Structural complexity and the acquisition of recursive locative PPs

Tyler Peterson,[1] Ana Pérez-Leroux,[2] Anny Castilla-Earls,[3]
Susana Béjar,[2] & Diane Massam[2]

[1]University of Arizona, [2]University of Toronto, [3]SUNY/Fredonia

1.   Introduction

This paper examines the experimental results that reveal the referential behavior of children and how they navigate the complexity and quantity of structure in recursive NPs. Modification introduces complexity in the syntax and the semantics of noun phrases. Recursive modification, in particular, seems heavily constrained in the speech of children. We explore the source of children's difficulty in this domain and whether the complexity introduced by NP modification relates to the task of restriction and the processing of descriptive content, or whether the syntax and semantics of NP embedding specifically introduces complexity in acquisition. However, these results also show that this complexity must be considered separately from the quantity of structure in meeting the referential task. As such, in this paper we focus on the results that show that children will at times produce overly-elaborated sentences (i.e. sentences with more structure) in response to the referential task which avoid recursive embedding, but introduce complexities in other ways. We concentrate on one specific type, which is the production of referential descriptions containing what we call a *DP-internal anaphoric dependency* (DPIAD). These dependencies are found in sentences where a pronoun within a locative PP modifier co-refers with a c-commanding NP within the same DP, such as *the alligator$_i$ with a bird on it$_i$*. An examination of DPIADs reveals how children negotiate the different linguistic strategies in meeting a referential task, and contributes to our understanding of the notion of semantic complexity in general.

2.   The study

The goal of this study was to compare the production of two types of doubly-modified definite descriptions, which differ minimally with respect to the attachment of a second modifier. The children were presented with the task of identifying a unique referent in the given visual contexts given in (1) and (2) (based on Pérez-Leroux et al. 2012).

(1)   *Which plate got broken?*
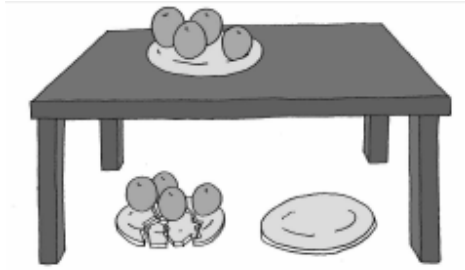
(2)   *Which bird got the worm?*
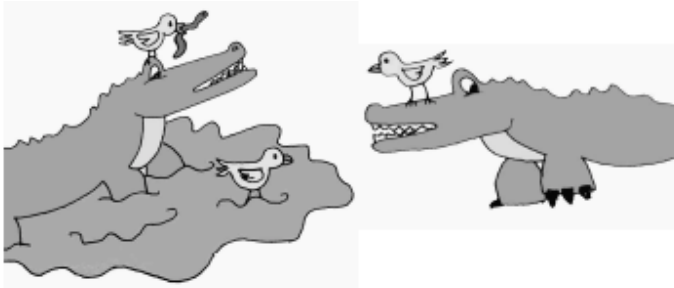
     The target responses constitute a minimal pair which allows us to separate the quantity of descriptive content from level of embedding. The correct response to the scenario in (1) involves *non-recursive modification*: the same (highest) referring NP is sequentially modified by two independent restrictive PPs, as shown in the bracketing in (3). The response to the scenario in (2) also involves *recursive modification*: the second modifier restricts the referent contained in the first PP modifier, as shown in (4):

(3)   *Non-recursive modification* (scenario in (1))
     The [plate$_{NP}$ [$_{PP}$ under the table] [$_{PP}$ with oranges]]

(4)   *Recursive modification* (scenario in (2))
     The [bird$_{NP}$ [$_{PP}$ on the alligator [$_{PP}$ in the water]]]

     Both the non-recursive and non-recursive NPs use the same relatively simple syntactic and semantic ingredients to build structure and meaning, while differing only in at what points merge, function application and predicate modification are applied in the derivation. The key area of activity is found in how the prepositions modify the head NP. Prepositions, generally considered to be a two-place relation between entities (cf. Barker 2011), can further compose with the nominal constituent. At this point, the nominal expression becomes part of a predicate, as evidenced by their use in copular constructions, as in (5a), or in restrictive NP modification, as in (5b):

(5)   a. The vase is [$_{PP}$ on [$_{DP}$ the table]]
      b. [$_{DP}$ The vase [$_{PP}$ on [$_{DP}$ the table]]]

The lowest nominal expression in (5b) is now part of a modifier. These differences result in two different referential targets, which are uniquely determined by the PP modifiers:

(6)   a.   *Referential target for the non-recursive DP* (scenario in (1)):
            The unique x such that x is a plate and x is under the unique y such that y is a table and x is with the unique z such that z are oranges
        b.   *Referential target for the recursive DP* (scenario in (2)):
            The unique x such that x is a bird and x is on the unique y such that y is an alligator and y is in the unique z such that z is the water

Within the acquisition literature there are two contrasting views regarding the referential behavior of children (Nadig & Sedivy 2002): on one hand five-year olds are fully sensitive to context in their production of modified NPs, using adjectives only when a competing referent was present in the context. On the other hand, there seems to be a productivity gap in children's production of complex NPs, as most of the nouns they produce spontaneously are simple DPs with no modification. As such, the main goal of this study was to determine what the source of children's difficulty in this domain is. More specifically, we rely on the contrast between recursive modification and double non-recursive modification to examine whether modification is the sole determinant of children's difficulty with complex DPs, or whether embedding plays a specific role. Our research questions were whether children are sensitive to the distinction between recursively and non-recursively modified DPs, and to what extent children differ from adults in this regard.

2.1   Study participants, materials and procedures

Fifty monolingual English-speaking children recruited from preschools in Fredonia, NY, and 13 adults from the same area participated in a referential elicitation task. Children were between the ages of 4;00 and 5;11, mean age 4;11, (SD = 5.8 months). The task, modeled after Pérez-Leroux et al. (2012), targeted the following types of complex NPs:

(7)   a.   Recursive genitives: *Batman's baby's lollypop*
        b.   Recursive comitative PPs: *The girl with a dog with a hat*
        c.   Recursive locative PPs: *The toothbrush in the cup on the shelf*
        d.   Recursive relational noun complement PPs: *The box of the cans of tomatoes*
        e.   Double non-recursive PP modification: *The hotdog with mustard on the blue table*

The scenarios were designed so two different modifier PPs were required to uniquely describe the target. Children were then presented a question of the form which x … ?, to prompt them to produce the target description.

In addition to the recursive elicitation task children were administered a battery of independent measures of cognitive and language development, including various subtests of the Non-Verbal Scale of the Kaufman Assessment Battery for Children Second Edition (KABC- II; Kaufman and Kaufman 2004); the Clinical Evaluation of Language Fundamentals Preschool 2 (CELF-P2; Wiig et al. 2004) and the Pearson Productive Vocabulary test

(PPVT-4); and a non-word repetition task (Dollaghan & Campbell 1998). They were also given a narrative elicitation task (Frog story) which was used to calculate the subordination index, number of functionally complete sentences (NU-TU), mean length of utterance in words (MLUW), and lexical diversity based on number of different word types (NDW).

The data was analyzed using two independent coding systems. Using the stimuli in (1) and (2), a description of their referential properties in terms of the success in identifying the target referent was coded as shown in (8):

(8)   *Semantic coding: Referential properties*

Code	Definition	Example
INCOMPLETE	Single NP or level 1 embedding	*the plate; the bird on the alligator*
ALTERNATIVE	Use of an appropriate but alternative description that does not include the target modifiers	*the one on the left; the smallest statue*
SEQUENTIAL	Sequences of responses that together include the 3 target predicates	*the bird on the alligator; the one in the water*
NON-EMBEDDED	Containing the head and the two target modifiers, semantically ok but non-embedded syntax	*the bird is on the alligator that is in the water; the worm inside the apple and the apple is on the plate and the worm is green*
TARGET	**Level 2 or 2 Level 1 N PP PP sequence: both target syntax and referential success**	***the plate with oranges under the table; the bird on the alligator in the water***
ANOMALOUS	Semantically anomalous or ungrammatical combinations of the target semantic predicates	*The bird in the water on the alligator*

A description of the syntax of the responses in terms of levels of embedding was coded as in (9):

(9)   Syntactic coding: Levels of Embedding

   a.   Single NP: [$_{NP}$ the bird]
   b.   Level 1: [$_{NP}$ the bird [$_{PP}$ on the alligator]]
   c.   2 Level 1: [$_{NP}$ the plate [$_{PP}$ with oranges] [$_{PP}$ under the table]]
   d.   Level 2: [$_{NP}$ the bird [$_{PP}$ on the alligator [$_{PP}$ in the water]]]
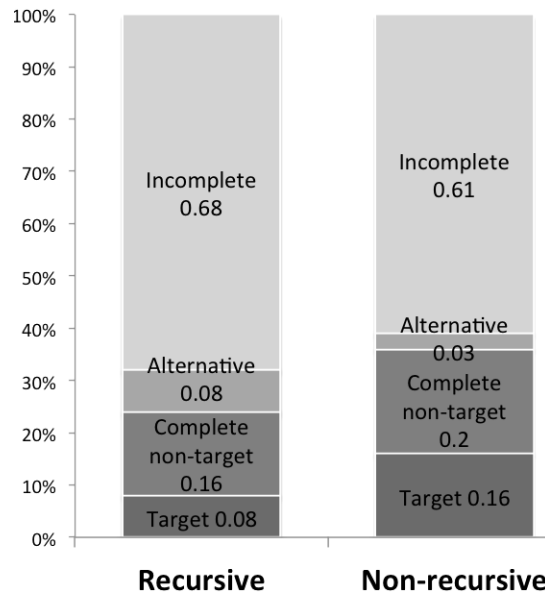
2.2   Results

The results in (10) and (11) show a comparison between locative PPs and double non-recursive modification for the production tasks in (1) and (2):

(10)   *Percentage of responses classified as to level of embedding*
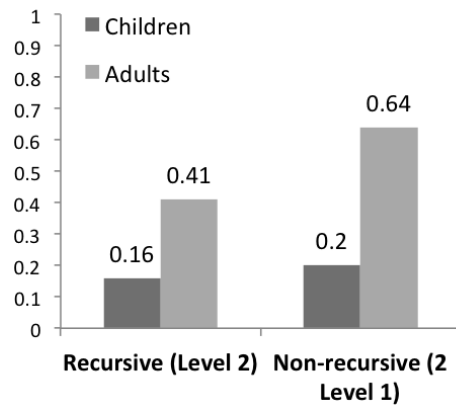
(11)   *Percentages of children's responses classified by referential coding*

Children resemble adults in that they find recursive modification more complex than non-recursive modification. More specifically, both groups often fail to notice the relevant two-way contrasts required by the visual contexts provided, and consequently fail to integrate all three referring expressions into their responses. Furthermore, both groups rely primarily on PP and relative clause embedding as the linking strategies used to connect the relevant referents (as was intended in the design of the elicitation task), and occasionally substitute semantically equivalent but overtly more elaborate strategies.

Children differ from adults in that they produce complex structure much less frequently, at a ratio as low as 4:1 for the non-recursive condition and 9:2 for the recursive condition. Additionally, children produce more descriptively complete but syntactically unintegrated responses than target responses (such as sequential, anomalous, and non-embedded), whereas adults, if they manage to produce the target referents, are more likely to produce a target response than not.

2.3 Over-elaborations

Complex NPs produced by children in elicited tasks show two forms of dissociation of syntax and semantics: complex reference achieved without syntactic complexity, or complex syntax that does not represent complex reference. Let's consider these in turn:

Utterances which achieve referential success coupled with flattening of the syntactic structure. Pérez-Leroux et al. (2012) note that children manage to produce descriptions of the target referent where the relevant NPs are not recursively embedded, but linked by coordination or by other clausal strategies that make the structure actually longer (in words) than the intended recursive target. A representative sample of these are given in (12):

(12)   a.   She looks like, um, the dog has the hat . . .(TRB 4;06)
            Target: The girl with a dog with a hat (from Pérez-Leroux et al 2012)
       b.   The worm inside the apple and the apple is on the plate and the worm is green. (DA, 5;10)
            Target: The worm inside the apple on the plate (From Pérez-Leroux et al. *under review*)
       c.   The big toothbrush is in the cup on the shelf. (MD, adult)
            Target: The toothbrush in the cup on the shelf (Pérez-Leroux et al. 2012, under review)

The example in (12a) relates the first nominal referent ('girl') to the rest by parataxis, and the second ('dog') to the third ('hat') by a clausal/predicative relation. The example in (12b) contains an NP modified by PP, and the subsequent modifier introduced by a coordinated clause. Clausal responses, while fully informative, are in fact pragmatically infelicitous, since the question under discussion is not 'what happened', but which bird did it. Only a DP response fits the question context optimally. A similar effect is observed in the task from Fig 2: while (13a) is a true of Fig 2, the response in (13b) with the DP is the optimal answer to the question.

(13)   Question: *Which bird got the worm?*

       a.   The bird is on the alligator in the water.   (True but non-target)
       b.   The bird on the alligator in the water (got the worm).   (True and target)

Utterances which do not meet the demands of the referential, but which in fact introduce additional layers of syntactic embedding. Children produce relatively more (overtly) over-elaborated responses than adults, who tend to produce the referentially complete but structurally simpler PP responses. We observed three types of expansions in (14):

(14)   Embedding without sufficient descriptive content

       a.   *RC (structurally separates the phases interpretable as target referents):*
            the bird **that's** on the crocodile **that's** in the water

b. *Expansion by self reference:*
the yellow plate with the apple **on it** (LM, 4;09)
c. *Expansion by inalienability:*
the birds that's **on top of the crocodiles** the crocodiles that's in water (AL4, 5;04)

Most of the examples involved the inanimate pronoun *it* with the two-place relational prepositions *on* and *in*:

(15)　a.　The one on the blue table with mustard on it　　　　　　　　(AL 5;04)
　　　b.　The one that has the smallest cup in it　　　　　　　　　　(SS, 4;05)
　　　c.　The blue bucket with the pencils in it.　　　　　　　　　　(LG, Adult)
　　　d.　the girl with the dog with the hat on it　　　　　　　　　　(AA, Adult)

However, there were instances of other pronouns, and other prepositions, which suggests it is not restricted to these:

(16)　a.　The one with the oranges in them　　　　　　　　　　　　(JD, 5;04)
　　　b.　um ... the one that has the flowers with her　　　　　　　　(AN, 5;11)
　　　c.　The pond with the ... with the flag boat in it.　　　　　　　(DA 5;10)
　　　d.　That one, the one with the flowers under it.　　　　　　　(OC 4;05)

While these structures are not highly frequent, they are consistent enough, representing 3% (9/288) of the total trials for the adults; and 5% of the total child trials (58/1200). Nor are they restricted to a few individuals: 28/50 children, and 5/13 adults produced one or more of these cases. Although children produce these structures at slightly higher rate than adults, age does not seem to be a factor in the production of these responses. The mean age of the children who produce them is the same as the overall sample of children in the study.
　　In this form of over-elaboration an additional PP modifier is added that further specifies the relation between the second modifier and the referent of the highest DP:

(17)　　[The box$_i$ [with pencils [in it$_i$]]]

In this construction, the pronoun is coreferent with the DP that contains it. As in its clausal or relative clause correlate, in (13) the same coreference holds between the the pronoun *it* and its antecedent *the box*:

(18)　a.　[The box]$_i$ has [pencils in it$_i$]
　　　b.　[The box]$_i$ that has [pencils in it$_i$]

To the best of our knowledge, these constructions are not described in the linguistic litera-ture. For the purpose of this paper, we refer to these cases as *DP-internal anaphoric depen-dencies* (DPIADs). These instances are problematic for standard representations since the pronoun seems bound to an antecedent that dominates it. As a point of entry in exploring

these cases, a plausible parallel may be drawn between these constructions and the syntax commonly assumed for restricted relative clauses:

(19) $[[_{DP} N]_i [_{CP} Op_i [... t_i]]]$

Under this line of reasoning we could treat DPIADs along the lines of a resumptive pronoun in relative clauses. However, replacing the trace with the pronoun is ungrammatical:

(20)   *This is $[[_{DP}$ the camel$]_i$ that he$_i$ likes Oscar]   (McKee & McDaniel 2001)

This is generally attributed to the restriction on an overt pronoun occurring in the same binding domain as the DP to which they corefer (Zaenen et al. 1981, Sharvit 2000). Furthermore, unlike DPIADs, resumptive pronouns do not involve anaphoric binding: they are in complementary distribution with a trace. Where they do appear, resumptive pronouns (at least in English) are conditioned in part by the 'linear distance' from their antecedent, either with the insertion of extra material that separates the pronoun and antecedent, as in (21a), or as a 'replacement' for a trace where coreference is is not possible due to island constraints, as in (21b) (McKee & McDaniel 2001):

(21)   a.   This is the camel$_i$ that **maybe, maybe, maybe, maybe** he$_i$ likes Oscar.
         b.   This is the girl$_i$ that **I don't know** what she$_i$ said.

As such, a resumptive pronoun analysis of DPIADs does not seem tenable. While it is not clear what the contribution of the extra PP is, it appears as an elaboration of the relationship between the head nominal and the embedded PP modifier that allows to further specify the spatial relationship expressed by the first preposition:

(22)   a.   The box with pencils (in it). (container)
         b.   The dog with the hat (on it). (cover)
         c.   The cake with jam on/in it. (container and cover)

Because prepositions like *in* and *on* are two-place relations, where these optional expansions emerge these two positions must be saturated, as in (23b). However, the internal argument of the preposition can only serve to restrict the reference of the higher DP, and not to create a relation between the subject of the preposition and another DP, as in (23c):

(23)   a.   The box$_i$ with pencils in it$_i$
         b.   The box with pencils (*in)
         c.   #The box$_i$ with pencils in it$_j$ (i.e. the room)

3.   Discussion: quantity vs. complexity

Our data introduces two novel observations about children's production of complex referential descriptions. The first observation is that children frequently fail to produce re-

sponses that meet the demand of the referential task. Because this difficulty goes beyond the demands of producing complex descriptions, we infer that children have a specific problem with integrating the various referential expressions into a recursive structure. The paratactic and clausal strategies they often employ flatten the structure and reduce the number of nominal expressions dominating other nominal expressions. Undeniably, children's reliance on such structural alternatives relates to their propensity, in speech production, to conduct sentence planning over shorter spans, when compared to adults (McDaniel et al. 2010). However, the recursively modified DP targets in our study are not particularly long, nor are they longer than their non-recursive counterparts. Thus, phrasal length, in the traditional sense, cannot properly account for our results.

The second observation is that children tend to introduce relative clauses in the target responses to the recursive condition. Again, this points in the same direction as the previous observation. Children produce longer, more overtly elaborated description, which have the consequence of decreasing the structural proximity between referential expressions. At this point, it would be useful to consider this result in light of Arsenijevic and Hinzen's (2012) discussion of how the forms of recursion arise from the way in which narrow syntax and interpretation interlock, rather than from narrow syntax itself. These authors note that, in natural language, instances of a category X directly dominating other instances of X are exceedingly rare. The norm is that referential expressions and propositions (DPs and CPs) dominate others of the same type only indirectly, mediated by a rigid sequence of functional categories. This brings us to a speculation about recursion, and the different interface behaviors of embedded and embedding categories. The parataxis and relative clause effects observable in our results suggest that the faculty of language (as revealed through the process of acquisition) is set to maximize the structural contour existing between a higher referential expression from lower expressions embedded under its domain.

One difference between children's and adults' responses to recursive trials stands out. In the recursive condition, children were more likely to provide the more elaborated RC or RC/PP mix responses than the structurally simpler PP responses. The ratio of pure PPs to responses containing relative clauses was of almost 1 to 7 for the recursive targets, compared to 8 to 3 for the non-recursive targets. This asymmetry in the distribution of linking types among children's responses to the two types of complex DPs is highly significant. In other words, despite the challenge involved in the production of recursively modified DPs, when children actually manage to compose these complex referential expressions, they produce more structure rather than less. This shows even more clearly that in the recursive conditions, children do not necessarily opt for the structurally and lexically simpler form. Instead of producing the simpler stacked NP/PP sequences – the target – they seem to favor forms that introduce additional structural distance between the relevant referential expressions, as schematized in the difference between (24a) and (24b):

(24) a. $[_{DP}$ N $[_{PP}$ P $[_{DP}$ N $[_{PP}$ P $[_{DP}$ N $]]]]$
      b. $[_{DP}$ N $[_{CP}$ C $\ldots[_{PP}$ P $[_{DP}$ N $[_{CP}$ C $\ldots[_{PP}$ P $[_{DP}$ N $]]]]]]]]$

These observations give us a sense of increased quantity without introducing complexity in the syntax and semantics: there are simply more operations of merge, function appli-

cation and predicate modification. However, expansion by self-reference – the DPIADs – introduces an additional semantic complexity not found in the other elaborations.

Arsenijevic & Hinzen (2012) differentiate embedded and embedding categories on the basis of the referential and predicative uses of a DP. This two-way split does not illuminate the contrast between recursive and non-recursive modification, as in either type, at the highest point of the derivation, all PP modifiers have become part of predicative content. While all the descriptive content in lower phases remains relevant at interface, the scopal possibilities have become fixed by the derivational history. As such, each step in embedding introduces a new layer of descriptive complexity; the reasons why are not evident within a framework in which earlier phases are rendered inert.

References

Arsenijevic, Boban, & Wolfram Hinzen. 2012. On the absence of X-within-X recursion in human grammar. *Linguistic Inquiry* 43:423–440.

Barker, Chris. 2011. Possessives and relational nouns. In *Semantics: An International Handbook of Natural Language Meaning, Part 2*, ed. Claudia Maienborn, Klaus von Heusinger, & Paul Portner, 1109–1130. Berlin: Walter de Gruyter.

McDaniel, Dana, Cecile McKee, & Merrill F. Garrett. 2010. Children's sentence planning: Syntactic correlates of fluency variations. *Journal Child Language* 37:59–94.

McKee, Cecile, & Dana McDaniel. 2001. Resumptive pronouns in English relative clauses. *Language acquisition* 9.2:113–156.

Nadig, Aparna, & Julie C. Sedivy. 2002. Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science* 13:329–336.

Pérez-Leroux, Ana Teresa, Anny Castilla-Earls, Susana Bejar, & Diane Massam. 2012. Elmo's sister's ball. the development of nominal recursion in children. *Language Acquisition* 19:301–311.

Pérez-Leroux, Ana Teresa, Anny Castilla-Earls, Susana Béjar, Diane Massam, & Tyler Peterson. under review . Strong continuity and children's development of NP recursion. In *Recursion in Brazilian languages and beyond*, ed. L. Amaral M. Maia T. Roeper, A. Nevins. Springer.

Sharvit, Yael. 2000. Resumptive pronouns in relative clauses. *Natural Language & Linguistic Theory* 17.3:587–612.

Zaenen, Annie, Elisabet Engdahl, & Joan Maling. 1981. Resumptive pronouns can be syntactically bound. *Linguistic Inquiry* 679–682.

Tyler Peterson, Ana Pérez-Leroux, Anny Castilla-Earls, Susana Béjar, Diane Massam
tylerpeterson@email.arizona.edu, at.perez.leroux@utoronto.ca,
anny.castilla@fredonia.edu, susana.bejar@utoronto.ca, diane.massam@utoronto.ca