Linearization and the Architecture of Grammar: A view from the Final-over-Final Constraint*

THERESA BIBERAUER[1,2], ANDERS HOLMBERG[3] & IAN ROBERTS[1]

[1]University of Cambridge, [2]Stellenbosch University and [3]Newcastle University

mtb23@cam.ac.uk, Anders.Holmberg@ncl.ac.uk, igr20@cam.ac.uk

This paper addresses the issue of the locus of linearization information in the context of a minimalist grammar. Contrary to what is arguably the dominant view in minimalist theorizing today, it is argued that linearization information must in fact be specified Narrow Syntax-internally. The imperative underlying this conclusion is an empirical skewing in the domain of word-order variation, in terms of which head-initial structures associated with a given projection line may only be (harmonically) dominated by head-initial structures, while head-final structures may be dominated either by head-initial or head-final structures in the same context – the so-called *Final-over-Final Constraint* (FOFC). It is argued that attested FOFC effects suggest that linearization information is in fact encoded in such a way – namely, by harnessing an already-required movement diacritic in accordance with Relativized Minimality, arguably a third-factor-imposed principle – that its NS-internal presence does not violate the Strong Minimalist Thesis of Chomsky (2001 *et seq.*). We also consider the question syntactic categories and their formal status, against this background.

1. Introduction

This paper is primarily concerned with linear ordering and its locus in the architecture of grammar. During the GB era, syntactic structure was generally assumed to involve both hierarchy and fixed linear ordering, with the former falling out as the consequence of a principle of UG (X-bar theory) and the latter following from the setting of universally given parameters (e.g. the Head Parameter). In the context of Minimalism, by contrast, there presently appears to be a fairly strong consensus that linear ordering is only established at PF (cf. i.a. Berwick & Chomsky 2008, Boeckx 2008 and Richards 2009). Further, it is often asserted that the language faculty exhibits an "LF bias", with the mapping from syntax to SEM conforming to the Strong Minimalist Thesis (SMT; cf. Chomsky 2001) in (1), while that from syntax to PHON does not.

* The research reported here is funded by AHRC Grant AH/E009239/1: "Structure and Linearization in Disharmonic Word Orders". For valuable input on the ideas central to this paper we thank in particular Angel Gallego, John Hawkins, Neil Myler, and Michelle Sheehan, and also the IGG35 and GLOW32 audiences. All usual disclaimers apply.

(1) The **Strong Minimalist Thesis**: Language is an optimal solution to legibility conditions, i.e. it is a "perfect", maximally efficient solution to the conditions imposed by the interfaces, SEM and PHON.

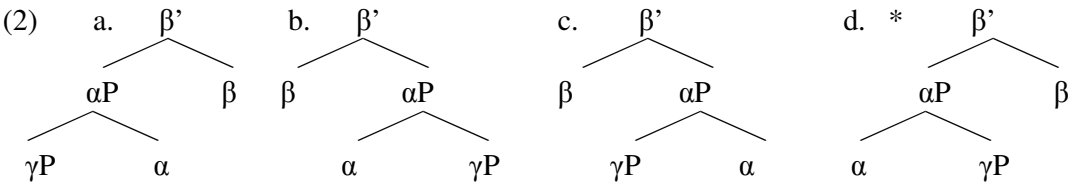
The purpose of this paper is to take issue with these views, and to argue, firstly, that the commonly held "good design" expectation that linearization only comes into play at PF cannot in fact be correct and, secondly, that this fact undermines the view that PF, unlike LF, simply has to "make do" with the incomplete structural information that is fed to it. Our argument is based on a striking empirical asymmetry in the domain of attested word-order patterns, one which is not amenable to a parsimonious "late linearization" account.

The paper is structured as follows: section 2 introduces the empirical facts and the word-order constraint that lead us to reconsider the role of linearization information in Narrow Syntax; section 3 outlines the proposed analysis of these facts; finally, section 4 considers the implications of the analysis and concludes.

2. A word-order asymmetry: introducing the Final-over-Final Constraint

Building on Holmberg (2000), Biberauer, Holmberg & Roberts (2007; BHR) observe a skewing in the word-order patterns attested in the world's languages: while both types of "harmonic" order (i.e. consistently initial and consistently final) are readily found, only one of the expected "disharmonic" types surfaces, namely that involving head-initial phrases dominating head-final ones. This asymmetry is schematized in (2):

(2)



Consistent head-final **Consistent head-initial** **Initial-over-Final** **Final-over-Initial**

Evidence of the skewing in the disharmonic domain comes from a wide range of structures, discussed in more detail in BHR and also Biberauer, Newton & Sheehan (2009a,b) and Biberauer, Holmberg & Roberts (2009). Here we mention only three: (i) the apparent crosslinguistic absence of VOAux orderings, and (ii) the absence of VO languages with initial complementisers, and (iii) the diachronic pathways which languages appear to follow during the process of word-order change. As the discussion in section 3.3 will show, these empirical facts can be understood as reflexes of a general constraint, which may be formulated as follows:

(3) **The Final-Over-Final Constraint (FOFC)**
For all heads {α, β, ..} on a single projection line, if α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final.

*2.1. *VOAux*

Many Germanic varieties permit both harmonic and disharmonic permutations of Aux, V and O. Thus spoken Afrikaans, for example, allows both of the orders in (4a,b) and English-influenced Kaaps additionally permits (4c); no variety of Afrikaans, however, allows (4d), the FOFC-violating order:

(4) a. … dat sy [$_{DP}$ 'n brief] *geskryf* **het** [OV-Aux]
that she a letter written has
'… that she has written a letter'
b. … dat sy **het** [$_{DP}$'n brief] *geskryf* [Aux-OV]
c. … dat sy **het** *geskryf* [$_{DP}$'n brief] [Aux-VO]
d. *… dat sy *geskryf* [$_{DP}$'n brief] **het** [VO-Aux]

This pattern is replicated in other Germanic varieties, both modern and historical, in Finno-Ugric, and also in Basque and Italian Sign Language (cf. Cecchetto 2009). More generally, it appears to be the case that VOAux structures are exclusively permitted in languages featuring non-inflecting auxiliary elements, commonly designated *particles* (cf. i.a. Dryer 2009b for discussion of Niger-Congo languages permitting this structure). That particle-containing VOAux structures should not be viewed as counter-examples to (3) is strongly suggested by data such as the following:

(5) a. yə- ca dɛyo **lɔ** [Bwe-Karen]
1SG-see picture ASP
"I am looking at a picture"
b. ce-dɔ mi **jə-khɔ** phi má nɔ (***jə-khɔ**)
3- say COMP 3- FUT take what
"What did he say that he would take?" (data from Dryer 2009a)

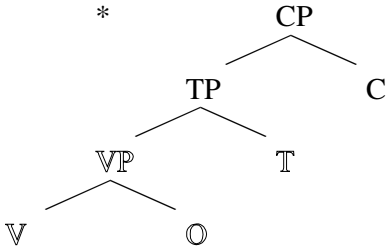
Here we see that non-inflecting particles expressing tense-aspectual (auxiliary) information necessarily occupy a very different position to that in which inflecting auxiliaries obligatorily surface. Significantly, the FOFC-violating final position is never available to the latter, the same pattern that we see in more familiar European languages, all of which feature non-particle auxiliaries. This suggests that inflecting and particle elements are formally distinct in a way that is crucially relevant for FOFC. We return to this point in section 3.3 below.

2.2 The cross-linguistic absence of VO languages with final complementisers

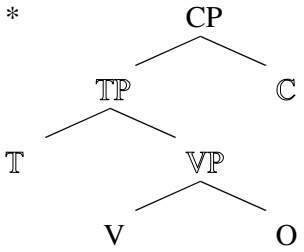
An oft-noted fact about the distribution of subordinating conjunctions (e.g. *that*) is that VO languages systematically lack sentence-final complementisers (Cs; cf. i.a. Hawkins 1990: 256-257, Dryer 1992: 102; 2009a). By contrast, many OV languages have initial Cs. According to the on-line *World Atlas of Language Structures* (*WALS*; Haspelmath et al. 2008), 54 clearly OV languages (out of a sample of 64) feature an initial adverbial subordinator ("because"), and this excludes familiar OV languages like German and Dutch, which, on account of their matrix V2 property, are listed as languages with "no dominant order"; only 2 VO languages are said to feature final adverbial subordinators, but see Newton (2008) for a critique of the descriptions underlying this classification. Evidently, then, there is a skewing in the distribution of complementisers.

On the assumption that C is on the projection line of V (cf. Grimshaw 1991),[1] the fact that VO languages systematically lack final Cs follows directly from (3), as the following diagrams show:

(6) a. *[$_{CP}$ [$_{TP}$ [$_{VP}$ V O] T] C] -- violates FOFC for α = V and β = T



b. *[$_{CP}$ [$_{TP}$ T [$_{VP}$ V O]] C] -- violates FOFC for α = T and β = C



Worth noting here is that it appears to be necessary to draw a formal distinction between subordinating elements of the type illustrated above and so-called *C particles* such as the force particles in i.a. the Chinese dialects (Paul 2009), the Northern Italian dialects (Munaro & Poletto 2006) and Gungbe (Aboh 2006). Firstly, the latter are very commonly restricted to matrix contexts and secondly, comparison of the placement of these elements indicates that subordinators consistently surface in the expected initial position, while the particles surface in apparently FOFC-violating positions. The examples from Vietnamese below illustrate:

(7) a. Tân mua gi **the**?
Tan buy what PRT
'What did Tan buy?'

b. Anh đã nói (**rằng**) cô ta không tin
PRN ANT say that PRN NEG PRT believe
'He said that she didn't believe (him)'

Once again, then, there are indications that particles are, in some sense, special. Pending further research into their properties, we leave them aside here.

2.3 Diachronic evidence

Absolute synchronic constraints are expected to have diachronic consequences. Thus if FOFC represents an absolute universal, we expect word-order change to proceed along certain pathways, specifically:

[1] In the sense that it is part of the clausal functional sequence, it is clear that C should count as part of the extended projection of V. The fact that complementisers are sensitive to verbal properties such as finiteness further reinforces this impression.

Head-final to head-initial ("OV" to "VO") change must proceed "top-down":
(8) [[[O V] I] C] → [C [[O V] I]] → [C [I [O V]]] → [C [I [V O]]].

Head-initial to head-final ("VO" to "OV") change must proceed "bottom-up":
(9) [C [I [V O]]] → [C [I [O V]]] → [C [[O V] I]] → [[[O V] I] C].

Any other route entails FOFC violations at intermediate stages. Biberauer, Newton & Sheehan (2009a,b) present case studies from the history of Germanic and Ethiopian-Semitic, showing that directionality changes in languages belonging to these families conform to the expected FOFC-determined pathways. Further, it appears that Niger-Congo languages that have undergone varying amounts of initial to final change have likewise done so in the predicted "bottom-up" fashion (cf. Nikitina 2008 for recent discussion); similarly, the Sami languages within Finno-Ugric, appear to have undergone the reverse change "top-down".

3. Accounting for the asymmetry: a closer look at linearisation

The previous section has shown that there appears to be a striking asymmetry in the domain of disharmonic word orders: while initial-over-final orders involving elements on a single projection line are relatively common, final-over-initial orders fail to surface. The same skewing is notably absent where structures involving elements associated with different projection lines occur together. Consider German (10):

(10) … dass sie gerne [$_{DP}$ ein Täschen Kaffee] trinkt
that she gladly a cup-DIM coffee drinks
"… that she enjoys drinking a cup of coffee"

Here a head-initial DP is dominated by a head-final VP, a pattern which is just as common among the world's languages as its inverse (cf. BNS 2009b for discussion and references). Since nouns define extended projections independently of the verbs with which they combine, (3) leads us to expect the absence of an asymmetry in this cross-categorial domain. A similar explanation may also account for the availability of head-initial PPs in V-final languages. Crucially for present purposes, the fact that the asymmetry in disharmonic word orders is not an across-the-board phenomenon suggests that an "external" account in terms of processing constraints is likely to be problematic (see, however, Cecchetto 2009, for a proposal along these lines[2]). Similarly, the observed facts and the added complication introduced by the non-total nature of the asymmetry make it unclear how a Head Parameter (HP)-based account would be able to rule out the problematic orders without stipulation. In the minimalist context, this is significant: as noted in the introduction, the view that linearization information, like phonological specifications more generally, has no place in NS and should therefore only be imposed at PF, possibly via a PF parameter, is widespread (cf. Biberauer 2008a and Richards 2009 for recent overview discussion and references). It is our contention, however, that the empirical facts mentioned above and the apparent universality of (3) fatally undermine the validity of this assumption. In what follows, we outline an analysis which facilitates a new and, importantly, FOFC-compatible perspective on linearization.

[2] John Hawkins (p.c.) confirms that his influential processing theory (Hawkins 1994, 2004) would not seem to offer a ready explanation for the observed facts.

3.1 Theoretical background: the Probe-Goal-Agree system

In terms of the Probe-Goal-Agree approach to syntactic derivations developed since Chomsky (2001), NS-internal movement is triggered by (generalized) EPP-features. Crucially, these features must be thought of as distinct from the formal features involved in Agree operations: unvalued features on a given head (loosely designated the *Probe*) may probe corresponding valued features on one or more heads (loosely designated the *Goal*) in their c-command domain, thereby effecting Agree operations. Importantly, feature valuation, which is the outcome of Agree, is therefore not dependent on the creation, via movement, of specific local configurations (Spec-Head or Head-Head), as was the case in earlier checking-based theories. Agree-driven movement, instead, only takes place if a given probe is associated with an EPP-feature. As pointed out by Pesetsky & Torrego (2001), movement triggers may therefore essentially be viewed as "a feature-of-a-feature", i.e. as diacritics.

Movement need, however, not always be Agree-related; phase-heads specifically are assumed to be able to trigger *non*-Agree-mediated movement of elements to their edge, via so-called Edge Features/EFs[3], which they may or may not "spread" to the heads they select. We assume these EFs to be identical to the generalized EPP-features mentioned above and henceforth represent movement triggers as ^.

3.2 Linearization and movement

The apparatus outlined in the previous section leads us to expect the following types of movement:

(11) a. Agree-driven movement: e.g. v [ϕ^], where v's ϕ-probe is associated with a movement trigger.
b. Non-Agree-driven movement: e.g. v^, where ^ is a free-standing trigger not specifically associated with any of a head's contentive features (see note 3).

Building on insights in Rizzi (2008), we suggest that there is in fact a third species of movement, namely Selection-driven movement (cf. also Holmberg 2000 and Julien 2002 for early proposals along these lines; Pesetsky & Torrego 2006 and Cecchetto & Donati 2009 also assume Selection to involve probing, although their proposal differs from that made here). For Rizzi, Agree may be viewed as Internal Search, whereas Selection is External Search; a head can therefore be thought to probe either the structure which it c-commands or the Lexical Array which is at that time feeding the derivation. Since the features triggering Internal Search are uncontroversially assumed to have the option, subject to parametric variation, of being associated with ^, the question arises why the same should not be possible for those triggering External Search. More specifically, if External Search is in fact driven by c-selection features, we might expect these features, like their Agree-triggering counterparts, also to have the parametrically determined option of being associated with ^ or not. Our contention is that this option does in fact exist, and that the non-availability of this mode of

[3] Worth noting here is that the phase head-related EFs discussed here should not be confused with the generalised Merge features, also designated *Edge Features*, ascribed to every lexical item in Chomsky (2006 *et seq.*): as languages do not differ in respect of the fact that their lexical items may undergo External Merge/EM, whereas they do differ in respect of whether already-merged, and thus EF-bearing, items can trigger movement (Internal Merge/IM), it may be necessary to draw a distinction (*contra* Chomsky 2006:17, 2008:144). We leave open the possibility that non-Agree-driven movement simply involves a head associated with two EFs, i.e. an EM-triggering EF which bears a further IM-triggering EF as a secondary feature.

movement, just like the non-availability of movement more generally (cf. Chomsky 2005), would have to be *stipulated*. Specifically, we propose that c-selection-driven movement triggers comp-to-spec movement, with the selected XP raising into the specifier of its selector.[4] Further, we argue that interpreting ^ in the usual manner as a trigger for *leftward* movement paves the way for a principled account of the FOFC constraint, one which has important implications for our understanding of the architecture of the grammar.

3.3 Accounting for the asymmetry

Assuming the movement options outlined in (11) plus the existence of Selection-driven movement, we are led to expect that languages may differ in respect of:

(a) whether given heads feature or lack ^; and, if ^ is present,
(b) whether ^ is free-standing or associated with Selection or with Agree features.

If ^-distribution were completely free, with heads being able to bear or lack ^ independently of one another and independently of the nature of ^'s association, the crosslinguistic occurrence of "harmonic" patterns would reduce to an unexplained coincidence: in the context of the system assumed here, "harmony" falls out from heads systematically lacking or bearing ^ (head-initial and head-final systems, respectively). Similarly, the discrepancy that lies at the heart of FOFC cannot be accounted for if heads may either bear or lack c-selection-related ^ independently of one another. To capture the observed skewings, a formal constraint on the distribution of (c-selected) ^ (henceforth: ^) along the lines of (12) appears to be necessary:

(12) If a non-lexical head X^n in the extended projection E of a lexical head L has ^ associated with its selection feature for a lower head X^{n-1}, then so does X^{n-1}.

In terms of (12), higher heads on a given projection line may only bear ^ if the lower heads on that projection line do so. If v is v^, V will therefore also have to be V^. Where auxiliaries are v-elements, this constraint ensures that VOAux orders will remain unattested, this order requiring v^ to select V, in contravention of (12). By transitivity, the same is true of cases where auxiliaries are T-elements. VOAux orders derived via c-selection-driven movement are therefore ruled out, as desired. By contrast, the inverse disharmonic order – AuxOV – is ruled in since this order results when V bears ^, but v lacks this diacritic. Assuming C to be part of the V's extended projection, VOC is likewise ruled out: this case would involve C^ dominating ^-less lower clausal heads, in contravention of (12). Frequently attested COV (cf. Latin, West Germanic, Turkish, etc.) is, however, expected to be permitted since this ordering entails (a) lower head(s) bearing ^, while C lacks this diacritic. (12), then, offers a formal account of the synchronically attested FOFC asymmetries discussed in sections 2.1 and 2.2 above. The diachronic facts noted in section 2.3 likewise follow since we expect (12) to determine the shape of all synchronic grammars including the successive grammars involved in change scenarios; intermediate grammars may not, therefore, entail a violation of (12).

(12) also allows us to understand why cases in which FOFC appears to be violated would arise. Consider first the case of V selecting DP. Unlike V and v/T/C, V

[4] C-selection-driven movement therefore violates anti-locality (cf. i.a. Pesetsky & Torrego 2001 and Abels 2003). To the extent that anti-locality still holds of *non*-c-selection-related movements, earlier arguments contra anti-locality stand, with the anti-locality violation in the c-selection case being justified in SMT terms as it entails that an already required movement diacritic (^) may also be harnessed to signal directionality information (see section 4 for further discussion).

and D are not part of the same projection line: while the former are all heads on the clausal projection line, – the main "spine" in a clausal structure – D represents a functional head associated with a distinct "spine", the extended projection of N. As such, (12) does not lead us to expect V^/v^ dominating D to be any more problematic than V/v dominating D^, precisely what the typological record suggests.

To the extent that Ps are lexical categories which therefore initiate extended projections independently of N and V, we are also led to expect both V^/v^ and P-containing grammars, i.e. grammars in which initial PPs are dominated by final V/vPs, and the inverse, non-FOFC-violating word order, i.e. initial V/vPs with final PPs. This seems to be correct (cf. Haspelmath et al. 2008), although it is worth noting that P-directionality does appear to harmonise very strongly with V-directionality (cf. also Dryer 1992, 2009a). Also worth noting in connection with Ps is the non-uniform nature of the elements ascribed to this class: while some Ps exhibit formal properties more typical of lexical categories, others appear to instantiate "light"/functional elements (cf. van Riemsdijk 1998, the contributions in Asbury et al. 2008, and much recent work by Peter Svenonius). This discrepancy is particularly evident in the context of circumpositional structures, where the difference in the linearization properties of the two Ps coincides with differences in their semantic, phonological and grammatical properties (essentially, one P is "light"/more functional head-like, the other is "strong"/more lexical head-like). The expectation in the present context is, then, that (more) lexical Ps may exhibit directionality at odds with the nominal and/or verbal heads they dominate and/or are dominated by – since they count as the bottom of a new extended projection – whereas functional Ps may not – since only lexical heads count as bottoms of projections. This prediction remains to be fully tested, but seems correct for familiar West Germanic circumpositional structures.

As noted in sections 2.1 and 2.2, there are numerous contexts in which so-called *particles* – roughly, non-inflecting "little words" – appear to violate FOFC.[5] The question is why this should be so? Given how little is presently understood about the formal properties of particles, it is difficult to answer this question with any degree of confidence. We, however, take data of the type illustrated in (5) and (7) to indicate that particles should not simply be viewed as exponents of the same functional heads as those realised by composite elements, e.g. an auxiliary which expresses not only tense, but also agreement and possibly mood or voice or evidentiality, etc., or a C-element which expresses not just subordination, but also finiteness and possibly also mood and/or evidentiality, etc. Particles, instead, appear to realise sub-heads associated with "cover-term" heads like v, T, C and D (consider, for example, Aboh's detailed (2004, 2006) illustrations of this state of affairs in Gungbe). If particles are indeed a (sub)type of functional head, however, we would expect them to be FOFC-respecting elements, which cannot therefore bear ^ in structures featuring lower heads lacking ^. Worth noting here is that two aspects of the formal realisation of particle elements suggest that a functional-head analysis may not be correct. To see this, it is necessary to clarify what is meant by 'lexical' and 'functional' category in the context of the theoretical framework adopted here.

Taking canonical cases like V and N as our point of departure[6], a reasonable

[5] Worth noting here is the fact that Greenberg (1963), similarly, observed that particles frequently emerge as outliers in the context of the typological word-order generalisations he established. He therefore disregarded them.

[6] In referring to *N* and *V* here, we leave aside the possibility that N and V may in fact be complex heads derived via prior merger with nominalizing and verbalizing heads (cf. Myler 2009 for discussion which

interpretation would be that lexical categories are elements which may only probe externally, i.e. whose NS-relevant featural make-up entails at most a categorial specification (see note 6) and c-selection features, with Internal Agree-related features being entirely absent; in turn, functional categories are those which may probe both internally and externally, in the manner usually assumed. Against this background, particles emerge as a rather unusual type of functional head since their formal realisation systematically fails to give any evidence of the Internal Agree operations they are expected to be involved in: although there is no requirement for PF to spell out NS-internal Agreement operations, it is clear from examples such as those given in (5) and (7) that it is meaningful to draw a distinction between agreement- and non-agreement-reflecting elements, with the latter exhibiting peculiar behaviour that seems to be absent with the former. The extent to which particles are involved in Internal Agree (i.e. standard probing) operations is therefore unclear. Secondly, it is notable that the semantics (and pragmatics) of particles is more idiosyncratic and difficult to pin down than that of non-particle functional elements, a state of affairs which to some extent resembles the situation with lexical elements. In the cases where this is less obviously true (e.g. Focus particles), it is worth noting that the particles in question appear to lack not only Internal Agree-triggering probes, but also c-selection-related (External Agree-triggering) properties: these elements may combine with XPs of any type. As such, the only possible "point of attachment" for ^ would be the EFs of the relevant elements (cf. section 3.1 above). Particles of this type, then, would not violate (12) since their ^ is not c-selection-related. Assuming particles not to be a homogenous category, with the absence of both internal and external probes not being definitive of the entire class, we are left with the question of how particles which do appear to c-select (i.e. those which seem to surface in fixed hierarchical positions) can be accounted for. The proposal here is that these particles may in fact be lexical rather than functional elements, i.e. elements bearing only c-selection, but not Agree-related probes. As lexical elements, they cannot count as part of the extended projection of the structure they dominate, with the result that (12) is once again respected. We leave the details of this matter to future research, noting here only that current theoretical assumptions may facilitate valuable, but until now, unexplored insights into the nature of syntactic categories.

What we have seen so far is that (12) appears to fare well in accounting for the asymmetry discussed in section 2. A question that arises given the discussion of particles, which highlights that head-finality need not only result from c-selection-driven movement, is how unattested orders derived via Agree- or EF-related movement (cf. section 3.2) can be ruled out. Why, for example, is it not possible for SVOAux orders to be derived via one of these alternatives modes of movement, neither of which violates (12)? Starting with EF-related movement, we follow BHR in assuming SVOAux orders derived in this way to be ruled *in*, but, crucially, to be associated with typical A' interpretive properties (focus, etc.). This seems to be correct (see BHR for discussion). As far as Agree-related movement (involving piedpiping) is concerned, we again follow BHR, this time ruling *out* this possibility on the grounds that the phase-based way in which derivations proceed, with material in the complement domain of phase heads undergoing Radical Spellout and therefore being unavailable for subsequent movement operations, precludes the generation of Agree-driven SVOAux structures (again, see BHR for discussion). This last case

---

suggests that verbalizing v and transitivity, etc-determining v cannot be one and the same entity). The point here is crucially that we assume lexical categories to be categorially specified.

highlights the important fact that FOFC effects are not uniquely the consequence of (12), but that they in part also follow from independently motivated architectural considerations.

4. Implications and conclusions

The preceding discussion has argued for a model of the language faculty in terms of which linearization information is present in NS. At first sight, this proposal might seem to conflict with (1) since a system in which purely hierarchical syntactic structure is converted into linear structures at the PF interface, i.e. at the point at which it is first *needed*, may intuitively seem like the most economically designed one; having the information available throughout the derivation, where it serves no purpose, seems unnecessary (another instance of "look ahead"), particularly if a principled mechanism can be uncovered to govern hierarchy-linearisation conversion.

We have argued, however, that a striking skewing in the typological record suggests that the language faculty cannot in fact be constructed in this way: head-directionality options are subject to the constraint in (12), which entails that the directionality of individual heads is dependent on the directionality of certain others. It is not clear how the effects of this constraint may be imposed at PF, other than by stipulation. One could, of course, interpret ^ as a PF diacritic, interpreted as signalling that the complement of a ^-bearing head be linearised to its left. This would, however, entail the postulation of a specifically PF-legible diacritic, i.e. a special device relevant to PF only. Furthermore, it would leave unexplained why ^ should signal leftward and not rightward linearization, with the question of whether this system implies that head-final structures are "more marked" than head-initial ones also arising. By contrast, (12) viewed as an NS-internal constraint, drawing on devices evidently required by the computational system, emerges as a very natural constraint.

Firstly, it references c-selection features and the movement diacritic (^), both of which are independently required, the latter to enable the computational system to construct both discourse-neutral and discourse-marked structures (cf. the so-called *duality of semantics* property of natural language syntax). Secondly, since movement is always leftward (cf. Abels 2008 for overview discussion), the fact that ^ should signal leftward placement of complements also follows directly: linearization is just another movement-derived effect. Thirdly, the constraint in (12) in its essence states that the distribution of ^ is subject to a "no skipping" constraint of the type that is familiar from Relativised Minimality (RM, Rizzi 1990, 2001): just like other movement types (Agree-driven, non-Agree-driven, etc.) cannot "skip" an intermediate position of the same type, the distribution of c-selection-related ^ cannot "skip" an intermediate c-selecting head within the same extended projection. All three of the movement types discussed in section 3.2 and 3.3 above therefore respect RM, which appears to be a syntax-internal constraint, though undoubtedly one imposed by third-factor considerations (cf. BHR 2009 for more detailed discussion). Furthermore, viewing linearization as the consequence of a species of movement operation opens the possibility of moving beyond the view that (partially) head-final languages are somehow "more marked" than their head-initial counterparts: if linearization-related movement is simply one species of movement, it is clear that the "calculation" as to how "marked" a language is in movement terms should include not just linearization movements, but movement operations *tout court*. In this context, the oft-observed fact that OV languages tend to be *wh*-in situ takes on a new significance.

The picture that emerges, then, is of a system which satisfies the SMT by

harnessing the movement diacritic it independently requires to facilitate "duality of semantics" not only to fulfil this design requirement on natural language, but also to solve a further design requirement, namely that of ultimately delivering linearised strings that can be externalised in a temporal dimension. Viewed from this perspective, the NS-internal presence of linearization information can be viewed as the reflex of a system which is constructed so as to make maximally economical use of its necessary components, the diametric opposite of the standardly held view. Evidently, many aspects of the proposal outlined here require more detailed, systematic working out, something which space constraints preclude in the present context (but cf. BHR 2009). The aim here was simply to outline some low- and high-level architectural consequences of rejecting the currently dominant minimalist view that linearization information has no place in syntax; if the ideas in this paper are on the right track, there would appear to be no need to assume the currently much-discussed "LF bias" in the context of the SMT.

References

Abels, K. (2003) Successive Cyclicity, Anti-Locality and Adposition Stranding. Ph.D. dissertation: UConn.

Abels, K. (2008) Towards a restrictive theory of (remnant) movement. *Linguistic Variation Yearbook* 7: 53-120.

Aboh, E. (2004). *The Morphosyntax of Complement-head Sequences: Clause Structure and Word Order Patterns in Kwa.* New York: OUP.

Aboh, E. (2006) Complementation in Saramaccan and Gungbe: The case of C-type modal particles. *Natural Language and Linguistic Theory 24*, 1-55.

Asbury, A., J. Dotlačil, B. Gehrke & R. Nouwen (2008) *Syntax and Semantics of Spatial P*. Amsterdam: Benjamins.

Berwick, R. & N. Chomsky (2008) Biolinguistics: Its current state of development. Unpublished ms: MIT.

Biberauer, T. (2008) Introduction. In T. Biberauer (Ed.) *The limits of syntactic variation*, (pp.1-72). Amsterdam: Benjamins.

Biberauer, T., A. Holmberg & I. Roberts (2007). Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC). In A. Bisetto & F. Barbieri (Eds.) *Proceedings of XXXIII Incontro di Grammatica Generativa*.

Biberauer, T., A. Holmberg & I. Roberts (2009). A hierarchical universal and its consequences. Talk given at GLOW 32 (Nantes).

Biberauer, T., G. Newton & M. Sheehan (2009a) Limiting synchronic and diachronic variation and change: the Final-Over-Final Constraint. To appear in *Language and Linguistics*.

Biberauer, T., G. Newton & M. Sheehan (2009b) Impossible changes and impossible borrowings. To appear in A. Breitbarth, C. Lucas, S. Watts & D. Willis (Eds.) *Continuity and Change in Grammar*. Amsterdam: Benjamins.

Boeckx, C. (2008) Approaching parameters from below. Unpublished ms: Harvard.

Cecchetto, C. (2009) Backward dependencies must be short. A unified account of the Final-over-Final and the Right Roof Constraints and its consequences for the syntax/morphology interface. To appear in T. Biberauer & I. Roberts (Eds.) *Challenges to Linearisation*. Berlin: Mouton.

Cecchetto, C. & C. Donati (2009). On Labeling: Principle C and Head Movement. To appear in *Syntax*.

Chomsky, N. (2001) Derivation by Phase. In: M. Kenstowicz (Ed.) *Ken Hale. A Life in Language*, (pp.1-52). Cambridge, MA.: MIT Press.

Chomsky, N. (2006). Approaching UG from below. In U. Sauerland & H-M. Gärtner (Eds.). *Interfaces + recursion = language?*, (pp.1-29) Berlin: Mouton de Gruyter.

Chomsky, N. (2008). On phases. In R. Freiden, C. Otero & M-L. Zubizarreta (Eds.) *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, (pp.133-166) Cambridge, MA.: MIT Press.

Dryer, M. (1992) On the Greenbergian Word-Order Correlations. *Language* 68: 81-138.

Dryer, M. (2009a) The branching direction theory revisited. In S. Scalise, E. Magni, & A. Bisetto (Eds.) *Universals of Language Today*, (pp. 185-207). Berlin: Springer.

Dryer, M. (2009b) Verb-Object-Negative Order in Central Africa. In N. Cyffer., E. Ebermann & G. Ziegelmeyer (eds.). *Negation patterns in West* Africa. Amsterdam: Benjamins.

Grimshaw, J. (1991) Extended projection. Unpublished ms: Brandeis University.

Haspelmath, M., M. Dryer, D. Gil & B. Comrie (Eds.) (2008) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library.

Hawkins, J. (Ed.) (1990). *Explaining Language Universals.* Oxford: Blackwell.

Hawkins, J. (1994). *A Performance theory of Order and Constituency.* Cambridge: CUP.

Hawkins, J. (2004). *Efficiency and Complexity in Grammars.* Oxford: OUP.

Holmberg, A. (2000). Deriving OV Order in Finnish. In P. Svenonius (Ed.) *The Derivation of VO and OV*, (pp.123-152). Oxford/New York: OUP.

Julien, M. (2002). *Syntactic heads and word formation*. New York: OUP.

Kayne, R. (1994). *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Munaro, N. and C. Poletto (2006) The diachronic path towards particles in the North-Eastern dialects. *Nordic Journal of Linguistics 28*: 247-268.

Myler, N. (2009) *Form, function and explanation at the syntax-morphology interface: agreement, agglutination and post-syntactic operations.* M.Phil. dissertation: University of Cambridge.

Newton, G. (2008) Final C elements: complementisers and particles. Handout: University of Cambridge.

Nikitina, T. (2008) Nominalization and word order change in Niger-Congo. Unpublished ms: Stanford.

Paul, W. (2009) Consistent Disharmony: Sentence-final particles in Chinese. To appear in T. Biberauer (Ed.) *Cambridge Occasional Papers in Linguistics*.

Pesetsky, D. & E. Torrego (2001) T-to-C: Causes and Consequences. In M. Kenstowicz (Ed.) *Ken Hale: A Life in Language*, (pp. 355-426). Cambridge MA: MIT Press.

Pesetsky, D. & E. Torrego (2006). Probes, Goals and Syntactic Categories. To appear in *Proceedings of the 7th Annual Tokyo Conference on Psycholinguistics*.

Richards, M. (2009) Two Kinds of Variation in a Minimalist System. Unpublished ms: Leipzig.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. (2001). Relativized Minimality effects. In: M. Baltin & C. Collins (eds.). *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 89-110.

Rizzi, L. (2008). *On Delimiting Movement*. Paper presented at Cambridge-Connecticut-Hyderabad-Nanzan-Tsingua-Hua Colloquium(Tsing-Hua, Taiwan).

Van Riemsdijk, H. (1998) Categorial feature magnetism: The endocentricity and distribution of projections. *Journal of Comparative Germanic Linguistics* 2: 1–48.