Stack-Sorting Grammar

David P. Medeiros
University of Arizona

Abstract

I propose that, within local domains corresponding to extended projections, typologically possible information-neutral word orders are limited to the stack-sortable (231-avoiding) permutations of a universal head-complement-specifier linear order. This proposal explains and unifies some well-known but previously unrelated word order universals, while successfully generating phenomena that challenge traditional approaches. Applications include Cinque's revision of Greenberg's Universal 20, the Final-Over-Final Condition, a modified Head Movement Constraint allowing attested Long Head Movement, English Affix Hopping, Germanic cross-serial subject-verb dependencies, and Icelandic Stylistic Fronting. Extending the system to multiple extended projections requires stack-sorting in cycles, expanding the set of allowed orders.

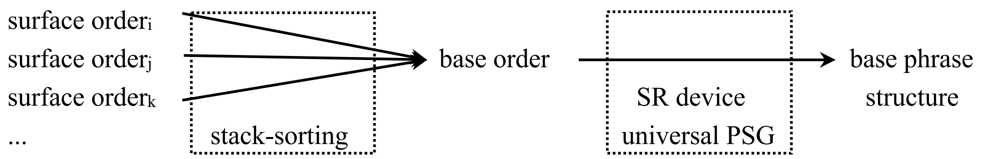
1. Introduction

This paper explores the following proposal about neutral word order possibilities in natural language.

(1) Typologically possible information-neutral surface word orders are stack-sortable (231-avoiding) permutations of a universal *head-complement-specifier* linear order.

The proper domain of application of (1) is a single extended projection (Grimshaw 1990); we delay consideration of the technology required to connect multiple domains, and the resulting elaboration of claims about possible orders, until section 8. Our goal is an account of information-neutral word order possibilities, setting aside wh-fronting, topic and focus movement, and the like.

Derivationally relating possible surface orders to an underlying base order is a familiar idea. But *contra* Kayne (1994), who assumes a universal *specifier-head-complement* linearization scheme, we take the base order to be *head-complement-specifier*.[1] More importantly, the derivational relationship is radically reconceptualized; word orders are fed to an invariant stack-sorting algorithm (Medeiros 2018), which aims to digest them into a uniform output, the base order. The base order is further processed to yield the constituent structure; we can implement this second stage of processing with a Shift-Reduce (SR) device. The algorithm fails to correctly sort word orders that contain a certain forbidden pattern, *231. By hypothesis, such orders are typologically unavailable because the universal parser transduces them into a sequence indigestible to the SR phrase processor, which requires and receives a universal order for attested neutral orders. The kind of architecture I have in mind can be sketched as (2).

(2)


Note, though, that (1) does not go so far as to endorse (2) as a claim about human sentence processing: it says local neutral orders are stack-sortable, not that humans actually stack-sort them. There is a strongly-

[1] I set aside questions about the adequacy of the X-bar theoretic (Chomsky 1970, Jackendoff 1977, Stowell 1981) notions of head, complement, and specifier. The reader should be alerted that head-complement-specifier base order has the counterintuitive property that heads are ordered top-down, while arguments and modifiers are ordered bottom-up. Consider an iterated X-bar structure [[Head$_1$ [[Head$_2$ Comp$_2$] Spec$_2$]] Spec$_1$], where 1 elements asymmetrically c-command 2 elements, and Head$_1$ takes [[Head$_2$ Comp$_2$] Spec$_2$] as its complement. The higher Head$_1$ precedes the lower Head$_2$ in the base order, but the lower Spec$_2$ precedes the higher Spec$_1$.

 Nordlyd 48.1: 59–91, Special issue on Extended Projections, edited by Peter Svenonius.
Published at the University of Tromsø – The Arctic University of Norway.
https://septentrio.uit.no/index.php/nordlyd https://doi.org/10.7557/12.7978

equivalent generative account (discussed in section 8.4) which also derives (1) but does not make use of stack-sorting. Nevertheless, I will largely keep to stack-sorting as a simplifying expository device.

However implemented, in the present proposal, a mapping procedure relates the base to a set of possible orders. These surface orders are the stack-sortable permutations of the underlying base order; they are counted by the Catalan numbers (1, 2, 5, 14, 42, ...), which grow much more slowly than all possible permutations, counted by the factorial n! (1, 2, 6, 24, 120, ...). The relevant orders can also be characterized as the 231-avoiding permutations of the base order, which is taken as the identity permutation *i* = 123... A 231-avoiding permutation is an order of elements in *i* not containing subsequence *...*b*...*c*...*a*..., for any subsequence ...*a*...*b*...*c*... in *i*.

The set of allowed surface orders is generally larger than what is realized in any given language.[2] But supposing that allowed orders must be stack-sortable accounts for a surprising range of word order universals. Beyond ruling out universally-forbidden word order patterns, the account also successfully generates syntactic phenomena that are problematic for many current syntactic theories. These results follow from (1) without invoking any further constraints or mechanisms, within or across languages. This contrasts with existing accounts covering the same empirical terrain, where the relevant effects are seen as unrelated.

The paper is structured as follows. In section 2, I define and illustrate stack-sorting. In section 3, I discuss Cinque's (2005) version of Universal 20 (Greenberg 1963), showing that we derive not just the same possible and impossible orders, but nearly-identical bracketed structures. I also discuss the apparent counter-examples to Cinque's typology in Shupamem raised by Nchare (2012), showing that all orders in that language that fall outside Cinque's typology involve focus, and are thus irrelevant to the generalization about neutral orders pursued here. Section 4 takes up the Final-Over-Final Condition (FOFC; Holmberg 2000; Biberauer, Holmberg, and Roberts 2014; Sheehan, Biberauer, Roberts, and Holmberg 2017, *i.a.*), showing that the present account derives every major case of FOFC as a consequence of our *231 theorem. I also consider apparent counterexamples to FOFC involving clause-final C-like particles (see Paul, this volume), pointing out that these may not falsify the account (the prototypical example of such an element is a question particle, which by its very nature does not occur in neutral expressions).

In section 5, I discuss Travis' (1984) Head Movement Constraint (HMC), showing that the core data taken to support the HMC is predicted by this account. At the same time, known exceptions to the HMC, in the form of attested Long Head Movement (LHM; Rivero 1991, Lema and Rivero 1991), are also shown to be generated by this system, obeying a novel generalization concerning the linear position of the subject. Section 6 presents a treatment of cross-serial relations in surface order, including cross-serial subject-verb dependencies in Dutch (Huybregts 1976; Bresnan, Kaplan, Peters, and Zaenen 1982) and Swiss German (Shieber 1985). These patterns are readily explained in the present framework without recourse to additional operations or constraints. I also provide an analysis of English Affix Hopping, an apparently problematic case of head lowering, where we reconstruct something remarkably like Chomsky's (1957) classic transformational analysis. Moving from large-bore phenomena to a detailed account of a particularly puzzling effect in one language, section 7 analyzes the possibilities and restrictions of Stylistic Fronting (SF; Maling 1990; Jónsson 1991; Holmberg 2000, 2006; Ott 2018, *i.a.*) in Icelandic, including its Subject-Gap Restriction and Accessibility Hierarchy. Section 8 extends the framework in various ways, including confronting the issue of cycles and observing that cycles expand the set of allowed orders. The final section summarizes and sketches broad conclusions.

[2] The account has significant implications for language acquisition, and makes the freer word order phenomenon much less mysterious. Space prevents exploration of these matters, but see Medeiros (2018) for relevant remarks.

2. Stack-sorting

I present a slightly edited version of Medeiros' (2018) algorithm,[3] itself a variant of Knuth's (1973) stack-sorting algorithm. Algorithm (3) rearranges input sequences into output sequences by two operations, Push and Pop, which take elements from input to stack, and from stack to output. Surface orders are the input; the stack-sortable inputs yield the uniform base order. Non-stack-sortable orders produce ill-formed output deviating from the uniform base, plausibly excluding them as surface orders: the procedure that automatically unscrambles allowed orders converts these forbidden orders into gibberish.

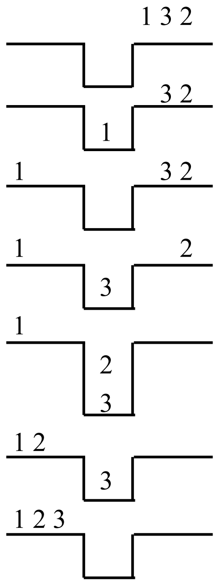
(3) Stack-Sorting Algorithm
While input is non-empty,
 If S < I, Pop.
 Else Push.
While Stack is non-empty,
 Pop.

(4) Definitions for (3)
I: next item in Input
S: item on top of Stack
x < y: x precedes y in the base order
Push: move I from Input to Stack
Pop: move S to Output

Stack-sorting can be represented as in (5). The input sequence (132) begins on the right; elements are pushed into the stack (center) before popping to the output at left. The sequence of operations is governed by the algorithm (3), as indicated at right of each step of (5).

[3] An important change here is the choice to represent the desired output (the identity permutation, in present terms) as an increasing sequence 123... . In Medeiros (2018), the nominal hierarchy [[[[N] AdjP] NumP] DemP] was instead numbered 4321. A consequence of that convention is that the forbidden permutation is then characterized as *213, rather than *231. The convention adopted here is in line with other work on permutations and sorting.

(5) Stack-sorting 132 into the identity 123



a. Input non-empty; S < 1 False: Push (1)

b. Input non-empty; 1 < 3 True: Pop (1)

c. Input non-empty; S < 3 False: Push (3)

d. Input non-empty; 3 < 2 False: Push (2)

e. Input empty; Stack non-empty: Pop (2)

f. Input empty; Stack non-empty: Pop (3)

g. Input empty; Stack empty.

In this way, we stack-sort the input permutation 132 (a surface order) into the desired output, the identity permutation 123 (representing the base order). Stack-sorting successfully rearranges many different permutations into the underlying identity. For example, if the input is 123, identical to the desired output, each element is pushed and immediately popped. If the input is 321, the mirror image of the output, the entire sequence is first pushed onto the stack and then popped, reversing its order. All of this, it should be emphasized, is done with the invariant algorithm (3); no order-particular instructions are required (the word order itself determines the actions to be taken, step by step).

However, not all permutations can be successfully stack-sorted. For example, among the six possible permutations of 123, five can be stack-sorted (123, 132, 213, 321, 312), but one cannot: 231. (6) demonstrates why 231 is not stack-sortable.

(6) Stack-sorting 231 fails to deliver the identity 123

a. Output: — ; Stack: — ; Input: 2 3 1 — Input non-empty; S < 1 False: Push (2)

b. Output: — ; Stack: 2 ; Input: 3 1 — Input non-empty; 2 < 3 True: Pop (2)

c. Output: 2 ; Stack: — ; Input: 3 1 — Input non-empty; S < 3 False: Push (3)

d. Output: 2 ; Stack: 3 ; Input: 1 — Input non-empty; 3 < 1 False: Push (1)

e. Output: 2 ; Stack: 1 3 ; Input: — — Input empty; Stack non-empty: Pop (1)

f. Output: 2 1 ; Stack: 3 ; Input: — — Input empty; Stack non-empty: Pop (3)

g. Output: 2 1 3 ; Stack: — ; Input: — — Input empty; Stack empty.

This produces illicit output order 213. This is not the desired output, the base order 123. In the context of language, we have failed to decode the surface word order into the universal base order, hence it will not be readable by the universal phrase structure grammar that builds constituent structure from the base order.

3. Universal 20

As our first empirical application, 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). More recent work has refined the empirical picture in this domain. I focus on the proposal of Cinque (2005) below.

3.1 Cinque's (2005) typology of attested nominal orders

According to Cinque (2005), 14 of 24 possible orders of these four elements are attested. I show all logically possible orders of these elements below, preserving Cinque's lettering scheme.

(7) Attested and unattested nominal orders, after Cinque (2005)

a. Dem Num Adj N
b. Dem Num N Adj
c. Dem N Num Adj
d. 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
l. N Adj Dem Num
m. *Dem Adj Num N
n. Dem Adj N Num
o. Dem N Adj Num
p. N Dem Adj Num
q. *Num Adj Dem N
r. Num Adj N Dem
s. Num N Adj Dem
t. N Num Adj Dem
u. *Adj Num Dem N
v. *Adj Num N Dem
w. Adj N Num Dem
x. N Adj Num Dem

Cinque shows that this pattern can be succinctly described by assuming a universal underlying base, built by a uniform sequence of External Merge operations, affected by phrasal movement, but not head movement or remnant movement (i.e. Internal Merge in the noun phrase must affect the noun, possibly pied-piping dominating structure).[4] His hierarchy is given in (8).

(8) [DemP ... [NumP ... [AdjP ... [N]]]]

The hierarchy in (8) is shorthand for a more articulated structure. Specifically, Cinque assumes the nominal modifiers are specifiers of associated functional phrases; he also posits interspersed agreement phrases, to host potential movements. Steddy and Samek-Lodovici (2011) provide full bracketed representations for each attested order, reproduced in (9).

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

(9) Bracketed representations from Steddy and 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$_{AgrXP}$]]]

We will return to this set of bracketed representations below, where we will find a corresponding but simpler kind of bracketed representation in the action of the stack-sorting procedure.

3.2 Universal 20 and Shupamem

There is considerable debate in the literature about the empirical status of Cinque's typology (Dryer 2018, Nchare 2012, Abels 2016, Steedman 2020, *i.a.*). A full discussion goes beyond the scope of this paper. But a crucial caveat is that the relevant typology concerns information-neutral orders. As we will see, this is the key to defusing an apparent counter-example to Cinque's characterization of allowed nominal orders.

A particularly sharp challenge is presented by the case of Shupamem (Grassfields Bantu), extensively discussed by Nchare (2012). Nchare documents 19 permitted orders of demonstrative, numeral, adjective, and noun in Shupamem. This set of orders excludes some of the orders permitted by Cinque, while including a number of others that fall outside Cinque's typology. I reproduce the list in (10) (after Nchare 2012: 134, *ex*. 10), alongside Cinque's for comparison. The examples illustrated in Shupamem permute the orders of demonstrative *ʃi* 'this' [*sic*], numeral *kpà* 'four', adjective *mìŋkɛ́t* 'dirty', and noun *pɔ́n* 'children' (note that postnominal modifiers are prefixed with noun class agreement *pí-*, which in the case of the demonstrative produces the form *pǐ*.

(10) Orders in Cinque (2005) and orders in Shupamem (Nchare 2012)

a. Dem Num Adj N	ʃì kpà mìŋkɛ́t pɔ́n
b. Dem Num N Adj	ʃì kpà pón pí-mìŋkɛ́t
c. Dem N Num Adj	*ʃì pɔ́n pí-kpà pí-mìŋkɛ́t
d. N Dem Num Adj	*pɔ́n pǐ pí-kpà pí-mìŋkɛ́t
e. *Num Dem Adj N	kpà ʃì mìŋkɛ́t pɔ́n
f. *Num Dem N Adj	kpà ʃì pɔ́n pí-mìŋkɛ́t
g. *Num N Dem Adj	*kpà pɔ́n pǐ pí-mìŋkɛ́t
h. *N Num Dem Adj	*pɔ́n pí-kpà pǐ pí-mìŋkɛ́t
i. *Adj Dem Num N	mìŋkɛ́t ʃì kpà pɔn
j. *Adj Dem N Num	*mìŋkɛ́t ʃì pón pí-kpà
k. Adj N Dem Num	mìŋkɛ́t pɔ́n pǐ pí-kpà
l. N Adj Dem Num	pɔ́n pí-mìŋkɛ́t pǐ pí-kpà
m. *Dem Adj Num N	ʃì mìŋkɛ́t kpà pɔ́n
n. Dem Adj N Num	ʃì mìŋkɛ́t pɔ́n pí-kpà
o. Dem N Adj Num	ʃì pɔ́n pí-mìŋkɛ́t pí-kpà
p. N Dem Adj Num	pɔ́n pǐ pí-mìŋkɛ́t pí-kpà
q. *Num Adj Dem N	kpà mìŋkɛ́t ʃì pɔ́n
r. Num Adj N Dem	kpà mìŋkɛ́t pɔ́n pǐ
s. Num N Adj Dem	kpà pɔ́n pí-mìŋkɛ́t pǐ
t. N Num Adj Dem	pɔ́n pí-kpà pí-mìŋkɛ̀t pǐ
u. *Adj Num Dem N	mìŋkɛ́t kpà ʃì pɔ́n
v. *Adj Num N Dem	míŋkɛ̀t kpà pɔ́n pǐ
w. Adj N Num Dem	mìŋkɛ́t pɔ́n pí-kpà ʃì
x. N Adj Num Dem	pɔ́n pí-mìŋkɛ́t pí-kpà ʃì

Nchare reports that Shupamem allows nineteen orders of these elements. In the present framework, there is no significance to the orders permitted by Cinque that do not happen to occur in Shupamem, namely (9c) Dem N Num Adj and (9d) N Dem Num Adj. However, the orders outside Cinque's typology that are available in Shupamem present a problem. Specifically, there are seven orders reported in Shupamem that are outside Cinque's typology: (9e) Num Dem Adj N, (9f) Num Dem N Adj, (9i) Adj Dem Num N, (9m) Dem Adj Num N, (9q) Num Adj Dem N, (9u) Adj Num Dem N, and (9v) Adj Num N Dem.

Crucially, though, Nchare counts non-neutral orders in Shupamem: "The revisited typology repeated in (10) will include focus as well as non focus orders" (Nchare 2012: 135). In fact, Nchare explicitly mentions focus in the derivation of all seven problematic orders. The derivation of (9f) is assumed to be a further movement applied to (9e), which itself "[...] can be derived if we assume that there is a phrasal movement of the numeral to the specifier of DP where it checks the focus feature under D" (*ibid*., 215). Order (9i) "[...] is derived by fronting the AP *mìŋkɛ́t* 'dirty' to the specifier position of DP to check its focus feature under D" (*ibid*., 218). Likewise in order (16m), "[...] the adjective undergoes a phrasal movement to a focus position" (*ibid*., 224). As for order (9q), Num Adj Dem N, Nchare's diagram (64q) shows the Num moving to spec of a D marked with a [+Foc] feature (*ibid*., 227). The same is true for orders (9u) and (9v) (*ibid*., 231); for the former, Nchare is explicit that there is "AP movement to spec-DP to check its focus feature" (*ibid*., 231).

Thus, the results of Nchare's (2012) investigation of ordering in Shupamem are entirely consistent with the present account. While discourse-information effects make other orders possible, this does not falsify the predictions made here, which keep solely to information-neutral ordering. Shupamem allows ten neutral orders, and all are within Cinque's typology.

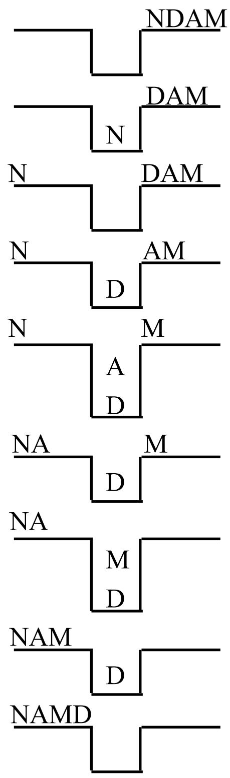
3.3 Stack-sorting reproduces Cinque's typology and bracketing

Returning to the present account, the stack-sortable orders turn out to be all and only the attested orders in (7), described by Cinque's (2005) version of Universal 20. Following his analysis in which demonstrative, numeral, and adjective are specifiers of functional phrases above the noun, their base order (and base-level phrase structure) in our head-complement-specifier format is (11).

(11) [[[[N] AdjP] NumP] DemP]

To illustrate stack-sorting nominal orders, I show the procedure for order (7p), N Dem Adj Num. This order is of some independent interest, as it exhibits cross-serial dependencies (see section 6). In (12), we see the stack-sorting process for this order, which produces the desired (11) as output (for space, I abbreviate the categories as **D**(em), (Nu)**M**, **A**(dj), **N**).

(12) Stack-sorting N Dem Adj Num (NDAM) into the identity N Adj Num Dem (NAMD)



a. Input non-empty; S < 1 False: Push (N)
b. Input non-empty; N < D True: Pop (N)
c. Input non-empty; S < D False: Push (D)
d. Input non-empty; D < A False: Push (A)
e. Input non-empty; A < M True: Pop (A)
f. Input non-empty; D < M False: Push (M)
g. Input empty; Stack non-empty: Pop (M)
h. Input empty; Stack non-empty: Pop (D)
i. Input empty; Stack empty.

As claimed, stack-sorting this order produces the desired head-complement-specifier base order (11). The sequence of stack-sorting operations is (13).

(13) Push (N), Pop (N), Push (Dem), Push (Adj), Pop (Adj), Push (Num), Pop (Num), Pop (Dem)

Consider the bracketed sequence we get from (13) by taking Push operations to be left brackets, while Pop operations correspond to right brackets, in each case labeled by the element they affect. Translating (13) into a bracketed representation in this way, we get (14). As we will see in a moment, this corresponds systematically to the bracketing in Cinque's derivation, as shown above in (9).

(14) (N N) (Dem (Adj Adj) (Num Num) Dem)

The nominal orders that can be stack-sorted into the base order (11) are exactly the same 14 attested orders Cinque (2005) describes; the remaining 10 unattested orders are non-stack-sortable. Collecting the Push and Pop sequences for each attested order and representing them as bracketed representations as in the conversion of (13) to (14), we get (15).

(15) Stack-sortable nominal orders and brackets read from Push-Pop sequence

a.	Dem Num Adj N	(Dem (Num (Adj (N N) Adj) Num) Dem)
b.	Dem Num N Adj	(Dem (Num (N N) (Adj Adj) Num) Dem)
c.	Dem N Num Adj	(Dem (N N) (Num (Adj Adj) Num) Dem)
d.	N Dem Num Adj	(N N) (Dem (Num (Adj Adj) Num) Dem)
k.	Adj N Dem Num	(Adj (N N) Adj) (Dem (Num Num) Dem)
l.	N Adj Dem Num	(N N) (Adj Adj) (Dem (Num Num) Dem)
n.	Dem Adj N Num	(Dem (Adj (N N) Adj) (Num Num) Dem)
o.	Dem N Adj Num	(Dem (N N) (Adj Adj) (Num Num) Dem)
p.	N Dem Adj Num	(N N) (Dem (Adj Adj) (Num Num) Dem)
r.	Num Adj N Dem	(Num (Adj (N N) Adj) Num) (Dem Dem)
s.	Num N Adj Dem	(Num (N N) (Adj Adj) Num) (Dem Dem)
t.	N Num Adj Dem	(N N) (Num (Adj Adj) Num) (Dem Dem)
w.	Adj N Num Dem	(Adj (N N) Adj) (Num Num) (Dem Dem)
x.	N Adj Num Dem	(N N) (Adj Adj) (Num Num) (Dem Dem)

Even more significant than deriving the same set of orders, we also generate a simplified version of the same bracketing that results from Cinque's derivational account. We can demonstrate the correspondence as follows. Starting with the standard bracketed representations in (9), keep only left brackets that immediately precede overt items, and their matching right brackets (we assume NP contains ...[N...]...). The result is identical to the bracketing in (15). Returning to our earlier example, order (7p) N Dem Adj Num has traditional representation in (9p), repeated in (16), now with [...N...] in NP.

(16) $[_{AgrWP}$ [...N...] $[_{WP}$ DemP $_{W}$ $[_{AgrXP}$ $[_{YP}$ AP $_{Y}$ $t_{NP}]$ $[_{XP}$ NumP $_{X}$ $[_{AgrYP}$ $t_{YP}]]]]]$

Keeping only the left brackets immediately preceding overt elements and their matching right brackets yields (17) (I suppress the labels, as these come out differently in the present account).

(17) [N] [DemP [AP] [NumP]]

The bracketing derived in the present account for order (7p) is (14), repeated in (18).

(18) (N N) (Dem (Adj Adj) (Num Num) Dem)

In this novel kind of representation, left brackets are positions of pronunciation. Leaving off right bracket labels, and writing overt elements following their corresponding left bracket positions, gives (19).

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

This is identical to the simplified standard bracketing for this order shown in (15); the interested reader may verify that this correspondence holds for every nominal order in (9) and (15).

As a final note, we may associate these elements with indices representing their relative order in the base (11): N = 1, Adj = 2, Num = 3, Dem = 4. If we translate the set of logically possible nominal orders into their sequence of base indices, we will see that the attested, stack-sortable orders are all 231-avoiding. Continuing with the example of order (6n) Dem-Adj-N-Num, the corresponding index sequence is 4213. This is 231-free.

Contrast this with all of the unattested orders; to pick one example, order (6e), *Num Dem Adj N. In terms of indices representing the base order of these elements, this is a 3412 order. This order contains the forbidden *231 contour in two different ways: the subsequences Num...Dem...Adj, 342, and Num...Dem...N, 341, are both forbidden subsequences.

3.4 Summary on stack-sorting Universal 20

This section has reviewed Cinque's (2005) version of Greenberg's (1963) Universal 20. We considered the empirical challenge posed by Nchare's (2012) description of Shupamem, discovering that all of the putative counter-examples to Cinque's typology presented by that language are in fact non-neutral focus orders, and as such irrelevant to the present study.

We have seen an application of the simple architecture of our universal stack-sorting grammar. Cinque's 14 attested neutral orders prove to be exactly the stack-sortable permutations of the base order, while his 10 unattested orders are its non-stack-sortable permutations. Furthermore, the surface bracketed structures we generate correspond directly to the syntactic structures derived in Cinque's account, though they are systematically simpler. This is remarkable, as Cinque's account represents a traditional External and Internal Merge derivation of the relevant orders and structures, obeying a parochial condition (movement in this domain can only be phrasal and must affect an XP containing the noun, excluding head movement and remnant movement). Here, by contrast, the same orders and structures are derived in quite a different way: the allowed orders are all and only the stack-sortable orders, and their bracketed surface structure is a record of the steps of their transformation into the base order.

In the next section, we turn to what appears to be an unrelated set of facts, the Final-Over-Final Condition, and show that it follows from the same architecture.

4. The Final Over Final Condition

In this section, I show that the present account explains 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 (10) to be compatible with FOFC.

4.1 Background: The Final Over Final Condition

FOFC prohibits configuration (20).

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

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

(21) FOFC in Finnish

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 (21b), 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 head-final over head-initial configuration. Biberauer *et al* (2014) list the following FOFC effects; these configurations are claimed to be robustly ungrammatical across languages (see below for discussion of a class of apparent counterexamples involving final C-like particles with head-initial clauses).

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

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)][5]

f. *Pol-TP-C   *[$_{CP}$ [$_{PolP}$ Pol TP] C]

FOFC effects obtain when the elements in question are in a head-complement relation. This well-known characterization is the key to its unification with the Universal 20 pattern within the present system.

*4.2 *231 predicts FOFC*

Consider a configuration with nested complementation: head α takes a complement headed by β, which in turn has complement γP. Given our head-complement-specifier scheme, the base order is then (23) α β γP, and the forbidden *231 permutation is (24) *β γP α.

(23) [α [β γP]]   Nested complementation base structure

(24) * β γP α   Forbidden 231 word order

Order (24) is traditionally described as a head-final phrase (αP) dominating a head-initial phrase (βP); this is exactly the configuration ruled out by FOFC (20), repeated as (25).

(25) *[$_{\alpha P}$ [$_{\beta P}$ β γP] α]

For example, if head Aux has complement headed by V, with complement Obj, the base order is Aux V Obj (26). We correctly exclude unattested *231 order *V Obj Aux (27).

(26) Aux [V [Obj]]   Base structure

(27) *V Obj Aux   Forbidden 231 word order

This generalizes to iterated head-complement structures of other kinds, reconstructing the predictions of FOFC. Let us emphasize again that we are restricting our attention to information-neutral orders, and, for the moment, to domains corresponding to single extended projections.

Arranging standard assumptions about clause structure into our assumed head-complement-specifier base order, major categories of a transitive clause are underlyingly as in (28).[6] Here, S and O signify external argument and internal argument (rather than the superficial grammatical functions).

(28) [C [Pol [T [Asp [*v* [V O] S]]]]]   Base structure for transitive clause

It is helpful to consider elements in the base order three at a time; we should find, for each such triple, five attested orders and one forbidden order. Drawing on order (28), understanding that the O position may be realized as clausal complement CP, we make the following predictions (among many others) about impossible neutral orders.

---

[5] 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. I do not adopt Roberts' analysis here. Instead, Universal 20 and FOFC are unified consequences of the *231 prohibition. This is confusing; the essential thing to keep in mind is that our way of ordering the base, head-complement-specifier, renders heads in top-to-bottom hierarchical order but specifiers in bottom-to-top order.

[6] The base structure in (28) is deliberately simplified, standing in for a more articulated structure. For example, our use of T for a variety of inflected verb endings below is clearly too crude, and we should distinguish T from Fin and Mood in, for example, the Icelandic examples in section 7 (thanks to a helpful reviewer on this point). We will also expand the aspectual categories for our English Affix Hopping examples.

(29) Selected forbidden 231 surface orders for base clause order (28)

a. *O S V
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 (30); I take adposition P to be a head with noun phrase complement NP (31).

(30) N O_N Base order for noun and complement object

(31) P NP Base order for P and nominal complement

Taken together, embedding the complement-taking nominal within an adpositional phrase, the base structure is (32) and the forbidden 231 surface order is (33). This accounts for the effect seen in the Finnish example (21).

(32) P [N O_N] Base order of noun with complement within PP

(33) *N O_N P Forbidden 231 surface order

In fact, setting aside (22e) (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 in (34).

(34) FOFC effects predicted here

a. *V O Aux see (27)
b. *V O C (29e)
c. *C TP V (29g)
d. *N O P (33)
e. *Pol TP C (29h)

Beyond reconstructing these FOFC effects, (29) contains other interesting predictions. If one basic clause order is to be ruled out, OSV appears to be the right choice (29a), as it is the rarest cross-linguistic order. Among 1376 languages recorded in WALS (Dryer 2013) as having a single dominant clause order, only four are reported to have this order (Warao, Venezuela; Nadëb, Brazil; Wik Ngathana, Australia; Tobati, Indonesia). That said, some mechanism going beyond the simple base-generation system here must be invoked for the handful of languages with OSV orders.[7] Another interesting prediction is (29i), taken up again in section 5 below as a reformulation of Travis' (1984) Head Movement Constraint.

4.3 Further extensions and challenges to FOFC

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. Following much recent

[7] One point worth mentioning is that full clauses generally have some discourse-information articulation: for example, subjects are often topics. If the initial O position in these languages has a discourse-information-linked character, it falls outside the predictions of the current account, which is restricted to neutral ordering.

cartographic work, we treat adverbials as specifiers of functional phrases, which will thus occur in the later portion of the base order. The base structure for the case Sheehan discusses is then [Aux [[V ...] Adv]] (35); unattested *V Adv Aux (36) is the forbidden *231 permutation.

(35) [Aux [[V...] Adv]] Base structure

(36) *V Adv Aux Forbidden 231 word order

Finally, a brief word is in order about a class of apparent counterexamples to FOFC involving clause-final C-like particles with head-initial lower structure. A full discussion would take us too far afield; see Paul (this volume) and references there. But a typical example of such an element is a question particle. Note that such elements do not fall under the purview of this theory, which is concerned with neutral ordering; questions, by their nature, are not neutral. If all of the relevant final C particles are non-neutral by nature, their occurrence with final-over-initial surface order does not falsify our predictions. I leave the matter for future research.

4.4 Interim summary on stack-sorting and FOFC

In this section, we have demonstrated that our basic theorem, *231, extends without additional machinery to cover the core empirical terrain described by the Final-Over-Final Condition. Specifically, the head-final over head-initial configuration banned by FOFC instantiates a *231 permutation of the underlying head-complement-specifier base order.

This is interesting for several reasons. First, FOFC has been held up as an instance of a purely syntax-internal constraint that is not explained by other factors. In this architecture, FOFC, as an instance of *231, is a necessary consequence of our system, rather than an additional constraint on movement that could have turned out otherwise. This result is also significant in that the account unifies Universal 20 and FOFC as instances of the same phenomenon. This is starkly at odds with the traditional treatment, where the two effects seem to have little to do with each other. Indeed, they seem to conflict; see the discussion in Roberts (2017) for relevant considerations.

In the following section, we turn to another application of our *231 theorem: it derives a version of the Head Movement Constraint, while allowing known exceptions.

5. The Head Movement Constraint and its exceptions

The present account also explains Travis' (1984) Head Movement Constraint (HMC), while correctly predicting some well-known exceptions. Travis argues that a head cannot move to a higher head position over an intervening governing head, formalizing this claim with her Head Movement Constraint (HMC)..

(37) Head Movement Constraint (HMC) (Travis 1984: 131)

An X^o may only move into the Y^o which properly governs it.

The HMC (37) rules out configurations like (38), where head Z^o has "skipped" intervening head Y^o and left-adjoined to X^o.

(38) *[... Z^o-X^o ... [... Y^o ... [... t_{Z^o}]]]]

Of course, merely requiring head movement to be short does not suffice to rule out (38); Z^o cannot move to Y^o first and then excorporate, moving without Y^o to X^o. That requirement is a stipulation that does not appear to follow from independent principles, and is quite different from phrasal movement, which appears to obey no such restriction (phrases may move successive-cyclically without picking up additional structure along the way). Indeed, head movement remains controversial, posing a number of challenges to standard accounts of syntactic movement, and has inspired a variety of analyses; see Dékany (2018) for a recent overview.

That the present account extends to these effects is surprising at first glance, as movement violating the HMC does not produce an impossible *231 order of the heads themselves. Instead, HMC-violating movement "skipping" an intervening head, as in (38), produces 312 order among the heads, readily generated by this system.

Consider a simplified version of our base clause structure (28). For ease of exposition, we focus on the core C-T-V categories; adding more elements does not affect the conclusions. First, we examine each possible permutation of C, T, V order (39). Next to each order permutation, we write the corresponding index sequence (123, etc), and, for clarity, the traditional description of the derivation of the order (e.g., V-to-T, indicating head movement of V to T).

(39) Permutations of C,T,V head order

a.	C T V	1 2 3	base order
b.	V T C	3 2 1	V-to-T-to-C
c.	C V T	1 3 2	V-to-T
d.	V C T	3 1 2	HMC violation, V-to-C skipping T: ok
e.	*T V C	*2 3 1	FOFC violation
f.	T C V	2 1 3	T-to-C

As indicated, our *231 principle only rules out the FOFC-violating order *T-V-C. The order violating the head movement constraint, V-C-T (Long Head Movement of V to C, skipping T) is generated without problems.

However, something interesting emerges when we consider the order of these elements with the addition of another element from later in the clause base structure, here an external argument S. We repeat the permutations in (40), now trying all possible surface positions of S with respect to the orders in (39); a * marks an impossible position (one that will produce the forbidden *231 contour in the surface order).

(40) Permutations of C,T,V head order with interposed S

a.	(S) C (S) T (S) V (S)	base order
b.	(S) V (*S) T (*S) C (S)	V-to-T-to-C
c.	(S) C (S) V (*S) T (S)	V-to-T
d.	(S) V (*S) C (*S) T (S)	HMC violation
e.	*T V C	FOFC violation
f.	(S) T (*S) C (S) V (S)	T-to-C

In (40), we generate: (40a) C T V (no head movement); (40b) V T C (full roll-up of heads obeying HMC, "V-to-T-to-C"); (40c) C V T (partial HMC-obeying movement, "V-to-T"); and (40f) T C V (partial HMC-obeying movement, "T-to-C"). The independently FOFC-violating order (40e) is ruled out already, and adding the subject anywhere has no effect (the forbidden subsequence persists regardless of additional material). Meanwhile, (40d) V C T ("V-to-C, skipping T"), the HMC-violating order, is ruled out only if a higher-index element (e.g., the external argument S) intervenes between V and T (or another higher head).

We have generated all the core cases of local head movement in the C-T-V system. Note, too, that we derive another important effect: obligatory surface adjacency of the head cluster. That is, a later element from the base order, such as the external argument, may never occur amid an inverted sequence of heads (because the heads form a 21 sequence; the later 3 between them produces the forbidden *231 permutation). In other words, all of the instances of head movement forming a "complex word", such as V-to-T, or V-to-T-to-C, necessarily occur adjacent, without the possibility of other material intervening. To see this more clearly, (41) repeats (40), substituting an underscore where arguments and adjuncts are impossible. Note in particular (41b, c, and f).

(41) Obligatory adjacency in permutations of C,T,V head order

a.	(S) C (S) T (S) V (S)	base order
b.	(S) **V __ T __ C** (S)	V-to-T-to-C
c.	(S) C (S) **V __ T** (S)	V-to-T
d.	(S) **V __ C __ T** (S)	HMC violation
e.	*T V C	FOFC violation
f.	(S) **T __ C** (S) V (S)	T-to-C

Most interesting of all, we arrive at a new claim: there is nothing wrong *per se* with the "improper head movement" case (39d) that violates the HMC. Instead, what we expect to be ruled out is a subsequence in which a later element from the base order, in particular the external argument S, occurs between the long head-moved V and a higher head like C or T in the surface order. Abstracting now from the simple C-T-V system to a larger set of clausal head positions, we rule out (42).

(42) *V Subj *v*/Aux/T

That is, the verb cannot precede an external argument which precedes some head above V (in fact, we have seen this prediction already, in (29i) 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. Importantly, understanding the HMC as actually reflecting the condition in (40) also allows us to account for Long Head Movement, a much-discussed violation of the HMC first described for Old Spanish by Rivero (1991) and Lema and Rivero (1991). The Breton example (43) illustrates the phenomenon.

(43) Lennet en deus Anna al levr
read.pprt has Anna the book
'Anna has read the book.'
(Roberts 2010: 194)

I take the base order for this example to be (44); (45) shows the numerical indexing of the surface order, which is indeed 231-free, and thus generable in this system (the treatment of the auxiliary Aux and associated affix –Fx will be discussed in the following section).[8]

(44)

T	Aux	–Fx	V	O	S
1	2	3	4	5	6

(45)

Lennet	en	deus	Anna	al	levr
V	–Fx	Agr	Aux+T	S	O
4	3	?	2+1	6	5

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

[8] The labeling and numbering here deserves some clarification. A reviewer points out that the first part of *en deus* is something like a subject clitic, while tense is expressed in the second portion. The rather tortured 2+1 is intended to indicate that it is not clear where or whether to segment these items. Meanwhile, I do not assign an index to the agreement element (agreement being neither interpreted nor universal in its presence or location).

(46) Bulgarian (Harizanov and Gribanova 2019:482)
  a. Bjah pročel knigata.
    had read the.book
    'I had read the book.'
  b. Pročel bjah knigata.
    read had the.book

(47) Bulgarian (Embick and Izvorski 1997:231)
  a. Bihte bili arestuvani ot policijata.
    would been arrest.PTCP by the.police
    'You would be arrested by the police.'
  b. Arestuvani bihte bili ot policijata.
    arrest.PTCP would been by the.police

Interestingly, in all these cases information-neutral Long Head Movement obeys the *V S T condition in (42).[9] That is, either the subject is placed after the entire verbal complex, as in Breton (43), or the subject is null, as in the Bulgarian examples in (46); (47) shows a prepositionally-marked passivized thematic subject, which again does not intervene between participle and higher heads to its right.

Summarizing, our *231 principle of neutral word order derives a version of Travis' (1984) Head Movement Constraint (HMC) that covers core cases of (V-to-)T(-to-C) movement, predicting obligatory surface adjacency of the "head cluster", while also allowing attested LHM as in Breton and Bulgarian. We discover a novel and apparently exceptionless generalization about when neutral LHM is possible: only when it obeys *V S T/Aux. No special principles or mechanisms are invoked; head movement, often seen as unlike other kinds of syntactic movement, falls together with Universal 20 and FOFC as an immediate consequence of our *231 principle.

6. Generating some well-known cross-serial dependencies

Thus far, we have mostly been concerned with ruling out typologically unattested orders. In this section, I show that our allowed orders include somewhat exotic constructions that have figured prominently in arguments that natural language grammars are mildly context-sensitive (Joshi 1985). Specifically, the architecture provides simple analyses of attested cross-serial dependencies, including unbounded crossing subject-verb dependencies in certain Germanic languages, as well as the more limited crossing pattern seen in English Affix Hopping.

These constructions have proven challenging to describe: standard upward, leftward movement does not generate them. In the case of English Affix-Hopping, we see apparent rightward lowering movement (of affixes associated with higher auxiliary verbs onto lower verbal stems). Meanwhile, the long-distance cross-serial dependencies below require something like Richards' "Tucking in" movement, failing to obey the Extension Condition, or a radically different notion of movement, as in Tree-Adjoining Grammar.

Our framework allows these patterns without additional devices. That is surely interesting, but it may not seem impressive if one has the mistaken impression that almost anything goes with this ordering scheme. A consideration of scale brings some perspective: the examples below contain about ten clause-level formatives; less than one percent of logically possible orders this size are stack-sortable. To put it simply, if stack-sortability were not a relevant condition, it would be surprising to encounter orders this size meeting the condition; the fact that every example we consider does so is striking indeed.

[9] Ian Roberts (p.c.) and Maria-Luisa Rivero (p.c.) observe that this appears to be true quite generally for Long Head Movement.

6.1 Cross-serial verb-argument dependencies

Bresnan *et al* (1982) discuss unbounded cross-serial subject-verb dependencies in Dutch (Huybregts 1976). Example (48), taken from Steedman (2000: 25), illustrates the phenomenon.[10]

(48) ...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 orders in Swiss German, which also show long-distance cross-serial case dependencies, as in (49).

(49) ... das mer d'chind em Hans es huus lönd hälfe aastriiche
... that we the children Hans the house let help paint
'...that we let the children help Hans paint the house'

Our stack-sorting grammar permits these orders.[11] I take the Dutch example (48) above to contain the categories in (50), abstracting away from internal structure of the object *de nijlpaarden* and segmenting a Tense suffix from inflected and non-finite verbs, even if realized as zero.

(50) ...omdat ik Cecilia Henk de nijlpaarden zag -Ø 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.

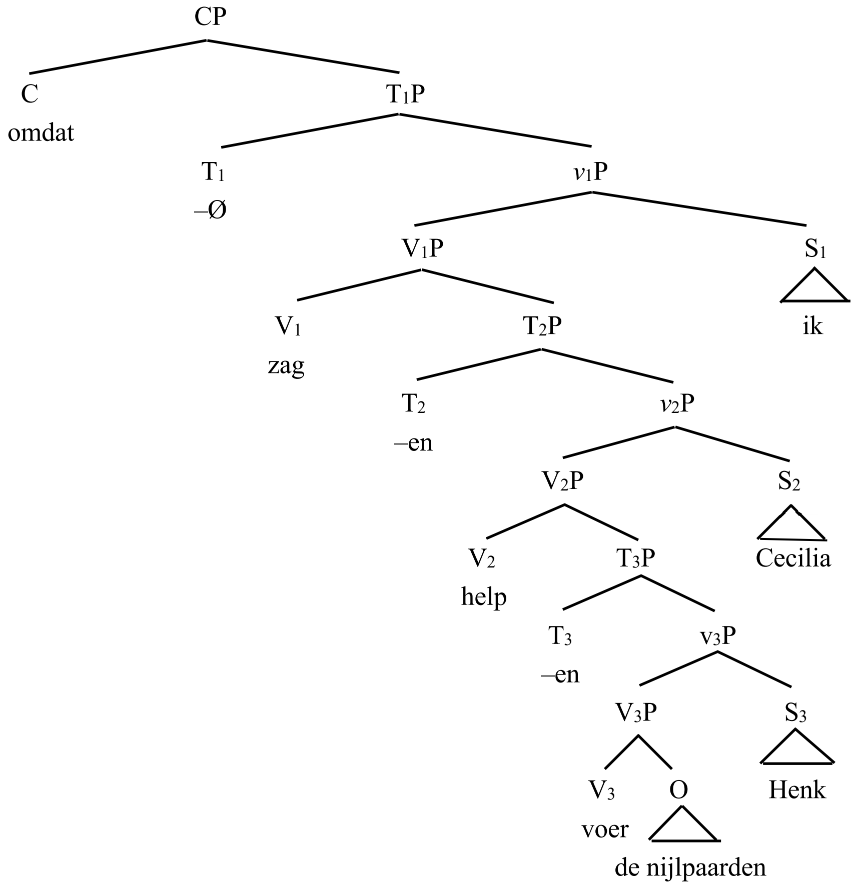
The categories in (50) form the base tree given in (51) below, given the assumption adopted throughout this work of head-complement-specifier base order.

---

[10] The lines above the example connect subjects to their corresponding verbs; note that in both (48) and (49), there is another relation crossing these, not shown, between the final verb and its direct object.

[11] 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.

(51) Base structure for (50)



Given this base structure, we can read off and number the base order (52), which allows us to identify the Dutch surface order as a permutation (53) of the base.

(52) omdat –Ø zag –en help –en voer de nijlpaarden Henk Cecilia ik
1 2 3 4 5 6 7 8 9 10 11

(53) ...omdat ik Cecilia Henk de nijlpaarden zag–Ø help–en voer–en
1 11 10 9 8 3 2 5 4 7 6

This permutation, 1-11-10-9-8-3-2-5-4-7-6, is 231-avoiding and thus stack-sortable.

6.2 Affix Hopping

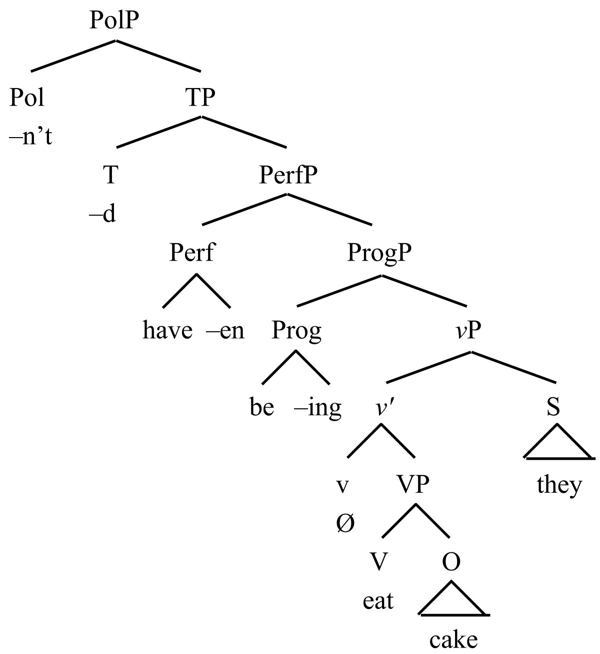
Another crossing configuration that has figured prominently in generative work is English Affix Hopping. Chomsky's (1957) analysis of Affix Hopping provided a strong argument for the necessity of transformational rules beyond the generative capacity of phrase structure systems. However, this pattern has not been easy to analyze with the tools available in later theories. For example, early Minimalist work proposed that the relevant pattern did not involve overt syntactic movement, but rather resulted from checking features on fully inflected lexical items inserted from the lexicon. By contrast, the present account allows a return to something very close to the original transformational analysis.

Sentence (54), *They hadn't been eating cake*, exhibits two instances of the phenomenon. As Chomsky (1957) noted, affixes group with the preceding auxiliaries in distribution and meaning, despite being separated by the intervening verb in surface order. This involves a more limited and local form of cross-serial relations than those illustrated for Dutch in the previous section.

(54) They have –d –n't be –en eat –ing cake
S Aux_1 T Pol Aux_2 $–Fx_1$ V $–Fx_2$ O
9 3 2 1 5 4 7 6 8

I take the base structure to be (55) below. Note that auxiliaries and their associated affixes are adjacent in the base, plausibly constituting pieces of a single category.[12] This base structure lets us read off the base order of the elements involved, and label the surface order as a permutation: 9-3-2-1-5-4-7-6-8. This permutation is 231-avoiding, thus stack-sortable.

(55) Base structure for (54)



In effect, we recover Chomsky's (1957) classic analysis of Affix Hopping: an auxiliary and its associated affix are introduced as pieces of a single lexical item, and a transformation "hops" the affix onto the following verbal element. Here, the relevant transformation is part of the standard reorganization of surface order into base order by stack-sorting; no special principles or extra machinery are required.

6.3 Wrapping up cross-serial dependencies

In this section, we have seen that the present account readily allows cross-serial dependencies that have proven challenging to capture in other frameworks. In particular, both the cross-serial subject-verb dependencies of some Germanic languages, as well as the familiar pattern of Affix Hopping in English, turn out to require nothing beyond the tools we have already developed. Both patterns obey our *231 condition, given a relatively uncontroversial understanding of their base structure; as such, they are expected to be typologically possible.

In the following section, we turn to another phenomenon that challenges standard conceptions of syntactic movement: Icelandic Stylistic Fronting.

[12] We could also treat auxiliaries and affixes as hierarchically adjacent heads, as in Harwood (2014).

5. Stylistic Fronting

This section examines Stylistic Fronting (SF; Maling 1990, Jónsson 1991, Holmberg 2000, 2006, Ott 2018, *i.a.*) in light of our stack-sorting architecture. The phenomenon presents a number of features that are puzzling from the point of view of standard theories of phrase structure and movement. Three properties in particular stand out in this regard: (i) SF is optional and information-neutral; (ii) SF can apply to a broad array of syntactic objects, both head-like and phrasal, including verbal participles, adverbs, negation, and argument NP or PP, dependent on the presence or absence of other such categories; (iii) SF is contingent on not having an overt subject in the typical subject position. I show that all of these facts are consistent with the *231 prediction of the present account. The exposition below relies heavily on the work of Ott (2018), from which I draw the majority of the examples I consider.

Since its first description by Maling (1990), SF has received various analyses. Platzack (1987) proposes that SF is movement to Spec, TP. Rögnvaldsson and Thráinsson (1990) pursue a similar analysis, analyzing SF as topicalization movement to Spec, TP. According to the treatment of Holmberg (2000), in SF T agrees with the subject, but moves a different category to fill Spec, TP at PF. Jónsson (1991) proposes to treat SF as head movement, adjoining to T. Bošković (2004), in turn, argues that SF moves to a null affixal F head above T. Finally, Ott (2018) argues that SF involves remnant movement, which gives a way to treat the phenomenon as strictly phrasal movement, even when only an overt head appears to move.

7.1 The subject-gap restriction

In descriptive terms, SF appears to involve movement of other categories (both heads and phrases) to the typical subject position, which must not contain the external argument. The subject-gap restriction closely parallels our *V S T prediction for long head movement. Bošković (2004) ascribes this to an affixation requirement on a null F head. Other analyses try to derive the subject-gap restriction by moving the stylistically-fronted element into Spec, TP, with considerable problems (notably, that that position is presumably filled by a trace when the subject is A-bar moved).

The basic phenomenon of the subject-gap restriction can be seen in (56): the negation element *ekki*, which is otherwise capable of being stylistically fronted, cannot be fronted if the subject position is filled. Note that, unlike English *–n't*, analyzed as an instance of the relatively high head Pol (Laka 1990), I treat Icelandic *ekki* as a negative adverb (in a rightward specifier position between the internal argument O and the external argument S).[13]

(56) Subject-gap violation (Bošković 2004: 40, *ex*. 5b)
*Ég held að ekki Halldór hafi séð þessa mynd. *1 **7 8 3** 2 5 4 6
I think that not Halldor has seen this film

(57) Base order for (56)

að	-i	haf	–ð	sé	[þessa mynd]	ekki	Halldór.
C	T[14]	Aux	–Fx	V	O	NegAdv	S
1	2	3	4	5	6	7	8

Note, though, that the subject gap can be the result of A-bar extraction, which should leave a trace, blocking movement to Spec, TP.

[13] See Roberts (2019: chapter 7) for extensive discussion of different cartographic sites for negation. Although the arguments must be reexamined in light of the different assumptions in this account, it is clear that Icelandic *ekki* cannot be analyzed as an instance of Pol here.

[14] This stretches the use of the category T to the point of abuse; a reviewer points out that the relevant suffix expresses subjunctive mood, and occurs outside the position for the genuine T marker. The point is well taken. Crucially, though, the base position for this element, like T, falls between C and Aux. In turn, the oversimplifications adopted here will only be problematic if they get the relative base order among overt elements wrong.

(58) SF of V participle with A-bar extracted subject (Ott 2018: 3, *ex*. 7a)

Hver heldur þú [$_{CP}$ að stoliðᵢ hafi tᵢ hjólinu]. ...C V –Fx Aux T O

who think you that stolen has the.bike 1 5 4 3 2 6

Other ways to satisfy subject-gap restriction are to have a postposed (necessarily indefinite) subject, or an impersonal.

(59) Postposed subject (Ott 2018: 4, *ex* 8b)

Keypt$_i$ hafa t$_i$ þessa bók margir stúdentar. V –Fx Aux T O S

bought have this book many students 4 3 2 1 5 6

(60) Impersonal (Ott 2018: 4, *ex*. 9a)

Keypt$_i$ hefur verið t$_i$ tölva fyrir starfsfólkið V –Fx$_2$ Aux$_1$ T Aux$_2$ –Fx$_1$ O PP

bought has been computer for the.staff 6 5 2 1 4 3 7 8

In the present framework, the Subject-gap restriction parallels the *VST prediction for long head movement discussed in section 5. That is, SF places a medial element in the base order left of the higher T head; having a subject (later in the base order than either) between them would necessarily create the forbidden *231 permutation pattern.

7.2 Promiscuity of SF and the Accessibility Hierarchy

One of the stranger properties of SF is that it seems to affect both heads, such as participles, and full phrases, such as argument NPs or PPs. This "promiscuity" is especially problematic from the point of view of standard approaches that draw a sharp distinction between head and phrasal movement. In our framework, however, this is actually an expected result. In (59-61), we see that with a verb taking a NP or PP argument, SF can affect the V, or the O/PP, but not both.

(61) SF of participle (Ott 2018: 17, *ex*. 43b)

Þeir sem **búið$_i$** hafa t$_i$ [$_{PP}$ í Ósló] 1 5 4 3 2 6

those that lived have in Oslo

(62) SF of PP (Ott 2018: 9 *ex*. 18d)

Þeir sem **[$_{PP}$ í Óslo]$_i$** hafa búið t$_i$ 1 6 3 2 5 4

those that in Oslo have lived

(63) SF of V and PP (Ott 2018: 10, *ex*. 22a)

*þeir sem **[$_{vP}$ búið í Óslo]$_i$** hafa t$_i$ ... *...1 5 4 6 3 2

those that lived in Oslo have

The index sequences to the right of the examples above come from the following base order (64). Thus, our *231 principle makes the right predictions: movement of just the verbal participle (61), or just the PP (62), produces a 231-free surface word order. Moving both together, however, produces a 231-like subsequence, correctly ruling out (63).

(64) Base order for (61-63)

sem –a haf –ið bú [í Óslo]

C T Aux –Fx V PP

1 2 3 4 5 6

If an adverb or negation is present, it can undergo SF, but it will "block SF of vP-internal material" (Ott 2018: 12).

(65) SF of negation (Ott 2018: 13, *ex*. 29a-b)

a. þegar búið$_i$ var t$_i$ að borða ... 4 3 2 1 5 6
when finished was to eat

b. þegar ekki$_i$ var t$_i$ búið að borða ... 7 2 1 4 3 5 6
when not was finished to eat

c. *þegar búið$_i$ var ekki t$_i$ að borða ... 4 3 2 1 7 5 6 ok
when finished was not to eat

(66) Base order for (65a-c)
–r va –ið bú að borða ekki
T Aux –Fx V T V NegAdv
1 2 3 4 5 6 7

Here, we correctly allow SF of the participle when no adverb or negation is present (65a), and SF of the negation (65b). Example (65c), ungrammatical in Icelandic, is actually 231-free, and so theoretically permitted in our system. In general, this sort of thing is as expected. That is, we do not expect any language to allow the full range of possible surface order permutations. However, we expect that the ungrammaticality of (65c) arises from a different source than (65a-b), which are ruled out for all languages by the *231 principle.

A PP can undergo SF, but only if negation is not present (67). If it is present, only negation can undergo SF (68b), blocking SF of the PP (68c).

(67) SF of PP (Ott 2018:13, *ex*. 30-31)
Þeir sem í Danmörku$_i$ hafa verið t$_i$ ... 1 6 3 2 5 4
those that in Denmark have been

(68) a. Þeir sem hafa ekki verið í Danmörku ... 1 3 2 7 5 4 6
those that have not been in Denmark

b. Þeir sem ekki$_i$ hafa t$_i$ verið í Danmörku ... 1 7 3 2 5 4 6
those that not have been in Denmark

c. *Þeir sem í Danmörku$_i$ hafa ekki verið t$_i$ *...1 6 3 2 7 5 4
those that in Denmark have not been

(69) Base order for (67-68)
sem –a haf –ið ver [í Danmörku] ekki
C T Aux –Fx V PP NegAdv
1 2 3 4 5 6 7

Given this base order for the above examples, we correctly generate the attested SF options and rule out SF of PP in the presence of negation (68c).

In verb-particle constructions, either the verb or the particle can undergo SF. Negation, if present, blocks both, being the only candidate for SF then. We correctly exclude particle movement in the presence of negation, but not movement of the participle.

(70) SF of particle and V participle without negation (Ott 2018: 24, *ex*. 57)

a. fundurinn sem fram$_i$ hefur farið t_i ... 1 6 3 2 5 4
the.meeting that forth has gone

b. fundurinn sem fariði$_i$ hefur t_i fram ... 1 5 4 3 2 6
the.meeting that gone has forth

(71) SF of particle and V participle blocked with negation (Ott 2018: 24, *ex*. 58)

a. *fundurinn sem fram$_i$ hefur ekki farið t_i ... 1 **6** 3 2 **7** **5** 4
the.meeting that forth has not gone

b. *fundurinn sem fariði$_i$ hefur ekki t_i fram ... 1 5 4 3 2 7 6 ok
the.meeting that gone has not forth

(72) Base order for (70-71)

sem -ur hef –ið far fram ekki
C T Aux –Fx V Ptcl NegAdv
1 2 3 4 5 6 7

In verb-particle constructions with an object, extraposition of the object (i.e., postposing of O to the end of the surface order) is a necessary condition for SF to apply. This is correctly captured by the present proposal: the order Ptcl ... O V in the ungrammatical (73a) is a *231 permutation; Ptcl ... V O order in the grammatical (73b) is not.

(73) SF in verb-particle constructions with object (Ott 2018: 26, ex. 64)

a. *það var þa sem ut$_i$ voru [$_{NP}$ einhverjir kettir] reknir t_i ...*1 6 3 2 7 5 4
it was then that out were some cats driven

b. það var þa sem ut$_i$ voru reknir t_i [$_{NP}$ einhverjir kettir] ... 1 6 3 2 5 4 7
it was then that out were driven some cats

Extraposition of the object is optional in verb-particle constructions without SF.

(74) Optional extraposition of object with verb-particle (Ott 2018: 27, fn. 37, citing Thráinsson)

a. það var þa sem það voru [$_{NP}$ einhverjir kettir] reknir ut ... 1 ? 3 2 7 5 4 6
it was then that EXPL *were some cats driven out*

b. það var þa sem það voru reknir ut [$_{NP}$ einhverjir kettir] ... 1 ? 3 2 5 4 6 7
it was then that EXPL *were driven out some cats*

(75) Base order for (73-74)

sem –u vor –ir rekn ut [$_{NP}$ einhverjir kettir]
C T Aux –Fx V Ptcl O
1 2 3 4 5 6 7

Again, the *231 principle gets the facts right, given the head-complement-specifier base order in (75).[15]

[15] I do not assign a base position, or corresponding index, to the existential or presentational expletive element *það*; I mark its position in the permutation with '?' (*sem* being 1; the cleft expletive *það* in the matrix is glossed 'it' following Ott). This is because the base order is a representation of thematic structure, and I assume that the expletive is inserted to satisfy some (presumably language-specific) surface-oriented predictive pattern. Many interesting questions arise here, which must be put aside for future work.

7.3 Summary on Stylistic Fronting

Stylistic Fronting is an example of neutral word order variation, our intended explanatory target. The framework correctly allows all attested SF configurations. At the same time, we rule out many (but not all) of the things SF can't do. This is as expected: the current approach aims to capture word order variation across languages, and as a result, overgenerates with respect to particular languages. Thus, not all restrictions on SF fall out here, the typical result. The crucial claim is that we *do* allow attested neutral expressions. This is indeed the case, for all instances of SF examined here.

In particular, some but not all aspects of the Accessibility Hierarchy are explained. We correctly capture the subject-gap restriction, which parallels the *VST prediction about long head movement derived in section 5. Most importantly, the account sheds new light on the "promiscuous" nature of SF, affecting both heads and phrases of a variety of categories. This is a central consequence of our approach: all neutral movement reflects *231 over a unified linear representation of the underlying hierarchy, including heads and phrases.[16] The hierarchy-order mapping we have developed collapses all varieties of neutral head and phrasal movement to a single mechanism. Different movement possibilities for different syntactic categories follow from the invariant order of the base and the *231 theorem.

In the next section, I take up some various loose ends that have been left aside.

8. Some loose ends

This section takes up some matters that have been left hanging in the discussion so far, which deserve some comment from the present perspective.

8.1 Cycles

We have kept throughout to a relatively simplistic view of phrase structure. In particular, we have avoided the topic of recursive embedding, outside of the treatment of cross-serial subject-verb dependencies in section 6. The astute reader may have noted that we ignored internal structure of argument NPs and PPs in the previous section, assigning them a single index and base position with respect to the clause.

A full treatment of the topic is beyond the scope of this paper. But it is immediately clear that some notion of cycle is required for this account to get the facts right. To see this, consider a classic problem for the Final-over-Final Condition: head-final VPs may embed head-initial DPs.

(76) German head-final VP with head-initial object (Biberauer *et al* 2014)

Johann hat [$_{VP}$ [$_{DP}$ einen Mann] gesehen].

Johann has a man seen

'Johann has seen a man.'

Simply treating all of the elements as part of a unified base order, [Det N] V should be a *231 order. Why, then, is it possible as a neutral order?

In the FOFC literature, the standard approach is to assume that the nominal is a separate cycle: a distinct hierarchy, or a different kind of extended projection. I adopt this solution, supposing that nominal and verbal cycles are disjoint for the purposes of the *231 condition.

We can sketch how this would work for (76). Suppose a single node within the verbal cycle can contain a pointer to a separately-computed nominal cycle. The nominal subtree internally obeys 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. Meanwhile, the embedding clausal cycle manipulates an atomic pointer to the nominal (O). Within the clausal cycle, the visible S-Aux-O-V order is 231-avoiding.

[16] It bears repeating that the head-complement-specifier base order disrupts the usual way of thinking about correspondences between surface order and hierarchy. In the base, heads are ordered top-down, while arguments and adjuncts are ordered bottom-up.

As stated at the outset, our claim that neutral word orders are restricted to stack-sortable orders applies to elements within a single extended projection. This makes a good deal of sense, especially if the more general claim about allowed orders is implemented with something like the stack-sorting architecture sketched in the introduction. This is because, if deployed as a parser, the stack-sorting algorithm requires a native hierarchy among elements to drive a pairwise comparison of relative base order, determining whether to Push or Pop at each stage. This comparison is, we assume, fixed for elements within a single projection; for example, T < V within a single clause.

But there is clearly no once-and-for-all relative hierarchy among elements from distinct extended projections. One clause may be embedded in a second clause, or vice versa; a nominal may contain a relative clause, or a clause may contain a nominal in an argument position. These distinct embedding possibilities produce distinct base orders of the relevant elements. Now, with respect to any particular expression, the base structure is whatever it is, with a unique order. From the point of view of the alternative generative account we have alluded to, the relevant base structure is available with all such embeddings specified, and no problem arises. But if this system is deployed as a parsing device, where the intended meaning (hence, base structure and order) is not known in advance, non-determinism is inevitable where distinct domains are joined.[17]

Many important questions arise at this point, especially about where such cycles are to be posited, under exactly what conditions. Here, the extensive discussion of this question in the FOFC literature is directly relevant. Further questions specific to this framework concern the effect of this additional machinery on possible surface permutations. I set aside these topics for future work, beyond asserting one interesting observation: in the limit, if cycles are allowed completely freely, the allowed orders are then the so-called *separable permutations*, characterized by avoidance of the permutations *2413 and *3142. Steedman (2020) derives the same permutation-avoidance pattern as a consequence of CCG, under very different assumptions (see discussion and references there).

8.2 What about verb clusters?

We have, to this point, avoided discussion of ordering restrictions in verb clusters. Some words are in order on the topic, especially given the close parallels drawn by other authors between order avoidance in verb clusters and the Universal 20 pattern in nominals. Somewhat surprisingly, perhaps, this account does not support this parallel. On the other hand, we do discover an explanation of impenetrability effects in verb clusters, which amount to the same effect seen in our novel prediction *VST about long head movement, and in our account of the subject-gap restriction in Stylistic Fronting.

The convention in the literature is to number the elements of a verb cluster to reflect their underlying scope hierarchy, with 1 scoping over 2, etc.[18] That numbering scheme aligns with the ordering among elements adopted in the present framework, at least insofar as we understand the relevant structure to be iterated instances of heads taking complements (more on this in a moment).

This brings to light an immediate (apparent) problem: several authors have tied the ordering restrictions of Universal 20 to the avoidance of 213 orders in verb clusters, which are indeed severely restricted, even if not, in fact, truly unattested (see especially Salzmann 2019 on the existence of verb clusters with this order in Swiss German). We expect, instead, to find that 231 order is ruled out, while 213 order ought to be possible.

There are a few possible resolutions of this tension. One possibility worth exploring, already mooted in Abels (2016), is that the relevant structure is not, in fact, a series of heads and complements, but rather has a different structure--as it has to, in his theory, to make the higher verbal elements “satellites” of the main verb. If the higher elements are, in the base structure, modifiers of the verb head, then 213-avoidance

---

[17] This is a good reason for languages to develop characteristic orders and marking strategies that resolve this inherent indeterminacy for parsing.

[18] In light of the discussion in the immediately preceding section, it is salient that verb cluster elements do not have a fixed hierarchy: for example, some clusters have modals scoping over auxiliaries, while auxiliaries scope over modals in other verb cluster constructions.

is the expectation. This is because our base order would then diverge from the standard system of numbering verb clusters (for example, a Mod-Aux-V cluster would have base order 123 = [[[V] AuxP] ModP]).

On the other hand, keeping the base structure as we have assumed elsewhere in the paper might be tenable after all. We then are left without a direct explanation for the extreme rarity of 213 orders, and revert to claiming that 231 orders are really the bad ones. Tackling the former deficiency first, note that 213 order, when translated to basic clause order, corresponds to the rarest non-231 order: OVS. Perhaps there is some independent reason why such orders (OVS, and 213 verb clusters) present problems. And attested 231 orders are restricted (see Svenonius 2007, Abels 2013), and in telling ways.

The best evidence to decide the issue, as far as I can see, is the pattern of impenetrability effects in verb clusters. In brief, in certain verb cluster orders non-verbal elements may interrupt the sequence, but other verb cluster orders are impenetrable to such interruptions. “In both Hungarian and West Germanic, the left-branching order (3>2>1) the sequence cannot be interlaced with XPs, while in the right-branching order (1>2>3) the sequence of verbs may be interrupted by XPs. In a mixed sequence, such as 1>3>2 illustrated above, the left-branching (i.e., right-headed) portion is inviolable (20b), (21b) while the right-branching (i.e., left-headed) part of the cluster (1>[3>2]) allows for intervening XPs, as expected.” (Bobaljik 2004: 140)

This set of facts is exactly what we predict if the base order for verb clusters follows the conventions for head-complement order in the base assumed elsewhere (i.e., the standard numbering correctly represents the relative base ordering among verb cluster elements). In the case of a three-verb cluster, the interrupting XP will be a 4 with respect to the 1, 2, 3 assigned to the verb cluster elements (it is a specifier or adjunct, necessarily base-ordered after all heads of the head-complement spine it attaches to). The observed impenetrability of 321 clusters follows at once: *3421, *3241 contain the forbidden 231 contour. A 123 cluster allows an intervening XP 4 without creating a 231 contour: 1423, 1243. Finally, the mixed order 132 permits interruption of the 1-3 portion (1432) but not of the 3-2 portion (*1342). This seems to speak strongly in favor of maintaining our assumptions about base order intact, even in verb clusters.

Why, then, are 231 orders possible at all? The only way out this framework offers, other than the alternate base structure considered and rejected above, is to appeal to additional machinery we need anyway: cycles. For the case in question, the idea is that the 231 order reflects two disjoint structures for the purposes of stack-sorting, each of which is 231-avoiding internally. Specifically, the 2 and 3 will have to be processed as an opaque embedded unit with respect to the level of structure containing the 1. We can represent the intended structure with brackets as [23]1.

There are many interesting aspects of this phenomenon that warrant further exploration, and important questions to answer about how the notion of cycle invoked here intersects with restructuring and the verb cluster phenomenon. Particularly sharp questions revolve around the structure and distribution of 231 orders, as well as about the status of 213 orders. I leave the matter for future work.

8.3 Is it significant?

The number of stack-sortable surface orders for a base order of length n is the nth Catalan number, drawn from the sequence (1, 2, 5, 14, 42, 132, ...). This quantity grows much more slowly than the number of logically possible orders, which are counted by the factorial function $n! = n(n–1)(n–2)...(2)(1)$. In (77) I show these quantities explicitly; the last column shows the nth Catalan number as a percentage of $n!$. This is a useful metric, measuring the relative rarity of stack-sortable orders as a function of length.

(77) Number of 231-avoiding orders (Cat(n)) compared to possible orders of n items (n!)

n	Cat(*n*)	*n*!	Cat(*n*) as % of *n*!
1	1	1	100%
2	2	2	100%
3	5	6	83.3%
4	14	24	58.3%
5	42	120	35.0%
6	132	720	18.3%
7	429	5,040	8.51%
8	1,430	40,320	3.55%
9	4,862	362,880	1.34%
10	16,796	3,628,800	.462%
11	58,786	39,916,800	.147%
12	208,012	479,001,600	.043%

Stack-sortability (231-avoidance) is a weak condition for shorter sequences: it allows both orders of two elements, and only rules out one of six logically possible orders of length three. For four items, more than half of the possible orders are permitted (14 of 24, as for Universal 20). However, as the length of the strings increases, 231-avoidance becomes highly characteristic and unlikely to arise by chance. In example (54), *They hadn't been eating cake*, we identified nine categories in the clausal domain; only about 1% of all logically possible orders of nine elements are 231-avoiding. If 231-avoidance were not a relevant condition for attested neutral word orders, it would be quite a coincidence to find any particular order of this length meeting the condition, much less one example after another.

As indicated earlier, it is natural for stack-sorting to apply in cycles determined by a single extended projection, in Grimshaw's (1990) sense. It is in such domains that there is a fixed hierarchy among elements (and corresponding fixed relative order in the head-complement-specifier base), which could be used to stack-sort surface word orders into the invariant inner form in online parsing. By contrast, at the boundaries of extended projections, relative embedding is not fixed once and for all, and thus relative order in the base is unknown to a purely local and invariant parsing device. Allowing cycles extends the set of allowed orders (notably to allow 231 orders as [23]1). But even in the worst-case scenario of allowing cycles absolutely anywhere, we would allow at most the set of separable permutations. I leave the details aside, but see Steedman (2020) for relevant remarks about this class of permutations.[19]

8.4 A generative account

This paper has focused on one way of implementing the claim that neutral word orders are limited to stack-sortable orders, in terms of a stack-sorting architecture. As mentioned, however, this is not a necessary commitment. There is an alternative, generative way of deriving the central *231 claim here; in fact, I provide two equivalent formulations.

The first method is suggested by the observation that the Push-Pop sequences for stack-sortable orders, read as left and right brackets as described previously, create all legal bracketings of a given size. Moreover, when we examine the labels on those brackets, we see that the right brackets always occur in the same order (the left-to-right order of right brackets is simply the base order). To illustrate these facts, I repeat our bracketings for the Universal 20 orders in (15) as (78), bolding right brackets.

[19] A significant feature of this larger class of permutations is that their avoided permutations, *2413 and *3142, are both mirror images and symmetric with respect to reversal of the order of the base structure (*i.e.*, interchanging 1s with 4s and 2s with 3s results in the same avoided permutations). This contrasts with the asymmetry with respect to the base order that characterizes the *231 pattern.

(78) Stack-sortable nominal orders and Push-Pop brackets

a.	Dem Num Adj N	(Dem (Num (Adj (N **N**) **Adj**) **Num**) **Dem**)
b.	Dem Num N Adj	(Dem (Num (N **N**) (Adj **Adj**) **Num**) **Dem**)
c.	Dem N Num Adj	(Dem (N **N**) (Num (Adj **Adj**) **Num**) **Dem**)
d.	N Dem Num Adj	(N **N**) (Dem (Num (Adj **Adj**) **Num**) **Dem**)
k.	Adj N Dem Num	(Adj (N **N**) **Adj**) (Dem (Num **Num**) **Dem**)
l.	N Adj Dem Num	(N **N**) (Adj **Adj**) (Dem (Num **Num**) **Dem**)
n.	Dem Adj N Num	(Dem (Adj (N **N**) **Adj**) (Num **Num**) **Dem**)
o.	Dem N Adj Num	(Dem (N **N**) (Adj **Adj**) (Num **Num**) **Dem**)
p.	N Dem Adj Num	(N **N**) (Dem (Adj **Adj**) (Num **Num**) **Dem**)
r.	Num Adj N Dem	(Num (Adj (N **N**) **Adj**) **Num**) (Dem **Dem**)
s.	Num N Adj Dem	(Num (N **N**) (Adj **Adj**) **Num**) (Dem **Dem**)
t.	N Num Adj Dem	(N **N**) (Num (Adj **Adj**) **Num**) (Dem **Dem**)
w.	Adj N Num Dem	(Adj (N **N**) **Adj**) (Num **Num**) (Dem **Dem**)
x.	N Adj Num Dem	(N **N**) (Adj **Adj**) (Num **Num**) (Dem **Dem**)

Taken together, these observations entail that there is a simple way to generate these labeled bracketed structures directly.

(79) Procedure for generating bracketed representations in (78)
- a. Generate all legal pairings of *n* brackets, for a base order of length *n*.
- b. Write the base order onto the right bracket labels, from left to right.
- c. Copy right bracket labels to the matching left brackets.

We get the same structure and derive the same ordering restrictions from (79) as we do with stack-sorting. We can also achieve the same result if we imagine the structures we freely generate to be *n*-ary branching tree data structures rather than strings of brackets, by following the procedure in (80).[20]

(80) Procedure for generating labeled tree representations
- a. Generate all trees with $n + 1$ nodes, for a base order of length *n*.
- b. Write the base order onto every node except the root in postorder.
- c. Read the word order from the tree in preorder.

These generative procedures are at least simple, but the implicit shift in perspective is considerable. This way of looking at things highlights the mathematical beauty of the set of allowed surface structures, which are all and only the set of legal bracketings, or *n*-ary trees of a given size. We may imagine that the base order is generated independently by a context-free phrase structure grammar or similar device. The determination of surface order is inherently a kind of interface effect; the base order is "refracted" through the lens of any tree structure whatsoever, and the stack-sortable orders emerge. Yet again, further exploration is set aside for reasons of space.

8.5 Gathering loose ends

This section has touched on a variety of extensions and lingering issues. We sketched why stack-sorting is naturally limited to extended projections, and considered how adding cycles changes our ordering predictions. We considered orders in verb clusters, concluding that we correctly predict the impenetrability

[20] See Feil, Hutson, and Kretchmar (2005) for an overview of permutations and tree traversals. See also Kural (2005) for a very different account of word order variation in terms of tree traversals.

effects in this domain. However, we also predict that 231 orders should be forbidden as (single-cycle) neutral orders, in contrast to the literature drawing an explicit parallel between the Universal 20 pattern and 213-avoidance in verb clusters. Moving on, we pointed out that stack-sortability would be a surprising property for even moderate length orders to have by chance, and it is striking to find it repeatedly in long orders. Finally, we outlined a simple but strange generative procedure yielding the same structures and orders by "refracting" the independently-determined base order through freely-generated branching structure.

9. Conclusion

This paper has introduced a novel framework for understanding neutral word order variation. The basic ingredients are twofold. First, we have claimed that there is a base structure shared by all languages, and that this structure is ordered heads first (i.e., in head-complement-specifier order, to use familiar X-bar terms). Second, we have claimed that typologically possible word orders are the 231-avoiding permutations of the base order. We have focused on a realization of the architecture as a stack-sorting transducer feeding an SR machine, though alternative formulations are possible.

Some immediate questions arise, when presented with any new formalism for language structure. Does it overgenerate, creating many configurations not found in human language? Or do its inherent limits correspond in an interesting way to observed word order universals? Does it undergenerate, failing to account for well-attested structures? (A particularly thorny case is presented by constructions with cross-serial dependencies, which challenge classic phrase structure theories.) Finally, and most importantly, we must ask about strong generation. Beyond merely generating attested orders and failing to generate unattested ones, does the formalism assign them the proper structure?

If we take the target of explanation to be the set of typologically possible neutral orders (rather than the orders permitted in just a single language), the present account performs well on these metrics. We capture a range of word order universals, including Universal 20, the Final-Over-Final Condition, and a version of the Head Movement Constraint: orders violating these generalizations cannot be parsed by our architecture. At the same time, we find that the system correctly allows a range of constructions that have been challenging to accommodate within other theories, including discontinuous constituency, cross-serial dependencies, head movement including apparent lowering, and so on. In certain cases, such as our treatment of Long Head Movement, we appear to have improved on existing accounts in our empirical coverage. Importantly, we have done more than simply capture the proper range of string orders; we also automatically assign a surface structure representation to each allowed surface word order that corresponds closely to existing descriptions.

These effects follow from our heads-first base order and a single principle regulating the hierarchy-word order mapping: *231. That principle is itself a theorem, a necessary consequence of the stack-sorting framework. One could not keep the same basic mechanisms here while deriving any other principles of permitted and forbidden orders and structures.

The architecture described here demonstrates that assuming a linearly ordered underlying representation permits a particularly simple account of the relationship between hierarchy and word order, unifying a range of previously unrelated ordering restrictions. At the same time, the present theory implicitly represents a retreat from understanding the details of individual languages, in particular with respect to how they select a subset of universally-possible orders. On this view, we are born knowing how to compute and comprehend language; what we learn is how to predict and produce particular languages.

References

Abels, Klaus, and Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1): 25–74. http://dx.doi.org/10.1111/j.1467-9612.2011.00163.x.

Abels, Klaus. 2013. On 2-3-1. In *UCL Working Papers in Linguistics* 25: 67–101.

Abels, Klaus. 2016. The fundamental left–right asymmetry in the Germanic verb cluster. *The Journal of Comparative Germanic Linguistics* 19(3): 179–220. https://link.springer.com/article/10.1007/s10828-016-9082-9.
Biberauer, Teresa, Anders Holmberg, and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45: 169–225. https://dx.doi.org/10.1162/LING_a_00153.
Bobaljik, Jonathan. 2004. Clustering theories. In *Verb Clusters: A Study of Hungarian, German, and Dutch*, edited by Katalin É. Kiss and Henk van Riemsdijk, pp. 121–146. John Benjamins, Philadelphia.
Bošković, Željko. 2004. PF merger in stylistic fronting and object shift. In *Minimality Effects in Syntax*, edited by Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel, pp. 37–71. Mouton de Gruyter, New York.
Bresnan, Joan, Ronald M. Kaplan, Stanley Peters, and Annie Zaenen. 1982. Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13(4): 613–635. Available at https://www.jstor.org/stable/4178298.
Chomsky, Noam. 1957. *Syntactic Structures*. Mouton de Gruyter, New York.
Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English Transformational Grammar*, edited by Roderick Jacobs and Peter Rosenbaum, pp. 184–221. Ginn and Co., Waltham, MA.
Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.
Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford University Press, New York.
Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36(3): 315–332. https://dx.doi.org/10.1162/0024389054396917.
Dékány, Éva. 2018. Approaches to head movement: A critical assessment. *Glossa: A Journal of General Linguistics* 3(1): 65. https://doi.org/10.5334/gjgl.316.
Dryer, Matthew. 2013. Order of subject, object and verb. In *The World Atlas of Language Structures Online*, edited by Matthew Dryer and Martin Haspelmath. Max Planck Institute for Evolutionary Anthropology, Leipzig. [http://wals.info/chapter/81, accessed 4 January 2021.]
Dryer, Matthew. 2018. On the order of demonstrative, numeral, adjective, and noun. *Language* 94(4), 798–833. https://doi.org/10.1353/lan.2018.0054.
Embick, David, and Roumyana Izvorski. 1997. Participle-auxiliary word orders in Slavic. In *Annual Workshop on Formal Approaches to Slavic Linguistics: The Cornell Meeting 1995*, edited by Wayles Browne, Ewa Dornisch, Natasha Kondrashova, and Draga Zec, pp. 210–239. Michigan Slavic Publications, Ann Arbor, MI.
Feil, Todd, Kevin Hutson, and R. Matthew Kretchmar. 2005. Tree traversals and permutations. *Congressus Numerantium* 172: 201–221.
Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, edited by Joseph Greenberg, pp. 73–113. MIT Press, Cambridge, MA.
Grimshaw, Jane. 1990. *Argument Structure*. MIT Press, Cambridge, Ma.
Harizanov, Boris, and Vera Gribanova. 2019. Whither head movement? *Natural Language & Linguistic Theory* 37(2): 461–522. https://dx.doi.org/10.1007/s11049-018-9420-5.
Harwood, William. 2014. Rise of the auxiliaries: A case for auxiliary raising vs. affix lowering. *The Linguistic Review* 31(2): 295–362. https://doi.org/10.1515/tlr-2014-0001.
Holmberg, Anders. 2000a. Deriving OV order in Finnish. In *The Derivation of VO and OV*, edited by Peter Svenonius, pp. 123–152. John Benjamins, Philadelphia.
Holmberg, Anders. 2000b. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31: 445–484.
Holmberg, Anders. 2006. Stylistic fronting. In *The Blackwell Companion to Syntax*, edited by Martin Everaert and Henk van Riemsdijk, pp. 532–565. Blackwell, Oxford.
Hrafnbjargarson, Gunnar Hrafn. 2004. Stylistic fronting. *Studia Linguistica* 58: 88–134. https://dx.doi.org/10.1111/j.0039-3193.2004.00111.x.
Huybregts, Riny. 1976. Overlapping dependencies in Dutch. *Utrecht Working Papers in Linguistics* 1:24–65.

Jackendoff, Ray. 1977. *X-Syntax: A Study of Phrase Structure*. MIT Press, Cambridge, MA.

Jónsson, Johannes Gísli. 1991. Stylistic fronting in Icelandic. *Working Papers in Scandinavian Syntax* 48:1–43.

Joshi, Aravind. 1985. How much context-sensitivity is necessary for characterizing structural descriptions: Tree-Adjoining Grammar. In *Natural Language Processing: Theoretical, Computational, and Psychological Perspectives*, edited by David Dowty, Lauri Karttunen, and Arnold Zwicky, pp. 206–250. Cambridge University Press, New York.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. MIT Press, Cambridge, MA.

Knuth, Donald. 1973. *The Art of Computer Programming, Vol III: Sorting and Searching*. Addison-Wesley, Reading, MA.

Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal Complexes*. MIT Press, Cambridge, MA.

Kural, Murat. 2005. Tree traversal and word order. *Linguistic Inquiry* 36(3): 367–387. https://dx.doi.org/10.1162/0024389054396890.

Laka, Itziar. 1990. *Negation in Syntax: on the Nature of Functional Categories and Projections*. Ph.D. thesis, MIT, Cambridge, MA.

Lema, José, and Maria-Luisa Rivero. 1991. Types of verbal movement in Old Spanish: Modal, futures, and perfects. *Probus* 3(3): 237–278.

Maling, Joan. 1990. Inversion in embedded clauses in Modern Icelandic. In *Modern Icelandic Syntax*, edited by Joan Maling and Annie Zaenen, pp. 71–91. Academic Press, San Diego, CA.

Medeiros, David P. 2018. ULTRA: Universal Grammar as a universal parser. *Frontiers In Psychology: Language Sciences* 9: 155. https://dx.doi.org/10.3389/fpsyg.2018.00155.

Nchare, Abdoulaye Laziz. 2012. *The Grammar of Shupamem*. Ph.D. thesis, New York University, New York.

Ott, Dennis. 2018. Stylistic Fronting as Remnant Movement. *Studia Linguistica* 72(1): 1–38. https://dx.doi.org/10.1111/stul.12054.

Platzack, Christer. 1987. The Scandinavian languages and the null-subject parameter. *Natural Language and Linguistic Theory* 5: 377–401.

Rivero, Maria-Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak. *The Linguistic Review* 8(2–4): 319–351.

Roberts, Ian. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. MIT Press, Cambridge, MA.

Roberts, Ian. 2017. The Final-over-Final Condition in DP: Universal 20 and the nature of demonstratives. In *The Final-over-Final Condition: A Syntactic Universal*, edited by Michelle Sheehan, Theresa Biberauer, Ian Roberts, and Anders Holmberg, pp. 151-186. MIT Press, Cambridge, MA.

Roberts, Ian. 2019. *Parameter Hierarchies and Universal Grammar*. Oxford University Press, New York.

Rögnvaldsson, Eiríkur, and Höskuldur Thráinsson. 1990. On Icelandic word order once more. In *Modern Icelandic Syntax*, edited by Joan Maling and Annie Zaenen, pp. 3–40. Academic Press, San Diego, CA.

Salzmann, Martin. 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. https://dx.doi.org/10.1007/s10828-019-09104-x.

Sheehan, Michelle. 2017. The Final-Over-Final Condition and adverbs. In *The Final-over-Final Condition: A Syntactic Universal*, edited by Michelle Sheehan, Theresa Biberauer, Ian Roberts, and Anders Holmberg, pp. 97-120. MIT Press, Cambridge, MA.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts, and Anders Holmberg (eds.). 2017. *The Final-over-Final Condition: A Syntactic Universal*. MIT Press, Cambridge, MA.

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

Stabler, Ed. 2004. Varieties of crossing dependencies: structure dependence and mild context sensitivity. *Cognitive Science* 28: 699–720 https://dx.doi.org/10.1207/s15516709cog2805_4.

Steddy, Sam, and Vieri Samek-Lodovici. 2011. On the ungrammaticality of remnant movement in the derivation of Greenberg's Universal 20. *Linguistic Inquiry* 42: 445–469. https://dx.doi.org/10.1162/LING_a_00053.

Steedman, Mark. 2000. *The Syntactic Process*. MIT press, Cambridge, MA.

Steedman, Mark. 2020. A formal universal of natural language grammar. *Language* 96(3): 618–660. https://dx.doi.org/10.1353/lan.2020.0043.

Stowell, Timothy. 1981. *Origins of Phrase Structure*. Ph.D. thesis, MIT, Cambridge, MA.

Svenonius, Peter. 2007. 1...3-2. In *The Oxford Handbook of Linguistic Interfaces*, edited by Gillian Ramchand and Charles Reiss, pp. 239–288. Oxford University Press, New York.

Travis, Lisa. 1984. *Parameters and Effects of Word Order Variation*. Ph.D. thesis, MIT, Cambridge, MA.