A parsimonious method for generating syntactic structure[1]

D. P. Medeiros, University of Arizona. Dec. 2020

Abstract

This paper reformulates (External) Merge as freely generating bare *n*-ary trees, labeled with a universal hierarchy by postorder traversal, and linearized by preorder traversal. Important word order universals follow: in several domains, all attested neutral orders are base-generated, while unattested orders match a systematic gap in generative capacity. The framework unifies Universal 20 (Greenberg 1963, Cinque 2005) and the Final Over Final Condition (Holmberg 2000, Sheehan *et al* 2017) as consequences. We also find simple analyses of cross-serial dependency constructions, including English Affix-Hopping (Chomsky 1957), and Dutch cross-serial subject-verb dependencies (Bresnan *et al* 1982). Other applications include a version of Travis' (1984) Head Movement Constraint allowing attested long head movement as in Breton.

Keywords: Merge, Final-Over-Final Condition, Universal 20, Cross-serial dependencies, Head Movement

1 Introduction

Chomsky describes the discrete infinite character of human syntax in terms of an abstract operation Merge. Merge takes as input lexical elements or syntactic objects already built, and outputs a structured expression containing its inputs, in a format determining semantic and phonological configurations. There are various ways of working out the details, but something like Merge seems indispensable in a generative model of syntax.

Attention has focused on implementing Merge as set formation, which provides for a rich theory of syntactic structure. That implementation, whatever its successes and *a priori* appeal,[2] is not the only possibility. If other reasonable implementations of Merge

[1] I would like to thank the following individuals for providing feedback at various stages: David Adger, Tom Bever, Noam Chomsky, Guglielmo Cinque, and Ian Roberts, to whom I am particularly indebted for patient help with topics in section 6, especially concerning head movement and the typology of clausal negation.

[2] There are several reasons to prefer a set-based implementation for Merge. One is the same reason that set theory is chosen as an axiomatic basis for mathematics: it is maximally conceptually sparse. Another reason is that sets are unordered, and semantic composition can be described in terms that eschew linear ordering. But see fn. 9.

make different predictions about syntactic phenomenology, the alternatives should be evaluated by their empirical successes in addition to their conceptual properties.[3]

1.1 The Duality of Semantics

In recent years, Chomsky has highlighted the need for syntactic theories to provide a basis for the duality of semantics: the existence, in natural language expressions, of two layers of meaning. One layer of meaning is the information-neutral thematic structure, including predicate-argument structure and selectional relations. Another layer of meaning concerns operator-variable structure, topic and focus, and the like. This cut should be tied to some syntactic distinction, such as a distinction in how Merge applies. If Merge joins disjoint syntactic objects, it is External Merge (EM). Where Merge applies to an object and one of its subparts, we have Internal Merge (IM).

> "The two types of Merge correlate well with the duality of semantics that has been studied from various points of view over the years. EM yields generalized argument structure, and IM all other semantic properties: discourse-related and scopal properties. The correlation is close, and might turn out to be perfect if enough were understood." (Chomsky 2007: 10)

The assumption of a universal ordering of EM is an essential component of the cartographic program (Rizzi 1997, Cinque 1999), there realized in terms of hierarchies dictating how lexical items are Merged into a bottom-up derivation. IM operations interleave with EM, (ultimately) yielding displacement. If EM applies in a common order, and syntactic structures are linearized the same way across languages (Kayne 1994), it follows that IM must be involved in deriving word order variation.

But languages plainly vary in word order even in information-neutral contexts. Information-neutral contexts, by definition, do not involve discourse or scopal properties. So what drives displacement in the derivation of neutral orders? Moreover, how can we explain the constraints on possible and impossible neutral word orders?

1.2 A preview of the framework

---

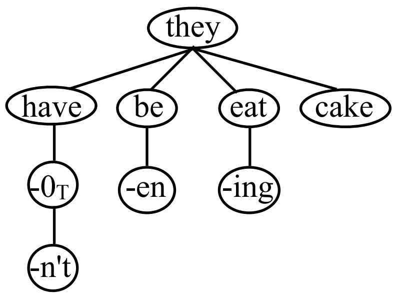
[3] Consider the choice between the real numbers and complex numbers for modeling physical phenomena. The reals seem conceptually inevitable, and are a strict subset of the complex numbers. However, complex numbers provide a better basis for understanding phenomena like electromagnetism, and with their greater complexity comes mathematical beauty (*e.g.*, in the context of the Fundamental Theorem of Algebra).

The account developed below uses a familiar tree-like notation, but in an unfamiliar way. As a first encounter with this system, I present an English sentence, with its segmentation into major morphemes, and its tree diagram. Below, S and O are pretheoretical terms for thematic subject and object; T is Tense, Neg is negation, Aux$_i$ and -Fx$_i$ are an auxiliary and the associated displaced affix.

(1) They haven't been eating cake.

(2) They have -0 -n't be -en eat -ing cake
S Aux$_1$ T Neg Aux$_2$ -Fx$_1$ V -Fx$_2$ O

(3)

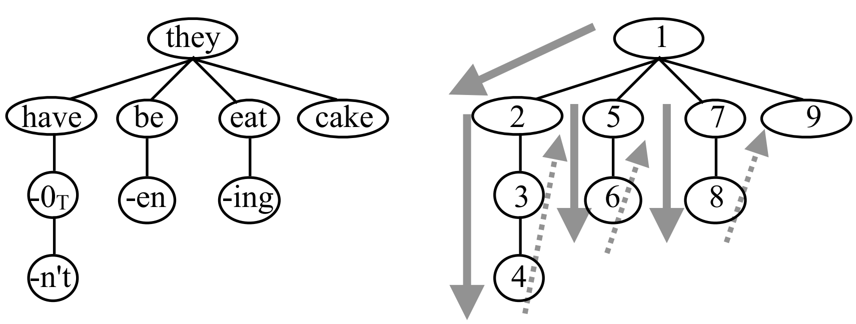


This is, admittedly, a very strange tree. It's certainly not a constituency tree. Nor is it a dependency tree, though it does share with dependency trees the property of having words or morphemes on non-terminal nodes. However, this particular tree has an interesting pair of properties:

(4) *Dual traversal property*
a. Surface word order is found by reading the tree left-to-right, and top-down (this is called *preorder traversal*).
b. An invariant representation of hierarchy is found by reading left-to-right and bottom-up (*postorder traversal*).

To illustrate the first property, I repeat tree (3) next to a copy of the same tree with individual nodes numbered in preorder. Grey arrows show the direction of the path.

(5)

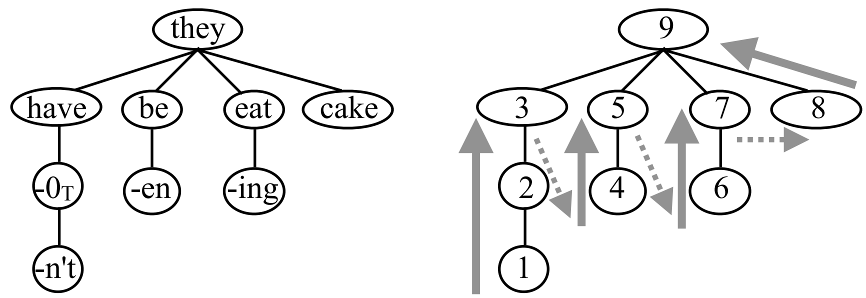


Assembling the words and morphemes from the nodes in the order visited by this path, we get string (6). As claimed, this is identical to the surface word order (2).

(6) They have -0 -n't be -en eat -ing cake

And as for the second property, (7) shows the postorder traversal path.

(7)

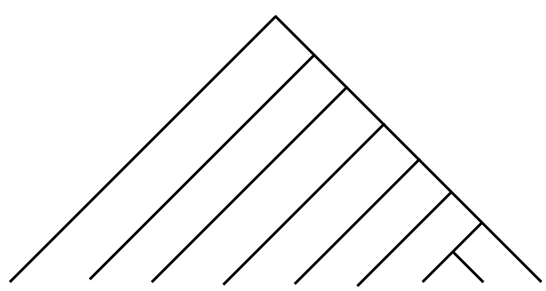


Following this traversal path, we find the following sequence:

(8) -n't -0 have -en be -ing eat cake they
Neg T Aux$_1$ -Fx$_1$ Aux$_2$ -Fx$_2$ V O S

It is perhaps not obvious, but (8) corresponds to a fairly standard clausal architecture. I show an unlabeled tree superposed on (8) to show semantic composition.

(9)

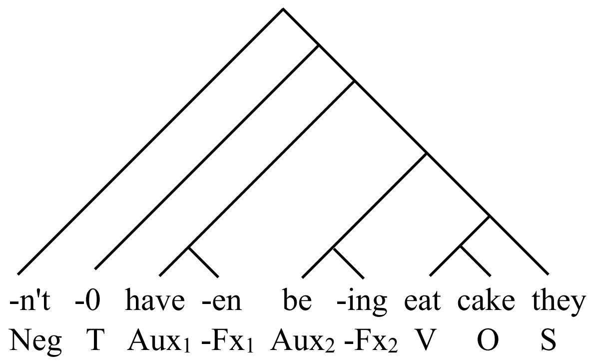


-n't -0 have -en be -ing eat cake they
Neg T Aux$_1$ -Fx$_1$ Aux$_2$ -Fx$_2$ V O S

With the exception of the rightward position of the S argument (and perhaps the position of negation), this is a familiar tree structure for the clause. While the subject is in what we might think of as a "rightward specifier", it is in the appropriate *v*P-internal layer (I omit *v* from the sequence because it is not overt). This representation corresponds closely to the universal structure produced by pure External Merge in standard frameworks. In the present theory, this postorder traversal sequence will be identical for all cross-linguistically allowed variations of word order for this meaning.

Before moving on, we might construct the interpretation differently for this example. The string format demonstrated in (8) readily supports treating auxiliaries and their associated displaced affixes as two pieces of a single lexical item (Chomsky 1957), in which case composition would proceed as indicated below (see section 5).

(10)



In the standard model, interpretation recapitulates the steps of the syntactic derivation, following it step by step. Here, by contrast, interpretation doesn't apply to the tree (3) directly, but rather to a single, invariant sequence (8) implicitly encoded in the tree. As we will see, arguably the same postorder traversal sequence is found in each different tree corresponding to a possible word order for this meaning. In this way, we can have different trees for each word order, but with an invariant meaning structure.

One might well ask, why these traversal methods? They are two of three standard tree traversal algorithms (the third, *inorder traversal*, plays no role here). Within the tree formalism described here, that may seem an arbitrary choice. However, the traversals used here are particularly natural and salient within an equivalent formulation over bracketed string representations, shown for this example in (11).

(11) [They [have [-0 [-n't -n't] -0] have] [be [-en -en] be] [eat [-ing -ing] eat] [cake cake] they]

Linking words with left brackets (12), we get the surface word order (6).

(12) [They [have [-0 [-n't -n't] -0] have] [be [-en -en] be] [eat [-ing -ing] eat] [cake. cake] they]
[They [have [-0 [-n't [be [-en [eat [-ing [cake
they have -0 -n't be -en eat -ing cake

If we associate words with right brackets (13), we get the hierarchical order (8).

(13) [They [have [-0 [-n't -n't] -0] have] [be [-en -en] be] [eat [-ing -ing] eat] [cake. cake] they]

-n't] -0] have] -en] be] -ing] eat] cake] they]
-n't -0 have -en be ing eat cake they

Put another way, in this system, given a bracketed representation like (11), we pronounce left brackets, and interpret right brackets.

1.3. Structure of this Paper

The remainder of this paper develops the representational format just introduced, and a corresponding theory of structure-building. The model developed here is deliberately only half a theory: I restrict attention to neutral word order variation. Obviously, this leaves aside some of the most interesting topics in syntax. Furthermore, the model developed is one of universal grammar. If an information-neutral configuration is attested in some language, or unattested in all languages, that fact falls within this theory's explanatory target. But the theory says nothing about the unavailability of constructions in some languages when those same constructions are allowed in others. That level of language-specific grammar is not the goal here, and must be explained in other ways.

In section 2, I discuss the noun phrase ordering restrictions known as Universal 20. I sketch a prominent account of the typological pattern due to Cinque (2005), and present an alternative method for generating the same structures and orders, involving an indexing scheme applied to free bracketing. Section 3 shows how the architecture is realized in terms of tree structures of the sort just previewed.

In section 4, I generalize the scheme for encoding hierarchy as a sequence, and show that the same machinery explains the Final-Over-Final Condition (Holmberg 2000, Sheehan *et al* 2017, *i.a*.). Section 5 shows how the account generates straightforward analyses of attested configurations containing cross-serial dependencies, including English Affix-Hopping (Chomsky 1957) and cross-serial subject-verb dependencies in languages like Dutch (Bresnan *et al* 1982) and Swiss German (Shieber 1985).

Section 6 provides a brief overview of some further applications of this architecture. I reanalyze some West Germanic data that had seemed to preclude the possibility of a unified hierarchy covering adverbial ordering as well as the relative ordering of verbal elements and arguments (Abels 2016), arguing that a single hierarchy may be viable after all. I show that the account provides the beginnings of an account of clitic-climbing constructions in Romance languages. And I derive a version of the Head Movement Constraint (Travis 1984) that captures familiar facts and allows known exceptions like "long" head movement in Breton. Along the way, I illustrate how the framework can provide simple analyses of several languages presenting a wide array of word order patterns. The last section concludes.

2 Generating Universal 20

As an example, consider possible and impossible neutral orders in the noun phrase, as described in Greenberg's Universal 20.

> "When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite." (Greenberg 1963: 87)

According to Cinque (2005), 14 of 24 possible orders of these four elements are attested. Cinque shows that this pattern can be succinctly described within the EM and IM framework. He assumes a universal underlying base, built by a uniform sequence of EM operations, affected by phrasal movement but not head movement or remnant movement (*i.e*. IM in the noun phrase must affect the noun, possibly pied-piping dominating structure).[4] This hierarchy is given in (14).

(14) [DemP ... [NumP ... [AdjP ... [N]]]][5]

In (15), I reproduce Cinque's (2005: 319-320, ex.6) list of attested (bold) and unattested orders of these elements. For attested orders, I include a simplified bracketed expression.

(15) O*rders of demonstrative, numeral, adjective, noun, after Cinque (2005)*

a. **Dem Num Adj N**	[Dem [Num [Adj [N]]]]
b. **Dem Num N Adj**	[Dem [Num [N] [Adj]]]
c. **Dem N Num Adj**	[Dem [N] [Num [Adj]]]
d. **N Dem Num Adj**	[N] [Dem [Num [Adj]]]
e. *Num Dem Adj N	
f. *Num Dem N Adj	
g. *Num N Dem Adj	
h. *N Num Dem Adj	
i. *Adj Dem Num N	
j. *Adj Dem N Num	
k. **Adj N Dem Num**	[Adj [N] [Dem [Num]]]
l. **N Adj Dem Num**	[N] [Adj] [Dem [Num]]

---

4 Cinque adopts Kayne's (1994) Linear Correspondence Axiom (LCA), which requires extra structure to provide landing sites for movement. Abels & Neeleman (2012) argue that the LCA is unneeded; the relevant constraint is simply that movement is leftward.

5 In Cinque's analysis, the nominal modifiers are all phrasal specifiers rather than heads. This is significant in light of the treatment of head-complement relations below.

m. *Dem Adj Num N	
n. **Dem Adj N Num**	[Dem [Adj [N]] [Num]]
o. **Dem N Adj Num**	[Dem [N] [Adj] [Num]]
p. **N Dem Adj Num**	[N] [Dem [Adj] [Num]]
q. *Num Adj Dem N	
r. **Num Adj N Dem**	[Num [Adj [N]]] [Dem]
s. **Num N Adj Dem**	[Num [N] [Adj]] [Dem]
t. **N Num Adj Dem**	[N] [Num [Adj]] [Dem]
u. *Adj Num Dem N	
v. *Adj Num N Dem	
w. **Adj N Num Dem**	[Adj [N]] [Num [Dem]]
x. **N Adj Num Dem**	[N] [Adj] [Num] [Dem]

The bracketed expressions in (15) can be derived from the bracketed expressions for Cinque's derivations provided by Steddy & Samek-Lodovici (2011) as follows. Find the left bracket immediately preceding each overt element (Dem, Num, Adj, N), and keep only those brackets and their matching right brackets. For these purposes, treat NP as [...N...], keeping [N]. The bracketed structures they provide for attested orders (*ibid*., 449) are in (16). Each modifier of the noun is a phrasal specifier of a dedicated functional head with a further agreement phrase above it to host movement.

(16) *Bracketed representations from Steddy & Samek-Lodovici (2011*)

a. [$_{AgrWP}$ [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ [$_{YP}$ AP $_{Y}$ NP]]]]]]

b. [$_{AgrWP}$ [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ NP [$_{YP}$ AP $_{Y}$ t$_{NP}$]]]]]]

c. [$_{AgrWP}$ [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ NP [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ [$_{YP}$ AP $_{Y}$ t$_{NP}$]]]]]]

d. [$_{AgrWP}$ NP [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ [$_{YP}$ AP $_{Y}$ t$_{NP}$]]]]]]

k. [$_{AgrWP}$ [$_{YP}$ AP $_{Y}$ NP] [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ t$_{YP}$]]]]]

l. [$_{AgrWP}$ [$_{AgrYP}$ NP [$_{YP}$ AP $_{Y}$ t$_{NP}$] [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{XP}$ NumP $_{X}$ t$_{AgrYP}$]]]]]

n. [$_{AgrWP}$ [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{YP}$ AP $_{Y}$ NP] [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ t$_{YP}$]]]]]

o. [$_{AgrWP}$ [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{AgrYP}$ NP [$_{YP}$ AP $_{Y}$ t$_{NP}$]] [$_{XP}$ NumP $_{X}$ t$_{AgrYP}$]]]]

p. [$_{AgrWP}$ NP [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ [$_{YP}$ AP $_{Y}$ t$_{NP}$] [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ t$_{YP}$]]]]]

r. [$_{AgrWP}$ [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ [$_{YP}$ AP $_{Y}$ NP]]] [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ t$_{XP}$]]]

s. [$_{AgrWP}$ [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ NP [$_{YP}$ AP $_{Y}$ t$_{NP}$]]] [$_{WP}$ DemP $_{W}$ [$_{AgrXP}$ t$_{XP}$]

t. [$_{AgrWP}$ [$_{AgrXP}$ NP [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ [$_{YP}$ AP $_{Y}$ t$_{NP}$]]] [$_{WP}$ DemP $_{W}$ t$_{AgrXP}$]]]

w. [$_{AgrWP}$ [$_{AgrXP}$ [$_{YP}$ AP $_{Y}$ NP] [$_{XP}$ NumP $_{X}$ [$_{AgrYP}$ t$_{YP}$] [$_{WP}$ DemP $_{W}$ t$_{XP}$]]]]

x. [$_{AgrWP}$ [$_{AgrXP}$ [$_{AgrYP}$ NP [$_{YP}$ AP $_{Y}$ t$_{NP}$]] [$_{XP}$ NumP $_{X}$ t$_{AgrYP}$] [$_{WP}$ DemP $_{W}$ t$_{AgrXP}$]]]

Cinque's analysis captures important facts: not just the possible and impossible nominal orders,[6] but their derivation as well, hence their bracketed structure. Any purported improvement on this account should preserve these descriptive successes, while either capturing additional empirical facts, or simplifying the theoretical apparatus.

It turns out that this array of orders (and their bracketed structure) admits a method of generation that appears simpler than Cinque's account (or that of Abels & Neeleman 2012, Steddy & Samek-Lodovici 2011, or related analyses[7]). This method imposes freely generated *n*-ary branching structure[8] on an arbitrary string of formatives, closely following Chomsky's assertion that Merge applies freely. The account generates all and only the attested orders and bracketed structures; once the bracketing is fixed in any of the legal ways, the assignment of hierarchy to the elements follows uniquely. This result is unexpected, but notable in its simplicity. Here is the procedure:

(17) *Generative procedure over strings*

a. Start with a string of unidentified formatives.

x x x x

b. Place a left bracket just before each formative.

[x [x [x [x

c. Place a matching number of right brackets to form any *legal bracketing*.

[x] [x [x] [x]]

---

[6] See Dryer (2018) for a different assessment of the typological facts, allowing some orders Cinque (2005) excludes, and explaining the pattern in quite a different way. The present account assumes Cinque's typology is accurate.

[7] Medeiros (2018) proposes an analysis for Universal 20 involving identical tree structures. But that work commits to a performance-level account of a universal parser; the present proposal is a pure generative account. Moreover, we reject Medeiros' (2018) claim that hierarchical order (here, postorder traversal sequence; there, the string output of stack-sorting) directly follows the order of composition. Instead, we adopt the weaker claim that the relevant linear order unambiguously determines composition; see below.

[8] The *n*-ary branching structure in question is a tree with linear order; put another way, this version of External Merge produces an ordered tuple of its operands. This loses the competition with set-based Merge for mathematical simplicity. But allowing serial order within syntactic representations plausibly draws on capacities other animals possess.

d. Scan *left-to-right*, indexing right brackets in increasing order.[9]

$[x]_1\ [x\ [x]_2\ [x]_3\ ]_4$

e. Copy indices from right brackets onto formatives following the corresponding left brackets.

$[x_1\ ]_1\ [x_4\ [x_2\ ]_2\ [x_3\ ]_3\ ]_4$

The indexing encodes the relative hierarchy of the formatives (see below), and the bracketed structure is the surface structure bracketing. In this case, we derive (17):

(18) [1] [4 [2] [3]]

Procedure (17) generates all and only attested nominal word orders, and their bracketed structure. Importantly, this does not simply repackage the Cinque-style EM and IM account. Identifying Merge with brackets (one pair of brackets represents the Merge of what the brackets enclose), there is a fixed number of such operations for all orders: exactly *n* for *n* formatives. In a standard framework employing EM and IM, for the same lexical input there are *n*-1 External Merges, and variable *k* Internal Merges. The present perspective also dissolves the question of what drives movement: the attested orders are simply the base-generable structures. There are no steps of movement, and no need to explain them.[10] Conversely, unattested orders are not ruled out by constraints on movement, but correspond to impossible bracketings; see below.

No binarity constraint applies here: brackets may enclose singletons, triples, etc., effectively permitting *n*-ary branching. Placing left brackets before each lexical element, and nowhere else, differs from standard practice; linguists would expect [[*ab*]*c*] to be a possible structure, but that is ruled out here. This does not mean that "left-branching" structure is impossible. Rather, structure traditionally analyzed as left-branching maps to a horizontal relation between nodes, while right-branching structure comes out as

[9] Linguists number hierarchies top down, from least to most embedded. Following that convention would index right brackets in the reverse of postorder traversal order. This leads linguists to characterize the forbidden permutation as *213, (*e.g.*, in the verb cluster literature). But this conflicts with well-established conventions in computer science and mathematics, where the *PostPre* permutations (see Feil *et al* 2005) here are the stack-sortable words (see Medeiros 2018), avoiding *231 permutations. I adopt the more general convention, at risk of confusion.

[10] This also means that we lose any obvious syntax-internal explanation for the relative typological frequency of different orders (*e.g.*, the harmonic orders N-Adj-Num-Dem and Dem-Num-Adj-N are the most common). See Cinque (2005), and references cited there.

vertically arranged nodes.[11] While this departs from the usual way of thinking about brackets and their relation to lexical elements, it yields the right orders and their structure at a stroke. Table 1 shows all possibilities generated with four string formatives.

Brackets	Formatives	Index rt brackets	Index formatives	Order	Nominal order
(((())))	(x(x(x(x))))	$(x(x(x(x)_1)_2)_3)_4$	$(x_4(x_3(x_2(x_1)_1)_2)_3)_4$	4321	a. Dem-Num-Adj-N
((()()))	(x(x(x)(x)))	$(x(x(x)_1(x)_2)_3)_4$	$(x_4(x_3(x_1)_1(x_2)_2)_3)_4$	4312	b. Dem-Num-N-Adj
((())())	(x(x(x))(x))	$(x(x(x)_1)_2(x)_3)_4$	$(x_4(x_2(x_1)_1)_2(x_3)_3)_4$	4213	n. Dem-Adj-N-Num
((()))()	(x(x(x)))(x)	$(x(x(x)_1)_2)_3(x)_4$	$(x_3(x_2(x_1)_1)_2)_3(x_4)_4$	3214	r. Num-Adj-N-Dem
(()(()))	(x(x)(x(x)))	$(x(x)_1(x(x)_2)_3)_4$	$(x_4(x_1)_1(x_3(x_2)_2)_3)_4$	4132	c. Dem-N-Num-Adj
(()()())	(x(x)(x)(x))	$(x(x)_1(x)_2(x)_3)_4$	$(x_4(x_1)_1(x_2)_2(x_3)_3)_4$	4123	o. Dem-N-Adj-Num
(()())()	(x(x)(x))(x)	$(x(x)_1(x)_2)_3(x)_4$	$(x_3(x_1)_1(x_2)_2)_3(x_4)_4$	3124	s. Num-N-Adj-Dem
(())(())	(x(x))(x(x))	$(x(x)_1)_2(x(x)_3)_4$	$(x_2(x_1)_1)_2(x_4(x_3)_3)_4$	2143	k. Adj-N-Dem-Num
(())()()	(x(x))(x)(x)	$(x(x)_1)_2(x)_3(x)_4$	$(x_2(x_1)_1)_2(x_3)_3(x_4)_4$	2134	w. Adj-N-Num-Dem
()((()))	(x)(x(x(x)))	$(x)_1(x(x(x)_2)_3)_4$	$(x_1)_1(x_4(x_3(x_2)_2)_3)_4$	1432	d. N-Dem-Num-Adj
()(()())	(x)(x(x)(x))	$(x)_1(x(x)_2(x)_3)_4$	$(x_1)_1(x_4(x_2)_2(x_3)_3)_4$	1423	p. N-Dem-Adj-Num
()(())()	(x)(x(x))(x)	$(x)_1(x(x)_2)_3(x)_4$	$(x_1)_1(x_3(x_2)_2)_3(x_4)_4$	1324	t. N-Num-Adj-Dem
()()(())	(x)(x)(x(x))	$(x)_1(x)_2(x(x)_3)_4$	$(x_1)_1(x_2)_2(x_4(x_3)_3)_4$	1243	l. N-Adj-Dem-Num
()()()()	(x)(x)(x)(x)	$(x)_1(x)_2(x)_3(x)_4$	$(x_1)_1(x_2)_2(x_3)_3(x_4)_4$	1234	x. N-Adj-Num-Dem

Table 1: From free bracketing to word orders. Columns show: brackets; with formatives included; with right brackets indexed; with formatives indexed; hierarchically numbered order; nominal order. These are the attested orders, according to Cinque (2005); see (15). I provide Cinque's lettering scheme for the orders for reference.

3 A closer look at the details

This section explores selected aspects of the account in greater depth. This includes defining the notions we will use in this account, describing the architecture in terms of trees and tree traversal algorithms, illustrating how the brackets for nominal orders in this account correspond to Cinque's derivations, and examining how the account excludes unattested orders.

3.1 Initial definitions

[11] A question for future research is whether the predicted asymmetry between X-Y and Y-X orders can be aligned with Wagner's (2005) observations about prosodic asymmetries correlated with linear order of predicates and arguments, and modifiers and heads.

Before we proceed, we define some notions that will be pivotal to what follows.

(19) *Definitions*

a. **Legal bracketing** A legal bracketing is a string consisting of left and right brackets meeting the following two conditions: (i) at each position in the string, the number of preceding right brackets may not exceed the number of preceding left brackets, and (ii) at the end of the string, the number of left brackets and right brackets are equal. So, for example, '( ) ( )' and '((( )))' are legal bracketings, while ') (' and '(( )' are not (failing condition (i) and (ii), respectively). The number of these strings of fixed length is a number from the Catalan sequence (1, 1, 2, 5, 14, 42, 132, ...). With no bracket pairs, we have one legal bracketing (the empty string); there is one choice for one pair of brackets, two choices for two pairs, 5 for three pairs, 14 for four pairs, etc.

b. **Dyck word** A Dyck word is a string over an alphabet of two symbols (A, B), meeting the following two conditions: (i) at each position in the string, the number of preceding Bs may not exceed the number of preceding As, and (ii) at the end of the string, the number of As and Bs are equal. Where A = '(', B = ')', these are the legal bracketings.

c. **Dyck tree** The Dyck trees are the set of *n*-ary branching rooted trees with a fixed number of nodes. These trees correspond one-to-one with Dyck words, and are also counted by the Catalan numbers. Each matched pair of symbols/ parentheses in the Dyck word maps to a single node in the corresponding Dyck tree; one pair of parentheses containing another maps to a dominance relation between the corresponding nodes in the Dyck tree.

d. ***n*-ary branching tree** A rooted, non-tangling tree. Here, each node may have 0 or more daughter nodes, which are linearly ordered with respect to each other. (We do not distinguish left and right daughter in unary-branching contexts, unlike common implementations of binary search trees.) We will use these trees quite differently than standard syntactic trees, writing lexical items onto terminal and non-terminal nodes alike (each lexical item is linked to a single node in the tree, obviating issues of labelling and chains). This is closer to the notion of tree used in computer science.

e. **Tree traversal** Tree traversals can be defined in terms of the priority of direction of travel from each node, with respect to three directions: **Root** (Up to the dominating node), **L** (Down to the leftmost daughter), **R** (Down to the rightmost daughter). By providing an ordering of these three directions, we recursively define a method for traversing a tree, starting at the root.

f. **Postorder traversal** One of three standard traversal algorithms, defined by the priority list (L, R, Root). Descriptively, this visits nodes in a tree left-to-right and bottom-up. In terms of the equivalent bracketed expression, this visits right brackets in left-to-right order.

g. **Preorder traversal** Another of the three standard tree traversal algorithms, defined by the priority list (Root, L, R). Descriptively, this visits nodes in the tree top-down and left-to-right. In terms of the equivalent bracketed string, this visits left brackets in left-to-right order.

h. **Permutation** Given a set of symbols, a permutation is a sequential arrangement of those symbols. Given a reference sequence taken as the identity permutation, we can describe other permutations of the same set of elements perspicuously in terms of a numerical sequence, where the identity permutation is the sequence *1 2 3 4 5...* . In what follows, the identity permutation is a language-invariant representation of the underlying hierarchy of an expression, and we will be considering surface orders as permutations of this basic sequence.

i. **Index**. The number associated with a position in the identity permutation, which we will use to refer to the relative positions of lexical elements within the invariant hierarchical order.

j. **Subsequence** Given a a sequence of symbols (a permutation), a subsequence is any linear arrangement of a subset of symbols from that sequence that preserves their relative linear order. For example, the permutation *1 2 5 4 3* contains *2 4* as a subsequence.

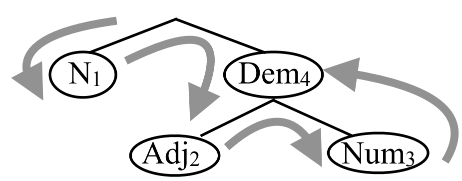
k. **Forbidden permutation** This work characterizes unavailable word orders in terms of a subsequence contour within their surface order (namely, *231). This condition does not refer to any three specific lexical items or hierarchical positions. Rather, given a linearly-ordered representation of the hierarchy as the identity permutation, we rule out any surface order permutation that contains a subsequence *bca*, where $a < b < c$ in the identity order. To be clear, this condition does not require adjacency of the elements in either the surface order or in the identity permutation/underlying hierarchical order. So, for example, given the identity permutation *1 2 3 4 5 6* ..., the order *4 1 6 2 3 5* contains the forbidden *231 contour (its subsequences *4 6 2* and *4 6 3* have mid-high-low index pattern).

l. **Hierarchical order** A hierarchical order is the identity permutation of the lexical elements in an expression. The idea is that expressions in different languages with commensurable lexical content share the same underlying hierarchical representation. In the present framework, this representation is a particular permutation of the elements involved. For example, the hierarchical order for a transitive clause (motivated in section 4) is C < Pol < T < V < O < Adv < S. Thus, *C Pol T V O Adv S* is the identity permutation *1 2 3 4 5 6 7*; C has index 1, Pol has index 2, etc. Clause orders in different languages are different permutations of this underlying hierarchical order. The English clause *...that they didn't often eat cake* arranges elements in the order *C S T Pol Adv V O* (taking the Neg head *-n't* to instantiate the Pol position; see below). In terms of numerical indices, the English clause is the permutation *1 7 3 2 6 4 5*. This permutation is 231-free; none of its subsequences form the forbidden *231 contour.

3.2 The procedure in terms of tree traversals

Procedure (17) equates to hierarchization (*i.e.*, labeling) of trees by postorder traversal, and linearization by preorder traversal. Postorder traversal visits nodes in the tree left-to-right and bottom-up. To illustrate, (20) shows 1423 nominal order (N-Dem-Adj-Num) in tree form. The direction of postorder traversal is indicated by large grey arrows; subscript indices record the order in which the nodes are visited.

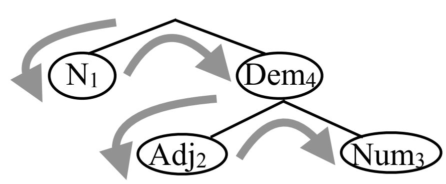
(20) *Postorder traversal*



As shown, postorder indexing allows the nodes to be mapped to a linear representation of the underlying syntax; in this example, we take the elements of the Universal 20 hierarchy bottom-up. (See section 4 for refinements in this linear hierarchy.)

Once the tree has been hierarchized this way, linear order is read off by preorder traversal, which goes top down, left-to-right. The path of preorder traversal is shown with grey arrows in (21); this path visits the nodes in surface order, N-Dem-Adj-Num.

(21) *Preorder traversal*



The notion of tree utilized here is the computer science data structure, which differs from traditional syntactic trees (notably, words are associated with all nodes). Figure 1 summarizes the action of this generative architecture over trees.

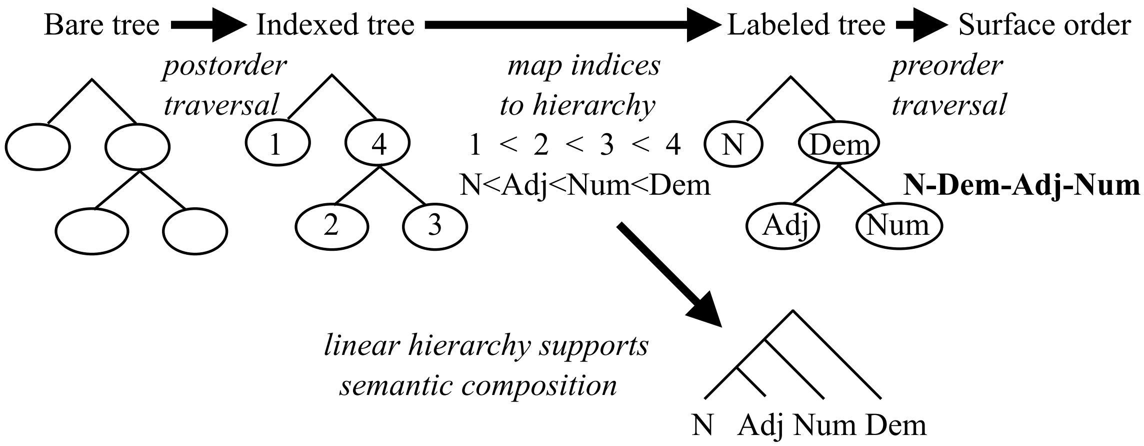


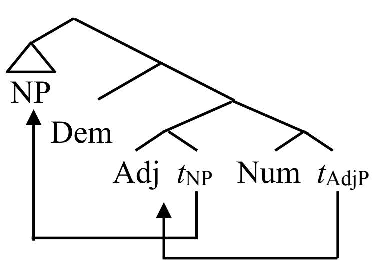
Figure 1: *Generating N-Dem-Adj-Num (1423) order*
Free Merge builds a bare *n*-ary tree. Postorder traversal indexes nodes. Indices map to hierarchical order (in this case, the hierarchy for Universal 20), yielding lexical labels on nodes. Preorder traversal of the labeled tree gives surface order; here, N-Dem-Adj-Num. Separately, hierarchical order supports semantic composition in familiar bottom-up order.

3.3 Correspondence with traditional bracketed representations

Returning to bracketed strings, the bracketing generated in this account closely matches that in Cinque's derivations. To illustrate the correspondence, we continue with the example of *1423* (N-Dem-Adj-Num) order. Translating to the Universal 20 hierarchy, the structure is (22).

(22) [N] [Dem [Adj] [Num]]

Illustrated below in (23) is a (simplified) Cinque-style derivation of this order.

(23)


In this derivation, the [Adj-NP] complex moves to precede Num, followed by subextraction of NP to a specifier position before Dem. In bracketed form, we have (24) :

(24) [[NP] [Dem [[Adj t_{NP}] [Num t_{AdjP}]]]]

Keeping only bracket pairs where the left bracket immediately precedes a lexical element (within the NP as well, i.e. NP ~ [N]), and ignoring traces, we get (25):

(25) [N] [Dem [Adj] [Num]]

As claimed, this simplified version (25) of Cinque's bracketing (24) is identical to expression (22) derived by the generative procedure in (17).

3.4 Unattested orders require impossible bracketing

Consider in more detail how unattested orders are ruled out. With a hierarchy of three elements (say, N=1, Adj=2, Dem=3), five of six logically possible orders are attested as neutral noun phrase orders. One permutation, *231 (*Adj-Dem-N, usually described as *213 according to linguists' convention; see fn. 9), does not occur as a basic noun phrase order. The present proposal explains this systematic gap.

Since left brackets occur immediately before each surface element, and nowhere else, we can begin to fill in what a *231 order would look like as a bracketed string.

(26) $[_2\ 2\ \ldots\ [_3\ 3\ \ldots\ [_1\ 1\ \ldots$

Right brackets are indexed left-to-right, so they occur in the sequence $]_1\ \ldots\ ]_2\ \ldots\ ]_3$. Furthermore, right brackets follow the left bracket and element they match. Therefore, the entire sequence of right brackets must follow the element 1. This gives us:

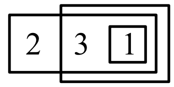
(27) $[_2\ 2\ [_3\ 3\ [_1\ 1\ ]_1\ ]_2\ ]_3$

This is not a legal (indexing of a) bracketing; the boundaries of bracketings 2 and 3 cross. To see this, we can think of brackets as denoting the edges of “boxes”. In generated orders, any pair of boxes may be in a containment relation, or be disjoint; they cannot overlap partially. Illustrating with 321 and 123 order and appropriate bracketing:

(28) $[_3\ 3\ [_2\ 2\ [_1\ 1\ ]_1\ ]_2\ ]_3$ $\boxed{3\ \boxed{2\ \boxed{1}}}$

(29) $[\ 1\ ]_1\ [_2\ 2\ ]_2\ [_3\ 3\ ]_3$ $\boxed{1}\ \boxed{2}\ \boxed{3}$

But the unattested *231 order entails overlapping boxes:

(30) $[_2\ 2\ [_3\ 3\ [_1\ 1\ ]_1\ ]_2\ ]_3$ 

Given the way procedure (17) works, unattested *231 order cannot be generated. Instead, the relevant bracketing must form a 321 order as in (28); bracketing determines hierarchy.

4 Generating the Final-Over-Final Condition

In this section, I show that the architecture developed to this point provides a ready explanation for another intensively studied word order universal, the Final-Over-Final Condition (FOFC; Holmberg 2000, Biberauer *et al* 2014, Sheehan et al 2017 *i.a.*). This is a surprising unification, as Universal 20 and FOFC appear to conflict; see for example Roberts (2017) on modifying the hierarchy for the noun phrase (14) to be compatible with FOFC.

4.1 Background: The Final Over Final Condition

FOFC prohibits configuration (31):

(31) $*[_{\alpha P}\ [_{\beta P}\ \beta\ \ \gamma P]\ \alpha]$

That is, a head-final phrase cannot dominate a head-initial phrase. The example below, from Finnish, illustrates the phenomenon.

(32) a. yli [rajan maitten välillä] $[P_1\ [N_1\ [[N_2]\ P_2]]]$
across border countries between
'across the border between countries'

b.*[rajan maitten välillä] yli $*[[N_1\ [[N_2]\ P_2]]\ P_1]$
border countries between across (Biberauer *et al* 2014: 187, *ex*. 29)

In the ungrammatical (32b), the outermost P_1 has its NP complement on the left, while the embedded nominal has its PP complement on the right. This is the banned *final-over-initial configuration. Biberauer et al (2014) list the following FOFC effects; these configurations are robustly ungrammatical across languages.

(33) a. *V-O-Aux $*[_{AuxP}\ [_{VP}\ V\ DP]\ Aux]$
b. *V-O-C $*[_{CP}\ [_{TP}\ T\ VP]\ C]$ or $*[_{CP}\ [_{TP}\ [_{VP}\ V\ O]\ T]\ C]$

c. *C-TP-V *[$_{VP}$ [$_{CP}$ C TP] V]
d. *N-O-P *[$_{PP}$ [$_{DP/NP}$ D/N PP] P]
e. *Num-NP-D(em) *[$_{D(em)P}$ [$_{NumP}$ Num NP] D(em)][12]
f. *Pol-TP-C *[$_{CP}$ [$_{PolP}$ Pol TP] C]

(Biberauer *et al* 2014: 196, *ex*. 46)

These canonical FOFC effects obtain when the elements in question are in a head-complement relation. This well-known characterization of the domain of this condition is the key to the unification of this class of word order constraints with the Universal 20 pattern pursued in the next subsection.

4.2 Refining the notion of hierarchical ordering

The account of Universal 20 offered in section 2 depends crucially on how the nominal hierarchy is mapped to freely-generated trees. This includes not just choosing post-order traversal, one of several standard tree traversal algorithms, but determining how to compress a representation of linguistic hierarchy into a sequence that can be mapped to the node traversal order.[13] In this regard, it is notable that fixed relations among syntactic elements seem to come in (at least) two flavors: selection and adjunction, or head-complement (and more generally, head-argument) and head-adjunct relations.[14]

Returning to the technical details of this framework, postorder traversal visits nodes/right brackets inside-out, left-to-right. It is natural to assign indices in the same order: the innermost leftmost right bracket/node is 1, the next is 2, etc. We define the hierarchical ordering relation '<' in the usual way with respect to this indexing of the traversal sequence; for example, 1 < 2.

In these terms, I propose that a head H and its adjunct A are mapped to this sequence such that H < A.

(34) H < A *Head-adjunct hierarchical order*

---

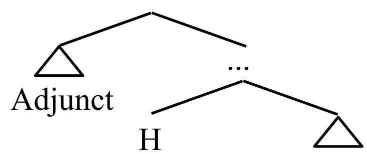
[12] See Roberts (2017) for motivation of this claim. D(em) here reflects an analysis where Dem originates low in the hierarchy, and in some languages moves to higher head D.

[13] The general idea of more or less "linear" syntactic representations is hardly new; see Starke (2004), Jayaseelan (2008), and references there.

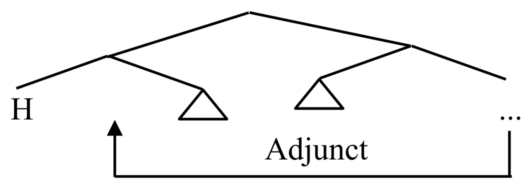
[14] Additional stipulations may be required to model conjunction, set aside here. But if we treat coordination asymmetrically with the mechanisms here, akin to [N PP] complementation (e.g., coordination of N heads would form [N [& N]] in traditional notation), we would predict an apparent typological gap in monosyndetic coordination (Haspelmath 2017) for order *&-N-N (an observation of Ryan Walter Smith).

This hierarchical ordering corresponds to a traditional tree structure in which the head is more deeply embedded than its adjunct, as in the abstract derivation of adjunct-head and head-adjunct order shown below.

(35) a.



b.



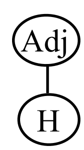
In Kayne's (1994) antisymmetric framework, adjunct-head order reflects the base structure unaffected by movement (35a), while head-adjunct order involves moving the head (here, as part of a phrasal movement) left of the adjunct (35b). In terms of bracketed structures, we have the following.

(36) a. [Adjunct [H ...]]

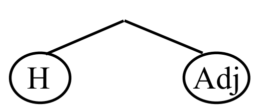
b. [[H ...] [Adjunct *t*]]

The representation of these orders with brackets and trees in the present framework is shown below (37). Nested bracketings (37a) make for vertical node arrangements in the case of adjunct-head order; disjoint bracketings are arranged horizontally (37b) for head-adjunct order.

(37) a. [Adjunct [H]]



b. [H] [Adjunct]



If H has several adjuncts A_1, A_2, with A_1 the closest in traditional representations, we will have $H < A_1 < A_2$. Restricting attention to a hierarchy comprised of a head and a series of adjuncts to that head, we will find *231-avoidance: *A_1-A_2-H . This pattern is seen in Cinque's version of Universal 20 (understanding demonstrative, numeral, and adjective as adjuncts)[15], and arguably in verb clusters.[16] For the Universal 20 case, I repeat the following hierarchy:

(38) N < Adj < Num < Dem *Universal 20 hierarchical order*

The next section takes up the matter of the hierarchical ordering of heads and their complements (and other arguments).

4.3 Hierarchical ordering extended to complementation

In standard analyses, heads and complements are in a symmetric hierarchical relationship. The present account provides no basis for such symmetry, and we must make a choice: heads must be hierarchically above, or below, their complements (because we are mapping syntactic hierarchy onto the necessarily-linear tree traversal sequence).

Suppose that head-complement relations obey the same *H < X* convention: head H and complement C map to the post-order traversal index sequence such that H < C.[17]

(39) H < C *Head-complement hierarchical order*

This will produce the basic phenomenology of the Final-Over-Final Condition (FOFC; Sheehan *et al* 2017) in structures characterized by head-complement relations.

To see this, consider a configuration with nested complementation: head α takes a complement headed by β, which in turn has complement γ. The hierarchical order is then (40) α < β < γ, and the forbidden *231 permutation is (41) *β-γ-α.

(40) α < β < γ *Nested complementation hierarchy*

[15] The relevant categories are indeed optional modifiers. But even under Cinque's analysis, in which they are specifiers of associated functional heads (f_{Dem}, f_{Num} etc.), we make the same prediction. Anticipating what follows, the relevant hierarchical order will be $f_{Dem} < f_{Num} < f_{Adj}$ < N < AdjP < NumP < DemP. Since the functional heads are not overt, this is effectively the same hierarchy assumed in (38).

[16] But see Salzmann (2019) on attested 213 (for us, 231; see fn. 10) verb clusters. See Abels (2016) on Universal 20 effects in other domains, and refinements.

[17] While H < Adjunct hierarchical order reflects traditional analyses, breaking head-complement symmetry this way is a stipulation. On the other hand, the distinction between head-adjunct and head-complement relations collapses (cf. Abels 2016 on the notion of "satellite"), in the sense that both obey the condition H < X.

(41) * β-γ-α *Forbidden word order*

Order (41) is traditionally described as a head-final phrase (αP) dominating a head-initial phrase (βP), exactly the configuration ruled out by FOFC (31), repeated as (42).

(42) $*[_{\alpha P}\ [_{\beta P}\ \beta\ \ \gamma P]\ \alpha]$

For example, if head Aux has complement headed by V, with complement Obj, the hierarchy is Aux < V < Obj (43). We correctly exclude unattested *231 order *V-Obj-Aux (44).

(43) Aux < V < Obj

(44) *V-Obj-Aux

Since the reasoning is about heads and complements (not just verbs and auxiliaries), we expect this to generalize to any head-complement chain, reconstructing the core of FOFC.

4.4 Further extensions of hierarchical ordering

What about structures with both adjuncts and complements? Sheehan (2017) argues that FOFC extends to certain adjunct relations. Concretely, parallel to the FOFC effect *V-Obj-Aux, *V-Adv-Aux is unattested. A full discussion is put aside, but note that this effect is correctly predicted here. This follows from the already-motivated hierarchical sequence, Aux < V < Adv (45); unattested *V-Adv-Aux (46) is the forbidden *231 permutation.

(45) Aux < V < Adv *Auxiliary, verb, adverb hierarchy*

(46) *V-Adv-Aux *Forbidden word order*

In existing models of syntax, complements are the closest element to the head; adjuncts are farther away. The same relation is encoded by our ordering, H < Comp < Adjunct: the complement is the unique closest element to the head. In the standard model, while H-adjunct relations involve asymmetric hierarchy (the adjunct is above the head), head-complement relations are symmetric. The present approach avoids this unwanted symmetry (by stipulation), with promising consequences for word order constraints.

Where a head H takes both arguments and adjuncts, I assume the relative hierarchy is H < Arg < Adj (47). If there are multiple arguments of a head, the complement is closest to the head: H < Comp < Arg' (48).

(47) H < Arg < Adj *Hierarchical order of head, argument, adjunct*

(48) H < Comp < Arg' *Hierarchical order of head, complement, argument*

In particular, for verb head V and complement object O, V < O. The same hierarchy holds for a verb and complement clause: V < CP. A ditransitive verb would have V < DO < IO (see Pearson 2000, Abels 2016 on the unattested status of *DO-IO-V). If there is an adverbial and an object, the hierarchy is V < O < Adv. (49) puts these together into a single ordering.[18]

(49) V < DO/CP < IO < Adv *Hierarchy of verb, objects, adverb*

Adding Tense and subject, the order is T < V < O < S. If we include little *v*: T < *v* < V < O < S. No overt item realizes little *v* in the majority of the examples considered below; I omit it for simplicity where it will not lead to confusion. If complementizer C and/or polarity head Pol is present, I assume the structure is as in (50).

(50) C < Pol < T < *v* < V < O < S *Hierarchical order for transitive clause*

This all may seem stipulative. Note, first, that the hierarchy in (50) is similar to standard proposals, *modulo* the unusual resolution of head-complement structures.[19] Once we have postulated an underlying hierarchy, this system makes systematic predictions about possible and impossible neutral word orders of the relevant elements. If an error is made in determining the hierarchy, a multitude of false predictions should follow through interactions with the rest of the ordering. But with the assumptions made so far, a range of familiar typological facts are captured.

We can consider elements belonging to a single hierarchy three at a time; we should find, for each such triple, five attested orders and one forbidden order. Drawing on order (50), understanding that the O position may be realized as clausal complement CP, we make the following predictions (among others) about impossible neutral orders.

(51) a. *O-S-V

[18] This is probably to be read as shorthand for a richer set of positions. The *Adv* position in (49) in particular is dubious; Cinque (1999) proposes a highly articulated hierarchy of adverbs interspersed with numerous heads. While the present work keeps to a minimal set of categories, finer-grained categorizations would not affect the ordering predictions here so long as the relative hierarchy of the categories involved is preserved.

[19] Note that the sequence (50) taken in descending order (S-O-V-T-C) is a cross-linguistically common clause order.

b. *CP-S-V
c. *O-S-T
d. *V-O-T
e. *V-O-C
f. *V-CP-T
g. *[C-TP]-V
h. *Pol-TP-C
i. *V-S-T

An adpositional phrase object O will be hierarchically ordered after a noun head N it complements; I take adposition P to be a head with noun phrase complement NP.

(52) N < O

(53) P < NP

This yields hierarchical order (54), with forbidden permutation (55).

(54) P < N < O *Hierarchical order for PP within PP*

(55) *N-O-P *Forbidden order*

This explains the typological gap illustrated in Finnish (30b) above, previously described with FOFC. In fact, setting aside (33e/56e) (we adopt Cinque's hierarchy for Universal 20 effects), we have reconstructed the list of canonical FOFC effects in Biberauer *et al* (2014: 196), repeated below.

(56) *FOFC effects predicted here*

a. *V-O-Aux	see (44)
b. *V-O-C	(51e)
c. *C-TP-V	(51g)
d. *N-O-P	(55)
e. *Num-NP-D(em)[20]	as Cinque (2005); see (38)
f. *Pol-TP-C	(51h)

Beyond reconstructing this core of FOFC effects, (51) contains other interesting predictions. If one basic clause order is to be ruled out, *OSV appears to be the right

---

[20] See Roberts (2017) for motivation of this claim. D(em) here reflects an analysis where Dem originates low in the hierarchy, and in some languages moves to higher head D.

choice (51a), as it is the rarest cross-linguistic order. That said, some mechanism going beyond the simple base-generation system here must be invoked for the handful of languages with OSV orders. Another interesting prediction is (51h), taken up again in section 6.3 below as a reformulation of Travis' (1984) Head Movement Constraint.

5 Generating some well-known crossing dependencies

Thus far, we have been concerned with ruling out typologically unattested orders. In this section, I turn to showing that the analysis of allowed orders extends to somewhat exotic constructions that have figured prominently in arguments that natural language grammars are mildly context-sensitive (see *e.g*. Joshi 1985). Bresnan *et al* (1982) discuss unbounded crossing subject-verb dependencies in Dutch (Huygbregts 1976). Example (57), taken from Steedman (2000: 25), illustrates the phenomenon.

(57) ...omdat ik Cecilia Henk de nijlpaarden zag helpen voeren
...because I Cecilia Henk the hippos saw help feed
'...because I saw Cecilia help Henk feed the hippos'

Shieber (1985) discusses similar word order facts in Swiss German, which also exhibits long-distance cross-serial case dependencies. Interestingly, the system already established can base-generate these orders.[21] I assume the example above contains the categories in (56), abstracting away from internal structure of the object *de nijlpaarden* 'the hippos' and segmenting a Tense suffix from inflected and non-finite verbs, even if realized as zero.

(58) ...omdat ik Cecilia Henk de nijlpaarden zag-0 help-en voer-en
C S_1 S_2 S_3 O_3 V_1 T_1 V_2 T_2 V_3 T_3

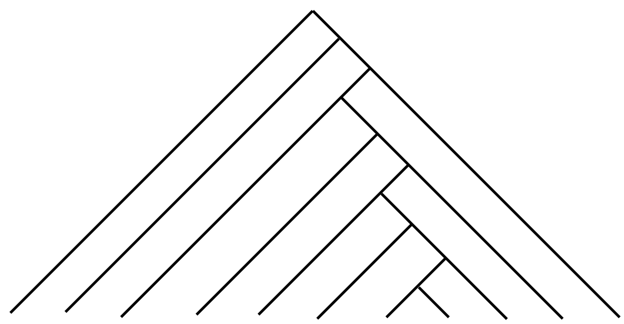
The categories in (58) will be rendered as a single linear hierarchy, which we assemble from the general clause ordering (50), together with the standard assumption

[21] Stabler (2004) discusses four different classes of cross-serial dependency constructions, with distinct formal properties. I restrict attention to the two classes in this section.

that complement clauses occupy the canonical direct object position; this allows us to integrate clausal complementation with the clause order (50) above.[22]

For single clausal embedding, $[_{CP1} \ldots [_{CP2}\ \_\_\ ]\ ]$, we have: $C < T_1 < V_1 < \underline{T_2 < V_2 < O_2 < S_2} < S_1$. Replacing O_2 with another embedded clause, we derive (57), the hierarchical order for sentence (58) above. I show postorder indices aligned to the hierarchy, on which a superposed tree shows bottom-up semantic composition.

(59) *Integrated hierarchy for (58) with postorder index and composition tree*



Composition tree

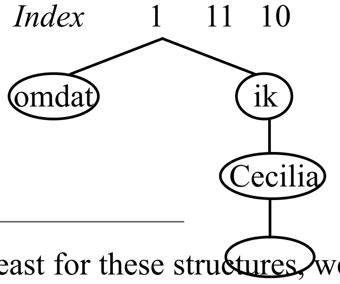
$C < T_1 < V_1 < T_2 < V_2 < T_3 < V_3 < O_3 < S_3 < S_2 < S_1$ Hierarchical order

$1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9 < 10 < 11$ Postorder index

Given this mapping from syntactic hierarchy to post-order index sequence, we can easily recover the tree structure corresponding to the Dutch surface order,[23] shown in (60).

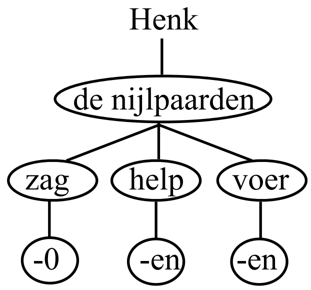
(60) ...omdat ik Cecilia Henk de nijlpaarden zag-0 help-en voer-en

Category	C	S_1	S_2	S_3	O_3	V_1	T_1	V_2	T_2	V_3	T_3
Index	1	11	10	9	8	3	2	5	4	7	6



[22] At least for these structures, we are implicitly developing a simple account of recursion by substitution. For this example, the account requires that the entire structure be generated at once; the clauses in this example cannot be cyclic domains, generated one at a time (see discussion in section 6). Other clauses may be; I leave fuller consideration of recursion in this architecture to future work, beyond the brief comments below.

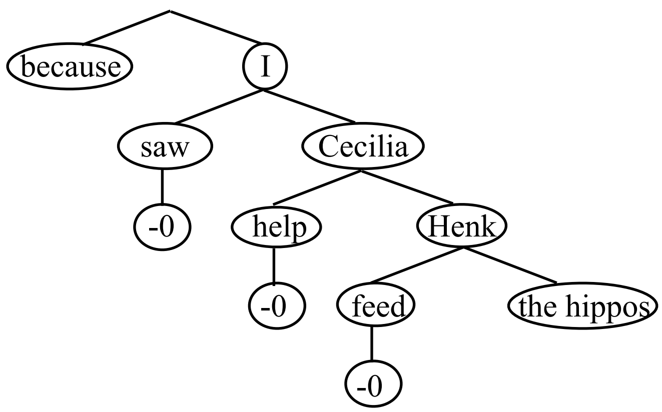
[23] An important question is whether these trees provide a basis for a successful theory of prosody (see also fn. 11). While it is promising that the trees derived here correspond closely to Cinque's derivations of nominal orders, I leave this question for future work. Unlike the nominal trees, the clausal trees in this section differ from standard analyses.



With the relevant syntactic hierarchy resolved as a universal linear sequence, we can readily represent other orders of the same elements, as in English (61).

(61) ...because I saw -0 Cecilia help -0 Henk feed -0 the hippos

C	S_1	V_1	T_1	S_2	V_2	T_2	S_3	V_3	T_3	O_3	*Category*
1	11	3	2	10	5	4	9	7	6	8	*Index*

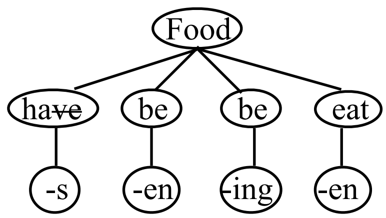


Finally, this architecture generates the more limited bounded crossing dependencies in English Affix-Hopping (Chomsky 1957), as seen in example (62).

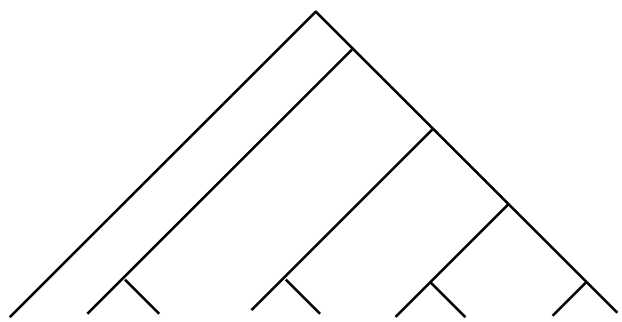
(62) Food ha-s be-en be-ing eat-en

As Chomsky noted, affixes group with preceding auxiliaries in distribution and meaning, despite being separated by the intervening verb in surface order. To accommodate this pattern, suppose auxiliary Aux and associated affix -Fx have hierarchical order Aux < -Fx < VP.[24] This generates (63), with composition tree (64).[25]

(63)	Food	ha~~ve~~	-s	be	-en	be	-ing	eat	-en	
	Obj	Aux_1	T	Aux_2	$-Fx_1$	Aux_3	$-Fx_2$	V	$-Fx_3$	*Category*
	9	2	1	4	3	6	5	8	7	*Postorder Index*



(64) *Hierarchical order and composition tree for English Affix-Hopping (60)*



Composition tree

-s have -en be -ing be -en eat food *Lexical items*

$T < Aux_1 < Fx_1 < Aux_2 < Fx_2 < Aux_3 < Fx_3 < V < Obj$ *Hierarchical order*

$1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9$ *Postorder index*

6 Further applications

In this section, I take up some further extensions to the framework, showing how they can account for some familiar phenomena. First, I reconsider Abel's conclusion that hierarchies must be relativized to different types of satellites, showing that for some core examples, the present framework allows a single hierarchy after all.

[24] One can read this as saying the affix (*e.g.*, -*ing*) is a head sandwiched between selecting auxiliary (*be*) and host verb (*cf*. Harwood 2013). Or auxiliary and associated affix might "fuse" for interpretation, mirroring Chomsky's (1957) analysis with a single lexical item (*be*+*ing*). The composition tree in (64) reflects the latter choice.

[25] It is unclear if passive movement of the object should be base-generated, or if it is obligatorily "real" movement. It is at least possible to generate with just this mechanism.

The following subsection illustrates analyses of clausal orders in several languages, suggesting that they permit a single "high" hierarchical position for negation. Next, I show how (a version of) the Head Movement Constraint is predicted here as well. Finally, I consider two well-known exceptions to FOFC, suggesting some simple analyses within the present framework, and sketching some cyclic effects.

6.1 A unified hierarchy after all?

Abels (2016) examines effects paralleling Universal 20 in other domains, focusing on verb clusters. He argues that verbal arguments and verb-auxiliary hierarchies cannot be unified into a single hierarchy, on the basis of examples like the following.

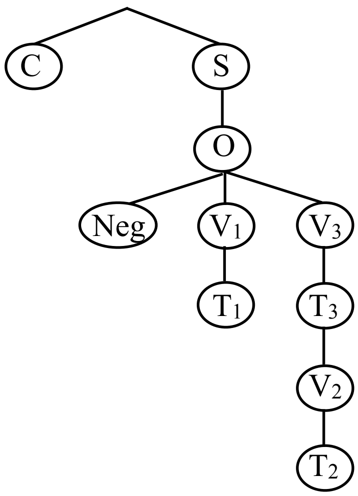
(65) a. ...dass er das Buch nicht hätte lesen sollen.
that he the book not had read should
'...that he shouldn't have read the book'
b. ...dass du mindestens bestätigen können musst, dass Fritz schwimmen kann
that you at.least certify can must that Fritz swim can
'...that you must at least be able to certify that Fritz can swim'
(Standard German; Abels 2016: 188 ex. 9)

In (65), the nominal direct object and complement clause appear on opposite sides of the verb. Together with the surface order of verbal elements, this seems to prevent any account of a consistent hierarchy that both preserves standard assumptions about constituency, and follows Abels' assumptions about movement options in the clause. Notably, the main verb is not adjacent to its complement in either order. However, the present account makes different assumptions, and readily generates the examples Abels discusses.[26] I illustrate with (66).[27]

(66) ...dass er das Buch nicht hätte lesen sollen
C S O Neg V_1 T_1 V_3 T_3 V_2 T_2
1 10 9 2 4 3 8 7 6 5

[26] The position of the object left of adverbials and negation is taken to reflect Scrambling, with interpretive consequences. As such, it is presumably not an information-neutral movement, and falls outside the generative system here. Note that such movement produces Obj-Adv-V order, an instance of *Arg-Adj-H, predicted to be unattested as a neutral order in light of hierarchy (45), H < Arg < Adj.

[27] Negation, like agreement, is thought not to adhere to a single, universal cartographic position (see, *e.g*., Cinque 1999: 126). For the suggestion that a unified high position for negation may be possible, as assumed for this example, see below.



In a similar vein, Abels (2016) argues from the examples in (67) that adverbs and auxiliaries cannot be placed into a consistent hierarchy obeying the generalization of Universal 20 he pursues. However, we find unremarkable analyses of these orders in the present system (the key difference again is how head-complement relations are represented). I show in (68) the categories indexed in hierarchical order; the reader may verify that the neutral orders (67b, c, d) are all 231-free, generated by the present architecture. Example (67a) contains a 231-like order, and falls outside the generative capacity of the current system. But this is an expected result; (67a) is an example of VP-topicalization, a non-information-neutral effect.

(67) a. schön singen hat er früher können — *Standard German*
beautifully sing has he formerly can — *VP-topicalization*
'He formerly used to be able to sing beautifully.'
b. ...dass er früener hat chöne schöön singe — *Zurich German*
that he formerly has can beautifully sing
c. ...dass er früher schön hat singen können — *Standard German*
that he formerly beautifully has sing can
d. ...dat hij vroeger prachtig heeft kunnen zingen — *Standard Dutch*
that he formerly beautifully has can sing

(68) a. schön singen hat er früher können — *Standard German*
A_3 V_3 T_3 V_1 T_1 S A_2 V_2 T_2 — *VP-topicalization*
7 6 5 2 1 9 8 4 3
b. dass er früener hat chöne schöön singe — *Zurich German*
C S A_2 V_1 T_1 V_2 T_2 A_3 V_3 T_3

1 10 9 3 2 5 4 8 7 6

c. dass er früher schön hat singen können *Standard German*

C S A_2 A_3 V_1 T_1 V_3 T_3 V_2 T_2

1 10 9 8 3 2 7 6 5 4

d. dat hij vroeger prachtig heeft kunnen zingen *Standard Dutch*

C S A_2 A_3 V_1 T_1 V_2 T_2 V_3 T_3

1 10 9 8 3 2 5 4 7 6

We are up to ten elements in the surface order, admitting very many possibilities, including thousands of orders that do not contain 231-like subsequences. But note that the number of 231-avoiding permutations forms a shrinking proportion of all orders, as the number of elements increases. That is, with just two elements, both possible orders are allowed; with three elements, 5 of 6 possible orders are generated, and we find 14 of 24 orders with 4 elements. With ten elements, there are 16,796 231-avoiding surface orders, among 10! = 3,628,800 possible orders. Put another way, the chance that a randomly selected order of ten elements is 231-avoiding, and thus generated by this system, is less than 0.5%. The chance that three randomly chosen[28] orders all fall into the generated orders is about 1 in 10 million. This should provide some confidence that we are describing the hierarchy accurately for these examples, provided the rest of the framework is on the right track.

6.2 More clausal orders, and a universal position for negation

We return to the issue of the placement of negation. Tentatively, let us hypothesize that negation is generated "high", in the cartographic zone corresponding to Laka's (1990) Pol head: C < **Neg** < T < V < O < S. I consider typical word orders with negation in a selection of languages below, which appear to be consistent with this hypothesis. For the overt elements appearing in these examples, the relevant hierarchy is Neg < T < Aux < -Fx < V < O < IO < S.

(69) *English*

a. She did not write it.

S T Neg V O

5 2 1 3 4

b. She has not written it.

S Aux T Neg V -Fx O

---

[28] Of course, examples (67b,c,d) are hardly randomly chosen orders; they are closely related languages with broadly similar orders, and we are applying to them a hierarchy designed to capture clausal ordering facts.

7 3 2 1 5 4 6

c. She will not have written it.

S T Neg Aux V -Fx O

7 2 1 3 5 4 6

(70) *Spanish*

a. Juana estaba cantándolo.

Juana be.past.3sg singing+3sg

'Juana was singing it.'

S Aux_1 T V $-Fx_1$ O

6 2 1 4 3 5

b. Juana lo estaba cantando. (de Andrade 2017)

S O Aux_1 T V $-Fx_1$

6 5 2 1 4 3

Note that this gives us the beginnings of an account of Romance clitic-climbing constructions. So long as the relative order of auxiliaries and verbs is fixed such that selecting forms precede the verbal head of their selected complements on the surface (*i.e.*, the typical Spanish order), and ignoring other elements, an object clitic could, in principle, be placed immediately before any verb stem or after any affix in the sequence, but not between stem and affix: (*lo*) *esta* (**lo*) *-ba* (*lo*) *cant* (**lo*) *-ando* (*lo*). Of course, this merely allows the possibility of this kind of ordering; describing the availability of this phenomenon for different verb classes and in different languages is not pursued here.

6.2.2 Selected word order issues in Irish

Turning from the SVO languages considered above to a VSO language, (71) is from Irish.

(71) *Irish*

Ní-or chuir sé isteach ar phost ar bith.

Neg.past put.past he in on job any

'He didn't apply for any job.' (McCloskey 2017)

Neg T V S P PP

1 2 3 6 4 5

It is not entirely straightforward to locate Tense in Irish. On the one hand, there are a set of verbal suffixes that differ in different tenses/moods. But at least historically, past tense was expressed with a preverbal particle *do*, triggering lenition. For present purposes, either location of T, before or after the verb, produces a 231-free surface order.

With clausal embedding, Irish again avoids 231 permutations of the hierarchical order. In (72), I locate T before the verb, although placement after the verb is also 231-avoiding.

(72) *Irish*
Deir siad gur ghoid na síogaí í.
say they C-[PAST] stole the fairies her
'They say the fairies stole her away.' (McCloskey 2001: 67, ex. 1)
T_1 V_1 S_1 C_2 T_2 V_2 S_2 O_2
1 2 8 3 4 5 7 6

Interestingly, we can analyze negative past complementizers like Irish *nár* as transparently reflecting exactly the syntax suggested by their morphology, Neg-C-T. This is 231-avoiding, even under embedding, as illustrated in (73).

(73) *Irish*
Creidim nár chuir sí isteach ar an phost.
I-believe C[Neg]-PAST put she in on the job
'I believe that she didn't apply for the job.' (*ibid*.: 75, ex. 26b)
T_1 V_1 (S_1) Neg_2 C_2 T_2 V_2 S_2 P PP
1 2 (10) 4 3 5 6 9 7 8

Other strange quirks of Irish word order fall into place under this account as well. This includes the curious placement of dative subjects to the left of what appears to be a negative non-finite complementizer *gan*.[29] The following example is a wh-question; setting aside the wh-moved adverbial *conas* (as outside the generative system developed here), the position of the dative subject left of the negative complementizer is unproblematic.

(74) Conas d' aonaránach gan a bheith ag braistint aonarach?
how to solitary-person NEG be[−FIN] feel[PROG] solitary
'How could a solitary person not feel solitary? (McCloskey 2001: 30)
~~Adv~~ P Subj C Neg T Aux Prog V Adj
~~8~~ 9 1 2 3 4 5 6 7

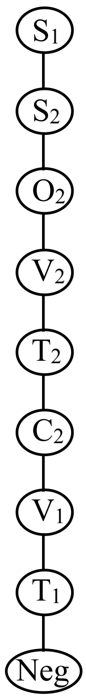
At least for this small selection of data, a single hierarchy, including a single position for clausal negation appears viable after all. Next, I illustrate the account for some sentences in VOS Malagasy and SOV Japanese.

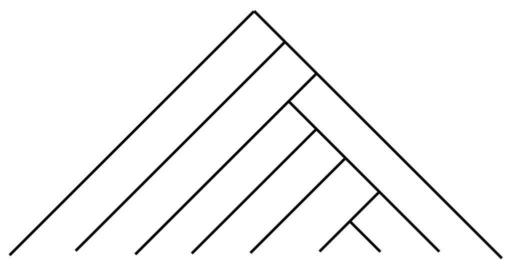
[29] Though see McCloskey (2017), and references cited there, on reasons to doubt that Irish *gan* is a C.

6.2.3 More clause orders: Japanese and Malagasy

Japanese presents an interesting case: standard word order with clausal embedding (75) corresponds to a vertical arrangement of nodes in this framework.

(75) *Japanese*
Bill-wa [John-ga sore-o katta ka] siranai.
'Bill does not know whether John bought it or not.' (Fukui 1995:115, ex.33c)
S_1 S_2 O_2 V_2 T_2 C_2 V_1 T_1 Neg
9 8 7 6 5 4 3 2 1



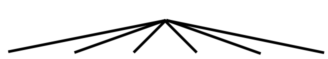


Neg < T_1 < V_1 < C_2 < T_2 < V_2 < O_2 < S_2 < S_1

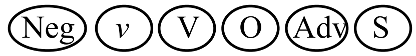
Malagasy, by contrast, presents a surface order corresponding to a nearly flat tree.

(76) *Malagasy*
a. Tsy manasa lamba intsony Rakoto 1 2 3 4 5 6
Neg wash clothes anymore Rakoto
b. Tsy manasa intsony ny lamba Rakoto 1 2 3 5 4 6
Neg wash anymore det. clothes Rakoto
'Rakoto doesn't wash clothes anymore.' (Rackowski 1998: 14, ex. 9a,b)

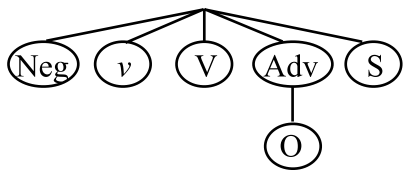
(77) Tsy manasa lamba intsony Rakoto
Neg *v* V O Adv S



1 2 3 4 5 6



(78) Tsy manasa intsony ny lamba Rakoto
Neg *v* V Adv O S
1 2 3 5 4 6



6.2.4 Pollock (1989) revisited

Next, let us take up the alternations considered in Pollock (1989), concentrating on the position of negation in English and French. French is a particularly interesting case, in that two distinct morphemes *ne* and *pas* express negation, in different surface structure positions. Holding the rest of our assumptions about hierarchy fixed, let us see if we can systematically narrow down the possibilities.

I assume that French *ne* is the Pol-related negation head Neg^{o}, located between C and T in the clausal hierarchy, as assumed for other languages above. In this context, the position of *pas* is treated as an unknown, coded for now with variable index *k*. I assume that English *-n't* has a fixed high position in Neg^{o}, and, provisionally, that English *not* is in the same position.

A first question is whether it is necessary to hierarchically distinguish *ne* and *pas*. That is, we might consider this a case of double exponence: perhaps *ne* and *pas* are adjacent in the hierarchy, and *ne...pas* has a split analysis, parallel to Affix-Hopping in the English verbal system. Or perhaps *pas* is an adverbial, in which case we expect its hierarchical position to be between object and subject.

As a first pass, we illustrate that the system here provides analyses of broad word order properties in English and French, including variation in the surface position of adverbials. I first provide an indexed hierarchy for the examples.

(79) *Hierarchy for French/English* (80)
[-ait pens [-er rencontr [son patron] [a la Sorbonne] demain] Jean]
-ed think to meet his boss at the Sorbonne tomorrow Jean
T_1 V_1 T_2 V_2 D_a N_a P_b D_b N_b Adv_2 S_1
1 2 3 4 5 6 7 8 9 10 11
(77c, e below have T_2 *would* = 3, S_2 *he* = 11, S_1 *Jean* = 12).

(80) a. Jean pensait rencontrer son patron a la Sorbonne **demain**.
11 2 1 4 3 5 6 7 8 9 10
b. Jean thought to meet his boss at the Sorbonne tomorrow
11 2 1 3 4 5 6 7 8 9 10
c. 'Jean thought he would meet his boss at the Sorbonne tomorrow.'
12 2 1 11 3 4 5 6 7 8 9 10
d. Jean pensait rencontrer **demain** son patron a la Sorbonne.
11 2 1 4 3 10 5 6 7 8 9
e. 'Jean thought he would meet his boss **tomorrow** at the Sorbonne.'
12 2 1 11 3 4 5 6 10 7 8 9
(Pollock 1989: 380, *ex*. 33, 34, incl. English translations)

The variations in adverbial positioning illustrated above are unproblematic; the orders in both languages are 231-avoiding. We turn now to examples that constrain the location of French negation. The basic clausal hierarchy in (81) covers these examples.

(81) Neg < T < Aux < -Fx < V < O < Adv < S
1 2 3 4 5 6 7 8

(82) Pierre ne mange pas.
Pierre ne eats not (*ibid*., 393)
S Ng V T ?
8 1 5 2 k

(83) Ne pas manger... (*ibid*., 394)
ne not to eat
1 k 5 2

(84) Jean (n') aime pas Marie. (*ibid*., 367)
S Ng V T ? O
8 1 5 2 k 6

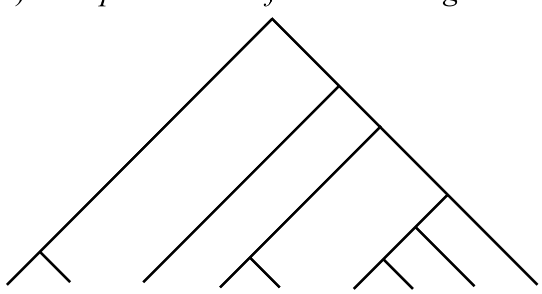
From (84), we can conclude that *pas* is not a head above Neg, else the *ne*...V...*pas* subsequence would be a 231 contour. In (83), we see *k* is not between index 2 (T) and index 5 (V), else a 231 sequence would arise in *pas manger*. We can also conclude that *pas* is not a phrasal modifier or operator hierarchically right of the subject: *k > S, else the subsequence S...*pas*...O would form a 231 permutation.

(85) shows possible positions of *pas* in the hierarchy (*i.e*., the index for *k*), with respect to the basic clausal hierarchical sequence (81). Here, every position before or after existing positions in the hierarchy is considered, and marked with * if ruled out.

(85) [*] Neg [ ] T [*] Aux [*] Fx [*] V [#] O [ ] Adv [ ] S [*]
1 ↑ 2 3 4 5 6 7 8
(i) (ii)

We can rule out the position between V and O as well, on semantic grounds assumed already (the complement is the first modifier). Summarizing, the options are: (i) *pas* is a head between Neg and T, with index ~1; or (ii) *pas* is in the zone corresponding to phrasal modifiers, consistent with analysis as an adverbial, between O and S, with index ~7. No other hierarchical locations are consistent with the data. Moreover, as far as I can tell the data in Pollock (1989) do not further decide the issue. It is interesting that both options supported in this framework correspond to classical analyses, either in terms of Laka's high Pol° head, or a low adverbial position often taken to demarcate the boundary of VP. Tentatively, I adopt an Affix-Hopping analysis whereby French *pas* combines (optionally, in modern varieties) with *ne* to form Neg° (option (i)). This is motivated by concerns of acquisition: because expressions of negation in other languages must be in Neg°, we should assume that all expressions of clausal negation are in this position, unless the data in some language clearly calls for another analysis. Insofar as *ne* and *pas* admit an analysis in terms of the single Neg° position (split, in this case, between two hierarchically adjacent morphemes), we should assume this analysis is correct.

(86) *Composition tree for French negation ne...pas*



ne < *pas* < T < Aux < Fx < V < O < Adv < S
1 2 3 4 5 6 7 8 9

6.2.5 On the placement of clausal negation across languages

Roberts (2019) surveys the placement of clausal negation across languages, concluding that three different positions for negation are found. I reconsider that conclusion here, arguing that the data appear largely compatible with the single Polº position pursued here. (87) summarizes the typological picture.[30]

(87) *Orders of Neg, Verb, Object, Subject*

a. S-O-V-Neg	49 languages
b. S-O-Neg-V	65 languages (optional doubled 108 more)
c. S-Neg-O-V	15 languages (optional/doubled 25 more)
d. Neg-S-O-V	11 languages (optional, doubling 27 more)
e. O-S-V-Neg	1 language (Kxoe)
f. *O-S-Neg-V	Unattested
g. *O-Neg-S-V	Unattested
h. *Neg-O-S-V	Unattested
i. V-S-O-Neg	1 language (Podoko), optional in 6 more.
j. V-S-Neg-O	1 language (Colloquial Welsh)
k. *V-Neg-S-O	Unattested
l. Neg-V-S-O	58 languages (doubled 17 more)
m. S-V-O-Neg	81 languages (optional/doubled 42 more)[31]
n. S-V-Neg-O	2 languages (incl. Colloquial French)
o. S-Neg-V-O	112 languages
p. Neg-S-V-O	10 languages (optional/doubled 24 more)
q. *O-V-S-Neg	Unattested
r. O-V-Neg-S	1 language (Selknam, maybe Hixkaryana?)
s. O-Neg-V-S	3 lgs (Mangarrayi, Tuvaluan, Ungarinju)
t. *Neg-O-V-S	Unattested
u. *V-O-S-Neg	Unattested
v. *V-O-Neg-S	Unattested
w. *V-Neg-O-S	Unattested
x. Neg-V-O-S	Attested

In the present account, there are four logical options for the placement of Neg within the fixed hierarchical order V < O < S. Each choice makes different predictions about which 10 orders among the 24 logically possible orders of all four elements should be ruled out. These predictions are shown in (88). At the top of each column is one

---

[30] Enormous thanks are due to Ian Roberts (p.c.) for his patient help and detailed discussion and clarification of the material in this section.

[31] Of this order, Roberts (2019: 594) notes, "The phenomenon is highly areal: the vast majority of these languages are found in a belt running from West Africa across Central Africa to Ethiopia, with the rest in South East Asia, Vanuatu, and Papua New Guinea (with one language, Mehri, in Oman/Yemen and one in North-Eastern Australia)."

possible relative hierarchy of Negation with respect to V, O, S, followed by a list of the 10 orders predicted to be impossible under that hierarchy.

(88) *Predictions of possible hierarchical positions of Neg*

	A	B	C	D
	Neg<V<O<S	V<O<Neg<S	V<Neg<O<S	V<O<S<Neg
	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
*2341	*VOSN x	*ONSV x	*NOSV x	*OSNV x
*2314	*VONS x	*ONVS	*NOVS x	*OSVN 1
*2413	*VSNO 1	*OSVN 1	*NSVO-many	*ONVS
*3421	*OSVN 1	*NSOV-many	*OSNV x	*SNOV-many
*3412	*OSNV x	*NSVO-many	*OSVN x	*SNVO-many
*1342	*NOSV x	*VNSO x	*VOSN x	*VSNO 1
*2431	*VSON 1	*OSNV x	*NSOV-many	*ONSV x
*3241	*OVSN x	*NOSV x	*ONSV x	*SONV-many
*3142	*ONSV x	*NVSO-many	*OVSN x	*SVNO-many
*4231	*SVON many	*SONV-many	*SNOV-many	*NOSV x

Column A in (88) is clearly the best match to the data among these choices. In each of B, C, and D, three or four very common orders (bold) are ruled out. In A, by contrast, for nine of the ten 231-containing orders, either no language or at most 1 language (Colloquial Welsh in the case of V-S-Neg-O, really (Neg-)V-S-Neg-O (*ni*) ... *ddim*, Podoko for V-S-Neg-O, and Kxoe for O-S-V-Neg) has the relevant word order.

However, the prediction that S-V-O-Neg should be ungrammatical is radically false; 81 languages in Dryer (2013) have this order. Is there some reason why this order should not obey our permutation-avoidance expectations? In German, this order reflects Scrambling or Object Shift, with semantic consequences. In present terms, the object is ignored with respect to checking for 231 subsequences (thereby obviating further problems having to do with *O-Adv-V order). Suppressing the object, we see that the S-V sequence is strictly descending, as required before a higher-index item. This predicts that T must follow V in this configuration: *S-T-V-Neg is a 231-containing order, while S-V-T-Neg is 231-free. Whether this sort of analysis can extend to other examples of SVON order in this large group of languages, I do not know. See Biberauer (2017) for relevant discussion of V-O-Neg order.

While these examples are suggestive, and the predictions about the typology of clausal negation appear broadly plausible, much more work would of course be required to determine if this framework might generalize to an adequate description of neutral word order across languages.

6.3 Deriving the Head Movement Constraint

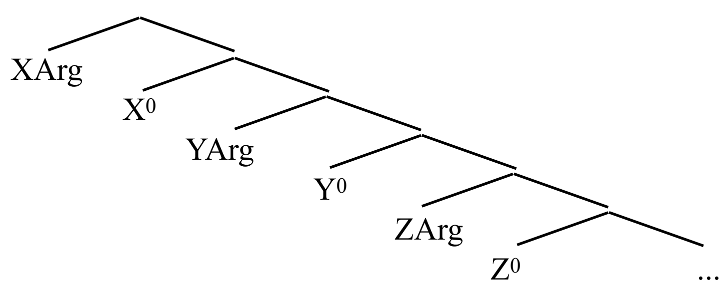
This framework also potentially explains Travis' (1984) Head Movement Constraint (HMC). This is surprising at first glance, as movement violating the HMC does not produce an impossible order of the heads themselves. Instead, HMC-violating movement "skipping" an intervening head produces 312 order among the heads (90d), readily generated by this system. Suppose the heads are hierarchically ordered as in (89).

(89) X < Y < Z
     1 < 2 < 3

(90) *Permutations of head order*
- a. X-Y-Z   123   base order
- b. Z-Y-X   321   multiple movement obeying HMC
- c. X-Z-Y   132   partial HMC-obeying movement
- d. Z-X-Y   312   HMC-violating order: not obviously bad
- e. Y-Z-X   *231   FOFC violation
- f. Y-X-Z   213   HMC-obeying movement starting high.

But now let us suppose phrasal satellites (arguments or adjuncts) are interspersed in the order (where XArg is a phrasal satellite associated with head X). A traditional representation of the base structure would be as below:

(91)



This corresponds to the abstract hierarchy below.

(92) X^0 < Y^0 < Z^0 < ZArg < YArg < XArg
     1   2   3   4   5   6

Now consider orders with head movement, as (90) above, but retaining the interspersed arguments.

(93) *Permutations of head order with interspersed arguments*
- a. XArg X^0 YArg Y^0 ZArg Z^0 6 1 5 2 4 3

b. XArg Z^0 Y^0 X^0 YArg ZArg	6 3 2 1 5 4	
c. XArg X^0 YArg Z^0 Y^0 ZArg	6 1 5 3 2 4	
d. *XArg **Z^0** X^0 **YArg** **Y^0** ZArg	*6 **3** 1 **5** **2** 4	HMC violation
e. (already ruled out)	*2 3 1	FOFC violation
f. XArg Y^0 X^0 YArg ZArg Z^0	6 2 1 5 4 3	

The HMC-violating order (93d) contains a 231 permutation. As we can see, what causes problems here is not the improper head movement itself, producing the subsequence Z^0 X^0 ... Y^0, but rather the subsequence Z^0 ... YArg ... Y^0.

Putting this in concrete terms, much work on head movement has focused on verb movement (V-to-T and T-to-C). Translating into these terms, suppose X = C, Y = T, Z = V. The subject is an argument of a head between T and V, such as 'little *v*' or voice (the modern understanding of Koopman & Sportiche 1991). In this system, what we expect to be ruled out is really the following order:

(94) *... V ... Subj ... *v*/voice/Aux/T

That is, the verb cannot precede a subject which precedes some head above V (in fact, we have seen this prediction already, in (51h) in section 4). As far as I know, that gives the right facts for V-to-T and T-to-C movement captured by the HMC. Understanding the HMC as actually reflecting the condition in (94) also allows us to account for "long" head movement in Breton, a much-discussed violation of the HMC.

(95) *Breton*

Lennet	en deus	Anna	al levr
read.pprt	has	Anna	the book

'Anna has read the book' (Roberts 2010: 194)

I take the hierarchy for this example to be (96); (97) shows the hierarchical indexing of the surface order, which is indeed 231-free, and generable in this system.

(96) T < Aux < -Fx < V < O < S

T	Aux	-Fx	V	O	S
1	2	3	4	5	6

(97)

Lennet		en deus		Anna	al levr
V	-Fx	Aux	T[32]	S	O
4	3	2	1	6	5

[32] Ordering T before the Aux, as suggested for Irish, produces a 231-free order as well.

Certain Slavic languages also allow fronting of a bare participle, as in Bulgarian.

(98) a. Sˇte saˇm **procˇel** knigata. 1 2 4 3 5
will have read.PTCP the.book
'I will have read the book.'
b. **Procˇel** sˇte saˇm knigata. 4 3 1 2 5
read.PTCP will have the.book
'I will have read the book.' (Harizanov & Gribanova 2018:17)

(99) a. Bihte bili **arestuvani** ot policijata. 1 2 4 3 5
would been arrest.PTCP by the.police
'You would be arrested by the police.'
b. **Arestuvani** bihte bili ot policijata. 4 3 1 2 5
arrest.PTCP would been by the.police
'You would be arrested by the police.' (Embick & Izvorski 1997:231)

Interestingly, in all these cases information-neutral long head movement obeys the condition in (94).[33] That is, either the subject is placed after the entire verbal complex, as in Breton, or the subject is null, as in the Bulgarian examples (98); (99) shows a prepositionally-marked thematic subject, which again does not intervene between participle and higher heads right of the participle.

6.4 A loose end: final particles

Section 6.2 sketched an analysis of Malagasy clause structure. Space prevents a consideration of many important properties in this language, especially concerning the rightward topic/trigger position. Picking just one further issue that arises for this language, consider the placement of the interrogative particle *ve*.

(100) a. Tsy manasa lamba mihitsy **ve** Rakoto? 1 2 3 4 5 **?** 6
Neg wash clothes at-all Q Rakoto
b. Tsy manasa mihitsy ny lamba **ve** Rakoto? 1 2 3 5 4 **?** 6
Neg wash at-all det. clothes Q Rakoto
'Does Rakoto not wash clothes at all?'
(Rackowski & Travis 2000: 120, ex. 6a,c)

It is easy to see that if the *ve* element is treated as a head high in traditional tree structures, these orders will be 231-containing. That is, *ve* is preceded by a series of

[33] Ian Roberts (p.c.) observes that this appears to be true quite generally.

ascending-index items, and would have a lower index than any of them. However, pursuing an observation of Medeiros (2018), the hierarchy format described here is closely related to postfix notation. In such notation, operators follow their operands; if we identify the meaning of *ve* as operator-like, we might guess that it appears at the end rather than the beginning of the hierarchical order.

(101) Neg < T < V < O < Adv < S < **Q**

With this conjectured hierarchy, the Malagasy examples are unproblematic.[34] This move may, in turn, shed light on one well-known class of exceptions to FOFC (see especially Biberauer 2017), involving VO order with clause-final particles, as illustrated in (102) for Mandarin.

(102) *Mandarin*
Hongjian xihuan zhe ben shu **ma**?
Hongjian like this CL book Q
'Does Hongjian like this book?'
(Li 2006:13, cited in Biberauer *et al* 2014: 199)

While the particles in question appear to contribute interrogative force, treating them as C heads leads to a violation of FOFC: *V-O-C is predicted to be impossible. However, suppose we extend the treatment of Malagasy *ve* to these elements in Mandarin as well. Then the relative hierarchy is V < O < Q, and the surface order in question is a 123 order; this order is allowed. I leave open the question of whether other instances of V-O-Q order might yield to such an analysis.

6.5 Cyclic effects

Finally, I turn to another well-known apparent counterexample to FOFC: OV languages such as German allow determiner-noun word order (103).

(103) *German*
Johann hat [$_{\text{VP}}$ [$_{\text{DP}}$ einen Mann] gesehen].
Johann has a man seen
'Johann has seen a man.' (Biberauer *et al* 2014)

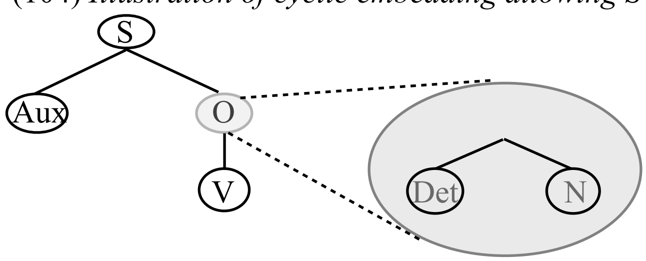
[34] On the other hand, we cannot treat Japanese *ka* in (75) the same way, as it would create a 231-containing order; *ka* must instead be a "high" C-like head.

It is standardly assumed that determiners are heads, so the partial hierarchy should be Det < N. Together with the verb-object hierarchy V < O, we get V < Det < N; *Det-N-V word order should be impossible, violating FOFC. I adopt the usual solution, supposing that nominal and verbal cycles are disjoint for the purposes of this condition.

One intriguing aspect of the present proposal is that a notion of *phase* is baked into the architecture. The tree traversal algorithms, which take the place of Transfer in a standard Minimalist model, cannot apply at each step of incremental construction of the bare trees here. Instead, they must apply to whole trees, or subtrees, mapping a hierarchy onto them and reading off linear order. If this process is recursive (trees may embed references to already-transferred subtrees), further ordering predictions follow.

We can sketch how this would work for (103). Suppose a single node within the verbal cycle can contain a pointer to a separately-computed nominal cycle. The nominal subtree is generated, hierarchized, and linearized by itself, internally obeying the permutation-avoidance condition. But the internal structure of the nominal is unavailable, and irrelevant, within the embedding verbal cycle; its already frozen word order is "plugged in" at the corresponding node. Det-N order itself is 231-free, as is S-Aux-O-V.

(104) *Illustration of cyclic embedding allowing S-Aux-[Det-N]-V order*



In this case, using the notion of cycles allows an order that otherwise falls outside this system's generative capacity (namely, the apparent 231 order [Det N] V). However, in the general case cyclic effects of this sort will tend to reduce ordering possibilities (because the elements of the embedded domain must occur in a continuous sequence free of any elements of the embedding domain).[35] Note that we do not, as a rule, need to treat every (notional) clause as a single cycle; indeed, doing so would prevent an account of the cross-serial subject-verb dependency examples discussed here. Treating clausal complementation configurations as a single cycle also aligns with the literature on FOFC, where the relevant effects are observed to hold across clause boundaries. I leave a fuller discussion of cyclicity in this architecture to future research.

[35] See Cinque (2020) for discussion of ordering facts within the nominal domain that appear to support this kind of analysis.

7 Conclusion

Implementing External Merge as an operation building bare *n*-ary branching trees, lexicalized and linearized by traversal algorithms, we derive and unify Universal 20 and FOFC permutation-avoidance patterns, and find simple analyses of cross-serial dependency constructions. Further applications include suggestions that a single hierarchical position for negation is possible across languages, an account of Romance clitic climbing constructions, and a version of the Head Movement Constraint that allows attested long head movement configurations. Strikingly, these effects follow from the structure-building system itself and single hierarchical ordering condition H < X, without additional constraints or mechanisms.

In this view, no additional operations create displacement in neutral orders; the typologically possible orders are all base-generated. This unification of movement with structure-building goes further than the view of movement as Internal Merge, where Internal Merge involves extra operations beyond the constant number of External Merges required to join the lexical items involved. Here, the same number of External Merge operations (bracket pairs) derives all neutral orders: exactly *n* such for *n* items.

That said, we still need actual movement in the present framework: effects like *wh*-movement and topic and focus displacement produce other orders.[36] However, the residue of actual movements under this account is the set of non-information-neutral transformations. This result aligns with Chomsky's suggestion that the duality of semantics is tied to the distinction between External Merge and Internal Merge: EM builds the base thematic structure, and IM induces discourse-information effects.

Raising our sights a bit, we have covered a good deal of ground with some rather minimal technical devices. However, there is reason for caution as well. The system presented here handles some or all of the work previously ascribed to head movement, A-movement (but not Scrambling), and the more nebulous phrasal movements deriving different neutral orders in recent accounts. The well worked out implementations of these phenomena in the standard theory underpin a large framework of further conclusions. Insofar as the present account can replace those earlier devices, conclusions which followed from details of the earlier formulation do not necessarily remain in force within this framework.

This is particularly troublesome with respect to cartographic issues. Considerations of word order have been a central component of cartographic arguments. While the vision of a single, unified hierarchy for all languages remains here, the highly articulated structure of that hierarchy, elaborated on the basis of arguments from traditional conceptions of movement and structure-building, must now be re-evaluated. I have suggested that clausal negation might admit a single "high" position; whether or not

[36] Thanks to David Adger for discussion on this point.

this proves ultimately viable, it illustrates how the generally more liberal constraints on word order imposed by this system might allow a simplification of the relevant cartography. This is one reason I have kept to a rather small inventory of categories throughout.

On the other hand, the present system provides a simple tool to determine if a single cartography is possible (check whether different neutral orders are 231-free, given a postulated hierarchy). As noted in the text, for even moderately long strings the permutation-avoidance pattern described here is quite unlikely to be found by chance. This fact also has important implications for acquisition, left unexplored here.

The theory developed here is a fragment. I have not demonstrated how this system generalizes to a full theory of word order, nor spelled out how real movement works, nor accounted for core grammatical phenomena such as coordination, ellipsis, binding, agreement, and so on. These are important topics, and much more work will be required to determine if they might find satisfying accounts within this framework.

References:

Abels, K., & Neeleman, A. (2012). Linear asymmetries and the LCA. *Syntax* 15(1), 25-74.

Abels, K. (2016). The fundamental left–right asymmetry in the Germanic verb cluster. *The Journal of Comparative Germanic Linguistics* 19(3), 179-220.

Biberauer, T. (2017). The Final-Over-Final Condition and particles. In Sheehan, M., Biberauer, T., Roberts, I., & Holmberg, A. (eds.), *The Final-over-Final Condition: A syntactic universal* (Vol 76), 187-296. Cambridge, MA: MIT Press.

Biberauer, T., Holmberg. A., and Roberts, I. (2014). A syntactic universal and its consequences. *Linguistic Inquiry* 45, 169-225.

Bresnan, J., Kaplan, R.M., Peters, S., & Zaenen, A. (1982). Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13(4), 613-635.

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (2007). Approaching UG from Below. In Hans-Martin Gärtner, H.-M., & Sauerland, U. (eds.). *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics.* Berlin: Mouton de Gruyter.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York: Oxford University Press.

Cinque, G. (2005). Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36(3), 315-332.

Cinque, G. (2020). Word order variation: towards a restrictive theory. Ms., University of Venice.

Dryer, M. (2018). On the order of demonstrative, numeral, adjective, and noun. *Language* 94(4), 798-833.

Feil, T., Hutson, K., & Kretchmar, R.M. (2005). Tree traversals and permutations. *Congressus Numerantium* 172, 201-221.

Fukui, N. (1995). *Theory of Projection in Syntax*. Stanford, CA: CSLI Publications.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J., (ed.), *Universals of language*, 73-113. Cambridge, MA: MIT Press.

Haspelmath, M. (2007). In Shopen, T. (ed.), *Language typology and syntactic description Vol. II: complex constructions*, 1-51. Cambridge, MA: Cambridge University Press

Holmberg, A. (2000). Deriving OV order in Finnish. in Svenonius, P., (ed.), *The Derivation of VO and OV*. Philadelphia: John Benjamins.

Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Laka, I. (1990). Negation in Syntax: on the nature of functional categories and projections. Ph.D. dissertation, MIT.

McCloskey, J. (2001). The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* (37): 67-100.

McCloskey, J. (2017). Ellipsis, polarity and the cartography of verb-initial orders in Irish. In Aboh, A., Haeberli, E., Puskás, G., & and Schönenberger, M., (eds.), *Elements of Comparative Syntax: Theory and Description*, 99–151. Berlin: De Gruyter.

Medeiros, D.P. (2018). ULTRA: Universal Grammar as a Universal Parser. *Frontiers In Psychology: Language Sciences 9*. https://doi.org/10.3389/fpsyg.2018.00155

Pearson, M. (2000). Two types of VO languages. In Svenonius, P. (ed.), *The derivation of VO and OV*. Philadelphia: John Benjamins.

Pollock, J-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20(3): 365-424.

Rackowski, A. (1998). Malagasy Adverbs. In Paul, I. (ed.), *The Structure of Malagasy: UCLA Occasional Working Papers in Linguistics 20*, 11-33.

Rackowski, A. & Travis, L. (2000). V-Initial Languages: X or XP-Movement and Adverbial Placement. In Carnie, A., & Guilfoyle, E., (eds.), *The Syntax of Verb-Initial Languages*, 117-141. New York: Oxford University Press.

Rizzi, L. (1997). The fine structure of the left periphery. In Haegeman, L. (ed.), *Elements of Grammar: Handbook of Generative Syntax*. Dordrecht: Kluwer.

Roberts, I. (2010). *Agreement and head movement: clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.

Roberts, I. (2017). The Final-over-Final Condition in DP: Universal 20 and the Nature of Demonstratives. In Sheehan, M., Biberauer, T., Roberts, I., & Holmberg, A. (eds.) *The Final-Over-Final Condition: A Syntactic Universal* (Vol 76), 151-186. Cambridge, MA: MIT Press.

Roberts, I. (2019). *Parameter Hierarchies and Universal Grammar*. New York: Oxford University Press.

Salzmann, M. (2019). On the limits of variation in Continental West-Germanic verb clusters. Evidence from VP-stranding, extraposition and displaced morphology for the existence of clusters with 213 order. *Journal of Comparative Germanic Linguistics* 22, 55-108.

Sheehan, M., Biberauer, T., Roberts, I., & Holmberg, A. (eds.). (2017). *The Final-Over-Final Condition: A Syntactic Universal* (Vol. 76). Cambridge, MA: MIT Press.

Sheehan, M. (2017). The final-over-final condition and adverbs. In Sheehan, M., Biberauer, T., Roberts, I., & Holmberg, A. (eds.) *The Final-over-Final Condition: A syntactic universal* (Vol 76), 97-120. Cambridge, MA: MIT Press.

Shieber, S.M. (1985). Evidence against the context-freeness of natural language. *Linguistics & Philosophy* 8(3), 333-344.

Stabler, E. (2004). Varieties of crossing dependencies: structure dependence and mild context sensitivity. *Cognitive Science* 28: 699-720

Steddy, S. & Samek-Lodovici, V. (2011). On the ungrammaticality of remnant movement in the derivation of Greenberg's universal 20. *Linguistic Inquiry* 42(3), 445-469.

Steedman, M. (2000). *The Syntactic Process*. (Vol. 24) Cambridge, MA: MIT press.

Travis, L. (1984). *Parameters and effects of word order variation*. Ph.D. Dissertation, MIT.

Wagner, M. (2005). *Prosody and recursion*. Ph.D. dissertation, MIT.

David P Medeiros
The University of Arizona
Department of Linguistics
medeiros@email.arizona.edu