

Appendix to

‘A syntactic universal and its consequences’

by Theresa Biberauer, Anders Holmberg and Ian Roberts

The paper will henceforth be referred to as BHR.

1. Deriving FOFC-violations

Consider the FOFC-violations discussed in section 2.7 in BHR, the list repeated here.

- | | | |
|-----|--------------|--|
| (1) | *V-O-Aux | *[AuxP [VP V DP] Aux] |
| | *V-O-C | *[CP [TP T VP] C] or *[CP [TP [VP V O] T] C] |
| | *C-TP-V | *[VP [CP C TP] V] |
| | *N-O-P | *[PP [DP/NP D/N PP] P] |
| | *Num-N-D(em) | *[D(em)P [NumP Num NP] D(em)] |
| | *PolP-C | *[CP [PolP Pol TP] C] |

1.1 *V-O-Aux

Taking Aux to be a realization of v, this order presupposes the distribution of the L-movement trigger \wedge in (2), which is impossible if $[\pm V \wedge]$ can only enter the derivation as a property of the head of the Extended Projection (condition (75a) in BHR).

- | | |
|-----|---|
| (2) | [_{VP} v [_{VP} V O]] |
| | [+V \wedge] [+V] |

We predict that a configuration such as (3), where Aux is a higher auxiliary, realizing T rather than v, will also be ill-formed, a violation of FOFC, regardless of whether VP is head-initial or head-final.

- (3) a. *vVOAux
 b. *vOVAux

(3a) would be ruled out as a case of (2), by condition (75a) in BHR. (3b) would be the result of the following distribution of [+V[^]] :

- (4) [TP T [vP v [VP V O]]]
 [+V[^]] [+V] [+V[^]]

Given that a functional head can have [+V[^]] only by virtue of it spreading from, ultimately, the head of the Extended Projection, this configuration violates Relativized Minimality, the intervening v blocking spreading of [^] to T from V.

The West Germanic 231 verb-cluster orders discussed in section 3.3 in BHR are very close to instantiating the structure in (3b), and hence the schema for the FOFC violation in (4). As suggested there, what seems to be going on in these IPP constructions is either that the infinitive is associated with a nominalised (or at least deverbalising) element, or that the 2-verb is structurally integrated with the 3-verb in a way that does not usually hold within Extended Projections.

1.2 *V-O-C

This order presupposes a distribution of the L-movement trigger in (5), ruled out for the same reason as (2), by condition (75a) in BHR.

- (5) [CP C ... V]
[+V[^]] [+V]

As discussed in BHR section 2, there are different varieties of this FOFC-violation, depending on whether v and T are initial or final: all of them are ruled out by condition (75a) in BHR.

We predict that a head-final C will be ruled out regardless of whether V is head-initial or head-final, if either v or T is head-initial.

- (6) a. *v-O-V-T-C
b. *T-O-V-v-C

(6a,b) will have the distribution of [^] in (7a,b) respectively. Both are ruled out by Relativized Minimality.

- (7) a. [CP C [TP T [vP v [vP V]]]]
[+V[^]] [+V[^]] [+V] [+V[^]]
b. [CP C [TP T [vP v [vP V]]]]
[+V[^]] [+V] [+V[^]] [+V[^]]

1.3 *C>TP>V

As discussed in BHR section 2.3, head-initial CP cannot occur in the preverbal object position in OV languages. What we find instead in OV languages is one of the following possibilities: (a) clausal complements are head-final and precede the verb, (b) clausal complements are nominalizations, or (c) clausal complements are head-initial and extraposed (Koptjevskaja-Tamm 1988, 1993).

Consider first the case of a lexical verb with a CP-complement in a harmoniously head-final language. The distribution of the L-movement trigger will be as follows:

$$(8) \quad \dots V \quad [_{CP} \ C \quad \dots V \] \\ [+V^{\wedge}] \quad [+V^{\wedge}] \quad [+V^{\wedge}]$$

Compare this with the case of an OV language with initial C:

$$(9) \quad \dots V \quad [_{CP} \ C \quad \dots V \] \\ [+V^{\wedge}] \quad [+V] \quad [+V^{\wedge}]$$

In both structures, the higher V is a lexical head entering the derivation with inherent \wedge . We propose that the system, nevertheless, interprets (8) and (9) as a single Extended Projection headed by the lower V; this is, in fact, a consequence of the definitions in (70) and (71) in BHR of spine and Extended Projection: the higher V is in the spine of the lower V, by virtue of sharing the value of $[\pm V]$. In the case of (8), this has no pernicious consequences; in terms of linear order, the structure ends up as consistently head-final. In the case of (9), though, it does have the consequence that the structure is a Relativized Minimality violation, equivalent

to, for instance, (4) and (7a,b). There are two solutions, as it were, to this dilemma: one is to fully nominalize the complement clause, introducing a nominal Extended Projection between the two V-projections, thus closing off the Extended Projection of the lower verb. The other solution is to add a nominal functional shell with an initial nominal head to the complement clause; the nominal then moves to the left of the verb given that we have V[^], “stranding” the visible residue of the CP (see Biberauer and Roberts 2008 for an implementation of this idea in terms of the Phase Impenetrability Condition, and Biberauer & Sheehan 2012 for an implementation appealing to Sheehan’s (2011, forthcoming a, b) “complement stranding” proposal). This accounts for Koptjevskaja-Tamm’s (1988, 1993) generalization, discussed above, that OV languages either postpose clausal complements or have nominalizations instead of embedded clauses.¹

1.4 *[N>O]>P

One case of *[N>O]>P discussed in BHR section 2.4 was (10), from Finnish, where the adposition and the head of the complement are in bold, and the well-formed examples all mean ‘across the border between the countries’.

- (10) a. **yli** [**rajan** maitten välillä]
 across border countries between
- b. **yli** [maitten välisen **rajan**]
 across countries between border
- c. ***[rajan maitten välillä] yli**

d. [maitten välisen **rajan**] yli

The permitted configuration is (11a), the disallowed, FOFC-violating configuration is (11b):²

(11) a. P [NP... N ...]
[-V[^]] [-V[^]]

b. P [NP... N ...]
[-V[^]] [-V]

If P is a functional head in the N-projection, then (11b) is a straightforward case of a functional head having an illegitimate [^], not inherited from the head of the Extended Projection. If P is a lexical head, heading its own Extended Projection, then (11) instantiates a situation much like C>TP>V, discussed above: a lexical head takes a complement with the same categorial feature. This is shown schematically in (12):

(12) X [Y^P Y]
[αV] [αV]

Our proposal was that, as far as the L-movement-trigger [^] goes, in this configuration the grammar cannot distinguish between the case where X is lexical and Y is the head of the complement of X, and the case where X is functional, with a categorial feature spreading to X from Y. Effectively they form a single Extended Projection, and (11b) is ruled out, explaining why (10c) is ill-formed, violating FOFC.³

1.5 * Num N D(em)

This order of constituents presupposes the following distribution of \wedge :

- (13) $[_{DemP} Dem [_{NumP} Num [_{NP} N]]]$
 $[-V^\wedge] \quad [-V] \quad [-V(\wedge)]$

The structure ends up violating FOFC whether N has \wedge or not. If it does, the structure violates Relativized Minimality. If it does not, it violates the condition that \wedge paired with $[\pm V]$ can only be introduced in the derivation by the lexical head.

1.6 *[Pol TP] C

The distribution of \wedge deriving the order [Pol TP] C would be (14):

- (14) $[_{CP} C [_{PolP} Pol [_{TP} T [_{vP} v [_{VP} V]]]]]$
 $[+V^\wedge] \quad [+V] \quad [+V^\wedge] \quad [+V^\wedge] \quad [+V^\wedge]$

This is ruled out as $[+V^\wedge]$ cannot be copied by C across Pol without violating Relativized Minimality. As discussed in section 2, many Indo-Aryan languages have borrowed a final complementizer. However, this is not the case for those languages which have an initial polarity particle. The explanation is that in those languages such a borrowing would not observe FOFC, now seen as a violation of Relativized Minimality.

2. Optional head-finality

In most of the cases discussed above, it is understood that head-final order is a matter of parametric specification. For example, languages have a head-final or head-initial V as a parametric option, formally encoded as \bar{A} paired with [+V], and \bar{A} spreads through the functional heads in the Extended Projection of V, hence the phrase-final property, along with the categorial feature [+V], as a parametric option, but subject to Relativized Minimality, with FOFC as a consequence. For instance, v might inherit \bar{A} and T might not, while the opposite situation is ruled out by Relativized Minimality. The choice of parameter values is constrained by UG in the sense that the expressions which are the result of the parametric choice must respect universal principles, including Relativized Minimality.

Head-final orders do not necessarily have to be the consequence of a parametric specification, though. There are languages where head-finality is an option, in certain Extended Projections, with a semantic/pragmatic effect. Finnish is a case in point: OV order is an option, with a defocusing effect roughly speaking, possible in clauses where a constituent has undergone focus fronting (for instance a wh-phrase); see Vilkuna (1995), Holmberg (2000), and the examples in (15) (= (11) in BHR):

- (15) a. Milloin Jussi olisi kirjoittanut romaanin? [Aux [VO]]
when Jussi would-have written a-novel
- b. Milloin Jussi olisi romaanin kirjoittanut? [Aux [OV]]
when Jussi would-have a-novel written

c. Milloin Jussi romaanin kirjoittanut olisi? [[OV] Aux]

when Jussi a-novel written would-have

‘When would Jussi have written a novel?’

d. *Milloin Jussi kirjoittanut romaanin olisi? [[*VO] Aux]

when Jussi written a-novel would-have

As noted in BHR section 2.1.2, FOFC is respected in this case, just as in languages with obligatory OV order.⁴ In a case like this, \wedge is not an inherent property of V, but can be optionally added to it. When it is, it is inherited along the spine of the Extended Projection, just like inherent \wedge . We take it that optional \wedge must be licensed by a discourse effect, along the lines of the Fox-Reinhart conjecture (see Chomsky 2001:34), and that optional \wedge may in principle be associated with any of the features $u\phi$, EF, or $[\pm V]$; see BHR (68) (cf. Biberauer and Richards 2006 and Biberauer 2010 for discussion of ϕ - and EF-related optionality).

A further point concerns different kinds of head-final systems. It is well-known that West Germanic differs from, for example, Japanese or Korean with regard to the amount of “leaking” of vP/VP-internal material into postverbal position which is allowed: Japanese and Korean allow almost no leaking, while Germanic requires finite CPs to leak and allows various other categories (some PPs, some non-finite clauses, “heavy” DPs, etc.) to do so. We will not address this issue here, since it is not germane to FOFC, aside from in one respect which we have already mentioned. In terms of the approach to linearization we have adopted above, we would say that in German \wedge is inherited through the clausal functional system as far as T. In Japanese, on the other hand, it is also inherited by C. This parametric difference gives rise to final complementisers in Japanese but initial ones in German. But it also has a further, indirect consequence, given FOFC: finite CPs in German must “extrapose”, i.e. leak,

in order to avoid a FOFC violation (see sections 2.3 in BHR and 1.3 for illustration and some discussion of the possible nature of “extraposition”). Clearly there is more to be said about different kinds of head-final systems in relation to FOFC, but we will leave this discussion aside for reasons of space (see Biberauer and Sheehan in press for discussion of a generatively oriented typology of OV languages).

3. A processing account for FOFC?

It may seem initially plausible that FOFC could have an explanation in terms of principles of online processing (see, for example, Cecchetto in press). Hawkins (1990a, 1994, 2004), in particular, has argued that Greenberg’s word-order universals and other cross-linguistic word-order tendencies are explicable in terms of principles of efficient processing. Human languages, Hawkins argues, have been shaped over time to make processing of linguistically communicated meaning as efficient as possible, and the preference for harmony is one effect of this; Hawkins refers to this as the “Performance-Grammar Correspondence Hypothesis” (see also Newmeyer 2005:119ff.). More precisely, the preference for cross-categorial harmony would be an effect of the human speech processor preferring shorter processing domains, given a dependency relation between two elements in a syntactic structure; see Hawkins (2004:33ff). In harmonic structures, the selecting head and the selected head are (typically) adjacent and so, in terms of left-to-right parsing, the distance between them is as small as it can be; in disharmonic structures, by contrast, there will be items intervening between the two heads. A challenge for a processing-based theory of word-order universals of this type is to explain why the FOFC-compliant disharmonic order is more common than the FOFC-violating order, which is non-existent. Hawkins (1994) is in part designed to explain it in terms of his Early Immediate Constituents principle (see Hawkins 1990b,

1994:95ff). In a recent paper, Hawkins (in press) argues that the FOFC-compliant disharmonic order is, in fact, only marginally more common than the FOFC-violating one, and he articulates a theory, based on Hawkins (2004), which on the one hand explains the strong preference for harmonic structure (in terms of minimization of processing domains) and the slight preference for FOFC-compliant disharmony over FOFC-violating disharmony. The latter would be an effect of a slight processing dispreference for head-final complementation (merging a phrase with a final head, as in $[_{VP} VO] Aux$, where VP merges to Aux) in general.

The typological data discussed by Hawkins (in press) do indeed show only a slight preference for FOFC-compliant orders over FOFC-violating ones. But he reaches this conclusion on the basis of certain disharmonic orders that do not fall under FOFC at all, such as disharmonic $[V [NP P]]$ vs. $[[P NP] V]$ order and disharmonic $[V [Poss N]]$ order vs. $[[N Poss] V]$ order. As discussed in BHR section 3.1, these do not fall under FOFC as the PPs and NPs in question are in different Extended Projections from V. So we do not predict a difference here; the fact that both orders are rare is due to the uncontroversial preference for harmonic orders; to the extent that FOFC is a constraint imposing partial harmony in relation to the distribution of head-final order, as discussed above, it is another manifestation of this general preference.⁵ On the other hand $[C [VP T]]$, another case of FOFC-compliant disharmony, is only slightly less common than the harmonic head-final order $[[VP T] C]$, while the FOFC-violating $[[T VP] C]$ is non-existent; see BHR section 2.3 and the references given there. So, while both disharmonic orders are much rarer than the harmonic orders, there is *pace* Hawkins, a difference between the two orders: the FOFC-compliant order appears to be merely rare (presumably owing to the general preference for harmony), while the FOFC-violating one is non-existent. See Sheehan (in press, b) for more detailed discussion of Hawkins's processing theory as an explanation of FOFC..

Another recent theory which addresses FOFC is Philip (in press). Like Hawkins (in press), and based largely on the same typological data, Philip denies that there is a significant difference between the two kinds of disharmonic order. Instead, both are rare because they violate a constraint, expressed in Optimality-theoretic terms as a universal ‘soft’ constraint which can be overridden by higher ranked constraints, but which appears to be universally highly ranked. The constraint in question requires heads to be adjacent, thus favouring harmonic structures, head-initial or head-final, over disharmonic ones. She discusses various cases of violations of head-proximity in terms of variation in the ranking of a small number of constraints. An interesting claim that Philip makes is that heads with semantic content can have their own ordering constraints, but purely relational heads (so called “linkers”) can’t. This accounts for some of the counterexamples to FOFC discussed above; Philip discusses the case of final negation and final question particles in VO languages. These are contentful categories which can have their own ordering restrictions. Purely subordinating complementizers have to obey the universal constraints, though, including head proximity. This, according to Philip, explains why [[C TP] V], an instance of FOFC-violating disharmony, as well as [V [TP C]], an instance of FOFC-compliant disharmony, are not just rare, but non-existent. But this overlooks the fact that [V [TP C]], although not itself a FOFC-violation, must contain one since either TP or some phrase it contains (vP or VP) will be head-initial in the complement to a final head (C or T). Therefore we expect this order too to be non-existent. It is also unexpected, for Philip’s approach, but not ours, that [C [VP T]], instantiating FOFC-compliant disharmony, is almost as common as the harmonic, head-proximity-respecting order [[VP T] C], while the FOFC-violating order [[T VP] C] is non-existent. The former does not violate FOFC, but the latter does.

References

For references, see BHR.

¹ The only exception to this generalisation are structures in which head-initial CPs have undergone scrambling, i.e. A'-movement. See BHR note 14 and the references given there.

² The offending phrase in (10c), *maitten välillä* 'between the countries' is an adjunct of the noun rather than a complement (see BHR section 2.4, including note 17). This makes no difference to the argument here. In the absence of \wedge on N, the adjunct is linearized postnominally, and the FOFC-violating structure/order (10c) is underivable.

³ We represent the complement of P in (11) as NP rather than DP, to remain neutral with regard to whether nominal arguments in an article-less language like Finnish are DPs or NPs. A complicating fact is that an initial demonstrative (or quantifier) on the complement of an adposition does not induce a FOFC-violation: (i) is perfectly well formed.

- (i) tämän rajan yli
this border across
'across this border'

Nor does the demonstrative alter the effect of a postnominal complement.

- (ii) *[tämän rajan maitten välillä] yli
this border countries between across

The demonstrative does not behave like a head. This could be because the demonstrative is in a Specifier position (SpecDP for example). See Roberts (forthcoming, a) and the references given there for relevant discussion.

⁴ It may be noted, though, that the FOFC-effect all but disappears when the verb is contrastively stressed and the object is a weakly stressed pronoun (Maria Vilkuna, pers. com.).

(i) (?)Milloin Jussi KIRJOITTANUT sen olisi?

when Jussi written it would-have

‘When would Jussi have WRITTEN it?’

Most likely this case falls under the exceptions discussed in BHR section 2.2: the verb and the weak pronoun have moved independently, and do not make up a constituent in their derived position.

⁵ See also our discussion of NPs in BHR section 2.1.4, where we noted Cinque’s (2005) observation that the order Dem>N>Num, which, given Hawkins’ assumptions, would be a case of FOFC-compliant disharmony with the structure [Dem [NP Num]], occurs in “few/very few languages”, while N>Num>Dem, which would be a case of FOFC-violating disharmony, apparently occurs in no language.