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; 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 (Bresnan *et al* 1982) cross-serial subject-verb dependencies.

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

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 on an earlier draft: David Adger, Tom Bever, Noam Chomsky, and Guglielmo Cinque.

[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.6.

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

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 so-called 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?

2 Generating Universal 20

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

---

[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).

"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's (2005) analysis, 14 of the 24 logically 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 excluding head movement and remnant movement (*i.e*. IM in the noun phrase must affect the noun, possibly pied-piping dominating structure).[4] This hierarchy is given in (1).

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

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). This method imposes freely generated *n*-ary branching structure[7] 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:

(2) *Generative procedure over strings*

---

[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] Note that 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.

[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] 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 syntax-internal representations plausibly draws on capacity other animals possess.

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 a *legal bracketing*.

[x**]** [x [x**]** [x**]** **]**

d. Scan the string *left-to-right*, indexing right brackets in increasing order.[8]

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

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 correct surface structure bracketing. In this case, we derive (3):

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

The simple procedure in (2) generates all and only the attested noun phrase word orders, and their bracketed structure. Importantly, this does not simply repackage the Cinque-style EM and IM account. In particular, 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 External Merge and Internal Merge, 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 is no notion of steps of movement, and no need to explain them.[9] Conversely, unattested orders are not ruled out by constraints on movement, but simply 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,

---

[8] 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 conventions in computer science and mathematics, where the *PostPre* permutations (see Feil *et al* 2005) here are the stack-sortable words, avoiding *231 permutations. I adopt the more general convention, at risk of confusion.

[9] This also means that we lose any obvious syntax-internal explanation for the relative typological frequency of different orders (for example, the harmonic orders N-Adj-Num-Dem and Dem-Num-Adj-N are the most common), on which see Cinque (2005).

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 a vertical relation among nodes.[10] 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	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	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	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	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	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	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	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	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	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	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	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	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	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	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).

3 A closer look at the details

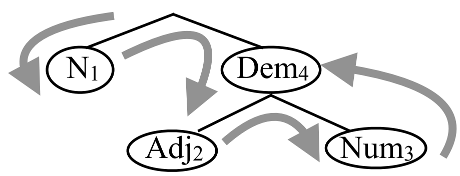
This section explores selected aspects of the account in greater depth. This includes describing the architecture in terms of trees and tree traversal algorithms, showing how the brackets for nominal orders correspond to Cinque's derivations, and examining how the account excludes unattested orders.

[10] 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.

3.1 The procedure in terms of tree traversals

The procedure (2) 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, (2) 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.

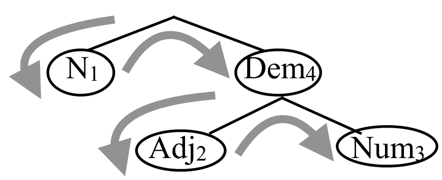
(4) *Postorder traversal*



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 (5); this path visits the nodes in surface order, N-Dem-Adj-Num.

(5) *Preorder traversal*



The notion of tree 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.

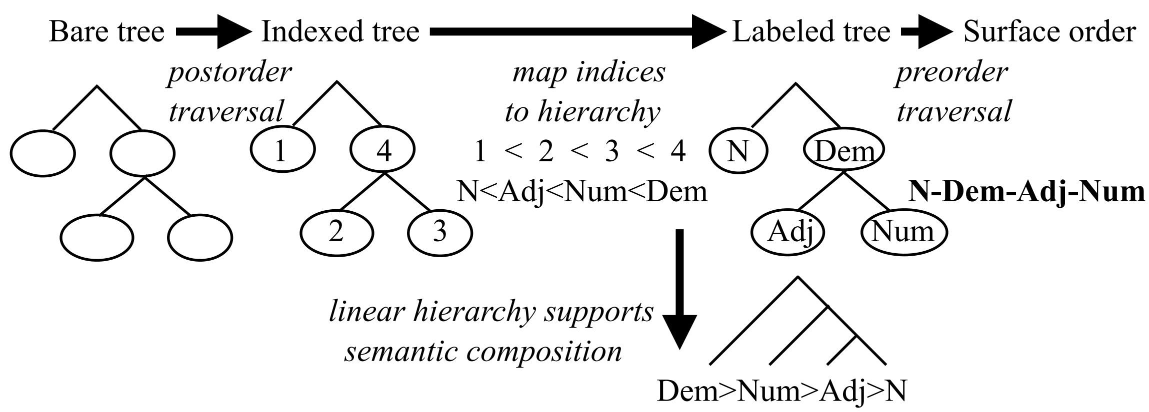


Figure 1: *Generating N-Dem-Adj-Num (1423) order*
Free Merge builds a bare *n*-ary tree. Postorder traversal indexes nodes. Indices map to hierarchical order (in this case, the hierarchy for Universal 20), yielding lexical labels on nodes. Preorder traversal of the labeled tree gives surface order; here,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

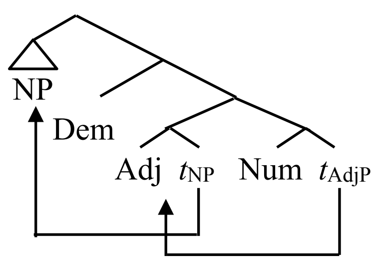
3.2 Correspondence with traditional bracketed representations

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* order. Translating to the Universal 20 hierarchy, the structure is (6).

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

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

(7)



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 (8) :

(8) [[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 (9):

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

As claimed, (9) is identical to expression (6) derived by the generative procedure in (2).

3.3 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. 7), 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.

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

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:

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

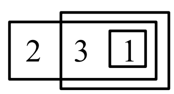
This is not a legal (indexing of a) bracketing; the boundaries of bracketings 1 and 2 cross. To csee this, we can think of brackets as denoting the edges of “boxes”. In generated orders, any pair of boxes may be in a containment relation, or be disjoint; they cannot overlap partially. Illustrating with 321 and 123 order and appropriate bracketing:

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

(13) $[\ 1\ ]_1\ [_2\ 2\ ]_2\ [_3\ 3\ ]_3$ $\boxed{1}\ \boxed{2}\ \boxed{3}$

But the unattested *231 order entails overlapping boxes:

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



Given the way procedure (2) works, unattested *231 order cannot be generated. Instead, the relevant bracketing must form a 321 order; bracketing determines hierarchy.

4 Generating the Final-Over-Final Condition

In this section, I show that the architecture developed thus far provides a ready explanation for another 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 (1) to be compatible with FOFC.

4.1 Background: The Final Over Final Condition

FOFC prohibits configuration (15):

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

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

(16) 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 (16b), 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.

(17) *V-O-Aux	$*[_{AuxP}\ [_{VP}$ V DP] Aux]
*V-O-C	$*[_{CP}\ [_{TP}$ T VP] C] or $*[_{CP}\ [_{TP}\ [_{VP}$ V O] T] C]
*C-TP-V	$*[_{VP}\ [_{CP}$ C TP] V]
*N-O-P	$*[_{PP}\ [_{DP/NP}$ D/N PP] P]
*Num-NP-D(em)	$*[_{D(em)P}\ [_{NumP}$ Num NP] D(em)][11]
*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 is a key insight in the unification pursued in the next subsection.

4.2 Refining the notion of hierarchical ordering

A crucial aspect of the account of Universal 20 in section 2 is how the nominal hierarchy is mapped to freely-generated trees. This includes not just choosing post-order

[11] 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.

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. In this regard, it is notable that fixed relations among syntactic elements seem to come in (at least[12]) two flavors: selection and adjunction, or head-complement (more generally, head-argument) and head-adjunct relations.

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 $H < A$. That corresponds to a traditional tree structure in which the head is more deeply embedded than its adjunct, a familiar analysis.

(18) $H < A$ *Head-adjunct hierarchical order*

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)[13], and arguably in verb clusters.[14] For the Universal 20 case, I repeat the following hierarchy:

(19) 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).

---

[12] 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).

[13] The relevant categories are indeed optional modifiers. But if one prefers Cinque's analysis treating them as 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 (19).

[14] But see Salzmann (2019) on attested 213 (for us, 231; see fn. 7) verb clusters. See Abels (2016) on Universal 20 effects in other domains and refinements.

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.[15]

(20) 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 H_a takes a complement headed by H_b, which in turn has complement C. The hierarchical order is then (20) α < β < γ, and the forbidden permutation is (21) *β-γ-α.

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

(22) * β-γ-α *Forbidden word order*

Order (22) is traditionally described as a head-final phrase (αP) dominating a head-initial phrase (βP), exactly the configuration ruled out by FOFC (15), repeated as (23).

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

For example, if head Aux has complement headed by V, with complement Obj, the hierarchy is Aux < V < Obj (24). We correctly exclude unattested *231 order *V-Obj-Aux (25).

(24) Aux < V < Obj

(25) *V-Obj-Aux

---

[15] 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's notion of "satellite"); both obey the condition H < X.

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-motivateded hierarchical sequence, Aux < V < Adv (26); unattested *V-Adv-Aux (27) is the forbidden *231 permutation.

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

(27) *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 (28). If there are multiple arguments of a head, the complement is closest to the head: H < Comp < Arg' (29).

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

(29) 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 Abels 2016). If there is an adverbial and an object, the hierarchy is V < O < Adv. (30) puts these together into a single ordering.

(30) 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 these examples; I omit it for simplicity. If complementizer C is present, I assume it takes TP as complement: C < T < V < O < S.

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

This all may seem stipulative. Note, first, that the hierarchy in (31) is identical to standard proposals, *modulo* the unusual resolution of head-complement structures. 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, an impressive range of familiar typological facts are captured.

We can consider elements belonging to a larger hierarchy three at a time; we should find, for each such triple, five attested orders and one forbidden order. Drawing on the order (31), with the understanding that the O position may be realized as clausal complement CP, we make the following predictions about impossible neutral orders.

(32) a. *O-S-V
b. *CP-S-V
c. *O-S-T
d. *V-O-T
e. *V-CP-T

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

(33) N < O

(34) P < NP

This yields hierarchical order (35), with forbidden permutation (36).

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

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

Setting aside the last two items in (17) (we have not treated polarity, and adopt Cinque's hierarchy for Universal 20 effects), we have reconstructed this list of canonical FOFC effects.

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.

Bresnan *et al* (1982) discuss unbounded crossing subject-verb dependencies in Dutch. Example (37), taken from Steedman (2000: 25), illustrates:

(37) ...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 facts in Swiss German, which also exhibits long-distance cross-serial case dependencies. Interestingly, the system already established can base-generate these orders.[16] I assume the example above contains the categories in (38), 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.

(38)	...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 (38) will be rendered as a single linear hierarchy, which we assemble incrementally. The work is done by the general clause ordering (31), together with the standard assumption that complement clauses occupy the canonical direct objct position; this allows us integrate clausal complementation with the clause order (31) above.[17]

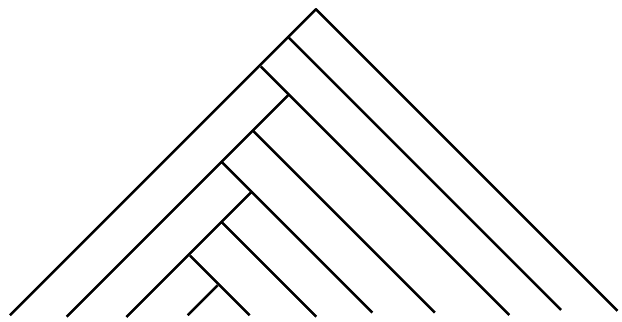
In what follows, I display hierarchical relations in descending order, as that yields a more familiar structure (*e.g.*, the clause hierarchy is S > O > V > T > C, a common word order). For single clausal embedding, $[_{CP} \ldots [_{CP} \underline{\ \ }\,]\ ]$, we have: $S_2 > \underline{S_1 > O_1 > V_1 > T_1} > V_2 > T_2 > C$. Replacing O_1 with an embedded clause, we derive (16), the hierarchical order for sentence (37) above. I show postorder indices aligned to the hierarchy, on which a superposed tree shows bottom-up semantic composition.

---

[16] 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.

[17] At least for these structures, we are implicitly developing a simple account of recursion by substitution. I leave fuller consideration of recursion in other domains to future work.

(39) *Integrated hierarchy for (37) with postorder index and composition tree*



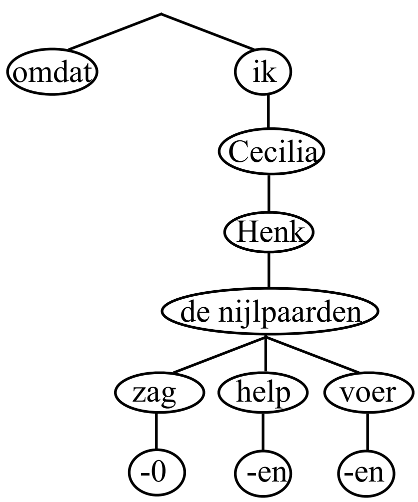
Composition tree

$S_3 > S_2 > S_1 > O_1 > V_1 > T_1 > V_2 > T_2 > V_3 > T_3 > C$ Hierarchical order

$11 > 10 > 9 > 8 > 7 > 6 > 5 > 4 > 3 > 2 > 1$ 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,[18] shown in (40).

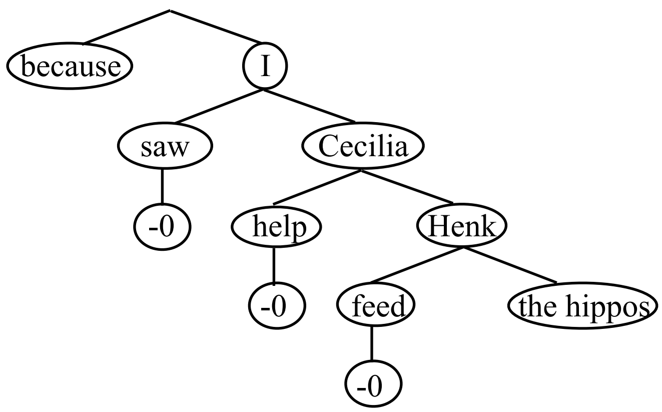
(40)	...omdat	ik	Cecilia	Henk	de nijlpaarden	zag-0		help-en		voer-en	
Category	C	S_3	S_2	S_1	O_1	V_3	T_3	V_2	T_2	V_1	T_1
Index	1	11	10	9	8	3	2	5	4	7	6
Brackets	( )	(	(	(	(	(	( ))	(	( ))	(	( ))))))



[18] An important question is whether these trees provide a basis for a successful theory of prosody (see also fn. 9). 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 in (41).

(41)	...because	I	saw	-0	Cecilia	help	-0	Henk	feed	-0	the hippos
Category	C	S_3	V_3	T_3	S_2	V_2	T_2	S_1	V_1	T_1	O_1
Index	1	11	3	2	10	5	4	9	7	6	8
Brackets	( )	(	(	( ))	(	(	( ))	(	(	( ))	( ))))



Finally, this architecture can generate the more limited pattern of bounded crossing dependencies that arises in English Affix-Hopping (Chomsky 1957), as seen in example (42).

(42) 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-Comp.[19] This generates (43), with composition structure (44).[20]

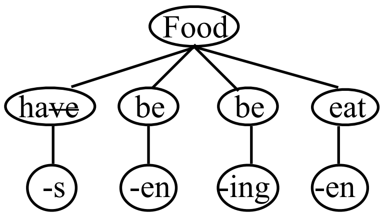
(43)	Food	ha~~ve~~	-s	be	-en	be	-ing	eat	-en
Cat.	Obj	Aux_1	T	Aux_2	$-Fx_1$	Aux_3	$-Fx_2$	V	$-Fx_3$

---

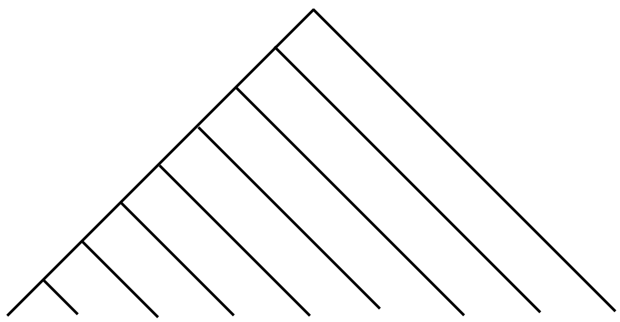
[19] One can read this as saying the affix (*e.g.*, *-ing*) is a head sandwiched between selecting auxiliary (*be*) and host verb. Or auxiliary and associated affix might "fuse" for interpretation, mirroring Chomsky's (1957) analysis with a single lexical item (*be*+*ing*).

[20] 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.

Index 9 2 1 4 3 6 5 8 7
Brackets ((( )) (( )) (( )) (( )))



(44) *Hierarchical order and composition tree for English Affix-Hopping (42)*



Food > eat > -en > be > -ing > be > -en > ha~~ve~~ > -s
Obj V -Fx_1 Aux_1 -Fx_2 Aux_2 -Fx_3 Aux_3 T
9 > 8 > 7 > 6 > 5 > 4 > 3 > 2 > 1

7.0 Conclusion

Implementing Merge as an operation building bare 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. 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.[21] However, the

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

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.

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.[22] 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., 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.

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.

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.

[22] One consequence, left unexplored here, is that a notion of *phase* is baked into this architecture. That is, 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 all at once onto them and reading off linear order. If this process is recursive (trees may embed references to already-transferred subtrees), further ordering predictions follow.

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. (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.

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.

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