyclic language change arise from children's lexical learning biases (Cournane 2019; Cournane & Pérez-Leroux 2020). In other words, the tendency for children to interpret comitative *with* as the conjunction *and* may be an important factor contributing to this common diachronic pattern. This particular learning bias is explainable if we consider that children use *with* as a verbal modifier from a young age: in a sentence like *the girl is playing with the boy*, both the subject NP and the verb modifier are engaging in the action, and it is possible that children extend this aspect of semantics to their interpretation of comitative PPs in complex NPs.

Taken together, these data allow us to evaluate our various hypotheses. Our first pair of contrasting hypotheses concerns whether referential accessibility results from a learning process or from implicit knowledge of hierarchical structure. In the former scenario, the expectation is a stage where children are willing to interpret PP modifiers as referentially accessible. In the latter scenario, knowledge of the referential inaccessibility of PP modifiers should be robust

⁹ This is not surprising, given that children have a certain degree of difficulty applying their implicit awareness of the ungrammaticality of agreement mismatches into explicit comprehension tasks, such as the ones in the present study.

and present at the stage where complex NPs enter the grammar. The evidence shows that young children consistently distinguish coordinate NPs from locative PP modifiers, supporting the view that this type of structural knowledge does not show a learning curve and is implicitly available to children. In the context of the learning question, the comitative data stands out as anomalous. The potential distinctiveness of the developmental path of *with* is the subject of our second hypothesis. We speculated that if children are responsible for regular cyclic changes, we might see that some children learning a comitative modification language will allow coordinate semantics for *with*. This documented historical process shows a striking parallel with our error data. The semantics of *with* can be assigned to two configurations: endocentric PP modification or non-endocentric comitative conjunction. The bimodal pattern we observe is what one would expect if structure assignment is a categorical choice. Therefore, we argue that the present results for comitatives lend support to the proposal that regular language change is driven by children's learning biases.

6 Conclusion

When children hear *the book on the table* or *the bear with the hat*, how do they integrate all of the component parts for interpretation? At the stage where they are beginning to produce complex NPs, the preschoolers in our study are reliably differentiating coordinated and locative-modified NPs. This is important evidence that children have insight into the semantic consequences of embedding. The property of referential inaccessibility under embedding (Arsenijevic & Hinzen 2012) seems a robust part of children's syntactic and semantic knowledge. This knowledge fails to exhibit a learning curve for the age range studied. This is a remarkable finding, given that prepositional linkers need to be learned, and given previous observations that children's knowledge of the referential inaccessibility of embedded CPs emerges later (de Villiers & de Villiers 2000; de Villiers 2018).

The asymmetry in our results (with target performance for locative prepositions, but chance performance for comitative *with*) rules out the possibility that children are simply relying on the form of the copula to guide their interpretations. Instead, the robust performance with unambiguous locative prepositions suggests that preposition meaning is the key. All three prepositions in our study, *in*, *on* and *with*, have basic spatial senses: containment, support, and adjacency, respectively. The comitative sense of *with* is used early by children; on that basis, we should expect equally good performance with all three lexical prepositions. At the same time, we noted that *with* has a diachronic record of grammaticalizing away from its prepositional qualities, giving rise to the comitative coordination construction. Children in our study did not simply fail to understand *with* modifiers and randomly choose an interpretation. Instead, we observed two dominant patterns: while quite a few children showed the adult pattern, most

children consistently treated *with* as a coordinating conjunction, ignoring the overt cue provided by the inflected singular copula.¹⁰

To conclude, in the domain of the noun phrase, children demonstrate knowledge that is unlikely to be incrementally learned (that embedded NPs are referentially inaccessible), built out of knowledge which is clearly learned (the basic inventory of lexical prepositions). The overall developmental patterns observed in our study support this inference. Children's comprehension of PP structure, as illustrated by the locative/coordinate contrast, shows that children's sensitivity to hierarchical structure goes beyond the mere distribution of forms, and directly shapes their sentence interpretation.

¹⁰ Ours is not the only work to tap into the potential linguistic overlap between coordinated and comitative interpretations. A study by Arunachalam, Syrett & Chen (2016) shows limitations in the comprehension of coordinate clauses. In an experiment that prompted them to choose between causative (one agent acting on a patient) and parallel events (two agents synchronously performing the same activity), children typically selected the causative event when presented with a transitive sentence with a novel verb (*X is verbing Y*). When presented with conjoined intransitive frames (*X and Y are verbing*), children and even some adults performed at chance. Children were able to use the lexical semantics of functional elements, in their study, adverbial modifiers, to augment the information provided by syntactic frames. Their basic results suggest the coordinate frame is not informative enough on its own. See also Gertner & Fisher (2012) and Noble et al. (2016) for related evidence on how contextual manipulations may aid coordinate comprehension.

Abbreviations

SUBJ = subject, MASC = masculine, FEM = feminine

Supplementary files

Supplementary file: Comprehension task data, scripts, and stimuli & instructions for production and comprehension tasks. DOI [10.17605/OSF.IO/BVKZ4](https://doi.org/10.17605/OSF.IO/BVKZ4)

Ethics and Consent

This study was approved by the University of Toronto Social Sciences, Humanities and Education Research Ethics Board, Protocol Reference #32990. Informed consent was obtained from each adult participant and the legal guardian of each child participant, and all research data has been anonymized.

Acknowledgements

We thank Susana Béjar, Diane Massam, Yves Roberge, and other members of the Complexity & Recursion Project for helpful discussion, Jill de Villiers for comments on the first version of this manuscript, and audiences at BUCLD and the Meeting of the Canadian Linguistic Association (CLA). We also thank Natalia Docteur, Mathusha Sanjeevan, Karoline Antonsen, Michelle Heo, and Cliona Duah for their assistance with recruitment, testing, and data analysis, Cata Morales for the illustrations, and Miguel Barreto for support with statistics.

Funding Information

This study was conducted with support from SSHRC Insight Grant 435-2014-2000 Development of NP complexity in children to A.T. Pérez-Leroux and Y. Roberge.

Competing interests

The authors have no competing interests to declare.

References

- Allen, Shanley. 2017. Polysynthesis in the acquisition of Inuit languages. In Fortescue, Michael & Mithun, Marianne & Evans, Nicholas (eds.), *The Oxford Handbook of Polysynthesis*, 449–472. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199683208.013.25>
- Arsenijevic, Boban & Hinzen, Wolfram. 2012. Recursion as a human universal and as a primitive. *Biolinguistics* 4(2–3). 165–173. DOI: <https://doi.org/10.5964/bioling.8781>

- Arslan, Burcu & Hohenberger, Anette & Verbrugge, Rineke. 2017. Syntactic recursion facilitates and working memory predicts recursive theory of mind. *PLoS ONE* 12(1). e0169510. DOI: <https://doi.org/10.1371/journal.pone.0169510>
- Arunachalam, Sudha & Syrett, Kristen & Chen, YongXiang. 2016. Lexical disambiguation in verb learning: Evidence from the conjoined-subject intransitive frame in English and Mandarin Chinese. *Frontiers in Psychology* 7. 138. DOI: <https://doi.org/10.3389/fpsyg.2016.00138>
- Aslin, Richard N. & Newport, Elissa L. 2014. Distributional language learning: Mechanisms and models of category formation. *Language Learning* 64(Suppl 2). 86–105. DOI: <https://doi.org/10.1111/lang.12074>
- Barrière, Isabelle & Kresh, Sarah & Aharodnik, Katsiaryna & Legendre, Géraldine & Nazzi, Thierry. 2019. Three streams of generative language acquisition research. In Ionin, Tanya & Rispoli, Matthew (eds.), *Three Streams of Generative Language Acquisition Research: Selected papers from the 7th meeting of Generative Approaches to Language Acquisition–North America*, 7–33. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/lald.63.02bar>
- Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Béjar, Susana & Massam, Diane & Pérez-Leroux, Ana T. & Roberge, Yves. 2020. Rethinking complexity. In Bárány, András & Biberauer, Theresa & Douglas, Jamie & Vikner, Sten (eds.), *Syntactic Architecture and its Consequences. Vol 1: Syntax inside the Grammar*, 15–24. Language Sciences Press.
- Bock, Kathryn & Miller, Carol A. 1991. Broken agreement. *Cognitive Psychology* 23(1). 45–93. DOI: [https://doi.org/10.1016/0010-0285\(91\)90003-7](https://doi.org/10.1016/0010-0285(91)90003-7)
- Brandt-Kobebe, Oda-Christina & Höhle, Barbara. 2010. What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua* 120(8). 1910–1925. DOI: <https://doi.org/10.1016/j.lingua.2010.02.008>
- Chien, Yu-Chin & Wexler, Kenneth. 1990. Children’s knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1(3). 225–295. DOI: https://doi.org/10.1207/s15327817la0103_2
- Chomsky, Noam. 2001. Derivation by phase. In Kenstowicz, Michael (ed.), *Ken Hale: A Life in Language*. 1–50. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2005. On phases. Ms. MIT.
- Collins, John. 2017. The copredication argument. *Inquiry* 60(7). 675–702. DOI: <https://doi.org/10.1080/0020174X.2017.1321500>
- Cournane, Ailís. 2019. A developmental view on incrementation in language change. *Theoretical Linguistics* 45. 127–150. DOI: <https://doi.org/10.1515/tl-2019-0010>
- Cournane, Ailís & Pérez-Leroux, Ana Teresa. 2020. Leaving obligations behind: Epistemic incrementation in preschool English. *Language Learning and Development* 16(3). 270–291. DOI: <https://doi.org/10.1080/15475441.2020.1738233>

- Culbertson, Jennifer & Newport, Elissa L. 2015. Harmonic biases in child learners: In support of language universals. *Cognition* 139. 71–82. DOI: <https://doi.org/10.1016/j.cognition.2015.02.007>
- Davies, Benjamin & Xu Rattanasone, Nan & Demuth, Katherine. 2020. Comprehension of the copula: Preschoolers (and sometimes adults) ignore subject-verb agreement during sentence processing. *Journal of Child Language* 47(3). 695–708. DOI: <https://doi.org/10.1017/S0305000919000680>
- Davies, Catherine & Katsos, Napoleon. 2010. Over-informative children: Production/comprehension asymmetry or tolerant of pragmatic violations? *Lingua* 120(8). 1956–1972. DOI: <https://doi.org/10.1016/j.lingua.2010.02.005>
- de Villiers, Jill G. 2018. Perspectives on truth: The case of language and false belief reasoning. In Syrett, Kristen & Arunachalam, Sudha (eds.), *Semantics in language acquisition*, 222–245. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/tilar.24.10vil>
- de Villiers, Jill G. 2021. With language in mind. *Language Learning and Development* 17(2). 71–95. DOI: <https://doi.org/10.1080/15475441.2020.1820338>
- de Villiers, Jill G. & de Villiers, Peter A. 2000. Linguistic determinism and the understanding of false beliefs. In Mitchell, Peter & Riggs, Kevin (eds.), *Children's reasoning and the mind*, 191–228. Psychology Press.
- de Villiers, Jill G. & Johnson, Valerie E. 2007. The information in third-person /s/: Acquisition across dialects of American English. *Journal of Child Language* 34(1). 133–58. DOI: <https://doi.org/10.1017/S0305000906007768>
- Deevy, Patricia & Leonard, Laurence B. & Marchman, Virginia A. 2017. Sensitivity to morphosyntactic information in 3-year-old children with typical language development: A feasibility study. *Journal of Speech, Language, and Hearing Research* 60(3). 668–74. DOI: https://doi.org/10.1044/2016_JSLHR-L-15-0153
- Durkin, Kevin. 1981. Aspects of late language acquisition: School children's use and comprehension of prepositions. *First Language* 2(4). 47–60. DOI: <https://doi.org/10.1177/014272378100200404>
- Educreations. 2016. Educreations Interactive Whiteboard v. 2.2.7. [Mobile application software]. Retrieved from <http://itunes.apple.com>.
- Eisenberg, Sarita L. & Ukrainetz, Teresa A. & Hsu, Jennifer R. & Kaderavek, Joan N. & Justice, Laura M. & Gillam, Ronald B. 2008. Noun phrase elaboration in children's spoken stories. *Language, Speech, and Hearing Services in Schools* 39(2). 145–157. DOI: [https://doi.org/10.1044/0161-1461\(2008/014\)](https://doi.org/10.1044/0161-1461(2008/014))
- Enge, Alexander & Friederici, Angela D. & Skeide, Michael A. 2020. A meta-analysis of fMRI studies of language comprehension in children. *NeuroImage* 215. 116758. DOI: <https://doi.org/10.1016/j.neuroimage.2020.116858>
- Fengler, Anja & Meyer, Lars & Friederici, Angela D. 2015. Brain structural correlates of complex sentence comprehension in children. *Developmental Cognitive Neuroscience* 15. 48–57. DOI: <https://doi.org/10.1016/j.dcn.2015.09.004>
- Fisher, Cynthia & Jin, Kyong-sun & Scott, Rose M. 2020. The developmental origins of syntactic bootstrapping. *Topics in Cognitive Science* 12. 48–77. DOI: <https://doi.org/10.1111/tops.12447>

- Fisher, Cynthia & Klingler, Stacy L. & Song, Hyun-joo. 2006. What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition* 101(1). B19–B29. DOI: <https://doi.org/10.1016/j.cognition.2005.10.002>
- Franchetto, Bruna & Santos, Mara. 2017. The ontology of roots and the emergence of nouns and verbs in Kuikuro. In Vapnarsky, Valentina & Veneziano, Edy (eds.), *Lexical Polycategoriality: Cross-linguistic, cross-theoretical and language acquisition approaches*, 275–306. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/slcs.182.10fra>
- Gerard, Juliana & Lidz, Jeffrey & Zuckerman, Shalom & Pinto, Manuela. 2018. The acquisition of adjunct control is colored by the task. *Glossa: a journal of general linguistics* 3(1). 75. DOI: <https://doi.org/10.5334/gjgl.547>
- Gertner, Yael & Fisher, Cynthia. 2012. Predicted errors in children's early sentence comprehension. *Cognition* 124(1). 85–94. DOI: <https://doi.org/10.1016/j.cognition.2012.03.010>
- Giblin, Iain & Zhou, Peng & Bill, Cory & Shi, Jiawei & Crain, Stephen. 2019. The spontaneous eMERGEance of recursion in child language. In Brown, Megan M. & Dailey, Brady (eds.), *Proceedings of the 43rd annual Boston University Conference on Language Development [BUCLD 43]*, 270–285. Somerville, MA: Cascadilla Press.
- Gleitman, Lila. 1990. The structural sources of verb meaning. *Language Acquisition* 1(1). 3–55. DOI: https://doi.org/10.1207/s15327817la0101_2
- Gleitman, Lila R. & Cassidy, Kimberly & Nappa, Rebecca & Papafragou, Anna & Trueswell, John C. 2005. Hard words. *Language Learning and Development* 1(1). 23–64. DOI: https://doi.org/10.1207/s15473341lld0101_4
- Goldin-Meadow, Susan. 1982. The resilience of recursion: A study of a communication system developed without a conventional language model. In Wanner, Eric & Gleitman, Lila R. (eds.), *Language Acquisition: The State of the Art*, 51–77. New York: Cambridge University Press.
- Goldin-Meadow, Susan & Yang, Charles. 2017. Statistical evidence that a child can create a combinatorial linguistic system without external linguistic input: Implications for language evolution. *Neuroscience & Biobehavioral Reviews* 81(Part B). 150–157. DOI: <https://doi.org/10.1016/j.neubiorev.2016.12.016>
- Gonzalez-Gomez, Nayeli & Hsin, Lisa & Barrière, Isabelle & Nazzi, Thierry & Legendre, Géraldine. 2017. Agarra, agarran: Evidence of early comprehension of subject-verb agreement in Spanish. *Journal of Experimental Child Psychology* 160. 33–49. DOI: <https://doi.org/10.1016/j.jecp.2017.02.010>
- Grigoroglou, Myrto, & Papafragou, Anna. 2019. Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology* 55(5). 951–966. DOI: <https://doi.org/10.1037/dev0000693>
- Grimm, Hannelore. 1975. On the child's acquisition of semantic structure underlying the wordfield of prepositions. *Language and Speech* 18(2). 97–119. DOI: <https://doi.org/10.1177/002383097501800201>
- Grinstead, John. 2016. Root infinitives in child language and the development of clause structure. In Lidz, Jeffrey & Snyder, William & Pater, Joe (eds.), *The Oxford Handbook of Developmental*

- Linguistics*, 341–366. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199601264.013.15>
- Hall, Erin. 2020. Child participation in linguistic changes in progress in Ontario English. Doctoral dissertation, University of Toronto.
- Haspelmath, Martin. 2007. Coordination. In Shopen, Timothy (ed.), *Language Typology and Syntactic Description (Vol 2.)*, 1–51. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511619434.001>
- Hollebrandse, Bart & Roeper, Thomas. 2014. Empirical results and formal approaches to recursion in acquisition. In Roeper, Thomas & Speas, Margaret (eds.), *Recursion: Complexity in Cognition, Studies in Theoretical Psycholinguistics*, 179–219. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-3-319-05086-7_9
- Johnson, Valerie E. & de Villiers, Jill G. & Seymour, Harry N. 2005. Agreement without understanding? The case of third person singular /s/. *First Language* 25(3). 317–330. DOI: <https://doi.org/10.1177/0142723705053120>
- Johnston, Judith & Slobin, Dan I. 1979. The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* 6(3). 529–546. DOI: <https://doi.org/10.1017/S030500090000252X>
- Katsos, Napoleon & Bishop, Dorothy V. M. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120(1). 67–81. DOI: <https://doi.org/10.1016/j.cognition.2011.02.015>
- Kidd, Evan & Cameron-Faulkner, Thea. 2008. The acquisition of the multiple senses of with. *Linguistics* 46. 33–62. DOI: <https://doi.org/10.1515/LING.2008.002>
- Klem, Marianne & Melby-Lervåg, Monica & Hagtvet, Bente & Lyster, Solveig-Alma Halaas & Gustafsson, Jan-Eric & Hulme, Charles. 2015. Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science* 18(1). 146–154. DOI: <https://doi.org/10.1111/desc.12202>
- Kouider, Sid & Halberda, Justin & Wood, Justin & Carey, Susan. 2006. Acquisition of English number marking: The singular–plural distinction. *Language Learning and Development* 2(1). 1–25. DOI: https://doi.org/10.1207/s15473341lld0201_1
- Koulaguina, Elena & Legendre, Géraldine & Barrière, Isabelle & Nazzi, Thierry. 2019. Towards abstract syntax at 24 months: Evidence from subject-verb agreement with conjoined subjects. *Language Learning and Development* 15(2). 157–176. DOI: <https://doi.org/10.1080/15475441.2019.1571417>
- Kuhl, Patricia K. 2004. Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5. 831–843. DOI: <https://doi.org/10.1038/nrn1533>
- Legendre, Géraldine & Barrière, Isabelle & Goyet, Louise & Nazzi, Thierry. 2010. Comprehension of infrequent subject-verb agreement forms: Evidence from French-learning children. *Child Development* 81(6). 1859–1875. DOI: <https://doi.org/10.1111/j.1467-8624.2010.01515.x>
- Legendre, Géraldine & Culbertson, Jennifer & Zaroukian, Erin & Hsin, Lisa & Barrière, Isabelle & Nazzi, Thierry. 2014. Is children's comprehension of subject–verb agreement universally late? Comparative evidence from French, English, and Spanish. *Lingua* 144. 21–39. DOI: <https://doi.org/10.1016/j.lingua.2013.05.004>

- Lidz, Jeffrey & Waxman, Susan & Freedman, Jennifer. 2003. What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition* 89(3). 295–303. DOI: [https://doi.org/10.1016/S0010-0277\(03\)00116-1](https://doi.org/10.1016/S0010-0277(03)00116-1)
- Limbach, Maxi & Adone, Dani. 2010. Language acquisition of recursive possessives in English. In Franich, Katie & Iserman, Kate M. & Keil, Lauren L. (eds.), *Proceedings of the 34th annual Boston University Conference on Language Development* [BUCLD 34], 281–290. Somerville, MA: Cascadilla Press.
- Lorimor, Heidi & Stephens-Hecker, Nola & Miller, Carol. 2019. When it comes to complex NPs, not all preschoolers agree. *Language Learning and Development* 15(4). 295–316. DOI: <https://doi.org/10.1080/15475441.2019.1621178>
- Lukyanenko, Cynthia & Fisher, Cynthia. 2016. Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition* 146. 349–370. DOI: <https://doi.org/10.1016/j.cognition.2015.10.012>
- Marcus, Gary F. & Vijayan, Sugumaran & Bandi Rao, Shoba & Vishton, Peter M. 1999. Rule learning by seven-month-old infants. *Science* 283(5398). 77–80. DOI: <https://doi.org/10.1126/science.283.5398.77>
- Martin, Alexander & Holtz, Annie & Abels, Klaus & Adger, David & Culbertson, Jennifer. 2020. Experimental evidence for the influence of structure and meaning on linear order in the noun phrase. *Glossa: a journal of general linguistics* 5(1). 97. DOI: <https://doi.org/10.5334/gjgl.1085>
- McKercher, David A. 2001. The polysemy of with in first language acquisition. Doctoral dissertation, Stanford University.
- Morgenstern, Aliyah & Sekali, Martine. 2009. What can child language tell us about prepositions? A contrastive corpus-based study of cognitive and social-pragmatic factors. In Zlatev, Jordan & Johansson Falck, Marlene & Lundmark, Carita & Andréén, Mats (eds.), *Studies in Language and Cognition*, 261–275. Cambridge Scholars Publishing.
- Nadig, Aparna S. & Sedivy, Julie C. 2002. Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science* 13(4). 329–336. DOI: <https://doi.org/10.1111/j.0956-7976.2002.00460.x>
- Noble, Claire & Iqbal, Faria & Lieven, Elena & Theakston, Anna. 2016. Converging and competing cues in the acquisition of syntactic structures: The conjoined agent intransitive. *Journal of Child Language* 43(4). 811–842. DOI: <https://doi.org/10.1017/S0305000915000288>
- Pérez-Leroux, Ana T. 2005. Number problems in children. In Gurski, Claire (ed.), *Proceedings of the 2005 Canadian Linguistics Association Annual Conference*, 1–12.
- Pérez-Leroux, Ana T. & Castilla-Earls, Anny P. & Béjar, Susana & Massam, Diane. 2012. Elmo's sister's ball: The problem of acquiring nominal recursion. *Language Acquisition* 19(4). 301–311. DOI: <https://doi.org/10.1080/10489223.2012.685019>
- Pérez-Leroux, Ana T. & Peterson, Tyler & Castilla-Earls, Anny P. & Béjar, Susana & Massam, Diane & Roberge, Yves. 2018. The acquisition of recursive modification in NPs. *Language* 94(2). 332–359. DOI: <https://doi.org/10.1353/lan.2018.0019>

- Pérez-Leroux, Ana T. & Roberge, Yves. 2018. A way into recursion. In Hollebrandse, Bart & Kim, Jaieun & Pérez-Leroux, Ana. T. & Schulz, Petra (eds.), *UMOP 41: Thoughts on Mind and Grammar (T.O.M. and grammar): A Festschrift in Honor of Tom Roeper*, 109–122. Amherst, MA: GLSA.
- Pinto, Manuela & Zuckerman, Shalom. 2019. Coloring Book: A new method for testing language comprehension. *Behavior Research Methods* 51. 2609–2628. DOI: <https://doi.org/10.3758/s13428-018-1114-8>
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Ramos, Elaine B. 2000. Acquisition of noun phrase structure in children with specific language impairment. Doctoral dissertation, University of Massachusetts–Amherst.
- Reeder, Patricia A. & Newport, Elissa L. & Aslin, Richard N. 2013. From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive Psychology* 66(1). 30–54. DOI: <https://doi.org/10.1016/j.cogpsych.2012.09.001>
- Rice, Sally. 2009. Growth of a lexical network: Nine English prepositions in acquisition. In Cuyckens, Hubert & Dirven, René & Taylor, John R. (eds.), *Cognitive Approaches to Lexical Semantics*, 243–280. Berlin, New York: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110219074.243>
- Roeper, Thomas. 2011. The acquisition of recursion: How formalism articulates the child’s path. *Biolinguistics* 5(1–2). 57–86. DOI: <https://doi.org/10.5964/bioling.8831>
- Roeper, Thomas & Snyder, William. 2004. Recursion as an analytic device in acquisition. In van Kampen, Jacqueline & Baauw, Sergio (eds.), *Proceedings of GALA (Generative Approaches to Language Acquisition) 2003*, 401–408. Utrecht: LOT.
- Saffran, Jenny R. & Aslin, Richard N. & Newport, Elissa L. 1996. Statistical learning by 8-month-old infants. *Science* 274(5294). 1926–1928. DOI: <https://doi.org/10.1126/science.274.5294.1926>
- Selkirk, Elizabeth. 1977. Some remarks on noun phrase structure. In Akmajian, Adrian & Culicover, Peter & Wasow, Thomas (eds.), *Formal Syntax*, 285–316. New York: Academic Press.
- Shi, Rushen & Emond, Emeryse & Badri, Sophia. 2020. Hierarchical structure dependence in infants at the early stage of syntactic acquisition. In Brown, Megan M. & Kohut, Alexandra (eds.), *Proceedings of the 44th annual Boston University Conference on Language Development [BUCLD 44]*, 572–585. Somerville, MA: Cascadilla Press.
- Shi, Rushen & Legrand, Camille & Brandenberger, Anna. 2020. Toddlers track hierarchical structure dependence. *Language Acquisition* 27(4). 397–409. DOI: <https://doi.org/10.1080/10489223.2020.1776010>
- Soderstrom, Melanie. 2008. Early perception-late comprehension of grammar? The case of verbal -s: A response to de Villiers & Johnson (2007). *Journal of Child Language* 35(3). 671–676. DOI: <https://doi.org/10.1017/S0305000907008616>
- Soderstrom, Melanie & White, Katherine S. & Conwell, Erin & Morgan, James L. 2007. Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy* 12(1). 1–29. DOI: <https://doi.org/10.1111/j.1532-7078.2007.tb00231.x>
- Stickney, Helen. 2009. The emergence of DP in the partitive structure. Doctoral dissertation, University of Massachusetts–Amherst.

- Sundara, Megha & Demuth, Katherine & Kuhl, Patricia K. 2011. Sentence-position effects on children's perception and production of English third person singular -s. *Journal of Speech, Language, and Hearing Research* 54(1). 55–71. DOI: [https://doi.org/10.1044/1092-4388\(2010/10-0056\)](https://doi.org/10.1044/1092-4388(2010/10-0056))
- Syrett, Kristen & Lidz, Jeffrey. 2009. QR in child grammar: Evidence from antecedent-contained deletion. *Language Acquisition* 16(2). 67–81. DOI: <https://doi.org/10.1080/10489220902769226>
- Tieu, Lyn & Romoli, Jacopo & Poortman, Eva & Winter, Yoad & Crain, Stephen. 2018. Children's comprehension of plural predicate conjunction. *Journal of Child Language* 45(1). 242–259. DOI: <https://doi.org/10.1017/S0305000917000137>
- Tomasello, Michael. 1987. Learning to use prepositions: A case study. *Journal of Child Language* 14(1). 79–98. DOI: <https://doi.org/10.1017/S0305000900012745>
- Weber-Fox, Christine M. & Neville, Helen J. 1996. Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 8(3). 231–256. DOI: <https://doi.org/10.1162/jocn.1996.8.3.231>
- Wiig, Elizabeth H. & Secord, Wayne A. & Semel, Eleanor M. 2004. *Clinical Evaluation of Language Fundamentals-Preschool*, 2nd edition. San Antonio, TX: Pearson Inc.

