Chapter 2

Rethinking complexity

Susana Bejar
University of Toronto

Diane Massam
University of Toronto

Ana-Teresa Pérez-Leroux
University of Toronto

Yves Roberge
University of Toronto

This paper addresses the nature of complexity of recursion. We consider four asymmetries involving caps on recursion observed in previous experimental acquisition studies, which argue that complexity cannot be characterized exclusively in terms of the number of iterations of Merge. While recursion is essentially syntactic and allowed for by the minimalist toolkit via Merge, selection, and labeling or projection, the complexity of recursive outputs arises at the interface.

1 Introduction

Watumull et al. (2014) (WHRH) discuss three criterial properties of recursion and argue that "by these necessary and sufficient criteria, the grammars of all natural languages are recursive" (p. 1). Phrases and sentences are defined recursively "in a stepwise strongly generative process creating increasing complexity" (p. 6). We focus here on this notion of complexity, since, from the perspective that recursive structures are the result of repeated applications of Merge opera-

Susana Bejar, Diane Massam, Ana-Teresa Pérez-Leroux & Yves Roberge. 2020. Rethinking complexity. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences I: Syntax inside the grammar*, 15–24. Berlin: Language Science Press. DOI: 10.5281/zenodo.3972830

tions, structures arising from similar derivational steps should all be derivationally equally complex. This squib sheds light on the nature of the complexity of recursion in human grammar through a theoretically-based exploration of four asymmetries observed in a series of experimental studies on the acquisition of self-embedding structures we have conducted in the last few years. Note that, while "self-embedding" often refers to complement structures, our use of the term generalizes over adjunction as well.

WHRH emphasize that recursion is an architectural property of the language faculty as opposed to a characterization of output structures, pointing to two correlates of this view: (i) recursion is an architectural universal, not an emergent property; (ii) the caps on recursion that are observable in output structures result from arbitrary external factors. Here, our work contrasts with WHRH in two points. First, we investigate recursion as a property of outputs while what matters for WHRH is the complexity of the recursive procedure itself.[1] Second, we examine caps on recursion in child language as a window into development of the language faculty. Nonetheless, we seek to explore the links between our studies and the positions articulated in WHRH. In particular, we examine the connection between children's capacity to produce self-embedding structures and the notion of complexity. We argue that while recursion is essentially syntactic and allowed for by the minimalist toolkit, via Merge, selection, and labeling or projection (cf. Hauser et al. 2002), the complexity of recursive outputs arises at the interface.[2]

The growth of grammatical competence gives rise to the ability to produce longer and more complex sentences. Although there is little consensus about what constitutes complexity (Culicover 2013; Roeper & Speas 2014; Trotzke & Bayer 2015; Newmeyer & Preston 2014; McWhorter 2011), most discussions agree that embedding increases complexity (Culicover & Jackendoff 2006; Givón 2009). However, in the narrow syntax, embedding by itself should not determine complexity, as it is given by recursive Merge. We argue that complexity, rather than being strictly correlated with recursive iterations of Merge, arises at the interface. Moreover, because recursive iterations of Merge can result in different varieties of recursively embedded output structures, some structural elaborations turn out to be more complex than others.

[1] This is not to say of course that the issue of the complexity of the recursive program is of no interest but the goal of our research is to identify the source of the difficulties that complex structures create for children (and adults).

[2] The view that recursion is in narrow syntax we share with WHRH and many others (e.g. Moro 2008; Nevins et al. 2009); however, it has also been proposed to be in the discourse (Evans & Levinson 2009; Koschmann 2010), or a consequence of phasal architecture and the interface (Arsenijević & Hinzen 2010).

At the outset, the language of young children does not include structurally elaborate expressions; various forms of structural elaboration emerge during the preschool years. Absence of a structure leads to the attribution of the property of complexity to that structure, but often without a clear notion of what complexity is. Here, we discuss four aspects of complexity in recursive structures that present challenges for a simple definition. We consider these issues in the context of recursive NP embedding, including conjunction, genitives, PP structures, and relative clauses. In previous work (Pérez-Leroux et al. 2012; Pérez-Leroux, Castilla-Earls, Bejar, Massam & Peterson 2018; Pérez-Leroux, Peterson, et al. 2018) we observed that recursive conjunction seems simpler than recursive PP modification, that sequential double modification is less complex than twice-embedded modification, and that the combination of relative clause and PP modification is somehow less complex than twice-embedded PP modification, at least in some of the languages studied. From this, we argue that complexity is not uniform, and that the complexity emerging from recursive embedding is a property of the interface, and not a property of narrow syntax.

We now turn to a discussion of four contrasts that shed light on the nature of complexity.

2 Coordination and modification

Children learn the basic ingredients required for NP elaboration quite early, including relevant functional elements (Brown 1973) and semantic relations (Bloom et al. 1975). Pérez-Leroux et al. (2012) investigated the points when children learn to iterate forms of NP elaboration. Using a referential task, we elicited twice-embedded genitives (1a) and modificational PPs (1b). Contexts were set up so twice-embedded modification was needed to disambiguate target referents from other competing referents. For instance, we need something like (1b) to uniquely describe the target in a scenario with two girls, each with a dog, where the only difference is a hat on one of the dogs. We controlled for whether children could produce utterances with three NPs, by testing coordination, as in (2), which matched the utterance length of the recursively embedded conditions.

(1) a. the boy's cat's tail
    b. the girl with a dog with a hat

(2) a boy, a bicycle, and a doll

Of key importance is the following result: children had no difficulties producing coordinate NPs, but had substantial difficulties with NP embedding. Two-

thirds of the younger children produced no NP embedding at all. This does not follow from current assumptions about coordination structures. Recently, the goal has been to integrate coordination into X-bar theory (contra, e.g. Jackendoff 1977), whether by adjunction (Munn 1993) or complementation (Johannessen 1998). Under this approach, coordinates are structurally equivalent to either of the twice-embedded structures in (1). This precludes a purely structural explanation of the relative difficulty of the PP and genitive recursive structures.

The NP embedding/coordination contrast is thus placed squarely in the domains of processing and/or semantics, i.e. interpretive complexity at the interface. Coordinating three NPs just augments a set. Embedding, via either adjunction or complementation, reformulates the description of a set. The descriptive content of lower referents serves to restrict the domain of the higher nominal.

3 Sequential and recursive PP modification

A subsequent study explored the next logical question (Pérez-Leroux, Peterson, et al. 2018). Does each step in embedding increase the complexity of the nominal structure? We set up a minimal comparison between two types of doubly modified structures involving locatives, relying on a similar referential task to the one previously employed, but contrasting two types of contexts. One condition required two PPs modifying the same head noun as in (3a), whereas in the other (3b), the head noun is modified by a PP, itself modified by a lower PP.

(3) a. the plate [under the table] [with oranges]
b. the bird [on the alligator [in the water]]

A detailed comparison of these two constructions reveals that, syntactically and semantically, they are equally complex, at least in principle. Their generation involves not only the same core operations (e.g. Merge, predicate modification), but also the same number of core operations. Given the formal parallels of the two constructions, we would expect comparable patterns of production. However, a strong asymmetry arises. Both children and adults produced twice-embedded PP modification at half the rates of double sequential modification. Since everything else is held constant, productivity can be interpreted as a reflection of less complexity. Given the comparability between the task and the structure, this suggests that depth of embedding results in more complex configurations. What might account for this difference? Again, we must look to the interface to explain this. Under the logic of phase theory, a phasally complete functional domain like DP should cease to function as a complex object (phase

impenetrability condition, PIC). While (3a) and (3b) are equivalent with respect to the number of phasal domains (assuming one views DP as a phase), in (3b), but not (3a), the referent of the head noun is restricted by an expression that is inaccessible under the PIC. In fact, the descriptive content of the lower phase *in the water* in (3b) was essential for success in the experimental task: other alligators lurked on land. We submit that this is the source of the added complexity of these structures, but note that this is not complexity in the narrow syntax – the narrow syntax freely generates such structures – the challenge rests in interpretive requirements at the interface.

4 PP/relative clause modification and recursive PP modification

A third observation in support of our view of complexity also originates from Pérez-Leroux, Peterson, et al. (2018). In lieu of the target PP modifiers (4a), speakers commonly substituted relative clauses (4b) and a mix of PP and relative clause constructions (4c).

(4) a. The one on the plate with the apple.
b. The bird that's on the crocodile that's in the water.
c. The one on the one on the crocodile's eyes that was in the water.

That adults were prone to use the more elaborate relative clauses (RCs) where simple PPs would do the work was a surprise. That children did so too was more so, given the extensive literature on children's difficulties with relative clauses (see references in Friedmann et al. 2009; Givón 2009). Interestingly, these expansions were particularly frequent when the target was a twice-embedded PP structure. There, the relative and mixed PP/relative strategies represented over 40% of adults' and children's target responses. This was true in English as well as in recent data from German preschoolers, obtained with the same methods (Lowles 2016). These responses are perfectly natural, and certainly successful in the context of our task. From a complexity perspective, they are perplexing – especially in the case of children – inasmuch as they constitute longer and structurally more elaborate constructions that, importantly, do not informationally add anything when compared to PP responses. The additional syntactic and semantic complexity introduced by RCs is not limited to the additional lexical material but is also due to the fact that they involve displacement and dependencies in syntax as well as additional semantic operations. Yet their use strongly suggests that the modification relation is not problematic. This leaves us with a mystery: Why should

children and adults frequently use the structurally more elaborate relative clause strategy to express modification?

If complexity is not computed in narrow syntax as the result of a number of recursive applications of Merge, then this result can be interpreted from a different angle. Several possibilities arise which differ with respect to how "detached" from the computational component the complexity issue really is. For instance, as early as 1963, Chomsky & Miller argued that the complexity of recursive self-embedding results from performance processes, not formal grammar. In contrast, Arsenijević & Hinzen (2012) note that instances of X directly dominating another instance of X are rare: the common strategy is for referential expressions to dominate others of the same type indirectly, via sequences of functional categories. For them this is a direct result of the phasal architecture of the computational component. Everything seems to function as if to create a structural contour between referential expressions in a phrase.

On a final note, our conclusion that complexity of recursive embedding does not reside in narrow syntax is supported by comparable data recently collected from French and Japanese (Bamba et al. 2016; Roberge et al. 2018). In these languages, children do not readily rely on the relative clause strategy; they incorporate it gradually, as one would expect. One possible explanation route is to link this cross-linguistic difference to uniformity in the directionality of embedding: French and Japanese are uniformly right- and left-embedding, whereas German and English mix branching directionality in their nominal syntax. If this is confirmed by further studies on additional languages, we would conclude that recursive PP embedding is not computationally more complex than any other applications of Merge and avoidance of twice-embedded PPs in our experiments must be accounted for by recourse to other considerations.

5 Genitives and PPs

The cases discussed so far implicitly follow a quantity metric, comparing the target structures in the two types of double-modification contexts with respect to the number of noun phrases, embedding steps, layers of functional structure, and steps required for semantic derivation. Let us now turn to qualitative differences. Do different types of NP embedding yield differences in complexity for reasons unrelated to structural metrics? Here we focus on possessive embedding (1a), which differs from comitative PPs (1b) in terms of directionality and case marking. Again, on minimalist assumptions about recursion, the answer should be no. However, accounts of acquisition difficulties often rely on notions of uniformity, and the basic typology of the target language. It is conceivable that in

English, a fundamentally right-branching and analytic language, the genitive *'s* construction might be constrained in acquisition. It is, after all, constrained in related languages. Roeper & Snyder's (2004) observation that the cognate possessive form in German does not iterate (i.e., German allows *NP's NP* but not *NP's NP's NP*) was the starting point in the study of the acquisition of recursive self-embedding structures. Such language differences prove that rule acquisition (i.e., possessive *-s*, in this case) is a learning step distinct from the acquisition of rule iteration (allowing multiple instances of the embedding process). The data in Pérez-Leroux et al. (2012) suggested a delay. First-level embedding appeared simultaneously for genitives and PP modifiers. Second-level of embedding was a distinct stage, attained first for PP modifiers. Since few children attained the second stage in the development of complex NPs, this was clearly worth further investigation. We recently elicited data on the production of recursive possessives and PPs in a group of seventy-one English-speaking children in Toronto (Pérez-Leroux et al. in preparation). While overall rates of production success were slightly higher for recursive comitative PPs, children did not acquire them earlier than genitives. In fact, the converse was true. Individually, more children could produce recursive sequences of possessive *-s* than of comitatives (*NP with NP with NP*) at a ratio of 5 to 1 compared to the converse pattern. This is due to the PP/RC trade-off described in §4. Possessives were rarely substituted by other forms, so a child could more easily embed possessives twice. We can safely conclude that the structurally distinct properties of the possessive construction do not constrain children's ability to iterate genitive embedding.

6 Conclusion

The notion of complexity – often loosely defined and used intuitively – is illuminated by the consideration of caps on recursion as observed in acquisition studies. Four cases were discussed, all pointing to the conclusion that complexity cannot be characterized exclusively in terms of the number of iterations of Merge. In closing, we return to WHRH and the view of recursion articulated therein. WHRH take complexity to correlate with iterations of the recursively defined generative structure-building procedure, with caps on recursion/complexity reducing to (arbitrary) extra-linguistic considerations. Couched in the traditional dichotomy, their focus is on competence. We argued that this view of complexity does not shed light on the nature of caps on recursion observed in the language acquisition studies reported here. However, we believe our results are consistent with the overall view of recursion articulated in WHRH. The absence of a correlation between complexity and recursive iterations of Merge is exactly what one

might expect if the recursive nature of grammar is an architectural universal and hence unlearnable (as WHRH say). Likewise, WHRH's view that caps on recursion/complexity must be understood in terms of conditions external to narrow syntax resonates with our findings, though it is not at all clear to us how external (or arbitrary) these really are. Our studies point to the need for future work to determine and articulate the nature of complexity at the interface.

Abbreviations

PIC	phase impenetrability condition	RC	relative clause

Acknowledgements

Authors are listed in alphabetical order. We are grateful to Anny P. Castilla-Earls, Erin Hall, Gabrielle Klassen, Erin Pettibone, Tom Roeper, Petra Schulz, Ian Roberts and two anonymous reviewers for helpful comments and discussion. We also gratefully acknowledge funding from the Social Sciences and Humanities Research Council of Canada (IG 435-2014-2000 "Development of NP complexity in children" to Ana T. Pérez-Leroux & Yves Roberge).

References

Arsenijević, Boban & Wolfram Hinzen. 2010. Recursion as a human universal and as a primitive. *Biolinguistics* 4(2–3). 165–173.

Arsenijević, Boban & Wolfram Hinzen. 2012. On the absence of X-within-X recursion in human grammar. *Linguistic Inquiry* 43(3). 423–440. DOI: 10.1162/LING_a_00095.

Bamba, Kaz, Midori Hayashi, Manami Hirayama & Ana T. Pérez-Leroux. 2016. On the development of complexity in Japanese nominal recursion. Paper presented at the Complexity in Acquisition Workshop, University of Toronto, 25 May 2016.

Bloom, Lois M., Patsy Lightbown & Lois Hood. 1975. *Structure and variation in child language* (Monographs of the Society for research in child development 160). DOI: 10.2307/1165986.

Brown, Roger. 1973. *A first language*. Cambridge: Harvard University Press.

Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In Duncan R. Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, vol. 2, 269–321. New York: Wiley.

Culicover, Peter W. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.

Culicover, Peter W. & Ray Jackendoff. 2006. The simpler syntax hypothesis. *Trends in Cognitive Sciences* 10. 413–418. DOI: 10.1016/j.tics.2006.07.007.

Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals. *Behavioral and Brain Sciences* 32(5). 429–492. DOI: 10 . 1017 / S0140525X0999094X.

Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1). 67–88. DOI: 10.1016/j.lingua.2008.09.002.

Givón, Thomas. 2009. *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam: John Benjamins. DOI: 10.1075/z.146.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579. DOI: 10.1126/science.298.5598.1569.

Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.

Koschmann, Timothy. 2010. The universality of recursion. *Lingua* 120(12). 2691–2694. DOI: 10.1016/j.lingua.2010.03.019.

Lowles, Alex. 2016. *The bird on the crocodile in the water: On the acquisition of double nominal modification in German*. Frankfurt am Main: Goethe University. (BA thesis).

McWhorter, John H. 2011. *Linguistic simplicity and complexity: Why do languages undress?* Berlin: De Gruyter.

Moro, Andrea. 2008. *The boundaries of Babel*. Cambridge: MIT Press.

Munn, Alan Boag. 1993. *Topics in the syntax and semantics of coordinate structures*. University of Maryland. (Doctoral dissertation).

Nevins, Andrew, David Pesetsky & Cilene Rodrigues. 2009. Pirahã exceptionality: A reassessment. *Language* 85(2). 355–404. DOI: 10.1353/lan.0.0107.

Newmeyer, Frederick J. & Laurel B. Preston (eds.). 2014. *Measuring grammatical complexity*. Oxford: Oxford University Press.

Pérez-Leroux, Ana T., Susana Bejar, Diane Massam & Yves Roberge. In preparation. Varieties of recursion. Ms., University of Toronto.

Pérez-Leroux, Ana T., Anny Castilla-Earls, Susana Bejar & Diane Massam. 2012. Elmo's sister's ball: The problem of acquiring nominal recursion. *Language Acquisition* 19(4). 301–311. DOI: 10.1080/10489223.2012.685019.

Pérez-Leroux, Ana T., Anny Castilla-Earls, Susana Bejar, Diane Massam & Tyler Peterson. 2018. Strong continuity and children's development of DP recursion. In Luiz Amaral, Marcus Maia, Andrew Nevins & Thomas Roeper (eds.), *Recursion across domains*, 296–313. Cambridge: Cambridge University Press. DOI: 10.1017/9781108290708.018.

Pérez-Leroux, Ana T., Tyler Peterson, Anny Castilla-Earls, Susana Bejar, Diane Massam & Yves Roberge. 2018. The acquisition of recursive modification in NPs. *Language* 94(2). 332–359. DOI: 10.1353/lan.2018.0019.

Roberge, Yves, Anna Frolova & Ana T. Pérez-Leroux. 2018. On recursive modification in child L1 French. *Languages* 3(1), 6. DOI: 10.3390/languages3010006.

Roeper, Thomas & William Snyder. 2004. Recursion as an analytic device in acquisition. In Jacqueline Van Kampen & Sergio Baauw (eds.), *Proceedings of GALA 2003*, 401–408. Utrecht: LOT.

Roeper, Thomas & Margaret Speas (eds.). 2014. *Recursion: Complexity in cognition*. Dordrecht: Springer.

Trotzke, Andreas & Josef Bayer. 2015. *Syntactic complexity across interfaces*. Berlin: De Gruyter.

Watumull, Jeffrey, Marc D. Hauser, Ian Roberts & Norbert Hornstein. 2014. On recursion. *Frontiers in Psychology* 4, 1017. DOI: 10.3389/fpsyg.2013.01017.