A syntactic universal and its consequences[1]

Theresa Biberauer, Anders Holmberg & Ian Roberts

University of Cambridge/Stellenbosch University, Newcastle University and

University of Cambridge

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

Abstract

This paper investigates the Final-over-Final Constraint (FOFC): a head-initial category cannot be the immediate structural complement of a head-final category within the same extended projection. First, we document the empirical evidence, logically possible but cross-linguistically unattested combinations of head-final and head-initial orders. Second, we formulate FOFC in terms of Extended Projections (Grimshaw 1991, 2001, 2005). Third, we reduce FOFC to the LCA, combined with a constraint on the formation of Extended Projections which we ultimately reduce to Relativised Minimality. Finally, we suggest that our approach, although it entails a minimal amount of linearization information in narrow syntax, nonetheless complies with the Strong Minimalist Thesis.

Keywords: linearization, relativised minimality, word order, universal, Extended Projection

[11] The work reported here is supported by AHRC Grant No. AH/E009239/1 ("Structure and Linearization in Disharmonic Word Orders"). A special thanks to the other two past and present members of the project group: Glenda Newton and, particularly, Michelle Sheehan. All or part of this material has been presented at: the International Conference on Linguistics in Korea, Seoul; the 33rd and 35th Incontri di Grammatica Generativa, held in Bologna and Siena respectively; the 26th West Coast Conference on Formal Linguistics, held at the University of California, Berkeley; the 18th Colloquium on Generative Grammar, held in Girona; the 6th Asian GLOW, held in Hong Kong; GLOW 32, hosted in Nantes; the Annual Meetings of the Linguistics Association of Great Britain, held at King's College, London and at the University of Essex; the 14th Welsh Syntax Seminar, held in Gregynog; the 23rd Comparative Germanic Syntax Workshop, held in Edinburgh; The Past Meets the Present: A Dialogue Between Historical Linguistics and Theoretical Linguistics conference, held in Taipei, Taiwan; the 10th Diachronic Generative Syntax conference, held at Cornell; the Ways of Structure Building Conference, held in Vitoria; the Deutsche Gesellschaft für Sprachwissenschaft workshop on complementation, held in Osnabrück; the 14th International Lexical Functional Grammar conference, held in Cambridge; and the 40th North East Linguistics Society Meeting held at MIT; and also at talks given at the English and Foreign Languages University in Hyderabad, Nanzan University, Nagoya, Senshu University, Tokyo, the University of Amsterdam, the University of Leiden, the University of Groningen, the University of Konstanz, the University of Essex, the Universitat Autónoma de Barcelona, and at various seminars and SyntaxLab meetings in Cambridge. We would like to thank the audiences at these presentations for their comments and criticisms, especially Enoch Aboh, Alastair Appleton, Norbert Corver, Joseph Emonds, Ángel Gallego, Chiara Gianollo, Elliott Lash, Terje Lohndahl, Iain Mobbs, Neil Myler, David Pesetsky, Henk van Riemsdijk, Sten Vikner and George Walkden.

Earlier versions of this material have appeared as Biberauer, Holmberg & Roberts (2007, 2008, 2009); BHR henceforth. The presentation and analysis here supersede those in BHR.

1. Introduction

In this paper, we investigate a language universal. Like much fruitful recent work, we build on the two principal currents of research on universals which have emerged in past fifty years or so: the Chomskyan tradition, in which the existence of language universals is deduced from the existence of an innate predisposition to language acquisition, and the Greenbergian tradition, in which universals, or at least strong tendencies to common patterning, are observed in increasingly wide-ranging surveys of cross-linguistic data.

In general, attempts to express generalisations emerging from Greenbergian word-order typology using the formal mechanisms of generative syntactic theory have suffered from two main drawbacks. First, the postulation of a simple Head Parameter, along the lines of "within an X', X° precedes/follows its complement YP" (to use the the vocabulary of X'-theory; see Lightfoot 1979, Hawkins 1983) is empirically inadequate. Many languages, including very well-studied ones such as those of the West Germanic and Chinese groups, as well as most of the more archaic Indo-European languages (Latin, Homeric Greek, Sanskrit, etc.), show variation in head-complement orders across categories, having both head-complement and complement-head orders in different categories. Indeed, according to Dryer (1992, note 17), such systems are numerically more preponderant than cross-categorially harmonic ones; see also Cinque (2010), who identifies more "word-order dyads" than is usual in the typological literature and observes that the fully harmonic types are in fact very rare or even nonexistent. Furthermore, as we will try to show here, an asymmetry exists within the disharmonic systems: certain types of disharmonic order are found, if rather infrequently, while others seemingly do not exist at all.

Second, it is not clear that the standard notion of constituent structure is necessary for expressing many of the Greenbergian word-order generalisations. Greenberg's (1963) Universals 2-5, for example, correlate prepositional vs. postpositional order, noun-adjective vs. adjective-noun order and noun-genitive vs. genitive-noun order with the basic word-order types SVO, VSO and SOV. These generalisations are syntactically simplistic: they are made purely on the basis of linear order or grammatical function. Furthermore, they are stated over languages ("If a language L has property A, then it also has B"), which is not even a coherent notion in

generative linguistics. A theory of syntax based on the notion of constituent structure in the Chomskyan tradition can and should express cross-linguistic generalisations of greater sophistication. In other words, we should look for hierarchical universals, in the sense of Whitman (2008: 234, 251). Such universals describe the relative position of two or more categories in a single structure, where this position follows from the underlying hierarchical arrangement of constituents. For example, according to Whitman, the fact that Specifiers appear universally to the left of the head they specify is an instance of a hierarchical universal (cf. i.a. Paul 1997, Pearson 2001 and Aldridge 2004 for evidence that apparently Spec-final VOS languages are also best analysed as Spec-initial). Whitman suggests that hierarchical universals are absolute, while implicational universals of the kind familiar since Greenberg (1963) do not have to be, and frequently are not.

The purpose of this paper is to introduce, motivate and explain the **Final-over-Final Constraint** (henceforth: *FOFC*), a generalisation that we take to instantiate a hierarchical universal in Whitman's sense. FOFC is universal constraint on phrase-structure configurations. Initially, we formulate FOFC as follows (see Holmberg 2000:124):[2]

(1) **The Final Over Final Constraint (FOFC)**

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.

Our account of this universal leads us to a number of striking and important conclusions regarding phrase structure, movement and linearization. These are: (i) that head-final order is derived by movement, supporting the basic proposal for this kind of order in Kayne (1994); (ii) that there is a single general movement-triggering

[2] Emonds (1976:19) put forward a similar constraint: phrases to the left of a head inside an XP have to be head-final under certain conditions (the Surface Recursion Restriction; cf. also Williams' 1981 Head-Final Filter). As Emonds points out (p.c.), the Surface Recursion Condition does not make identical predictions to FOFC, but it is somewhat similar. One difference is that Emonds' constraint rules out cases like **proud of his children man*, which FOFC does not, as the AP is not a complement of the Noun. We are grateful to Norbert Corver and Henk van Riemsdijk for drawing our attention to this and to Joseph Emonds for helpful discussion and references. David Pesetsky (p.c.) further drew our attention to Hale, Jeanne & Platero (1977:385), who observe a case of FOFC in Papago/Tohono O'odham, and propose a surface structure constraint on the distribution of a #-boundary in order to account for it (see their (25), p. 391).

diacritic "feature" (possibly reducible to a variety of Edge Feature in the sense of Chomsky 2008, although in most of our exposition we will keep our notion of movement trigger terminologically distinct for reasons that will become apparent): when associated with φ-features, this diacritic triggers A-movement, when associated with the head of a phase, it triggers A'-movement, and when associated with the c-selection features of a head, it gives rise to what we will call linearization-movement (L-movement); (iii) partly following from (ii), that the head-initial/head-final choice is encoded in the presence or absence of the movement-triggering property on heads; (iv) that the L-movement trigger is a property of Extended Projections, and may be projected up the tree through the Extended Projection of the lexical head, each occurrence of the movement trigger on a given head requiring movement of the structural complement of that head into its specifier; and, most importantly, (v) that this projection relation is subject to Relativized Minimality (Rizzi 1990, 2001). Our general conclusion is that FOFC is an effect of the interaction between (iii), (iv), and Relativized Minimality. We conjecture that this is why FOFC is universal and ubiquitous: it reflects one of the deepest formal properties of natural language, and arguably a "third-factor" economy principle in the sense of Chomsky (2004, 2005, 2008).

Our conclusion suggests that linearization information must be present in narrow syntax. It has been proposed in the recent minimalist literature (cf. i.a. Berwick & Chomsky 2008, Boeckx 2008, Richards 2009) that this state of affairs is incompatible with the claim, enshrined in the Strong Minimalist Thesis (SMT), that language optimally meets interface requirements. Since linear order is not required in syntax, the inference is that the maximally economical situation is that no such information is allowed. However, as we will see, the syntactic encoding of linearisation information harnesses the same formal element that is responsible for A and A'-movement. Furthermore, the distribution of this element is constrained by Relativised Minimality, a well-established constraint on syntactic operations/relations. We take this to mean that our hypothesis does not constitute an SMT violation.

The paper is organised as follows: in §2, we present FOFC and provide the principal empirical motivation for it; in §3, we present and account for a range of apparent counterexamples; in §4, we present our theory of linear order and show how FOFC can be derived from it; in §5, we apply the analysis to the data introduced in §§2 and 3, and in §6, we refine both our approach to OV orders and some of the

technical notions introduced to account for FOFC in earlier sections. §7 concludes the paper with a speculation as to how much needs to be assumed about Universal Grammar itself in order to obtain our results. Our conclusion is that, while some of the mechanisms we invoke, including Relativised Minimality, can be plausibly taken as "third-factor" mechanisms in the sense of Chomsky (2004 *et seq.*), others, including the Linear Correspondence Axiom (LCA) (Kayne 1994) and the trigger for movement, are part of UG. More generally, we argue, as already noted, that our analysis is compatible with the Strong Minimalist Thesis (SMT).

2. The Final-over-Final Constraint (FOFC)

As mentioned in the Introduction, our key proposal is that FOFC, as given in (1), is a universal. The import of the formulation of FOFC in (1) is that it rules out structures like (2), where αP is the complement β and γP is the complement of α:

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

Our principal empirical claim, then, is that configurations instantiating the schema in (2) are not found in the world's languages. We now present the evidence for this.

1.1 Empirical motivation for FOFC

*1.1.1 *(S)VOAux*

Our initial observation comes from comparative Germanic. Looking across Germanic varieties, both synchronically and diachronically, we observe a very wide range of word orders, particularly at the clausal level and in VP. If we consider the three elements Aux,[3] V and O, we find all possible permutations of these, with one very striking exception: the order VOAux is not found. This fact has often been noted; see, among others, Travis (1984:157-8), den Besten (1986), Kiparsky (1996: 168-171), Pintzuk (1991/1999), Hróarsdóttir (1999, 2000), Fuss & Trips (2002). Let us look at these orders one by one.

[3] "Aux" may either be an auxiliary or a verb capable of triggering clause-union/restructuring. Following Cinque (2004), we take both types of element to instantiate clausal functional heads. As a cover term, we label the position of these elements as I or T here.

First, O V Aux (*(John) the book read has*) is readily found. (3) illustrates this from German, extrapolating from main clauses, as is standard practice, in order to avoid the confound introduced by the verb-second phenomenon:

(3) ... dass Johann **das Buch gelesen hat**
that John the book read has
'… that John has read the book'

This order is found, primarily in subordinate clauses of various kinds, in German, Dutch, Afrikaans, all German, Dutch/Flemish and Afrikaans dialects, Old English (OE) and Old Norse (ON). This is usually thought to derive from head-final order in both IP and VP.

Second, we find the order O Aux V (*(John) the book has read*). This corresponds to what has been known, since Evers (1975), as "verb-raising" order in Dutch:

(4) a. Dutch:
... dat Jan **het boek wil lezen**
that John the book wants to-read
'... that John wants to read the book'

b. Old English (OE):
... þe æfre on gefeohte **his handa wolde afylan**
who ever in battle his hands would defile
'... whoever would defile his hands in battle'
(*Ælfric's Lives of Saints* 25.858; Pintzuk 1991: 102)

This order is also found in Standard Afrikaans and many nonstandard West Germanic varieties, but not in Standard German.

Third, we observe the order Aux O V (*(John) has the book read*). At least since Haegeman & van Riemsdijk (1986), this order has been known as "verb projection raising":

(5) a. West Flemish (Haegeman & van Riemsdijk 1986):
... da Jan **wilt een huis kopen**
that John wants a house buy-INF
'... that Jan wants to buy a house'

b. Zürich German (Haegeman & van Riemsdijk 1986):
... das de Hans **wil es huus chaufe**
that the Hans wants a house buy-INF
'... that Hans wants to buy a house'

c. OE:
... þæt hie **mihton** swa bealdlice **Godes geleafan bodian**
that they could so boldly God's faith preach
'...that they could preach God's faith so boldly'
(*The Homilies of the Anglo-Saxon Church* I 232; van Kemenade 1987:179)

This order is also found in Middle Dutch (Hoeksema 1991), Old High German (Behagel 1932), ON (Hróarsdóttir 1999: 203ff.), and in numerous non-standard varieties of Flemish, of Swiss and Austrian German and also of Afrikaans (see Wurmbrand 2006 and Schmid 2005 for discussion and overview). Note that we appear to have the mirror-image of the FOFC configuration here, in that we plausibly have a head-initial IP (with the order Aux > VP) and, as complement to the Aux in I (or v), a head-final VP. Initial-over-final orders are readily attested then, while final-over-initial, schematised in (2), is not. This is the central asymmetry that we observe.

A rarer, but still attested order is V Aux O (*(John) read has the book*). This has often been described as "object extraposition". Here we illustrate with "PP-extraposition" in Dutch and a DP in final position in OE:

(6) a. Dutch:
... dat het lijk **gevonden werd in de kast**
that the corpse found become in the closet
'... that the corpse was found in the closet' (Zwart 1997:39)

b. OE:

... þæt ænig mon **atellan mæge ealne þone demm**

that any man relate can all the misery

'... that any man can relate all the misery'

(*Orosius* 52.6 – 7; Pintzuk 2002: 283, 16b)

Where the "extraposed" element is a PP, CP or, more marginally, a heavy DP, this order is also found in German, as well as many German and Dutch dialects. It is also found in ON (Hróarsdóttir 1999:201-2).

Finally, we find Aux V O (*(John) has read the book*). This is of course the head-initial order, different variants of which are found in Modern English and throughout Modern North Germanic. It is also found in Dutch and OE (where it has been analysed as "verb raising" combined with "object extraposition"):

(7) a. Dutch:

... dat het lijk **werd gevonden in de kast**

that the corpse become found in the closet

'... that the corpse was found in the closet'

b. OE:

.. þæt he **mot ehtan godra manna**

.. that he might persecute good men

'.. that he might persecute good men'

(*Wulfstan's Homilies* 130.37 – 38; Pintzuk 2002: 282, 13b)

At first sight then, it seems that all possible word orders are found: that, across the range of varieties, synchronically and diachronically, anything goes. But this is not the case. The crucial observation is that **VOAux is not attested.**[4] The missing order is

[4] Walkden (2009) points out that this order is attested in West Flemish:

(i) ... da Valère **willen dienen boek lezen eet**

that Valère want-INF that book read-INF has

'... that Valère has wanted to read that book'

In this example, *willen dienen boek lezen* is arguably a head-initial complement (of unclear category, perhaps vP or TP) of head-final *eet*. However, it is unclear that *eet* here is really a finite verb/auxiliary, in that the negative element *en*, which cliticises to all finite verbs, is unable to cliticise onto it and it is

the one that instantiates the FOFC schema in (2) above for α = V, β = Aux. In other words, the missing configuration is that in (8):

---

unable to appear in its past-tense form in this order (Wurmbrand 2006:332). Wurmbrand (2006:240) observes that verb clusters with 231 order of this type only occur in West Flemish when the auxiliary is non-finite. So it is clear that the status of apparently FOFC-violating final *eet* is different to that of properly finite auxiliaries (cf. the discussion of particle elements in §3).

Similarly, Afrikaans appears to show 231 orders:

(ii) ... dat hy dit moes gedoen het
that he it must done have
'... that he must have done it'

But in main clauses *moes*, not *het*, is fronted:

(iii) Hy moes dit gedoen het
he must it done have
'He must have done it'

Since only finite verbs/auxiliaries are fronted under V2, this shows that *het* is not finite. We conclude that Afrikaans structures of this type therefore instantiate 132 rather than 231 orders.

So-called "linking verbs" in Afrikaans also appear to give rise to 231 orders:

(iv) ... dat hy die boek loop koop het
that he the book walk buy has
'...that he went to buy the book'

In V2 clauses lacking auxiliaries, both verbs, however, appear to front, suggesting that they act as a unit (de Vos 2006 refers to this as "quirky verb second"):

(v) Hy loop koop die boek
he walk buy the book
'He goes to buy the book'

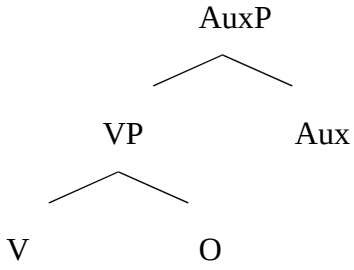
The two verbs also front together in Predicate-doubling and VP-fronting structures (cf. Biberauer 2009b), further supporting the idea that the verbs corresponding to 2 and 3 in this particular string do not function independently as is the case in verb clusters more generally. So again, it appears that we are not dealing with a genuine 231 order here and, if this is the case, there is also no FOFC violation.

A final relevant case involves so-called *Infinitivus Pro Participio* (IPP) structures like (vi), which are available in Afrikaans, West Flemish and Swiss German (cf. Schmid 2006 for discussion):

(vi) ... dat hy haar hoor kom het
that he her hear come has
'... that he has heard her come'

In these cases, the final auxiliary does indeed appear to be the finite element (it undergoes fronting under V2: *Hy het haar hoor kom* "he has her hear come"), but we observe the unexpected form of the verb it takes as its complement – infinitival *hoor* rather than past participial *gehoor*. Pending better understanding of IPP structures, we leave this case to future research.

(8) *



Moreover, this is not restricted to Germanic. The same gap exists in Finnish (Holmberg 2000: 128) and, more generally, in the Finno-Ugric languages that exhibit both initial and final structure (Marit Julien, p.c. to Michelle Sheehan), and in Basque (Haddican 2004: 116).

Holmberg (2000: 128) shows that Finnish is basically VO. But, under specific conditions, where the matrix C is [+focus] or [+wh] (see Holmberg 2000 for details), both AuxOV and OVAux orders are allowed. VOAux, however, is never allowed. The paradigm is illustrated in (9):

(9) a. Milloin Jussi olisi kirjoittanut romaanin? [Aux-V-O]
when Jussi would-have written INDEF-novel
'When would Jussi have written a novel?'

b. Milloin Jussi olisi romaanin kirjoittanut? [Aux-O-V]
when Jussi would-have INDEF-novel written
'When would Jussi have written a novel?'

c. Milloin Jussi romaanin kirjoittanut olisi? [O-V-Aux]
When Jussi INDEF-novel written would-have
'When would Jussi have written a novel?'

d. * Milloin Jussi kirjoittanut romaanin olisi? [*V-O-Aux]
when Jussi written INDEF-novel would-have

Latin is generally analysed as an OV language with rather free word order, especially in literary classical texts (Harris 1978, Vincent 1988, Pinkster 1990, Salvi 2004, Devine & Stephens 2006, Clackson & Horrocks 2007, Ledgeway forthcoming).

Given the fairly synthetic nature of its verbal morphology, it is uncertain that Latin had auxiliaries. However, one candidate construction is the perfect passive, formed from the perfective participle of the verb and *esse* "to be" (another is the future perfect passive). Citation forms of this tense generally give it in the expected head-final order, e.g. *amatus sum* (loved-Nom.sg.m. I-am, "I have been/was loved"). One might think that, since the participle is passive, VOAux orders will not in any case be possible, quite independently of FOFC. However, it is known that Latin allowed impersonal passives in which the logical direct object could appear in the accusative (see Keenan 1985); moreover, A-movement of the object was not obligatory in passives, as in Modern Italian (Burzio 1986). In that case, VOAux in this construction becomes a possibility, where "O" refers to the logical object, which may bear either nominative or accusative case. An initial, non-exhaustive survey of the manuals (Gildersleeve & Lodge 1997, Ernout & Thomas 1953, Kühner-Stegmann 1955, Salvi 2004, Devine & Stephens 2006) suggests that FOFC is respected. Owing to the great freedom of word order observed in Classical Latin (essentially attributable to scrambling combined with subextraction of nominal dependents from DP; cf. Salvi 2004, Devine & Stephens 2006), we find, alongside the expected head-final order, the order Aux V. In the latter case, we readily find material interspersed between Aux and V, while in the former case there is a strong tendency for the V Aux sequence to be uninterrupted. Thus we find three patterns:

(10) a. navis parata fuit [V Aux]
ship-NOM.SG.F. prepared-NOM.SG.F. be-PAST.3.SG
'The ship was prepared'
(Livy 3, 26, 11; Kühner-Stegmann 1955, I:164)

b. erit parata navis [Aux V]
be-FUT prepared-NOM.SG.F. ship-NOM.SG.F
'The ship will be prepared'
(Plautus, *Miles Gloriosus* 921; Kühner-Stegmann 1955, I:165)

c. Gallia est omnis divisa
Gaul-NOM.SG.F. be-PRES.3SG all divided-NOM.SG.F.
in partes tres [Aux X V]
in parts.ACC three.ACC
'All Gaul is divided into three parts'
(Caesar, *De Bello Gallico* 1,1,1; Kühner-Stegmann 1955, I:164)

In (10a,b), *navis* is probably the derived subject of the passive, but in (10c) *omnis* may be a stranded quantifier (cf. Sportiche 1988; the position of *omnis* here does not correspond to the first-merged object position if objects are always leftward-moved in Latin). Hence we find AuxOV.[5]

[5] However, two provisos must be added here. First, there are examples of apparent VOAux order with this construction:

(i) a. nocte ac die continuatum incendium fuit
night-ABL and day-ABL continued-NOM/ACC.SG.N fire-NOM/ACC.SG.N be-PAST.3.SG
'the fire went on night and day'
(Livy 26, 27, 4; Kühner-Stegmann 1955, I:164)

b. si mihi tributus honos fuit
if me-DAT. paid-NOM.SG honour-NOM.M.SG be-PAST.3.SG

'if honour was paid to me'
(C. Fam., 15, 4, 16; Kühner-Stegmann 1955, I:167)

In (ia), it is not possible to tell what case *incendium* is since 2nd-declension neuter nouns have syncretic nominative and accusative forms. But it is possible that *continuatum incendium* is the subject here, giving the meaning "Night and day, there was continuous fire". In (ib), *honos* is unambiguously nominative, and so again *tributus honos* could be the subject. The other possibility, which we cannot exclude, is that both of these DPs occupy the object position of the participle, Nominative Case being licensed under long-distance Agree with T in (ib). In that case, we seem to be dealing with true VOAux order here. However, participles can be scrambled, both clause-internally and to the left periphery, as shown in (ii):

(ii) Restitutus est mihi dies festus
given-back-NOM.M.SG be-PRES.3.SG me-DAT day-NOM.M.SG feast-NOM.M.SG
'It made my day a red-letter day after all'
(Salvi 2004: 129; Cicero *Ad Atticum* 12.4.1)

Examples like (10b) may therefore also involve some form of scrambling of the participle, or a constituent containing it. Therefore we are not forced to conclude that the examples in (i) instantiate the structure in (8). Thanks to Chiara Gianollo for discussion of the Latin data.

It is standardly observed that the Romance languages have developed auxiliaries where Latin had synthetic forms, e.g. in the active perfect. Some of these auxiliaries may have emerged in Late Latin. This is true of the future/conditional auxiliary based on Latin *habere* "have", which expressed deontic modality and futurity in post-3rd-century Latin, according to Benveniste (1968). Benveniste also notes that, in this usage, *habere* was always adjacent to and immediately followed the main verb (which was almost always a passive infinitive). If this observation is fully accurate, then once again no FOFC violation will be found here. Benveniste notes that this fact must have contributed to the later reanalysis of this auxiliary as a verbal affix. See Roberts & Roussou (2003, Chapter 2) for discussion.

Haddican (2004: 116) observes the absence of FOFC-violating VOAux structures in Basque:

(11) a. Jon-ek ez **dio** Miren- i egia esan [Aux-O-V]
Jon-erg not AUX Miren-dat truth say-perf
'Jon has not told Miren the truth.'

b. Jon-ek ez **dio** esan Miren -i egia [Aux-V-O]
Jon-Erg not AUX say-Perf Miren-dat truth
'Jon has not told Miren the truth.'

(12) a. Jon-ek Miren-i egia esan **dio** [O-V-Aux]
Jon-erg Miren-dat truth say-perf AUX
'Jon has told Miren the truth.'

b. *Jon-ek esan Miren- i egia **dio** [*V-O-Aux]
Jon- erg say-perf Miren-dat truth AUX

Biberauer, Sheehan & Newton (2010) have shown that VOAux orders are also unattested in language-contact situations. For example, in South Africa there is extensive contact between Afrikaans, an OV language with head-final order in IP and VP, and English, which, of course, has head-initial IP and VP. In the variety most heavily influenced by English, Kaaps, spoken by the so-called Coloured population in the Cape, we find a range of possible orders in subordinate clauses (where V2 is generally inoperative). However, the one order that we do not find is VOAux:

(13) a. … dat ek [$_{VP}$ R1400 van die Revenue gekry] het [O-V-Aux]
that I R1400 from the Revenue got have
'… that I got R1400 from the Receiver of Revenue'

b. … dat ek het **[**$_{VP}$ R1400 van die Revenue gekry] [Aux-O-V]

c. … dat ek het [$_{VP}$ gekry R1400 van die Revenue] [Aux-V-O]

d. *… dat ek [$_{VP}$ gekry R1400 van die Revenue] het [*V-O-Aux]

Our first piece of evidence for FOFC, then, stems from the cross-linguistic absence of VOAux order, notably in mixed systems of the West Germanic languages, Old Norse (Hróarsdóttir 2000) and the Western Finno-Ugric languages; this order also seems to be largely absent in Latin.

Drawing on the typological research discussed in Dryer (1992), Holmberg (2000: 134-135) further discusses the cross-linguistic distribution of the form expressing volition ("want") and V and O in VP, indicating that only 4 languages at first sight appear to permit the FOFC-violating [VO]-WANT order. Upon closer inspection, however, it emerges that these languages also permit OV orders in certain contexts and that VO plus final WANT strings appear not to occur. Even in languages that exhibit the means to potentially violate FOFC, we do not observe FOFC violations, then.

Cross-linguistically, we therefore see that a mix of patterns is found, and, notably, that disharmonic orders of the "verb-projection raising" (AuxOV) type are attested; but the mirror-image of verb-projection raising seems to be entirely missing. This kind of typological gap is striking, especially when attested in unrelated families, and calls for an explanation. FOFC is not an explanation, but at least subsumes this gap under a broader generalisation.

1.1.2 The cross-linguistic distribution of complementisers

Our second piece of evidence for FOFC also concerns clause-level syntax. This is the observation, originally due to Hawkins (1990a: 256-7), that sentence-final complementisers are not found in VO languages (see also Dryer 1992: 102; 2009b: 199-205, Hawkins 2004, Kayne 2000: 320-321). Cross-linguistically, we find OV languages with both initial and final complementisers. Latin is generally taken to be an OV language (see Clackson & Horrocks 2007, Ledgeway forthcoming, Salvi 2004, Vincent 1988 and the references given there), and has initial complementisers, as the following examples show (taking *ut* and *quod* to be complementisers):

(14) a. Ubii Caesarem orant [$_{CP}$ ut sibi parcat]
Ubii-NOM Caesar-ACC beg-3PL C selves-DAT spare
'The Ubii beg Caesar to spare them'

b. Accidit perincommode [quod eum nusquam vidisti]
happened-3SG unfortunately C him nowhere saw-2SG
'It is unfortunate that you didn't see him anywhere'
(see Roberts 2007: 162-3 for sources and discussion)

On the other hand, Japanese is an OV language with final complementisers:

(15) Bill-ga [$_{CP}$ [$_{TP}$ Mary-ga John-ni sono hon-o watasita] **to**]
Bill-NOM Mary-NOM John-DAT that book-ACC handed that
itta (koto)
said (fact) '
'Bill said that Mary handed that book to John' (Fukui & Saito 1998:443)

And of course we can readily find VO languages with initial complementisers, English being a good example.

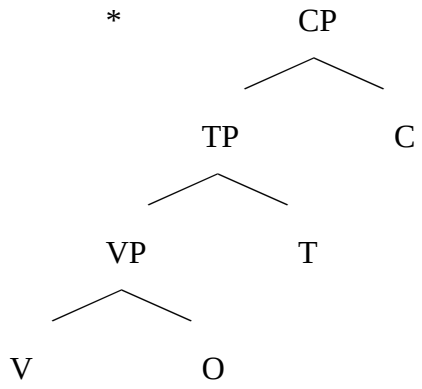
But the fourth logical possibility, VO languages with final complementisers, appears not to be attested. The *World Atlas of Language Structures (WALS)* does not have a specific map for the order of general clausal subordinators in relation to the clause they introduce, and so we cannot directly look for evidence there. However, the order of "adverbial subordinators" such as *although, when, while* and *if,* in relation to the clauses they introduce is covered (Map 94; Dryer 2008b). It is possible that some, if not all, of these elements are complementisers; however, Dryer (2008b:10) is explicit on the point that "care was taken not to include general markers of subordination", i.e. the clearest cases of complementisers. Nonetheless, in the 599 languages investigated, the skewing is evident: 279 languages have VO and initial subordinators and 85 have OV and final subordinators. Placing the subordinators in C, then, we observe cross-categorial harmony in the majority of cases. More importantly for our purposes, there is a very clear asymmetry in the disharmonic orders: 54 languages have OV and initial subordinators, but *only two* are said to show the combination of final subordinators with VO: Buduma (Afro-Asiatic) and Guajajara (Tupi-Guaraní). Newton (2007), however, shows that final Cs in both languages appear to introduce OV rather than VO clauses, i.e. both in fact exhibit the more generally attested disharmonic order rather than the FOFC-violating one in this context. Dryer also notes that subordinating suffixes are found, particularly in OV

languages. In fact, there is only one VO language with subordinating suffixes (the Australian language Yindjibarndi), as against 51 OV languages with subordinating suffixes. We can observe, then, that in this respect the data is significantly skewed.[6] The vanishingly small number of counterexamples clearly requires closer investigation, but the overall asymmetry in the distribution of logically possible combinations of orders is very clear.

In terms of the schema in (2) there are in fact two ways of ruling out final Cs in VO systems. On the one hand, we could have a head-initial VP inside a head-final IP and CP:

(16) a. *$[_{CP}\ [_{IP}\ [_{VP}\ V\ O\ ]\ I\ ]\ C\ ]$

b. *



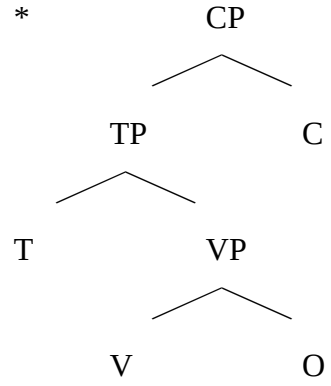
This instantiates the schema in (2) for α=V and β=T, and so constitutes a FOFC violation of the same type as the VOAux orders considered in the previous section.

Alternatively, we could have a head-initial TP inside a head-final CP:

(17) a. *$[_{CP}\ [_{TP}\ T\ [_{VP}\ V\ O\ ]]\ C\ ]$

[6] There are a number of other combinations which are somewhat indeterminate in relation to our concerns. For example, Dryer defines a "mixed" category for subordinators, which includes the combination of initial and final, as well as clause-internal (often second position) and suffixal. There are 30 VO languages with mixed subordinators, and clearly these need to be investigated. Furthermore, some languages are taken to have no dominant order among OV and VO: 4 of these have mixed subordinators, 4 have suffixal subordinators and 2 have final subordinators. The numbers are small in every case, but again these cases should be investigated more closely.

b. *



This structure instantiates (2), and hence violates FOFC, for α = T and β = C. This is the second piece of evidence in favour of FOFC.

We can in fact make a further observation in this connection, which constitutes the basis of a further piece of evidence for FOFC: OV languages with initial complementisers systematically extrapose their CP complements (the significance of this is in relation to FOFC was first pointed out by Sheehan 2008a). We can in fact see this in the Latin examples in (15), where the subordinate CP remains in postverbal position, apparently the typical order in Latin (cf. Devine & Stephens 2006:124-5). This is also true of German, where finite CPs must be postverbal, while raising complements, which we standardly take to be TPs, are not (cf. Biberauer & Roberts 2008):[7]

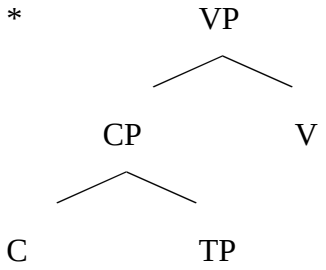
(18) a. Er weisst, dass sie kommen.
he knows that they come
'He knows that they're coming.'

b. ... dass Hans sich zu rasieren schien.
that Hans self to shave seemed
'… that Hans seemed to shave himself.'

[7] There are contexts in which it is possible for initially-headed CPs to surface preverbally in West Germanic and other OV languages, namely where these CPs have undergone A'-movement, either to clause-initial position (cf. Koster 1978, Alrenga 2005, on sentential subjects) or to a clause-internal position, which may superficially appear to be the complement position. As Barbiers (2000) notes, however, this preverbal position can be shown to be higher than that associated with the unmoved preverbal complement, CPs located in this position being to the left of VP-adverbs and satisfying the same diagnostics as scrambled elements.

The same is true in a great diversity of other OV languages, including Afrikaans, Bengali, Dutch, Hindi, Iraqw, Mangarrayi, Neo-Aramaic, Persian, Sorbian, Turkish, etc. (cf. Biberauer, Sheehan & Newton forthcoming for further details and exemplification; Dryer 2009a). This oddity of word order superficially appears to be a FOFC-avoidance strategy. If the head-initial CP were to appear in the complement position of the head-final V we would have a structure like (19):

(19) *



This structure violates FOFC for α=V and β=C. In fact, more generally, OV languages tend to have either postverbal finite CP clausal complements, or preverbal nominalised clauses (cf. Koptjevskaja-Tamm 1988, 1993, Givon 2001). We will see in §3 that preverbal nominalisations are exempt from FOFC.

A related point, which may be of some importance, is that failing to overtly realise the complementiser does not appear to be a strategy facilitating (non-scrambled; cf. note 7) preverbal head-initial CPs (Josef Bayer, p.c.). This can be seen in Hindi, a language which, like English, allows complementisers to delete (cf. Bayer 2001: 15), as shown in (20):

(20) a. He knows (**that**) they are coming

b. usee (**yah**) maluum hai [**ki** vee aa rahee haiN]
3s-DAT this known is that 3PL come PROG are
'He/she knows that they are coming'

c. *usee [(**ki**) vee aa rahee haiN] maluum hai
3s-DAT that 3PL come PROG are known is
'He/she knows that they are coming' (Davison 2007:177)

Question particles initially also appear to be good candidates for C-elements. As we shall see below, though, VO languages quite readily allow final Q-particles. We will therefore consider this case, along with clausal force-related particles more generally, when we look at other real and apparent counterexamples to FOFC in §3 below.

Coordinating conjunctions are another kind of clause-introducer. Zwart (2005) investigates 214 languages and finds no "true" final co-ordinating conjunctions in head-initial languages (cf. his Table 3, and also Zwart 2009a). We will return to this case in §5 below.

The absence (or extreme rarity) of VO... C orders cross-linguistically is our second piece of evidence for FOFC.

1.1.3 The nominal domain

Turning from the clausal to the nominal domain, we find two further pieces of evidence for FOFC, one direct and one indirect. The direct evidence comes from Finnish nominals (see Holmberg 2000). As a predominantly head-initial language, Finnish has postnominal complements. Finnish also has postpositions. An NP consisting of a noun with a PP complement will typically look like (21):

(21) a. käynti nurkan takana [NO[8]]
visit corner-GEN behind
'a visit around the corner'

b. [$_{NP}$ käynti [$_{PP}$ nurkan takana]]

Finnish also has some prepositions. As it happens, 'before' is a preposition, while 'after' is a postposition (the Finnish prepositions take a complement with partitive case, while postpositions take a complement with genitive case).

[8] *O* signifies the complement (PP) of the noun here. We leave aside the various potential complications regarding noun-complementation that have been raised in the literature (cf. Kayne 2008, Arsenijević 2009).

(22) a. ennen sotaa [PN]
before war-PAR
'before the war'

b. sodan jälkeen [NP]
war-GEN after
'after the war'

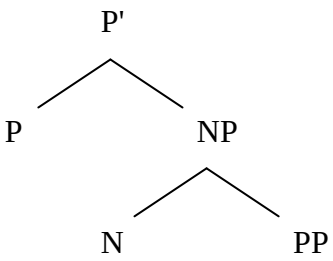
Consider the following minimal pair, where the complex NP (21) is merged with either of these two adpositions.

(23) a. ennen käyntiä nurkan takana [P-NO]
before visit- PAR corner behind
'before the visit around the corner'

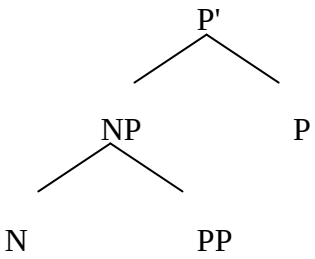
b. *käynnin nurkan takana jälkeen [*NO-P]
visit corner-GEN behind after

(23a), with the structure (24a), respects FOFC, while (23b), with the structure (24b) violates FOFC, for α=N and β=P.

(24) a.



b.



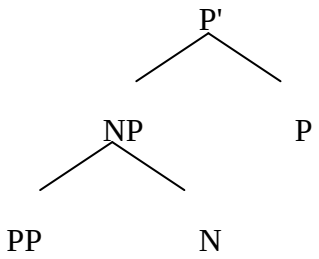
The PP complement of a noun can sometimes, albeit somewhat marginally, be preposed.

(25) ? nurkan takana käynti [ON]
corner-GEN behind visit
'a/the visit around the corner'

The resulting phrase can be merged with a postposition, as predicted, as it will not violate FOFC, the resulting structure being (26b). (26a) may be marginal, but still contrasts sharply with the ungrammatical (23b).

(26) a. ?nurkan takana käynnin jälkeen [ON-P]
corner-GEN behind visit-GEN after
'after the/a visit around the corner'

b.



The second piece of evidence for FOFC in nominals is more indirect, stemming from Cinque's (2005) account of Greenberg's Universal 20. Greenberg (1963:87) stated his Universal 20 as follows: "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". Thus Greenberg stated that these adnominal elements appear in the following orders in relation to each other and to N:

(27) a. Dem > Num > A > N
b. N > Dem > Num > A
c. N > A > Num > Dem

Cinque (2005: 315) observes that, since Greenberg's original proposal was put forward, "[(27a)] remains virtually unchallenged, while [(27b,c)] has proven both too restrictive and too permissive". He discusses the various counterexamples in detail in his paper, and we will not linger over them here as, for the most part, they are not directly relevant to our concerns (except one class, showing unexpectedly "low" demonstratives – see below).

For simplicity, let us disregard adnominal APs, and hence the relative order of A and N (AP is probably not a unified category and the possibility of reduced relatives is a confound; cf. Cinque 2005: 315-316, n.2; forthcoming a,b). We thus assume the following universal first-merged order (see also Shlonsky 2004:1482):

(28) $[_{DemP}$ Dem $[_{NumP}$ Num $[_{NP}$ { N (A) }]]]

Very much in the spirit of Cinque (2005), the attested orders now follow from the possibility of NP and/or NumP movement, along with standard conditions on movement such as the A-over-A Condition and the Strict Cycle. Most important for present purposes, FOFC accounts for a gap in the set of possible but unattested derivations.

Let us now consider the various surface orders that can be derived from (28). The first possibility is the very simple one that nothing moves, giving a surface harmonic head-initial order (*these three boys*):

(29) Dem Num NP

This order is found in "very many languages" according to Cinque (among them English). There is nothing further to say about this order here.

The second possibility involves NP-movement, but only as far as SpecNumP. This gives rise to surface Dem>N>Num order (*these boys three*):

(30) $[_{DemP}$ **Dem** $[_{NumP}$ **NP** [**Num** (NP)]]]

According to Cinque (2005:320), this order, with the adjective preceding the Noun is found in "few/very few languages," with the adjective following the Noun is found in "many languages".

The third order involves a further step of NP-movement to SpecDemP, with stranding of NumP, giving surface N>Dem>Num (*boys these three*):

(31) [$_{\text{DemP}}$ **NP** [**Dem** [$_{\text{NumP}}$ ((NP)) [**Num** (NP)]]]]

Again, this order is attested in "few/very few languages," perhaps owing to the markedness of NP's not pied-piping at all, but merely undergoing successive-cyclic movement through NumP to DemP, as Cinque (2005:321) suggests.

The fourth possible derivation involves "roll-up," i.e. pied-piping at each step, giving NP-movement followed by NumP pied-piping. This derivation creates a surface order which is the mirror image of that produced by the first-merged order of elements, i.e. N>Num>Dem (*boys three these*):

(32) [$_{\text{DemP}}$ [$_{\text{NumP}}$ **NP** [**Num** (NP)]] [**Dem** ([$_{\text{NumP}}$ NP [Num (NP)]])]]

This order, with AP intervening between NP and Num, is very common, as Greenberg originally observed.

FOFC is relevant to the derivation where NP does not move within NumP, but NumP raises to SpecDemP. This gives surface Num>N>Dem (*three boys these*):

(33) [$_{\text{DemP}}$ [$_{\text{NumP}}$ **Num NP**] [**Dem** ([$_{\text{NumP}}$ Num (NP)])]]

This structure instantiates the schema in (2) for α=Num, β=Dem, and so violates FOFC. FOFC underlies the observation that, if NumP moves, NP must first move to its Spec. In this respect, we can answer Cinque's (2005:325) question: "Why is movement of phrases other than NP unavailable?" This movement leads to a FOFC violation.[9]

[9] There is another case of illicit movement which gives the unattested surface order Num>Dem>N (*three these boys*), see in particular Cinque (2005:319). This is ruled out on the basis of the first-merged order of elements in (28). If NP moves first to SpecNumP, then Num' must move to give the surface order, and we follow standard assumptions in taking intermediate-level projections to be inert (see Chomsky 1995:249). If NP does not move, then either a non-constituent portion of Num (Specifier and head), or just the head Num, moves. The former is again ruled out on the standard assumption that movement can only affect constituents; the latter may be ruled out by a general ban on head-movement (Chomsky 2001), or by constraints on head-movement. If NP moves to some position above and to the left of NumP and NumP then moves, the derivation is ruled out by the Strict Cycle.

So we see that FOFC may play a role in a movement-based account of Greenberg's Universal 20 of the kind proposed by Cinque (2005) (or, more precisely, an account of the possible orders of some adnominal modifiers, since Universal 20 is not entirely empirically correct). However, the order in (34) is attested in a number of languages. Notably, it is found in all the Celtic languages, where the general order is Num>N>A>Dem:

(34) a. y tair plaid arall hyn [Welsh]
the three parties other these
'these three other parties' (Willis 2006:1831)

b. na trì leabhraichean mòra seo [Scottish Gaelic]
the-PL three books big-PL this
'these three big books' (D Adger (p.c.))

This order is also found in Semitic and other languages, and Num>A>N>Dem is found in various creoles (Bislama, Berbice Dutch, Sranan - Haddican 2002, cited in Cinque 2005:320, Notes 15, 16). However, we can observe that the problem cases of the kind in (34) feature an initial definite determiner. Moreover, there appears to be a morphosyntactic relation between an initial determiner and the low demonstrative. This is particularly clear in some Semitic languages, where the demonstrative appears preceding the (otherwise) initial determiner as well as in the low position, as in Moroccan Arabic shown in (35a), or where the definiteness marking "spreads" to N, A and the low demonstrative, as in the Hebrew example in (35b):[10]

(35) a. had ar rajeel haada [Moroccan Arabic]
this the man this
'this man' (Kaye & Rosenhouse 1997: 300)

[10] This is not true for Basque, where the determiner is low, and seems to occupy the same position as the Celtic demonstratives. There are number of other languages which show this order; see Cinque (2005:320, Note 16) for details and references.

b. ha-kufsa ha- gdola ha- zot [Hebrew]
the-box the-big the-this
'this big box'
(Berman 1997:327; see also Benmamoun 2000, Shlonsky 2004)

This evidence suggests that we should replace the label Dem in (28) with D, as follows:

(36) $[_{DP}$ D $[_{NumP}$ Num NP]]

In many, perhaps most, languages with both determiners and demonstratives Dem is in complementary distribution with D (e.g. in English), Dem either occupies D or SpecDP. In certain languages, such as Modern Greek (where we find *aftos o anthropos* "this-Nom the-Nom man-Nom" for "this man (nominative)"), Dem clearly occupies SpecDP as there is no complementary distribution with D. Moroccan Arabic appears to combine this with a low demonstrative, while Hebrew allows a copy of the article to appear in D and in the low position (and in fact in several other positions in DP). French examples like *cette voiture-ci/là* ("this car here/there") may also feature demonstrative elements in both positions (cf. Bernstein 2001). So we observe a range of variation regarding the realisation of these positions. What is clear that UG makes available at least two Dem-/D-positions, one initial in DP, and one in a low post-nominal position. A possible approach to analysing this situation, which we will not pursue here, is that the two positions correspond to the edge or head of the two phases which make up the extended nominal projection (on analogy with the biphasal structure of the clause proposed in Chomsky 2000, 2001, see Julien 2005:4).

The important conclusion from this brief excursus into the positions of demonstratives is that, if the structure in (36) is the correct one, with Dem appearing in SpecDP, D or lower in NP, then it is clear that the Num>N>(A>)Dem order observed in Celtic and elsewhere does not derive from NumP movement into SpecDP as the Determiner is clearly initial in these languages. In fact, it is more likely that there is no phrasal movement going on in DP in these languages, although the exact position of the postnominal Dem remains to be established. Hence we do not have a FOFC violation here.

1.1.4 Morphology: Myler (2009)

This section draws heavily on the investigation of FOFC in morphology in Myler (2009). We will report the results of his investigation, although we do not in every case agree with his conclusions.

If we adopt the general approach to morphology articulated in Marantz (1997), typified by the slogan that within the word the structure is "syntax all the way down," then we would expect that FOFC holds at and below the word level. There is a certain amount of evidence that this is the case, but we can also observe certain exceptions which, we will argue, are largely systematic in nature. This supports Marantz's thesis, as well as providing further evidence for FOFC. There are four cases to investigate: (i) the general cross-linguistic suffixing preference, observed by Hawkins & Gilligan (1988); (ii) the incidence of FOFC in inflectional morphology, (iii) the incidence of FOFC in derivation and (iv) in compounding.

Hawkins & Gilligan (1988, henceforth H&G) observe the "suffixing preference". In particular, picking up the observation in Greenberg's Universal 27, they observe that[11] "[l]anguages with VO and/or Pr+NP word orders in their syntax regularly have prefixes and/or suffixes in their morphology. But in a suggestively large number of cases, languages with OV and/or NP+Po have suffixes only" (219). H&G look at a database of 16 morphological categories in approximately 200 languages. The general conclusion shows a clear suffixing preference in general, and a much more marked preference in OV languages as compared to VO languages: "the prefixing/suffixing ratios in VO and Pr+NP languages are 34%/66% and 33%/67% respectively, i.e. roughly 2 to 1 in favor of suffixing; and 13%/87% and 16%/84% in OV and NP+Po languages respectively, i.e. between 5 to 1 and 7 to 1 in favor of suffixing" (230). Looking at the morphological categories shows a further skewing according to category, especially in OV languages. In (37) the percentage of prefixing in OV languages is given for 11 affix-types (from H&G's Table 9, p. 234); this illustrates the fact that OV languages strongly prefer suffixes (recall that chance distribution would lead one to expect an average of 25% here, since there are in principle four possibilities in play: prefixing vs suffixing and VO vs OV):

[11] "Pr" abbreviates "Preposition" and "Po" abbreviates "Postposition."

(37)

a.	Gender:	0%
b.	Case:	0%
c.	Indefiniteness:	0%
d.	Nominalizing:	3%
e.	Definiteness:	8%
f.	Plural:	2%
g.	Mood:	2%
h.	Tense:	1%
i.	Aspect:	9%
j.	Valence:	6%
k.	Causative:	7%

Five affix-types fall outside this pattern; according to H&G with these, "anything goes" (225): possessive, subject-marking, object-marking, negation and voice. We will see below that person-marking and negation in particular seem to behave in a systematically exceptional way in a number of respects.

The evidence in *WALS* supports H&G's results. Combining the data from Map 26 (Dryer 2008c, 894 languages) with that on OV vs VO order (Map 83: Dryer 2008a, 1,370 languages), there is relevant data for a total of 834 languages (60 of the 894 languages in Dryer 2008c are classified as having "no dominant order" in relation to OV vs VO). Of these, 248 combine "strong" suffixation (defined by Dryer as more than 80% of inflectional affixes being suffixes) with OV order, 91 combine strong suffixation with VO order, 47 show strong prefixation (more than 80% of inflectional affixes being prefixes) and VO order, while a mere 5 show strong prefixation and OV order (see Myler 2009:6-7 for more discussion and percentages).[12] This result is all the more striking as Dryer's data includes several of the affixal categories which H&G concluded fell outside the general suffixing preference: subject- and object-markers (including adverbal clitics), possessives and negatives.[13]

The suffixation preference in inflection can in fact be accounted for by FOFC. If prefixes are heads, then a prefixed word is a head-initial category. If the word-

[12] Of the remainder, 116 languages have little affixation, 108 are weakly suffixing (less than 80% of affixes are suffixes), 117 are equally prefixing and suffixing, 96 are weakly prefixing (less than 80% of affixes are prefixes).

[13] The five exceptional languages are: Chin (Tiddim), Chipewyan, Navajo, Slave and Tanacross (with the exception of Chin, which is Sino-Tibetan, these are all Amerindian languages).

phrase boundary is transparent, then a prefixed word constitutes a FOFC violation in a head-final phrase, i.e. in a system with head-final syntax. This situation does not exactly instantiate the schema in (2) since γ, the complement of α, is not phrasal, but that detail is irrelevant if we take bare phrase structure seriously. The structure in (38), which is the case of a prefixed word in a head-final phrase, violates FOFC:

(38) $[_{\beta P}\ [_{W}$ Prefix Root] β]

If we assume cyclic affixation and that each affix heads the category it immediately attaches to (see Williams 1981, Kiparsky 1982, Marantz 1997, Myler 2009, etc), then another FOFC-violating structure, this time entirely within the word, is illustrated by (39):

(39) *[[Prefix Root] Suffix]

We will look at some cases of this in derivational morphology below. In inflection, Myler (2009:13f.), on the basis of the survey of inflectional morphology in 530 languages in Julien (2002, Appendix 2), finds 82 apparent cases of (39), i.e. 16% of the total. Myler argues quite convincingly that the majority of these exceptions fall into patterns, in that in the vast majority of cases the Prefix or Suffix in (39) is one of the following classes of elements: an argument-adding morpheme, usually causative (42 languages), a "high" mood marker (17 languages), negation (10 languages), or discourse particles (7 languages). Aside from these cases, only 6 languages appear to violate FOFC in their inflectional system, hence if we can find an account of the cases just listed we would have a good case that FOFC is largely respected in inflectional morphology.[14] Myler (2009: 18-19) also excludes agreement markers, on the grounds that, following Chomsky (1995), these are not heads. This move is supported by Julien's (2002:235) observation that agreement morphemes are found in a very wide range of positions in the word (cf. also many of the contributions in Harbour, Adger & Béjar 2008 for other arguments against the idea that uninterpretable Agr features are syntactically represented as functional heads with a fixed syntactic position).

[14] Igbo, Oromo (East Cushitic), Mokilese (Central-Eastern Malayo-Polynesian), Hua (East New Guinea Highlands), Russian and North Puebla Nahuatl. See Myler (2009:32-4) for discussion of these cases.

There is a striking parallel here between the exceptions found by Myler and H&G's results as described above. The affixes which failed to show the suffixing preference in H&G's data were possessive, subject-marking, object-marking, negation and voice. If we take possessive-marking to be a form of agreement, then we can observe that agreement-markers generally form exceptions and account for this on the grounds that agreement morphemes are not heads. H&G did not consider discourse particles, and so there is no possibility of a comparison with Myler's results in this respect though it is worth noting more generally that discourse particles also frequently exhibit apparently FOFC-violating behaviour in the syntactic domain (see §3 for discussion). That leaves "high" mood markers as the one exceptional case observed by Myler and not by H&G. We will see below that negative and one important class of discourse particles, interrogative markers, fall outside FOFC at the clausal level, and we will offer an explanation for this in §5.1.2. That explanation may also carry over to the mood inflections discussed by Myler.

We conclude, with Myler (2009:33) that FOFC holds in inflectional morphology, and that it can account for the suffixing preference in inflection (see Myler 2009:33-37) for a discussion of why post-syntactic operations of the kind frequently adopted in DM, such as Lowering and Local Dislocation, do not affect this conclusion). We will return to the question of the suffixing preference in §3.

Concerning derivational morphology, Myler (2009:40-1.) observes that FOFC comes very close to deriving the Right-hand Head Rule (RHR), originally proposed by Williams (1981:248), which states that "[i]n morphology, we define the head of a morphologically complex word to be the right-hand member of that word". Myler points out that the RHR can be broken up into two parts, as follows:

(40) a. Where suffixes appear, they project.
b. Where prefixes appear, they do not project. (Myler's (5), p. 40)

Myler observes that, given Kayne's (1994) Linear Correspondence Axiom (LCA), which requires all structures, and in DM this includes the internal structure of words, to be head-initial, a suffix must trigger movement of its complement. In order for this to happen, the suffix must project. Prefixes, on the other hand, do not trigger movement of their complements. One way to ensure this is to treat them as non-projecting. In this case, FOFC is derived. Myler concludes "[t]he RHR is thus

reconceptualised as one of many strategies for avoiding the possibility of FOFC-violations" (41).[15]

There are, however, many cases of derivational morphology which appear to violate FOFC. Nominals derived from denominal and deadjectival verbs, where the verbalising morphology is a prefix, are a case in point, as in (41):

(41) a. $[_N\ [_V$ be $[_N$ head]] ing]

b. $[_N\ [_V$ en$[_A$ noble] ment]

Here we see cases where the prefix projects, converting the noun or adjective into a verb (although it does not trigger movement of its complement, and hence surfaces as a prefix; cf. Julien 2002 for discussion). Further suffixation of the nominalising morphology creates a structure which clearly instantiates (39), and as such is a FOFC violation. We will return to these cases when we discuss the Category Proviso in §3 below.

Turning to compounds, we can immediately observe that productive synthetic compounding in English obeys FOFC:[16]

(42) a. $[_N\ [[_N$ can] $[_V$ open]] er]

b. $[_N\ [[_N$ rocket] $[_N$ scient]] ist]

c. $[_A\ [[_N$ ear] $[_V$ split]] ing]

The affixes characteristic of synthetic compounds are suffixes, and so the complement to this affix must be head-final, given FOFC. We observe that English, although of course a head-initial language in VP, shows head-final order in these compounds; this can be seen as a necessary FOFC-compliance strategy, in that VO order (**open-canner*, etc.) would violate FOFC for α=V and β=affix. Other types of verb-object

[15] Another way to analyse prefixes, proposed by Julien (2002) for a range of cases, is to take them as heads taking the phrase headed by the suffix as complement. Again, this gives rise to FOFC-compliant structures.

[16] We abstract away from the details of the derivations that give rise to the (simplified) bracketings in (42), which, i.a. do not reflect the standard DM view that lexical roots lack a categorial specification. Note that, if we instead take both *can* and *open* to be acategorial lexical roots, then they are able to form a complex head in line with the proposal in Roberts (forthcoming) that incorporation is possible just where the incorporee's formal features are properly included in those of the incorporation host: in this case, neither head has any formal features and the constraint is satisfied. Similarly, $[_{Root}$ *can open*] is able to incorporate into *–er*$_N$, since its features (none) are properly included in those of *–er* (N).

compound, which lack a suffix, can show VO order: *pick-pocket, cut-purse*, etc. (although this is not obligatory: cf. *corkscrew*, *fly-spray*, etc.). Also, verb-object compounds in Romance languages, although less productive than English synthetic compounds, lack a suffix and show VO order: French *ouvre-boîte* "open-can" ("can-opener"), *tue-mouches* "kill-flies" ("fly-killer"), *tire-bouchon* "pull-cork" ("corkscrew"); Italian *gira-dischi* "turn-records" ("turn-table"). Finally, we can observe that, in certain semantic fields, English features both compounds involving head-initial phrasal expressions and head-final compounds, e.g. *history of science* and *rocket science*. Only the latter can enter into synthetic compounds, as observed by Ackema & Neeleman (2004: 164ff.); this is a further FOFC effect. Note further the contrasts in (43), also attributable to FOFC playing a role in compounding:

(43) a. the boy has red hair → he is [red-hair]ed
b. the boy is red of hair → *he is [red of hair]ed
c. gold-mining → gold-mining expert
d. mining of gold → *mining of gold expert

It appears, then, that FOFC plays a role in compounding.

However, Myler (2009:54ff.) points out examples like the following, which appear problematic for the general thesis that FOFC constrains compounding:

(44) a. an [[I-couldn't-care-less] attitude]
b. the [[man-of-the-match] award]
c. an [[easy-to-please] customer]
d. the [[Final over Final] Constraint][17]
e. the [[channel four] news]

Myler (2009) suggests that the complex modifiers in these examples are able to undergo "Renumeration" in the sense of Johnson (2002) and Harley (2009), thus being structurally opaque ("spelled out" in the sense of Uriagereka 1999, Nunes & Uriagereka 2000) at the point at which they are combined with the element which

[17] Thanks to Gertjan Postma for pointing out this rather embarrassing example to us. As Myler (2009:64) observes, this is not the first morphological constraint to violate itself in its name.

becomes the head of the compound. We will not go into this question here, although we will return briefly to Myler's proposals in §3 (on renumeration in relation to FOFC more generally, see §5.1.1 below and the discussion and analysis in Biberauer & Sheehan 2010).

In this section we have seen that FOFC holds in inflectional morphology, and holds, albeit vacuously, in derivational morphology. Its status in relation to compounds is uncertain, although we will argue in §3 that it holds for both synthetic and root compounds.

1.1.5 Diachronic evidence

FOFC is a constraint on synchronic grammars. However, this does not mean that it is not relevant to syntactic change. Since we take it to represent a universal constraint on synchronically possible word orders, we predict that no system can change into a FOFC-violating system. FOFC-violating systems fall outside UG, and therefore outside of the range of possible outcomes of syntactic change. This is a consequence of the general fact that, as Kiparsky (2008: 23) puts it: "If language change is constrained by grammatical structure, then synchronic assumptions have diachronic consequences."

More specifically, if FOFC is an absolute universal, then word-order change must proceed along certain pathways. Change from head-final to head-initial order in the clause must go "top-down", in that CP must be affected first, followed by IP, followed by VP, as follows:

(45) [[[O V] I] C] → [C [[O V] I]] → [C [I [O V]]] → [C [I [V O]]].

Conversely, head-initial to head-final change must go "bottom-up", starting at VP, then affecting IP and then affecting CP:

(46) [C [I [V O]]] → [C [I [O V]]] → [C [[O V] I]] → [[[O V] I] C].

Any other sequence of changes in either case will lead to an intermediate synchronic system which violates FOFC at some stage. Consider, for example, what would happen if, starting from a uniformly head-final system like the first one shown in the

series in (45), VP changed headedness first. This would give rise to an [[[V O] I] C] system; as we saw in §2.1.1.1 and §2.1.1.2 above, such systems are not found. If they were possible outcomes of natural processes of change, presumably such systems would be found; FOFC explains their absence synchronically and, therefore, diachronically.

Direct diachronic evidence concerning these trajectories of change is not easy to come by, given the general paucity of long-term attestation of many of the world's languages. Such evidence as we have concerning the earliest stages of Germanic supports this, though. The earliest attested stages of Germanic (Gothic, Old English and Old Norse) show C-IP order:

(47) a. **ef** han hefði þat viljað fága
if he has it wanted clean
"if he had wanted to clean it"
(ON: *Finn*; Hróarsdóttir 1999: 203)

b. .. **þæt** hie mihton swa bealdlice Godes geleafan bodian
that they could so boldly God's faith preach
"..that they could preach God's faith so boldly"
(OE, from (5c) above: *The Homilies of the Anglo-Saxon Church* I 232; van Kemenade 1987:179, 7b)

c. .. domjandas thata **thatei** ains faur allans **geswalt**
.. thinking this that one for all dies
".. thinking this, that one may die for all"
(Gothic: Longobardi 1978, Ferraresi 1991:30-35)

Cf. also the Latin examples in (14). These languages all have apparently mixed order in IP and VP (see above for OE and ON; Ferraresi 1997 on Gothic; Devine & Stephens 2006, Harris 1978:18ff., Ledgeway forthcoming, Salvi 2004 and Vincent 1988:59ff., on Latin). Later IP and VP became head-initial in English, Mainland Scandinavian and Romance. There is in fact evidence from the history of English that order in IP changed from head-final to head-initial before that in VP (examples from Biberauer, Newton & Sheehan 2009:8; Biberauer, Sheehan & Newton 2010):

(48) a. *Head-initial TP, head-final VP:*

Þat ne haue [$_{VP}$ noht here sinnes forleten].

Who neg have not their sins forgiven

"who have not forsaken their sins" (*Trinity Homilies* 67.934)

b. *Head-initial TP, head-initial VP:*

oðet he habbe [$_{VP}$ iȝetted ou al þet ȝe wulleð]

until he has granted you all that you desire

"until he has granted you all that you desire" (*Ancrene Riwle*)

Biberauer, Sheehan & Newton (2010) also show how OV to VO change from Latin to French follows the same pattern. It also holds for Finnish and Saami. Finno-Ugric languages further east are strictly head-final, with O > V > T > C order, reflecting the original Finno-Ugric, and indeed Uralic, pattern (Abondolo 1998). But Finnish and Saami, the westernmost languages in the family, have C > TP and T > VP, while S Aux O V is found in Saami (Marit Julien, p.c. to Michelle Sheehan). Finnish has VO, except for cases like (9).

Further evidence comes from Niger-Congo languages that have undergone a VO to OV change that is limited to VP (cf. Nikitina 2008 for recent discussion and references), and the Ethiopian Semitic languages, which have undergone a change from the typical Semitic head-initial pattern to a largely head-final pattern under the influence of Cushitic. See Biberauer, Newton & Sheehan (2009) and Biberauer, Sheehan & Newton (2010) for detailed discussion of these and other case studies corroborating the above pathways.

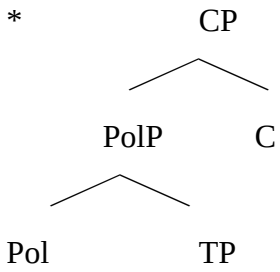
FOFC affects change in another way too, in that it also restricts borrowing options, i.e. change triggered by "external" factors; see Biberauer, Newton & Sheehan (2009) and Biberauer, Sheehan & Newton (2010) for discussion of a case study focusing on the borrowing/innovation of a clause-final complementiser. They report that, among Indo-Aryan languages which had borrowed a final complementiser, only those languages not featuring an initial question-marking polarity marker (Pol) developed a final complementiser. The relevant data is summarised in (49) (based partly on information in Davison (2007)):

(49)

Type	Position of Pol	Position of C	Languages
A	Initial	Initial Only	Hindi-Urdu, Panjabi, Kashmiri, Sindhi, Maithili, Kurmali
B	Final/Medial	Initial and Final	Marathi, Gujarati, Assamese, Bangla, Dakhini Hindi, Oriya, Nepali (plus some North Dravidian languages, i.e. Brahui)
C	Final/Medial	Final Only	Sinhala (plus most Dravidian languages)
D	**Initial**	**Final**	**Unattested in the area**

If Pol is hierarchically below CP, as proposed by Laka (1994) and Holmberg (2003), the gap in (49D) follows from FOFC, since the structure would be as in (50):

(50) *



This structure instantiates the schema in (2) for α=Pol and β=C, and as such is a further example of FOFC. So here we observe that FOFC may constrain borrowing; see again Biberauer, Newton & Sheehan (2009) and Biberauer, Sheehan & Newton (2010) for more discussion and further case studies.

1.1.6 Conclusion and Summary

We see evidence of FOFC in the absence of certain logically possible clausal word-order patterns (§§2.1.1.1, 2.1.1.2), similar effects in the nominal domain (§2.1.1.3), in morphology (§2.1.1.4) and in diachrony (§2.1.1.5). See also Cecchetto (2009) for evidence that FOFC holds in Italian Sign Language. We maintain that this suffices to

take the constraint seriously, as a possibly universal constraint on disharmonic structures.

A final typological observation supports this idea. It is clear that FOFC makes no predictions about purely harmonic systems, either head-initial or head-final. Regarding mixed systems, in the simple form given up to now (this will in fact be modified in the next section), FOFC predicts that head-final orders will predominate in the lower parts of a structure, and head-initial orders will predominate in the higher parts of a structure. This is because it is a direct consequence of FOFC that, looking up the tree from bottom to top, the order can switch from final to initial, but not vice-versa (or, equivalently, looking down the tree, the order can switch from initial to final and not vice-versa).

Now, one of the central consequences of the antisymmetric theory of phrase structure put forward in Kayne (1994) is that structures are right-recursive. Adopting this, we can equate "bottom" with "rightmost" and "top" with "leftmost," similarly "relatively high" implies "relatively to the left" and "relatively low" implies "relatively to the right." In these terms, FOFC makes two very general typological predictions, as follows:

(51) a. the "high" parts of the structure tend to be left-headed, even in disharmonic languages (cf. the left periphery in the sense of Rizzi 1997, which is often visible as such in otherwise head-final systems such as West Germanic, Latin, Basque, etc.).[18]

b. the "lower" parts of structure tend to be right-headed (cf. the suffixing preference in morphology discussed in §2.1.1.4; cf. also Zwart 2009b,c)

These observations appear to hold as strong tendencies. In particular, we do not observe the opposite tendencies: there is no tendency for a "high" (non-particle, see Note 18 and §3) right periphery, and there is no prefixing preference in morphology.

[18] Where "high" structure is realised in the form of particle elements – e.g. the force particles in the Chinese varieties (Paul 2009) or in Gungbe (Aboh 2004, 2006) – this generalisation does not apply. We will discuss this, and other (apparent) counterexamples to FOFC in the next section.

Neither are things neutral: instead, the tendencies are as predicted by a constraint like FOFC.

We summarise the FOFC violations we have observed in (52):

(52)

*VOAux	*[$_{AuxP}$ [$_{VP}$ V DP] Aux]
*VOC	*[$_{CP}$ [$_{TP}$ T VP] C]
*CTPV	*[$_{VP}$ [$_{CP}$ C TP] V]
*NOP	*[$_{PP}$ [$_{DP/NP}$ D/N DP] P]
*NumPD(em)	*[$_{D(em)P}$ [$_{NumP}$ Num NP] D(em)]
*[Prefix+Root] + Af	
*PolPC	*[$_{CP}$ [$_{PolP}$ Pol TP] C]

We will look more closely at some of these cases in what follows, and in certain cases revise our assumptions about the precise structures involved. For now, though, it can be taken as a convenient summary of the observations made so far, and the common pattern underlying them.

Having presented a range of empirical evidence for FOFC, we now move to the apparent counterevidence.

3. Counterexamples and Provisos to FOFC

Like any statement at this level of generality, we expect to find counterexamples to FOFC. The existence of apparent counterexamples is to be welcomed, in that they show that FOFC is in principle falsifiable. If we are to maintain FOFC in as a full a degree of generality as possible, however, we will have to address these counterexamples. This can be done in one of three main ways, in descending order of desirability. We can: (a) show that the counterexample is purely apparent, and that once the data are properly understood, there really is no problem; (b) show that the counterexamples fall into a well-defined analytical class, and refine the basic statement accordingly; (c) show that the counterexamples fall into a well-defined phenomenal class, and seek to characterise this in analytical terms. Of these, (a) is obviously unproblematic. However, some counterexamples do not fall into class (a). Counterexamples that can be shown to form theoretically coherent exceptions are also ultimately unproblematic, as long as the relevant refinement of the generalisation does

not weaken it unduly, and is not arbitrary in nature: indeed, such counterexamples can be highly useful in pointing to the correct theoretical characterisation of the generalisation. Finally, counterexamples of class (c) are the most recalcitrant; nonetheless, as long as these cases form coherent classes, they may ultimately be reducible to either (a) or – more likely – (b), pending greater theoretical understanding of both the generalisation and these classes of data.

It seems that all the counterexamples to FOFC that we have encountered so far can be understood in one of these three ways. In class (a), we find examples of structures which appear to violate FOFC in terms of the weakly generated string; once we look at the hierarchical relations, however, there is no violation. In class (b) we find two types: structures involving A'-movement, and structures which straddle two extended projections (in the sense of Grimshaw 1991, 2001, 2005); these observations will lead to a reformulation of the generalisation. In class (c) we find a range of structures in VO languages involving final particles of various kinds.

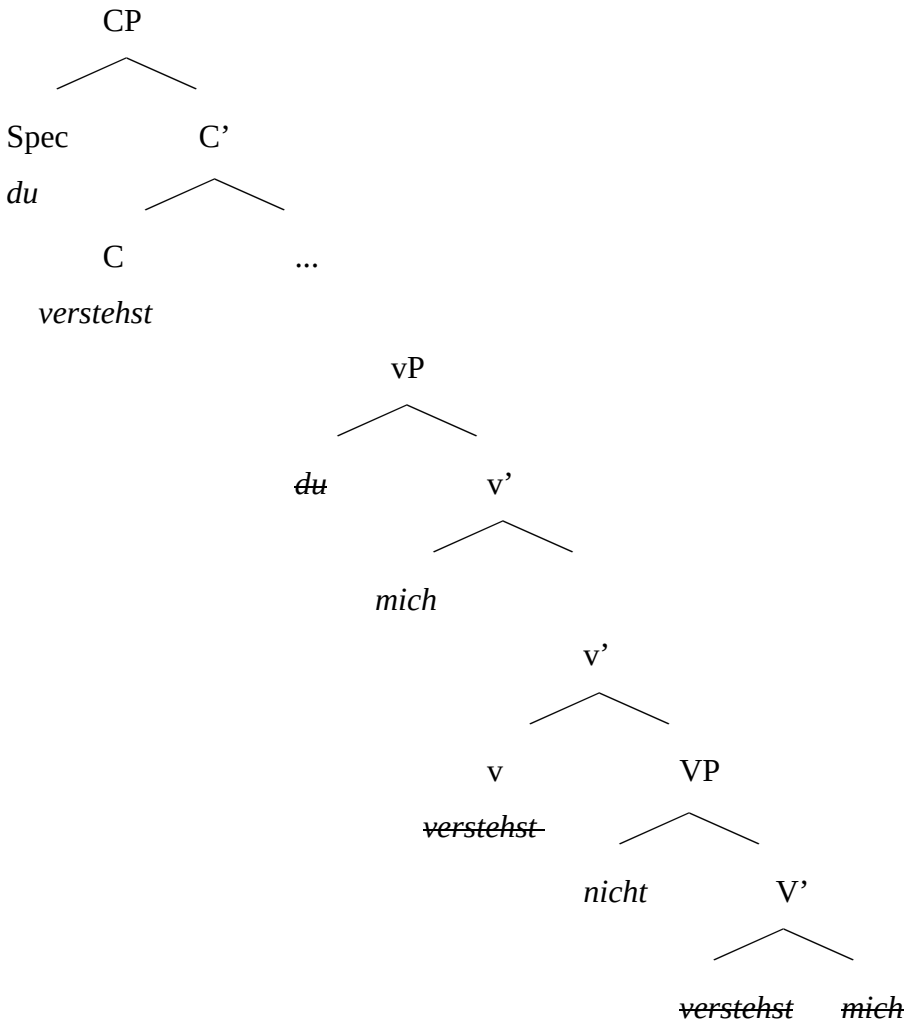
An example of the first type of apparent counterexample to FOFC comes from structures in Germanic involving "low" final negation, like (53):

(53) a. Du verstehst mich (einfach) **nicht** [German]
you understand me simply not
'You (simply) don't understand me'

b. Jag såg den **inte** [Swedish]
I saw it not
'I didn't see it'

These examples look as though they instantiate the order VONeg. If Neg is a head (as often thought since Pollock (1989), but see below), then this might seem to be a FOFC violation in that the verb and object, which might be thought to instantiate a head-initial VP together, precede the Neg head. However, it is generally agreed that these structures feature a combination of verb-movement (most likely to C, in order to meet part of the V2 requirement) and object-shift out of VP to the left of negation. Thus the verb and object move separately and to separate target positions. One likely partial structure for (53a) is (54):

(54)



Here we take it that V has moved to C (see den Besten 1983) and the object has moved to Spec,vP (Chomsky 2001); other analyses of these operations are of course possible, but the point for our purposes is that there is general agreement that V and the object do not form a single phrasal constituent in their derived positions, with the negative being located in a lower hierarchical position than that occupied by these elements, and so FOFC does not apply here. This kind of counterexample is thus easily dealt with and need detain us no longer. If the suggestion made in Note 5 that the apparent FOFC-violating Latin orders discussed there involve scrambling of the participle, or of a constituent containing it, is correct, then these examples also fall into this class.

Another type of counterexample which arguably falls under class (a) involves “circumpositions”, found in West Germanic, and in the Gbe lanagues of West Africa. These are illustrated in (55):

(55) a. auf den Berg hinauf [German]
up the-ACC mountain DIR-up
'up the mountain (path)'

b. onder de brug door [Dutch]
under the bridge through
'under the bridge (path)'

c. in die huis in [Afrikaans]
in the house in
'into the house'

Similar constructions are found in the Gbe languages (Aboh 2004):

(56) Àsíbá zé kwɛ́ dó távò lo jí
Asiba take-PERF money PREP table DET POSTP
'Asiba put the money on the table'

Here it appears that we have a head-initial PP in the complement of a postposition, in violation of FOFC. However, the postpositions in these constructions appear to be a rather non-uniform set of elements: adverbial or particle-like intransitive prepositions (see Svenonius 2003; forthcoming). In his analysis of these elements in the Gbe languages, Aboh (2004:120) refers to these elements as "fake postpositions", and points out that "the Gungbe postnominal morphemes are 'light Ps' in the sense that they fail to assign case, instead they function as a nominaliser head". If these observations are correct, then these constructions are not true counterexamples to FOFC.

Three important types of counterexamples fall under class (b), and we will introduce a special proviso on FOFC for each one. In fact, all three cases are illustrated by a simple example like (57):

(57) $[_{CP}\,[_{DP}$ which man] $[_{C}$ did] $[_{TP}$ you see …] ?

Here, just looking at the configuration, we apparently have a FOFC violation since we have an instantiation of (2) for α=D, β=C. Clearly, we need to allow such simple cases not to violate FOFC. We now show that there are in fact three provisos that must hold of FOFC: the Category Proviso, the Satellite Proviso and the A'-Movement Proviso. All three provisos are relevant to (57), as we shall see.

The first is very straightforward, at least on the face of things, and concerns complements to verbs in OV languages. A head-initial DP or PP may be immediately dominated by a head-final VP in many OV languages, e.g. German, as in (58):

(58) a. Johann hat [$_{VP}$ [$_{DP}$ einen Mann] gesehen].
John has a man seen
'John has seen a man'

b. Johann ist [$_{VP}$ [$_{PP}$ nach Berlin] gefahren].
John is to Berlin gone
'John has gone to Berlin'

All other things being equal, the examples in (58) instantiate the schema in (2) for α = D/P and β = V, and as such violate FOFC. However, they are obviously grammatical. Initially, we account for this class of counterexamples by introducing the Category Proviso,[19] as follows:

(59) In (1) and (2), α and β are non-distinct in categorial features.

Without going into detail concerning the precise nature and inventory of categorial features (a matter we will return to briefly below), we take it to be uncontroversial that DP and PP are categorially distinct from V in (58). This suffices then to exempt these cases from the effects of FOFC. FOFC should now be restated as follows:

[19] Thanks to Neil Myler for coining this term.

(1') **The Final Over Final Constraint (FOFC):**

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, where:

(i) α and β are non-distinct in categorial features (Category Proviso).

In order to retain the idea that the data reviewed in §2.1.1 falls under FOFC, we must assume that the clausal functional categories v, T and C, are non-distinct from one another and from V, and that the functional categories making up the nominal (as well as the Finnish postposition in (23b)) are non-distinct from one another and from N. We will return to this point when we introduce the notion of Extended Projection below.

The Category Proviso also has implications for how FOFC applies in morphology, as pointed out by Myler (2009). The following two configurations, repeated here from §2.1.1.4, are relevant for the morphological application of FOFC:

(38) $[_{\beta P}\ [_{W}$ Prefix Root] β]

(39) *[[Prefix Root] Suffix]

(38), we suggested, allows us to see the regularity of application of the suffixing preference in OV languages in terms of FOFC, since the Root must be the head of the word in a head-final phrase, and (39) holds as a general constraint on complex words. The Category Proviso leads us to add that this is true only where β in (38) and Suffix in (39) are categorially non-distinct from Prefix. For inflectional morphology, if we make the standard assumption that inflections do not differ in category from the stems they attach to (see Spencer (1991:193f.)), then Prefix and Root are categorially non-distinct in (38) and (39); if β is a functional head of the same categorial type as W, then FOFC applies and the suffixing preference is explained. The same holds if Suffix is of the same categorial type as Prefix and Root in (39). Assuming that V and O are categorially distinct, however, we cannot in these terms derive the fact noted above in

§2.1.1.3 that the suffixing preference holds so much more strongly in OV languages than in VO languages.[20]

Concerning derivational morphology, the Category Proviso may in fact explain why FOFC largely fails to apply here (see Myler 2009:53 and §2.1.1.3). If derivational morphology typically (Spencer 1991:9) involves category-changing affixes, then FOFC will not apply in (39), since Suffix and Prefix will always be categorially distinct – cf. (41), repeated here:

(41) a. $[_N [_V$ be $[_N$ head]] ing]

b. $[_N [_V$ en $[_A$ noble] ment]

However, if W is inserted as a complement to β in a given functional domain, then Prefix and β may not differ in category although Prefix and Root will, since we are dealing with derivational, i.e. category-changing, morphology. Hence (40) should still be ruled out. Hence FOFC may in fact derive the slight suffixing preference for derivational morphology (see Myler 2009:39).[21]

The second proviso has to do with the distinction between spines and satellites. The notion of the clausal or nominal spine is fairly natural, designating the relevant lexical projection (NP or VP) and all the functional categories which select that projection or which select one another (there may be a connection between our expository notion of spine as introduced here and the notion of phase in Chomsky 2000, 2001; we return to this point in the conclusion). Slightly more formally, we can

[20] The natural proposal here is that Julien's (2002) account of agglutination in OV languages is correct, and that complex words are built up in these languages by "roll-up" phrasal movement into the specifiers of heads which are realised as affixes. Although the Category Proviso on FOFC still applies, on Julien's view there is nonetheless a structural difference between complex words in OV agglutinating languages and VO languages, whether agglutinating or not. In this context, a more refined survey of G&H's data is required in order to determine whether their results hold differently in agglutinating OV languages and fusional ones.

[21] Myler (2009:41) cites the Afrikaans Noun *gewerkery* ("excessve/boring work") as a counterexample to FOFC where the derivational suffix does not change category, as it derives the Noun from the nominalisation *gewerk* ("work" (n.)): *[$_N$ [$_N$ ge [$_V$ [werk] Ø]] ery]*. However, it is unclear that we could not derive the noun *gewerk* from a category-neutral root (note that *werk* itself is made verbal by a null affix on Myler's analysis), and then convert *gewerk* into a concrete noun with a null affix, or the abstract pejorative noun by the suffix *–ery*. In both cases, the nominaliser acts on an acategorial root, and escapes FOFC due to the Category Proviso.

Another case Myler gives as a possible FOFC violation is *[$_A$ [$_A$ anti [$_N$ war]] ish]*. But *ish* is scarcely an affix in colloquial English, being able to stand alone under ellipsis (*Is John tall? Ish.*)). It may be a kind of adverbial adjunct to the word.

define spine as follows (this definition is inspired by the notion of g-projection in Kayne (1983:225), but certainly not identical to it):

(60) *Spine:*

A sequence of nodes $\Sigma = \{\alpha_1 \ldots \alpha_i \ldots \alpha_n\}$ forms a spine iff:

i) α_n is a lexical head H^{min};

ii) α_i is H^{-min}, a projection of α_n;

iii) for all $\alpha_{m>i}$, α is a head H' which c-selects either H or some $\alpha_j \in \Sigma$, or α is a projection of some $\alpha_j \in \Sigma$.

What (60) says is that a spine is made up of a lexical head, any non-minimal projections of that head and any functional category whose head either selects the lexical head or selects a member of the spine. We now define satellite as follows:[22]

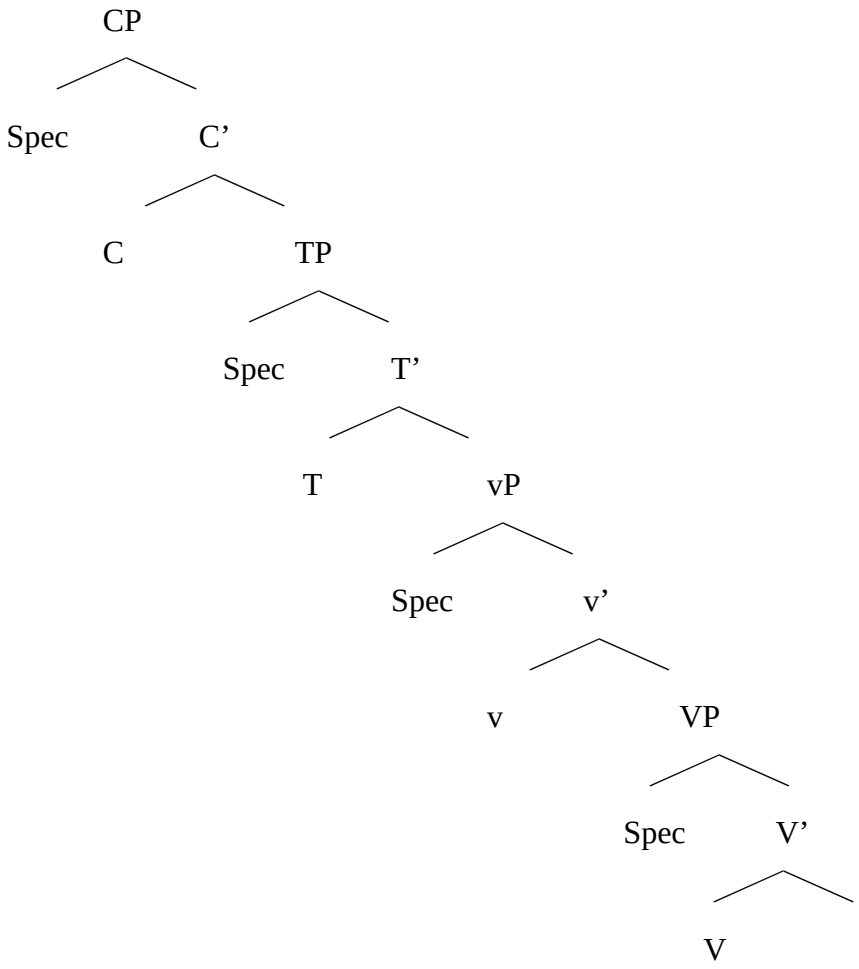
(61) *Satellite:*

All nodes β such that β is contained in $\alpha_i \in \Sigma$ and $\beta \notin \Sigma$ are satellites.

Let us apply these definitions to the structure of the clause adopted by Chomsky (2000, 2001):

[22] See Note 33 for a definition of containment.

(62)



Here we have the spine Σ consisting of V, a lexical head and so a member of Σ by (60i), V’ and VP, projections of V and so members of Σ by (60ii), v’ and vP, functional categories whose head, v, c-selects VP, and so members of Σ by (60iii), and T’, TP, C’ and CP as functional categories whose head selects a member of Σ, and therefore themselves members of Σ by (60iii). The functional heads v, T and C are also members of the spine by (60iii). On the other hand, all categories contained in members of Σ but not themselves members of Σ are defined as satellites by (61). This means that all the specifiers are excluded from the spine. Hence, for example, the subject is not part of the spine. Similarly any category adjoined to a member of the spine is defined as a satellite. Finally, the complement of V is also not part of the spine.

Let us now look at a real example of how these definitions apply in relation to FOFC. Consider the preposed genitive structures found in English and elsewhere in Germanic:

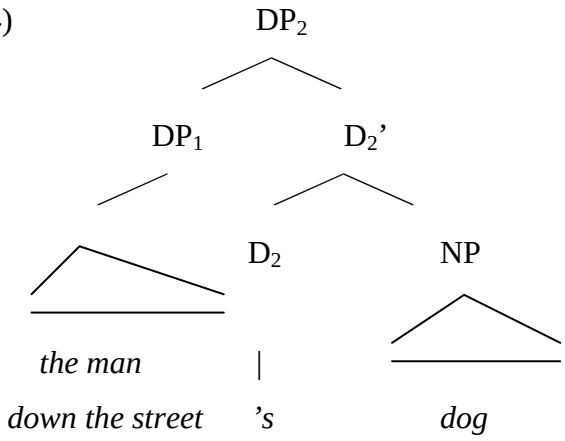
(63) a. [[that man]'s] hat

b. [[the man down the street]'s] dog

c. [[die meisie van Stellenbosch] se] wyn [Afrikaans]
the girl from Stellenbosch POSS wine
"the girl from Stellenbosch's wine"

We would naturally assign a structure like (64) to (63b), abstracting away from Num and other possible functional categories in the DP occupying the space between D and NP:

(64)



(The numerical subscripts on the D-nodes have no theoretical significance, they merely serve to distinguish the two DPs that we find here). Here DP_1, *the man down the street*, appears in the Specifier of the D_2-head *'s* (see Abney 1987 on the position of *'s* in terms of the DP hypothesis). This creates a structure which violates the schema in (2) for $\alpha=D_1$ and $\beta=D_2$, and so we expect a FOFC violation. However, the structure is grammatical. Clearly, the Category Proviso will not help us here, since α and β are categorially identical. However, DP_1 is merged as the specifier of DP_2; this makes it a satellite in relation to the spine of the projecting nominal. This becomes clear if we apply the definitions in (60) and (61) to (64): the spine Σ here consists of the lexical head N, any non-minimal projections of N (it is fact not clear that there are any in this case, but we leave the matter open), and any functional categories whose head c-selects a member of Σ (e.g. N), i.e. D_2' and DP_2. Satellites are those categories not in Σ, but contained in a member of Σ; hence we see that DP_1 is a satellite. So we now add a second proviso to FOFC, the Satellite Proviso:

(65) α is not a satellite of β.

The Satellite Proviso in (65) exempts structures like (64) from FOFC. FOFC is now stated as follows:

(1'') **The Final-Over-Final Constraint (FOFC):**

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, where:

(i) α and β are non-distinct in categorial features (Category Proviso);

(ii) α is not a satellite of β (Satellite Proviso).

In (64), DP_1, as a possessor, may have been first-merged as the specifier of DP_2 (although this is not obvious; cf. Longobardi 1996, Julien 2005:197ff. and others have proposed that the possessor always raises to this position). In "Passive in NP" constructions of the kind shown in (66), however, one would naturally take the pre-'*s* DP to have raised to that position from an argument position inside NP:

(66) the city's destruction by the emperor

Similarly, in the "active" counterpart to (66), we might take the agent argument to have raised from Spec,nP:

(67) the emperor's destruction of the city

In both of these cases, the Satellite Proviso saves the structure from FOFC. This contrasts with the VOAux case discussed in §2.1.1.1, where VP or vP is fronted; as we saw above, neither VP nor vP is a satellite in relation to the clausal spine, both are part of the clausal spine.

The Satellite Proviso may also be relevant in connection with the discussion of compounding in relation to FOFC in §2.1.1.4. There we observed that the compounds in examples like (44), repeated here, apparently violate FOFC:

(44) a. an [[I-couldn't-care-less] attitude]

b. the [[man-of-the-match] award]

c. an [[easy-to-please] customer]

d. the [[Final over Final] Constraint]

e. the [[channel four] news]

We noted that Myler (2009) proposed that these compounds are effectively reanalysed as lexical items by the process of "renumeration". If, as a consequence of renumeration, they can be treated as (non-lexical) heads, then they are defined as satellites by the definitions in (60) and (61), and hence fall under the Satellite Proviso as given in (65). We may then be able to maintain that compounding in general obeys FOFC; see §2.1.1.4. But since the relation between the complex bracketed constituent in (44) and the head noun is one of modification, we can treat these elements as satellites in any case. Therefore they will be exempted from FOFC by the satellite proviso. More generally, we take it that in root compounds the first element is a satellite in relation to the second, and hence they do not violate FOFC. In synthetic compounds, on the other hand, the first element is a complement of the same (null) category as the second; hence these compounds obey FOFC. We can now strengthen the conclusion of §2.1.1.4 above: FOFC holds of morphological processes generally; it holds clearly in inflectional morphology, as Myler (2009) demonstates; it holds vacuously in derivational morphology owing to the Category Proviso; it holds vacuously of root compounds owing to the Satellite Proviso, and it holds of synthetic compounds (following the suggestion in in Note 16, we assume that the stem *can-open* has no category in a synthetic compound like *can-opener*, and so cannot be categorially distinct from the nominal suffix *-er*).[23]

The third proviso concerns movement types. There is clear evidence that in many languages, including English, VP (or vP) can be fronted to SpecCP, an operation

[23] This appears to contradict our suggestion for the reanalysis of Myler's Afrikaans nominal *gewerkery* in Note 21. In that case, we said that the nominaliser *–ery* acts on an acategorial root, and escapes FOFC due to the Category Proviso. The difference between the two cases seems to be that *gewerk* can surface as a word, while *can-open* cannot. It may be possible to maintain the following: possible words are morphological phases, and FOFC always holds phase-internally (all of the exceptions to FOFC we review clearly involve phase boundaries, but not all phase boundaries are boundaries for FOFC; FOFC in fact holds across many types of phase boundaries, see also the brief discussion of "phase-linking" in §7). Hence *gewerk* forms a phase, and the Category Proviso holds, as stated in Note 21. But *can-open* is not a phase, and so no Proviso holds, and so FOFC holds for synthetic compounds. This implies that acategorial elements can be subject to the Category Proviso, which is the natural interpretation of the non-distinctness clause in (1''). We defer further discussion of the relation between FOFC and phases until the concluding section.

involving either topicalisation or focalisation (we refer to this operation as VPF henceforth, without prejudice as to the exact category which is fronted). That this movement is A'-movement can be seen from the fact that the landing site is clearly in the left periphery and the discourse effect of topicalisation, as well as the fact that the structure is island-sensitive:

(68) We expected John to eat the pies, and ...

a. ... eat the pies (we all saw that (John affirmed that)) he did

b. ?* ... eat the pies I heard the rumour that he did

c. * ... eat the pies I met a man who told me that he did

d. * ... eat the pies I went to London before he did

e. * ... eat the pies that he did amazed me

As things stand, these structures involve a FOFC violation in English: they involve movement of a head-initial non-satellite (vP or VP) into a specifier position in the left periphery (most probably via intermediate SpecvP positions, thus entailing structures formally identical to the barred VOAux structures discussed above). One relevant piece of evidence is that VPF is impossible in finite embedded clauses without overt *that*:

(69) We expected John to eat the pies, and we all saw *(that) eat the pies he did.

In this respect (69) parallels (70):

(70) I didn't know *(that) never had Peter danced so well.

(70) is a case of CP-recursion, with the auxiliary of the embedded clause moving to the lower C-head. It was observed in Rizzi & Roberts (1989) that the higher C in a CP-recursion structure does not readily tolerate *that*-deletion, and this is what we observe in (70) (compare *I know (that) John left*, where there is no CP-recursion and *that*-deletion is fully optional). If (69) likewise features CP-recursion, with an empty lower C, then we explain the impossibility of *that*-deletion in the higher C. And then it follows that the VPF targets the lower SpecCP here (this argument is based on the

one given in Kayne (1994:28-29)). But now we have a FOFC violation in the lower CP:

(71) ... that [$_{CP}$ [$_{VP}$ eat the pies] C [$_{TP}$ ...

(71) violates FOFC for α=V and β=C.

Similar cases are found with VPF in other languages. Here is an example from Sardinian:

(72) Tunkatu su barkone asa (Jones 1988: 339)
shut the window have-2SG
'It's shut the window you have!'

Assuming that the VP (or vP, or PrtP) *tunkatu su barcone* is raised to a SpecCP position here (see Jones for discussion and arguments to this effect), then we have the same configuration as in (70).

In the Mainland Scandinavian languages, we see the violation more directly, since these are V2 languages and so VPF to SpecCP is accompanied by movement of a finite verb or auxiliary to C (still assuming a standard analysis of V2; see above):

(73) Åt pajerna såg vi/ har vi sett att han gjorde [Swedish]
ate pies-DET saw we have we seen that he did
'Eat the pies we saw/have seen that he did'

Swedish VPF is clearly A'-movement: it is movement to SpecCP, it has a discourse effect (broadly comparable to that of English VP-fronting), and is island sensitive:

(74) Vi misstänkte att Johan skulle äta pajerna, och ...
we suspected that John would eat pies-DET and
'We suspected that John would eat the pies, and ...'
a. ...åt pajerna hörde vi [att han berättade för Elsa [att han gjorde]]
eat pies-DET heard we that he told to Elsa that he did
'... eat the pies we heard that he told Elsa that he had'

b. ?*… åt pajerna hörde jag ett rykte att han gjorde
eat pies-DET heard I a rumour that he did

c. *… åt pajerna mötte jag en man som sa att han gjorde
eat pies-DET met I a man who said that he did

d. *… åt pajerna for jag till London innan han gjorde
eat pies-DET went I to London before he did

e. *... åt pajerna att han gjorde förvånade oss inte
eat pies-DET that he did surprised us not

So we exempt A'-movement from FOFC, by adding the A'-Movement Proviso:

(75) αP has not been A'-moved to SpecβP.

Note that VPF illustrates the pure case of the A'-Movement Proviso, since, in the configuration in (71), VP is neither categorially distinct from C nor a satellite of the clause.

So now we have the following statement of FOFC:

(1''') **The Final-Over-Final Constraint (FOFC):**
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, where:
(i) α and β are non-distinct in categorial features (Category Proviso);
(ii) α is not a satellite of β (Satellite Proviso);
(iii) αP has not been A'-moved to SpecβP (A'-Movement Proviso).

Returning to the standard cases of wh-movement in (57), we can now see why they do not produce FOFC-violating structures:

(57) $[_{CP}\,[_{DP}$ which man] $[_{C}$ did] $[_{TP}$ you see …] ?

Here, just looking at the configuration, we apparently have a FOFC violation since we have an instantiation of (2) for α=D, β=C. But all three provisos are relevant: the DP is categorially distinct from CP, it is a satellite and it is A'-moved.

Clearly, the three provisos to FOFC that we have proposed, although independently motivated, as we have seen, are often redundant. We can in fact collapse the Category Proviso and the Satellite Proviso under the following definition of Extended Projection, inspired by, but distinct from, the proposals in Grimshaw (1991, 2001, 2005):

(76) The Extended Projection of a lexical head L (EP(L)) is the sequence of categories EP = $\{\alpha_1 \ldots \alpha_i \ldots \alpha_n\}$ such that:

i) α_i is in the spine defined by L;

and for each pair of heads $<H_i, H_{i+1}>$ in EP

ii) H_i c-selects H_{i+1};

iii) H_i is categorially non-distinct from H_{i+1}.

What (76) says is that the Extended Projection of a lexical head, for example V, is the spine defined by V in terms of (60), i.e. the projections of V, all the functional material which selects VP, and all the functional material which selects functional material which selects VP, up to the first categorially distinct element.[24] Thus, in a standard clause such as that shown in (62), all the projections of v, T and C are in the Extended Projection of V. Moreover, so are the heads of these projections. Finally, where CP is selected by a higher V, that V and the entire further clausal functional structure associated with that V will be in the Extended Projection of the lower V. However, if CP is selected by N, that element, and its Extended Projection, will not be in the Extended Projection of V, since they are categorially distinct from V. This approach has the consequence that a head-initial CP cannot be directly selected by a head-final V, or a FOFC violation (C=α, V=β in terms of (2)) results. This derives part of the generalisation noted in §2.1.1.2: in OV languages embedded clauses are either postverbal or preverbal but nominalised.[25] If the preverbal nominalisations are DPs, then there is no FOFC violation thanks to the Category Proviso. We return to the question of the postverbal CPs in OV languages and how they come to be postverbal in §5.1.1 below.

[24] Functional heads may have features not shared by lower elements in their Extended Projection, but these features would be either probing features (unvalued φ-features, etc) or substantive features (tense, aspect, etc). The categorial features are shared, however, and these define the Extended Projection.

[25] In fact, we allow a third possibility: OV languages with final complementisers allow preverbal, non-nominalised CP complements. This is what we see in Japanese, as in (15).

The definition of Extended Projection just given allows the following formulation of FOFC:

(1′′′′′) **The Final Over Final Constraint (FOFC):**

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, where:

(i) α and β are in the same Extended Projection;

(ii) αP has not been A'-moved to SpecβP.

This is an improvement on (1′′′). We will integrate the A'-Movement Proviso more fully in §6.

So we have seen three types of "principled counterexamples" to FOFC, which have led us to restate the generalisation in a sharper way, ultimately as (1′′′′′). However, there remain some constructions which appear to constitute less principled counterexamples. Here we introduce and comment on them; we will return to them sporadically throughout the rest of our presentation.

One very prominent class of counterexamples involves sentence-final particles in otherwise head-initial languages. These are found in genetically unrelated languages in two main geographical areas: East Asia and Central Africa. The following are representative examples:

(77) a. Ni yao kan zhe-ben shu **ma**? (Mandarin, Aldridge 2009)
you want read this-CL book Q
'Do you want to read this book?'

b. Kɔ̀kú yrɔ́ Kòfí **à**? (Gungbe, Aboh 2004: 318)
Koku call-PERF Kofi Q
'Did Koku call Kofi?'

It is very tempting to analyse these particles as Cs. If so, we would have instances of final Cs in VO languages, and hence counterexamples to the generalisation put forward in §2.1.1.2 above (FOFC is violated for α=V and β=C here).

Although many cases involve putative C-elements like those in (77), final particles occur in phrases of all types: Aspect, Mood, Negation, Polarity, Specificity. Often, the languages with these types of particles are "repeat offenders", with multiple FOFC-violating elements (see Dryer 2009b).

We can, however, immediately make two observations. First, we do not find this kind of order with true subordinating Cs, and so the generalisation in §2.1.1.2 is at least partially intact and hence in need of explanation. Second, we see positional discrepancies in languages with particle and non-particle counterparts of "the same" category of element:

(78) a. Tân mua gi **the**? (Vietnamese)
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)'[26]

(79) a. yə- ca dɛyo **l**ɔ (Bwe-Karen)
1SG-see picture ASP
'I am looking at a picture'

b. ce-ɗɔ mi **jə-kh**ɔ phi má nɔ (***jə-kh**ɔ)
3- say C 3- FUT take what
'What did he say that he would take?' (cf. Dryer 2009a)[27]

(78) shows that in Vietnamese the sentential particle in final position is distinct in form from the complementiser-like element in medial position. In (79), we see that

[26] Cf. also Taiwanese *kong* (Simpson & Wu 2002), Mandarin *shuo* (Wang et al. 2003), and Cantonese *waa* (Yeung 2006), as well as the sentence-final particles in certain Northern Italian Dialects (Munaro & Poletto 2006).

[27] Cf. also Duffield (2001) on Vietnamese *được*, Cheng & Sybesma (2003) on so-called "forked modality" in Chinese, and Simpson (2004) and Enfield (2002) on final possibility modals more generally.

Bwe-Karen has an uninflected final aspectual particle which is distinct in position and in morphological properties from the medial inflected tense marker. What these examples suggest is that particles are distinct from standard auxiliaries and complementisers (cf. also Greenberg 1963: 66-7), and, more generally, that inflecting and particle elements are formally distinct in way that is crucially relevant for FOFC.

In fact, it is not difficult to see that these particles display a number of unusual properties. First, as we have just seen, there is evidence that particles differ from “full” elements expressing similar meanings (clausal force, tense, aspect, negation, definiteness, etc.). In general, they seem unable to be syncretic: for example, subordinators express only subordination, and interrogatives express only interrogativity; hence we do not find particles corresponding to English *if*, which expresses both. Similarly, the Bwe-Karen aspect particle above only expresses aspect and not agreement as well. This point is further illustrated in Biberauer, Newton & Sheehan (2009), and also in connection with specific languages in i.a. Aboh (2004, 2006), Paul (2009), Adger, Harbour & Watkins (2009). Particles of this type frequently appear to have what we might think of as very strict (syntactic and semantic) selectional restrictions, being licit only in structures containing elements whose properties do not clash with the very specific ones of the particle in question (cf. for example Adger et al. 2009:68-82 for discussion of what they designate *selective particles* (see below) in Kiowa).

Particles may, however, also differ from "full" elements in the opposite direction, i.e. in apparently imposing *less strict* selectional constraints than those typically associated with "full"-element counterparts. In this connection, Adger et al. (2009:69) draw a distinction between *selective* and *non-selective* particles, with the former being restricted to very specific environments and the latter surfacing in less specifically delineated contexts. Examples of particles that exhibit *non-selective* behaviour include negation, polarity, topic and focus (cf. for example Biberauer 2009a for discussion of Afrikaans polarity particle *nie*$_2$, which may surface as a left-peripheral final element in CPs, vPs, DPs, PPs, and APs; Aboh 2004:240-241 and 290ff. for discussion of the Gungbe topic and focus particles *yà* and *wɛ*, which exhibit similar behaviour).

In connection with negation, it is worth noting that, if we follow the approach to negation instigated by Pollock (1989) and take negation to be a head (with any of VP, TP or CP as its complement), then SVO structures featuring a final negation element

of the sort found in Central Africa (cf. Dryer 2009b) and in the Pacific region discussed by Reesink (2002) may well be counterexamples to FOFC.[28] However, negative markers do not obviously belong to any given syntactic category, and can occupy a wide range of positions in the clause. In fact, Cinque (1999: 126) remarks that "the evidence points to generating a NegP on top of every adverb-related functional projection". Negation, then, may well fall outside of Extended Projections. We will return to this point in §5.1.3 below.

Significantly, Adger et al. (*ibid.*) observe that there is an important difference in the positional behaviour of selective and non-selective particles in Kiowa: while the former surface in fixed preverbal positions which appear to respect a version of the Cinque hierarchy (in effect, behaving like the complementiser and auxiliary in (79b) and (80b) respectively), the latter necessarily surface postverbally in less rigidly constrained positions. Apparently, then, the distinction between selective and non-selective particles may also correlate with positional differences, in terms of which non-selective particles exhibit behaviour which may not always appear to respect hierarchically imposed structural constraints. This point requires detailed investigation in future research. What we suggest here is that there are indications that non-selective particles may in fact fall outside of FOFC owing to the fact that they are in a sense acategorial, i.e. incapable of projecting structure of a particular category) and thus not part of the Extended Projection of the category that they combine with. The same may be true of apparently FOFC-violating particles which can independently be shown not to be fully integrated into the head-initial structure at the end of which they surface (cf. Biberauer & Cyrino 2009 on Brazilian Portuguese final *não*, which, just like English polarity tags, as in **I own anything/a red cent, right?*, cannot license a Negative Polarity Item). To the extent that FOFC-violating selective particles also fail to project and thus do not become part of the Extended Projection of the structures in which they surface, similar argumentation can be extended to these

[28] As noted above in connection with SVONeg orders in German, this is only true insofar as it can be shown that the final negation elements in these languages are structurally higher than the languages' head-intial verbal phrases. In the case of many of the languages discussed by Dryer (2009b) in particular, it would indeed appear to be the case that the final Neg-element is structurally higher than the verb and the object, thus posing a problem for FOFC.

Also worth noting here is that negative morphemes constitute one of the exceptions to FOFC inside the word observed by Myler (2009); see §2.1.1.4.

elements (cf. i.a. Toivonen 2003, Svenonius 2008 for the view that particles generally may fail to project). We develop this idea further in §5.1.2 below.[29]

In this section we have seen that there are three main types of counterexamples to FOFC. Some are simply misleading surface orders, readily analysable in a FOFC-compliant way. Some fall under various rather systematic provisos: the Category Proviso, the Satellite Proviso and the Movement Proviso. We have seen that the first two can be subsumed under a variant of Grimshaw's (1991, 2001, 2005) notion of Extended Projection, although we continue to stipulate that A'-moved categories are exceptions to FOFC. Finally, we have seen some more recalcitrant counterexamples, mainly involving final particles in otherwise VO languages. We have tentatively suggested that the notion of Extended Projection may be relevant here, although why and how remain unclear. What is clear in this connection is that the categorial specification of particles is an issue independently of FOFC. If we can capture the properties of these elements in formal terms, then we would perhaps be able to develop an account of why particles appear to violate FOFC, whereas their "full" counterparts do not do so.

We put to one side the recalcitrant cases for now, although we will return to them in §§5.1.2 and 5.1.3. What we want to do now is show how FOFC, stated as in (1"") can be derived from general principles of UG.

4. Linear order and movement

In the previous section, we arrived at the following formulation of FOFC:

(1"") **The Final-Over-Final Constraint (FOFC):**

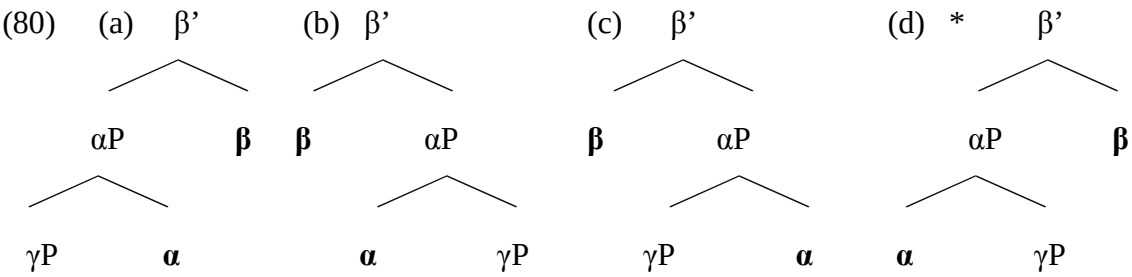
If α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final, where:

(i) α and β are in the same Extended Projection;

(ii) αP has not been A'-moved to SpecβP.

[29] To the extent that they can be characterised as representing a "high" mood category, often connected to illocutionary force, Mood particles correspond to one of Myler's classes of counterexamples to FOFC in morphology (see §2.1.1.4). The same is true of illocutionary particles (cf. Myler's "discourse particles") and negation (see previous note). We return to this point below.

In order to see what FOFC is saying, consider the logically possible complementation combinations among head-initial and head-final categories:



Consistent head-final (harmonic) | **Consistent head initial** (harmonic) | **Initial-over-Final** (disharmonic) | **Final-over-Initial** (disharmonic)

As (80) says, three of the four logically possible combinations are allowed and one is nonexistent (within a single Extended Projection and leaving aside A'-moved categories). In fact, the harmonic configurations in (80a,b) are very common, while (80c) is somewhat less common but still occurs. In other words, harmony is preferred (as has often been observed: Greenberg 1963, Dryer 1992, Baker 2008), but disharmony is allowed. Crucially, though, *only one kind of disharmony* is allowed. These observations set a challenge for any account of linearization: the theory of linearization should predict (a) the preference for harmony and (b) the fact that only one disharmonic order is allowed (i.e. FOFC).

In this section, we will briefly consider and reject an explanation based on the Head Parameter. We will then outline our own proposal, which relies on Kayne's (1994) Linear Correspondence Axiom (LCA).

4.1. The Head Parameter

We can formulate the Head Parameter, fairly standardly, as in (81):

(81) A head {precedes/follows} its complement in X'.

This formulation covers both the versions of the Head Parameter based on X'-theory (as in Koopman 1984, Travis 1984, for example), as well as more recent approaches (such as Fukui & Saito 1998), and also more recent approaches to linearization where

this parameter is seen as a derivative option at PF, or as a facet of the mapping to PF (as in Richards 2004, Fox & Pesetsky 2005).

We contend that (81) does not provide the basis for a principled explanation of FOFC. The existence of disharmonic systems means that the Head Parameter must be relativized to categories. In German, for example, DP and CP are head-initial, while TP and VP are head-final. But if the Head Parameter can be set independently for each category, then the fact that only one type of disharmonic combination is allowed cannot be stated; what prevents α being set to head-initial and β to head-final, where α is the complement of β? It seems that the Head Parameter cannot on its own predict FOFC. Note also that it cannot predict cross-categorial harmony, if relativised to categories.[30]

In sum, the Head Parameter of (81) is either too strong or too weak. If not relativised to categories, it can in fact predict both FOFC, and, falsely, "inverse FOFC" as is also ruled out and cross-categorial harmony, but then it is too strong in that it predicts that systems like German (along with Latin, Chinese and many others) should not exist. If relativised to categories, it can account for the attested disharmonic systems, but then predicts neither the specific constraint on FOFC-type disharmony nor cross-categorial harmony. So we abandon this kind of approach.

4.2 FOFC and processing

It may seem initially plausible that FOFC could have an explanation in terms of principles of online processing. 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

[30] If we combine a relativised Head Parameter with the assumption that acquirers operate on the basis of an acquisition strategy akin to Roberts' (2007) Generalisation of the Input principle, cross-categorial harmony could, of course, be seen to fall out (see Roberts & Holmberg 2010 for discussion). Importantly, however, Generalisation of the Input cannot play a role in ensuring that only the attested forms of disharmony are acquired.

two elements in a syntactic structure; see Hawkins (2004: 33ff). In harmonic structures, as in (80a,b), 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 the disharmonic (80c,d), by contrast, there will be items intervening between the two heads. A challenge, for this kind of processing-based theory of word-order universals is to explain why the FOFC-compliant disharmonic order (80c) is still relatively common, while the FOFC-violating order is rare, or even non-existent, as we would claim, once certain systematic exceptions are accounted for. Hawkins (ibid.) is well aware of this difference between (80c,d), and his theory is, in part, designed to explain it (in terms of his Early Immediate Constituents principle; see Hawkins (1990b, 1994: 95ff, 2004: 103ff), although Hawkins 2010 expresses scepticism regarding the difference in the cross-linguistic incidence of (80c) and (80d)).

There is, thus, at least one extant theory of processing which addresses the problem posed by FOFC. We will, however, not discuss Hawkins's theory here (but see Sheehan forthcoming). In part, this is for reasons of space, but also because we think there is at least one good reason to doubt that processing holds the key to FOFC: if we are right (following Myler 2009 in this respect) that FOFC applies word-internally, then this will all but rule out the possibility that FOFC is ultimately a matter of syntactic parsing along the lines of Hawkins (ibid.), because in that theory the minimal unit of analysis is the word. For Hawkins, complexity (or weight, or distance) in parsing is a matter of number of words and constituents larger than words, not of the number or arrangement of sub-word units such as morphemes or syntactic features (with the exception of certain components of words in agglutinating languages) (John Hawkins, p.c.). This is, in itself, not an uncontroversial hypothesis, and we acknowledge that a detailed discussion of processing and FOFC, taking into account the proposals in Hawkins (2010) and addressing processing theories other than Hawkins's, is called for. However, we defer this discussion to future publications.

4.3 An LCA-based account of linearization

As is well-known, Kayne's (1994) LCA ties asymmetric c-command to linear order, specifically precedence. We state the LCA as follows:[31]

(83) **The LCA**

If α asymmetrically c-commands β, then α precedes β.

We define c-command, standardly, as follows:[32]

(84) a. α c-commands β iff β is contained in the sister of α.

b. α asymmetrically c-commands β iff α c-commands β and β does not c-command α.

Consider the standard X' structure, as in (85):

(85) $[_{XP}\ \alpha\ [_{X'}\ X\ \beta\ ]]$

Given (83) and (84), the specifier, α, precedes the head X, since that head is contained in α's sister. As long as β has internal structure, X will precede it, since X asymmetrically c-commands anything contained in β (and containment dependencies cannot "cross"). If β has no internal structure, X and β cannot be ordered as either they c-command each other, or there is no c-command relation (this depends on

---

[31] We formulate (83) as a one-way conditional, as the biconditional version runs into difficulties with complex left branches. For example, in (i), γ precedes β but does not asymmetrically c-command it:

(i) [[α γ] β]

This is allowed by the formulation in (83). See Uriagereka (1999) for a discussion of structures like (i) in relation to the LCA. There is an obvious similarity between (i) and the FOFC-violating structure in (2): where α is a head, (i) instantiates the FOFC-violating structure in (2). Sheehan (2009) pursues a general account of FOFC based on this observation.

In our formulation of the LCA, we depart from Kayne's original formulation but follow the basic ideas of bare phrase structure in taking α and β to be potentially both terminal nodes and lexical items; in particular, we do not regard lexical items as constituents of categories, i.e. we do not assume a head-terminal distinction. This leads to the well-known "deepest sisters" problem, which we will briefly discuss below. To some extent, the residue of the LCA that we are assuming is the central idea that phrasal categories are linearised as Spec-initial and head-initial, unless movement of some kind takes place. This is the essential core of the LCA that we believe to be necessary for an account of FOFC (thanks to Iain Mobbs for discussion of this point).

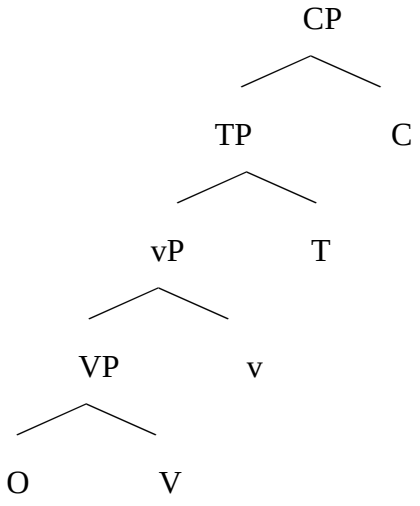
[32] If "contain" is irreflexive all c-command is asymmetric, and (84b) is not needed.

whether "contain" is reflexive; see Note 32). Further specifiers and adjuncts will be ordered amongst themselves and will always be to the left of the "core" X' containing X and β, given the definitions in (83) and (84).[33]

To the extent that movement is Internal Merge, and Merge must always apply at the root, then a moved element will always asymmetrically c-command its copy. The LCA then guarantees that movement is always leftward. It is worth noting that surface linear order is, quite independently of LCA-related assumptions, very often in part the result of movement of one kind or another (A-movement, A'-movement, etc). The proposal here, as in Kaynian work more generally, is that surface head-final order is also always the result of movement: in order to precede a given head, a complement must move to a position where it asymmetrically c-commands that head (Kayne 1994:47-48).

Furthermore, as also originally pointed by Kayne (1994:52-53), consistent head-final order is derived by "roll-up" (successive leftward movement of complements and categories containing moved complements).[34] The derived structure of a "roll-up" derivation in CP looks as follows (ignoring copies of movement):

(86)

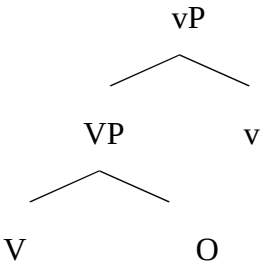


[33] Following Chomsky (1986:12) (who follows May 1985), we define inclusion, dominance and exclusion as follows:
(i) α includes β iff some segment of α contains β.
(ii) α dominates β iff every segment of α contains β.
(iii) α excludes β iff no segment of α contains β.
The notion containment, more precisely immediate containment, can be reduced to set-membership via Merge: α immediately contains β iff $\beta \in \alpha$, the set formed by merge of β with γ (where possibly $\gamma = \alpha$). We now add: α contains β iff either α immediately contains β or α immediately contains γ such that γ (immediately) contains β.

[34] This is the derivation of head-final orders we concentrate on here. But it is not the only way to derive surface head-final orders, as these can also result from Agree-driven movement. We will examine this latter case in relation to FOFC in §6.

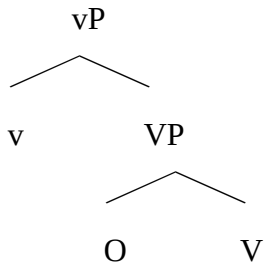
This is clearly a harmonically head-final structure, instantiating the possibility seen in (80a). Equally clearly, a harmonically head-initial structure arises where no complement-movement takes place. Most importantly for our purposes, we can now see that disharmonic orders result when some complements, and/or elements contained in those complements, undergo movement and others don't. If, as again standardly assumed, movement of an XP is always triggered by a property of some head, then what FOFC shows is that the distribution of the movement-triggering property is constrained, so that only one type of disharmony is allowed. More precisely, a typical FOFC violation will arise when a superordinate head triggers movement of its complement, but inside that complement the head does not trigger movement of *its* complement. Suppose, for example, that v triggers VP-movement but V does not trigger object-movement. Then we have the structure in (87):

(87)



If v can contain an auxiliary, then this gives surface VOAux order, the FOFC violation discussed in §2.1.1.1. More generally, (87) is clearly an instantiation both of (2) and of (80d). On the other hand, if V triggers movement of its complement, and v does not, we arrive at the FOFC-compliant disharmonic structure shown in (88):

(88)



Hence, if a superordinate head does not trigger movement, but the head of its complement does, the result is permissible, non-FOFC-violating disharmony. Note that this means that disharmonic languages are always partially harmonic: there is a

node in the disharmonic Extended Projection, such that above that node they are harmonically head-initial, and below that node they are harmonically head-final.

We are now in a position to take a crucial step forward in understanding FOFC. It emerges from the above discussion that head-final order can be derived by complement-movement, as long as, when iterated, complement-movement starts at the bottom of the tree, and iterates monotonically up the tree. The iterations can stop at any point, though, as long as the stopping is definitive, i.e. as long as complement-movement does not "start again" in a higher position (in the same Extended Projection).[35] This is an informal, movement-based statement of FOFC. We now have to make the statement more formal, and explain exactly why complement-movement should be constrained in this way.

4.4. Our proposal

It should be clear from the previous section that our account of FOFC relies on movement, and in particular on the way in which movement is triggered. Accordingly, we adopt the following idea:

(89) Movement is triggered by a general movement-triggering feature. We use ^ (caret) as a symbol of this feature.

We take ^ to be a purely formal, arbitrary diacritic. In itself, it has no semantic content, and no connection to phonological or morphological properties beyond simply causing movement (this may, however, under certain circumstances result in configurations associated with specific morphophonological effects; morphophonological effects are not therefore excluded). Moreover, although it can be seen as a kind of formal feature, ^ differs in several important respects from formal features like φ-features. Unlike φ-features, which are arguably best seen as attribute-value pairs, it has no internal structure, it cannot be valued or in any obvious way

[35] Cinque (2005:326) makes a similar observation in connection with the internal order and structure of DPs. As we mentioned in §2.1.1.3, his analysis requires him to ban movement of all DP-internal categories unless NP moves first. Hence his question: "Why is movement of phrases other than NP unavailable?" (325). We saw in §2.1.1.3 that this is connected to FOFC.

"checked off", and, as already mentioned it has no semantic or morphophonological effects.[36]

The idea that movement is triggered by a purely formal diacritic is widespread in the current literature. In different versions, and with different notations, it appears in Müller & Sternefeld (1993), Chomsky (2000, 2001, 2008), Pesetsky & Torrego (2001) and Roberts & Roussou (2003); the idea of a "spell-out" diacritic associated with certain positions is also found in the representational system proposed in Brody (1995).

Very much in the spirit of Müller & Sternefeld (1993), we take it that the properties of different types of movement depend on the features that ^ is associated with. Where the movement-trigger ^ is associated with the uninterpretable ϕ-features of an active Probe, it gives rise to A-movement; in this respect it functions exactly as the EPP features of Chomsky (2000, 2001).[37] Where ^ is associated with the Edge Feature of a phase head, it triggers A'-movement (see Chomsky 2008:144).[38] Finally, and most important for our purposes, where ^ is associated with the c-selection feature of a lexical root, then 'Linearization movement,' movement of the complement as seen in the previous section (L-movement henceforth), takes place.[39] Examples of the different types of movement triggers are:

---

[36] This raises the question of the status of ^ at the interfaces. It may be that LF simply ignores this element, since it has no denotation; unlike ignoring an unvalued/uninterpretable φ-feature, this has no deleterious effects. In PF, ^ has an effect, in that the head associated with it must have a category in its specifier (although that category may be a copy).

[37] This extends to expletives, given that, following Richards & Biberauer (2005:120-5), we take it that expletives are merged in SpecvP and raise to SpecTP in languages like English (cf. also Deal 2009).

[38] 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 (2007 *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 2007:17, 2008:144; cf. Kandybowicz 2008a, 2009 for a proposal along these lines). 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. Given that movement-triggering EFs trigger *leftward* movement (cf. the discussion in §4.4), this would lead us to expect Merge also to involve merger to the *left.* We leave systematic consideration of this possibility to future work.

[39] Head movement is not triggered by ^, either because it is not part of core syntax (Chomsky 2001: 37-38), or because it is the consequence of a particular type of Agree relation, that where the Goal is defective in relation to the Probe in that its formal features are included in those of the Probe (see Roberts forthcoming where this idea is developed). Hence ^ only triggers phrasal movement. In terms of the proposal below, this restriction to phrasal movement (or the pied-piping requirement, to put it another way) follows if ^ is seen as an instruction to remerge a category inside a projection of a head. In Roberts' account of head-movement, an incorporated head is not the sister of the head it incorporates with, since the incorporation does not involve Merge.

(90) a. $T_{[u\phi, \wedge]}$ triggers movement of the goal of the probe [uϕ] to specTP

b. $C_{[EF, \wedge]}$ triggers A'-movement;

c. $V_{[\_\_ D, \wedge]}$ triggers movement of the complement DP of V to specVP.[40]

We construe c-selection, following Rizzi (2008) and Cecchetto & Donati (forthcoming), as External Search, i.e. the case where an item searches in the Numeration for a needed item (in contrast with Agree, which is Internal Search for such an item). The categorial feature imposes a constraint on what can be merged as the complement to the selecting head (in this respect, it acts like the strict subcategorization features of Chomsky 1965); for lexical heads at least, there is clearly a relation between c-selection and s-selection/θ-role assignment (see Grimshaw 1981, Pesetsky 1982), but we take no view on that here. The movement-triggering feature ^ requires the selected element to be merged twice, once as complement and once as specifier to the head bearing the feature.[41]

We can now state FOFC in terms of movement more formally:

(91) If a head α_i in the Extended Projection E of a lexical head L has ^ associated with its selection feature for a lower head α_{i+1}, then so does α_{i+1}.

We can guarantee that (91) holds if we stipulate that ^ (as a pure linearizarion feature) can only originate on the lexical head L of an Extended Projection E, and can "spread" to any higher projection. In that case, however, it must "spread" monotonically from the lexical head L of E through all the heads of E, stopping anywhere, but skipping no head in E, possibly reaching the highest head of E.

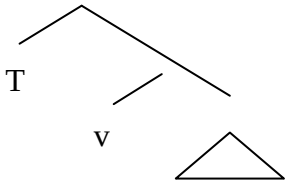
In these terms, (92a, b, c) are possible distributions of selection-based ^ in the

[40] Following Chomsky (2008:139), we could think of ^ as connected to the Edge Feature EF that a lexical item must have in order to be able to Merge; see Note 38 for discussion. However, we prefer to connect ^ to c-selection features, as there may be cases where selection for a complement of one category triggers movement while selection for a different one does not: one possible example is English synthetic compounds, briefly discussed in §2.1.1.4. In a compound like *truck-driver*, *drive* c-selects N, or Ø, and triggers movement, giving *truck-driver*, or c-selects D and does not, giving *drive trucks*. Further, certain lexical items may have EF but no selection features; we return to this below.

[41] 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 against anti-locality stand, with the putative 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 (cf. Biberauer, Holmberg & Roberts 2009).

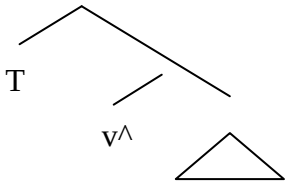
extended VP. No other distribution is possible:

(92)(a)



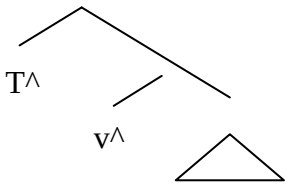
V^ XP

(b)



V^ XP

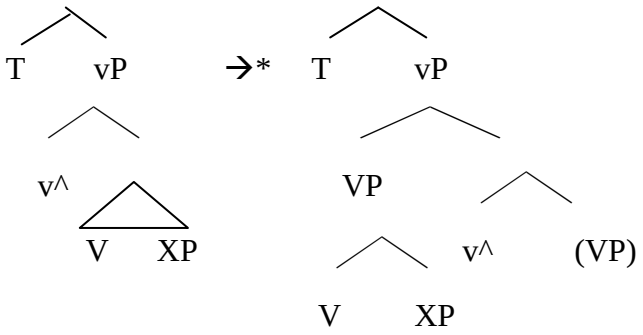
(c)



V^ XP

(92a) gives T > v > XP > V, i.e. a head-final VP, with all other categories in the Extended Projection of V being head-initial (various Niger-Congo languages instantiate this option cf. Koopman 1984, Baker 2005, Kandybowicz 2008b and Nikitina 2008; Dryer 2008a further mentions Kisi (Atlantic; Niger-Congo), Nuer (Western Nilotic; Sudan), Dinka (Western Nilotic; Sudan) and Dongo (Ubangian, Niger-Congo) – see Dryer's discussion for details and references). (92b) gives T > XP > V > v, a head-final VP and vP, but not TP (cf. Kru, discussed in Nikitina 2008 appears to instantiate this option). (92c) gives XP > V > v > T: all categories are head-final, this is of course the 'roll-up' case (cf. Julien 2002 for discussion of languages which exhibit this option). Finally, and most importantly, (93) is ruled out:

(93)



This structure gives rise to a FOFC violation. For example, if we assume, as is fairly standard, that V moves to v and Aux is in T, then (93) gives the surface order VOAux.

So we see we have the elements of an account of FOFC, but we are now faced with four questions: (A) Why should ^ be associated with head-final order? (B) Why does the movement requirement, i.e. the diacritic ^, "spread" monotonically? (C) Why does L-movement have to "start at the bottom", i.e. from the lexical head of the Extended Projection? (D) What is the nature of the "spreading" operation? Let us look at these in turn.

Concerning question (A), given that linear order is essentially a binary option (left vs right, precedes vs follows, etc.), the simplest encoding is to mark just one option. Distinguishing three types of movement along the lines in (90) entails that the diacritic involved in signalling linearization information is in fact the movement diacritic. The LCA in (83) guarantees that a head lacking a c-selection-related ^ will always be spelled out in a position preceding its complement. The presence of c-selection-related ^, which triggers comp-to-spec movement, will therefore signal head-final order as it results in a hierarchical structure in which a former complement asymmetrically c-commands its selecting head.

Concerning question (B), we propose that this is due to Relativized Minimality, in the sense of Rizzi (1990, 2001). We have already stated that the L-movement-triggering ^ is associated with the c-selection features of lexical heads. Now, we can assume that every head has a selection feature selecting its complement (we will qualify this assumption slightly below). If the spreading operation "skips" a head, we have a violation of Relativized Minimality. As we saw in the previous section, FOFC derives from just this kind of "skipping"; hence, in the last analysis, FOFC derives from Relativized Minimality (RM), in conjunction with the idea that

only a lexical head can inherently have ^, i.e. that spreading "starts from the bottom" (we will subject the latter idea to further scrutiny directly).

To see precisely how this works, we must look more closely both at RM and at the nature of spreading. Let us first look at RM. According to Rizzi (2001), RM holds in the configuration:

(94) ... X ... Z ... Y ...

Here "Y cannot be related to X if Z intervenes and Z has certain characteristics in common with X. So, in order to be related to X, Y must be in a minimal configuration with X, where Minimality is relativized to the nature of the structural relation to be established" (Rizzi 2001:89). Rizzi goes on to define Minimal Configuration as follows:

(95) Y is in a minimal configuration (MC) with X iff there is no Z such that
  (i) Z is of the same structural type as X, and
  (ii) Z intervenes between X and Y.

Since the version of the theory we are adopting here is feature based, and it is not clear that bare phrase structure offers a clear notion of "structural type", we interpret "same structural type" in (95i) as "same featural type" (this of course brings RM closer to Chomsky's (1995) Minimal Link Condition). In these terms, we see that spreading of ^ from Y to X across Z where all of X, Y and Z are heads endowed with c-selection features violates RM. On the other hand, if ^ spreads monotonically from Y to Z to X then RM is not violated, as Y is then not functioning as an intervener. Hence FOFC violations are RM violations caused by non-minimal spreading of the movement-triggering diacritic ^ within a single EP. For this reason the effects of FOFC are universal and ubiquitous.

Concerning question (C), the following considerations support the idea that L-movement must "start at the bottom". First, we can note that, since the only feature a lexical root bears is its c-selection feature,[42] then the only movement such a head *can*

[42] In this connection, we take the view, contra Marantz (1997), that the clausal lexical array features a verb rather than an acategorial root: the latter is part of an earlier sub-lexical derivation

trigger is L-movement. Second, the symmetrical nature of Merge and the associated "deepest sisters" problem alluded to in the previous section (the observation that in the structure $[_X{}^{-min}\ \ X\ \ \beta\ ]$, where β has no internal structure, X and β cannot be ordered by the LCA) means that a lexical head *must* choose: either ^ or not ^ . The choice of ^ can be seen one way of creating "symmetry-breaking" movement in the sense of Moro (2000): PF will "know" how to linearise the bottommost pair where V^ is present, since the presence of the diacritic results in the creation of a Specifier, with the result that the complement in the bottommost pair asymmetrically c-commands the head. In these terms, we can view the alternative, "head-initial" symmetry-breaking option as head-movement, at least as far as the immediately superjacent "little *x*".[43] These alternatives do not have to be exclusive; they represent independent properties of X and β which may break the symmetry induced by Merge. Hence we may well find systems in which, say, both V moves to v and the direct object moves to SpecVP, but we will not find a system in which neither V nor the object moves.[44]

Turning now to question (D) above, the key to understanding the "spreading" relation lies in our definition of Extended Projection, repeated here:

(76) The Extended Projection of a lexical head L (EP(L)) is the sequence of categories EP = $\{\alpha_1 \ldots \alpha_i \ldots \alpha_n\}$ such that:

i) α_i is in the spine defined by L;

and for each pair of heads $<H_i, H_{i+1}>$ in EP

ii) H_i c-selects H_{i+1};

iii) H_i is categorially non-distinct from H_{i+1}.

---

which combines the acategorial root with the verbaliser (hence we may have acategorial roots inside compounds, as mentioned several times above, but not in the clausal (or nominal) syntax).

[43] This is allowed in terms of the approach to head-movement sketched in Note 39 since the lexical root has just the same categorial features as the little *x*, and no others (on this view, then, theta-related properties cannot be encoded featurally, contra Hornstein 1999 et seq.). Hence its features are a subset of those of *x*, and so it counts as a defective Goal in relation to *x*, and so head-movement is allowed (in fact it is required according to Roberts (forthcoming), but that claim would have to be relaxed, given the present considerations).

[44] This may be an empirically incorrect position. For example, if Huang (2007) is right, Mandarin lacks V-to-v movement. An alternative would be to regard VO order as the PF default, if V bears no diacritic. This is not directly compatible either with the LCA the form in which we have adopted it or, as pointed out in the previous Note, Roberts' approach to head-movement. We leave these questions open here, since they do not bear directly on FOFC.

We construe “spreading” as the projection of categorial features involved in the creation of an EP. If we include ^ among the categorial features, we exclude the following configuration within a single EP:

(96) *X^ [$_{YP}$ .. Y …]

This is in fact enough to derive FOFC, as this is an instantiation of what (91) rules out, and, as stated there, (91) restates FOFC in terms of EP. However, this entails the rather unintuitive idea that ^ is a categorial feature. Furthermore, we do not want to imply that, for example, T or C have no features not shared with v and V. Clearly, we have to clarify the notion “categorial feature” in (76). We can divide formal features into the following classes: categorial (including c-selection), φ, Case and ^ (cf. Chomsky 1995). Of these, φ and Case are involved in the Agree relation, and as such associated with functional rather than lexical heads. Categorial features are associated with both lexical and functional heads, and the intuition behind the notion of Extended Projection is that these are shared between a lexical head and its functional superstructure, as captured in the definition in (76). The movement trigger ^, though, has no intrinsic association with either: if associated with probing features, as we have said, it triggers A-movement, and if associated with c-selection features it triggers L-movement. So we can say that ^ and categorial features are the non-probing formal features. We could restate (76) in these terms, but in fact this would be redundant: since the EP intrinsically involves features of a lexical head (but cf. Neeleman & van de Koot 2002 for a different approach), and since lexical heads can’t be probes, the only features which can form an EP are exactly the non-probing formal features.

We can in fact give a more satisfactory definition of EP, which among other things takes into account the considerations just raised. But first, let us look at how RM interacts with our definition of EP in deriving FOFC violations. Consider first the case of (97):

(97) *V O v

A case in point would be a construction with VO order combined with a head-final overt light verb, as in the ungrammatical version (98b) of the sentence (98a) in Gojri (Bukhari 2009):

(98) a. kiren-nɛ bakro xariid liyo
kiren-ERG he-goat buy take-PF
'Kiren bought a he-goat (for herself).'
b. *kiren-nɛ xariid bakro liyo.
Kiren-ERG buy he-goat take-PF

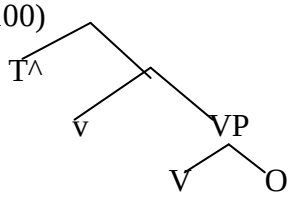
This case is ruled out because V has no ^ and so v cannot bear this feature and be part of the Extended Projection of V (on the possibility of v triggering A-movement of VP to its Specifier, see §6).

Now consider (99):

(99) *[$_{vP}$ v VP] T^

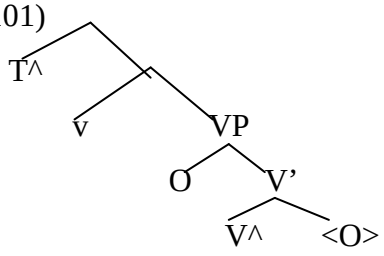
(99) is ruled out whether VP is head-initial or head-final. Consider first the case where VP is head-initial, schematically (100):

(100)



Here, T has selection-associated ^, but v and V have no ^. Now consider the version of (99) where VP is head-final, schematically (101):

(101)



To derive (99), T would need selection-associated ^, which it can only get if this feature can "spread" from V over v. V has ^, but both EP and RM rule out spreading of ^ from V across v to T. Consequently, (101) is underivable.

In (101) we can see the redundancy between EP and RM. We can eliminate this redundancy by reformulating the definition of EP as follows:

(76') The Extended Projection of a lexical head L (EP(L)) is the sequence of categories EP = $\{\alpha_1 \ldots \alpha_i \ldots \alpha_n\}$ such that:

i) α_i is in the spine defined by L;

ii) for each pair of heads $<H_i, H_{i+1}>$ in EP, H_i c-selects H_{i+1};

iii) each pair of heads $<H_i, \ldots, H_{n>i}>$ in EP is RM-compliant in formal features.

The crucial change here is (76'iii), which requires every pair of heads in EP such that the first is higher in the spine than the second to be RM-compliant in formal features. What this means is that all the heads in the EP have to be in a single minimal configuration in Rizzi's sense: so for any pair of heads H_i and $H_{n>i}$ there can be no head intervening in a given EP which is of the same featural type as H_i and $H_{n>i}$, which we take to mean "has c-selection properties", and is distinct in formal features from either H_i or $H_{n>i}$. This rules out both the .. *X*^ .. *Z* .. *Y*^ .. and the .. *X* .. *Z*^.. *Y* .. cases of FOFC (with (96) as a subcase of the latter, assuming EPs always contain more than two heads). Note that we allow a sequence of heads of the type .. *X Y*^.. , where *X* minimally asymmetrically c-commands *Y*^, precisely because there is no intervening head and so RM is vacuously satisfied. So we allow for "upward spreading" of ^ to an arbitrary point in the Extended Projection. On the other hand, the sequence .. *X*^ *Y* .. is not allowed. This is so because if ^ fails to continue to spread to the next head up, we have .. *Z X*^ *Y* ..., which, as we have just seen, is an RM violation, or, if it percolates, we have .. *Z*^ *X*^ *Y* …, also an RM violation. But we must be able to distinguish this last case from the possible .. *Z X*^ *Y*^… . In order to do this, RM must be understood in terms of non-distinctness of features, so that Z is non-distinct from Y^, but Y^ is distinct from Z.

The revised definition of EP in (76') eliminates the redundancy between EP and RM by partially defining the former in terms of the latter, and, since categorial features are one type of formal feature, subsumes the Category Proviso under RM.

Other cases comparable to (101), and ruled out in the same way, are (102a,b,c):[45]

(102) a. $[_{TP}$ T vP] C

b. $[_{NumP}$ Num NP] D

c. $[_{DP}$ D NP] P

d. $[_{CP}$ $[_{PolP}$ Pol TP] C]

These are all ruled out whether the embedded complement is head-initial or head-final: the ^ of the embedded lexical head cannot be spread across the intermediate functional head (T, Num, or P). These are all cases of .. *X*^ .. *Z* .. *Y*^ .. . This covers all the cases of FOFC listed in (52), except for the morphological case. FOFC follows from the definition of EP in (76'), combined with our version of the LCA and the assumption of the generalised movement trigger ^.

A further relevant point is that the diacritic ^ can in principle appear twice on a single head. For A- and A'-movement, this can be seen in examples like the following:

(103) a. Who did John [v appear [<John> to [v talk to <who>]] ?

b. Who were you [v asked <you> [<who> C to [v talk to <who>]]]] ?

Here the matrix v in both cases must have both [uφ] features (although these must be defective in order to facilitate further movement to SpecTP) marked ^ as a way to attract the subject, and a further occurrence of ^ which attracts the wh-element. This is in line with Chomsky's (2008) proposal that A- and A'-movement proceed along quite distinct lines. The same must be true of the linearizing ^, as A- and A'-movement are found in head-final languages (if perhaps rather less readily than in head-initial ones). In fact, we already indicated this in (90), where we indicated A-movement-triggering T as $T_{[u\phi,\ ^]}$, A'-movement-triggering C as $C_{[EF,\ ^]}$, and linearization-triggering V as $V_{[\_\_\ D,^]}$. So the matrix v in (103) has the features [uφ^, EF^, __V]. In a head-final language, the c-selection feature [__V] would also have ^.

[45] On synthetic compounds, see Note 23; the configurations in (38) and (39) (**[Prefix+Root] + Af* and **[Prefix + Root] β*), where β is a head-final root, are also cases of the .. *X*^ .. *Z*.. *Y*^ .. type.

What about the possibility that ^ spreads vacuously from one feature bundle to another? Consider, for example, a lexical verb with a CP-complement in a harmoniously head-final language:

(104) … V^(^) [$_{CP}$ C^ …

C inherits ^ from the lower lexical verb, via the rest of the lower Extended Projection. The higher V is in the Extended Projection of the lower one (see the definition in (76') above) and so can inherit ^ too. As a lexical head in a head-final language, it also introduces its own ^. So, do we have two occurrences of ^ in (104)? It is unclear how both could correspond to distinct occurrences of the movement trigger, as then each would presumably have to trigger its own case of complement-to-specifier movement, and given binary branching, this is impossible. Furthermore, since it is unclear how the two instances of ^ could be taken to correspond to a single element, we take it that there cannot be two occurrences of ^ here. We therefore adopt the following:

(105) Only one occurrence of ^ can be introduced on a lexical root per Extended Projection.

Since only lexical roots can introduce linearization ^, (105) has the effect that only one ^ can be introduced per Extended Projection. Hence it is impossible for two instances of linearization ^ to appear on a single head. This single occurrence of ^ will be introduced at the most deeply embedded merged pair, as a parametric option.

In this connection, we can make a further observation concerning CP-initial, verb-final languages. We have already noted Koptjevskaja-Tamm's (1988) observation that clausal complements in OV languages tend to be either preverbal nominalisations or postverbal. We can relate this observation to FOFC in the following way. The relevant configuration is (106):

(106) V^ .. C … T^

As just noted, the system cannot tell if the occurrence of ^ on V is an occurrence of that on T or a distinct one, introduced by the lexical verb. Furthermore, (106) is

equivalent to (107) as the system is unable to count occurrences of ^:

(107) V^^ … C … T^

We can see that (107) is an RM violation, as there has been a case of copying of ^ from T to V, "skipping" the intermediate C. Since (106) and (107) are equivalent, the same is true of (106). Hence, head-initial CPs in preverbal position in V-final languages are ruled out by FOFC.[46] Koptevskaja-Tamm's observation indicates other ways in which propositional categories can be embedded in V-final languages. Clearly the preverbal nominalisation is a nominal (perhaps containing a CP), and therefore exempted from FOFC, as we have seen. We will propose below that postverbal clausal complements in these languages involve a nominal layer too; see §5.1.1.

It is clear that both the presence of ^ on a given lexical root, and the extent to which it spreads through the Extended Projection of that root, are subject to parametric variation. In the clause, numerous West African languages have head-final VP, but head-initial vP, TP and CP (cf. the cases mentioned earlier in this section), while Latin, West Germanic languages and many Indo-Aryan languages, for example, have head-final VP, vP, and TP, but head-initial CP, and Japanese is head-final throughout the Extended Projection of V. If parametric variation is encoded in the features of lexical items, how can we parametrise the ^-copying operation? One possibility would be to allow any functional head to bear ^ as a free lexical choice at the level of UG, and state the account of FOFC just presented as a condition on EPs. EPs of the kind defined in (76') instantiate the only postulable options available to the child. On this view, the RM-based distribution of ^ therefore falls out as the consequence of a presumably third-factor-imposed constraint on acquisition: children acquiring (disharmonic) languages do not postulate RM-violating configurations. This amounts to a representational account of FOFC.

A derivational account, interestingly, is hardly different. The simplest approach seems to be to regard ^ as part of the feature-composition of the root of the

[46] The only exception to this generalisation are structures in which head-initial CPs have undergone scrambling, i.e. A'-movement (cf. Barbiers 2000 for discussion of this phenomenon in Dutch, and Bayer 1995, 2001 on the nature of the preverbal placement options in a Indo-Aryan languages; cf. Biberauer & Sheehan 2010 for discussion of CP-placement in the light of FOFC, and an analysis in slightly different terms from what is assumed here).

Extended Projection which can be shared, just as categorial features of the root must be shared. It is then possible in principle for any head in the Extended Projection to share this feature with the root; the parametric choices are of course subject to UG principles and so a choice which gives rise to the configuration of the type .. *X* .. *Y*^.. *Z* .. or *X*^ .. *Y* .. *Z*^ .. will be impossible, since it can never give rise to a well-formed derivation (with RM seen as a derivational condition on building EPs, subject to (76ˈiii)).

We thus assume that the acquisition device will never posit the combination of parameter values that leads to a FOFC-violating structure, even though the lexical entries of the individual lexical and functional items are independent of one another and so this possibility is available in principle. In this sense, the range of possible parameters, i.e. possible combinations of lexical entries, is restricted by the nature of EPs. This is quite a familiar state of affairs: for example, it is impossible for a noun to be lexically specified as having the "exceptional Case-marking" property (**our belief (of) John to be smart*); although the reason for this is not understood, it clearly involves an illicit interaction between what are in principle independent properties (the c-selection features of N and the probing capacity of n).

Head-final orders do not necessarily have to be the consequence of a parametric specification, however; 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; see Vilkuna (1995), and the examples in (9), repeated here:

(9) a. Milloin Jussi olisi kirjoittanut romaanin? [Aux-V-O]
when Jussi would-have written a-novel

b. Milloin Jussi olisi romaanin kirjoittanut? [Aux-O-V]
when Jussi would-have a-novel written

c. Milloin Jussi romaanin kirjoittanut olisi? [O-V-Aux]
when Jussi a-novel written would-have
'When would Jussi have written a novel?'

d. * Milloin Jussi kirjoittanut romaanin olisi? [*V-O-Aux]
when Jussi written a-novel would-have

As noted in §2.1.1.1, FOFC is respected in this case, exactly as in languages with obligatory OV order. In a case like this, ^ is not an inherent property of V, but can be optionally added to it and to higher heads in the EP of V. We take it that optional ^ must be licensed by a discourse effect, along the lines of the Fox-Reinhart conjecture (see Chomsky 2001:34), and that optional ^ may in principle be associated with any of the components highlighted in (90) (cf. Biberauer & Richards 2006 and Biberauer 2010 for discussion of phi- and EF-related optionality), as long as it respects the RM-based definition of EP in (76′). We tentatively treat the movements in (9) as A-movements. On the relation between A-movement triggered by $v_{[\varphi\wedge]}$ and FOFC, see §6.

To conclude, we account for FOFC by assuming:

(108) (i) the LCA (as in (83));
(ii) ^ as the general movement-triggering diacritic;
(iii) the notion of Extended Projection defined in (76′).

5. Admissible types of disharmonic order

In this section, we will revisit the (apparent) counterexamples to FOFC noted in §3, showing how they, for the most part, can be understood straightforwardly in terms of the formal characterisation developed in the preceding section.

5.1. Apparent counterexamples to FOFC

5.1.1 Head-initial satellites in head-final languages

As we observed in §3, an obvious apparent counterexample to FOFC involves head-initial DP in OV languages, as in (58a), repeated here:

(58a) Johann hat [$_{VP}$ [$_{DP}$ einen Mann] gesehen].
John has a man seen
'John has seen a man'

Here, DP is a complete extended projection (of N). The selecting verb, being categorially distinct from D, starts a new Extended Projection (D is plausibly a distinct lexical array, which is then renumerated in the manner assumed by Johnson 2002; see also Biberauer & Sheehan 2010). Because of this there is no interaction between the headedness of the two. More precisely, FOFC can be violated here, as it only holds within Extended Projections. To the extent that PP- and AP-complements likewise constitute Extended Projections independent of the main clausal spine, apparent exceptions involving head-initial PPs and APs in OV languages are also accounted for.

What about C-initial CPs in OV languages? Recall that the definition of Extended Projection given in (76) includes C and the higher V in the same Extended Projection. This has the consequence that head-initial CPs are not found in preverbal position in these languages:

(109) *… [$_{VP}$ [$_{CP}$ C [$_{TP}$ .. T^]] V^] …

(109) does not conform to FOFC; this is a case of the structure in (106/107) discussed in the previous section.

We have already observed that OV languages are subject to Koptjevskaja-Tamm (1988, 1993)'s observation that clausal complements are either preverbal nominalisations/non-finite clauses or postverbal CPs (for us, this applies only to OV languages with head-initial CPs). As we observed in §3, the former pattern straightforwardly falls under the Category Proviso (i.e. FOFC as formulated in (1'''), using the notion of Extended Projection). What of the languages with postverbal CPs? Latin, German and Hindi, among many others, are languages of this type, as we mentioned in §2.1.1.2.; see the examples in (14), (18) and (20) given there.

We suggest, following Biberauer & Sheehan (2010), who build on Rosenbaum's (1967) proposal (see also recent proposals by Kayne 2009 and Koopman & Sportiche 2008, Manzini 2009), that this is because the CPs here are embedded under head-initial nominals (which, for simplicity, we represent as DP here; cf. Note 47). One might expect this to lead to the order in (110):

(110) .. [$_{VP}$ [$_{DP}$ D CP] V^ <DP>]

However, at this point the Phase Impenetrability Condition (PIC) becomes relevant. This is stated in (111) (cf. Chomsky 2000: 108):

(111) In a phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

The edge of the phase is defined as material outside H': Specifier(s) of H and any category adjoined to HP (Chomsky 2001:13). What this entails is that material in the complement of a phase-head (*v*, C or D) becomes inaccessible once the phase in question (i.e. *v*P, CP or DP) has been completed. Thus material located in the complement of the phase head will be spelled out once the phase is completed. In (110) the CP complement of D is thus spelled out as soon as DP is completed, and therefore becomes immediately inaccessible to further operations. Let us refer to this operation as "radical spell out" (see Biberauer & Roberts 2005 *et seq.*).

Assuming then that complements of phase heads are radically spelled out in the sense just defined, this means that where we have V^, the DP will undergo selection-driven movement, as in the case of DP-complements, while CP surfaces in postverbal position.

(112) Er hat ... [$_{VP}$ [$_{DP}$ es ~~CP~~] vergessen [$_{DP}$ ~~es~~ CP] (CP spelled out postverbally)
he has it forgotten that …

(Below we will refine the account of object-movement in German, but this will not substantially change the point being made here). Hence CP is "leaked", or apparently extraposed, i.e. it occupies a postverbal position on the surface. The cross-linguistic frequency of this pattern attests to the frequency with which finite clausal complements are realised as CPs contained in head-initial DPs.[47]

[47] These DPs ought to act as rather weak "extraposition islands" (see Cinque 1991), which they do, when the pronoun is overtly realised, but not otherwise. A possible account of this fact is that overtly realised nominals in fact systematically differ from their covert counterparts in representing Ds which are incompatible with an EF-related ^, i.e. a movement trigger attracting material to their specifier; covert Ds, by contrast, bear EF which may be associated with ^. This proposal draws a parallel between the types of Ds found in clausal contexts – edge- vs non-edge-bearing Ds – and the types of Cs which McCloskey (2001, 2002) postulates for Irish. See Biberauer & Roberts (2008) for further discussion. See also Biberauer & Sheehan (2010) for an alternative account which relies on the notions of layered derivation and renumeration rather than the PIC.

5.1.2. Sentence-final particles in VO languages; syncategorematicity

We mentioned this as an apparently recalcitrant form of FOFC violation in §3. In the cases where the sentence-final particles appear to be C-elements, then clearly these give rise to a violation for α=V, β=C. Moreover, such particles are very common. Of 246 VO languages with interrogative particles listed in *WALS*, 135 languages have clause-final interrogative particles (cf. Tokizaki & Kuwana 2009). Mandarin Chinese is a typical example, shown in (113) (from Aldridge 2009):

(113) Ni yao kan zhe-ben shu **ma**?
you want read this-CL book Q
'Do you want to read this book?'

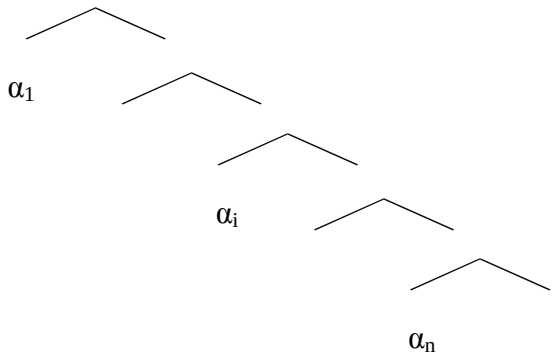
There are two options for these cases, both of which are probably found in different languages. Some sentence-final particles in VO languages are actually clause-initial elements in an elided clause. The structure of (113) under this analysis is (where ~~strikethrough~~ indicates PF-deletion/non-realisation):

(114) $[_{PrtP}\ CP_1\ [_{Prt'}\ Prt\ \sout{CP_2}\ ]$

Aldridge (2009) argues that this is the historical origin of sentence-final *ma* in Mandarin. It is very clearly also the consideration which underlies the availability of final *bu* in elided A-not-A questions (cf. Huang 1982): in that case, we presently still see an alternation between the "full" and the elided structure.

Another possibility is to think that certain kinds of element may fall outside the Extended Projection they appear to be part of. If they are neither selected nor selecting, they appear interspersed among the heads making up the Extended Projection without interfering in the formation of that Extended Projection. In terms of the definition in (76'), suppose we have (EP(L)), the Extended Projection of a lexical head L, $\{\alpha_1 \ldots \alpha_i \ldots \alpha_n\}$. The structure formed by the c-selection relations among these elements looks follows:

(115)



Nothing here prevents the occurrence of a category β interspersed among the members of the Extended Projection provided (a) that this element is categorially distinct from the rest of the Extended Projection, and (b) it neither selects a member of the Extended Projection nor is selected by one. Such an element will not be subject to the RM clause of the definition of Extended Projection in (76'), since it is not part of the Extended Projection. To the extent that functional hierarchies are built up by c-selection relations, this means that such elements will occupy no fixed place in the functional hierarchy, and its position may therefore be cross-linguistically variable. The presence of such a β has two consequences. First, it may introduce a disharmonic order; being outside the Extended Projection, it may have its own linearisation feature – or fail to – and hence be anomalous in relation to the general word-order pattern of the language. Second, assuming that there is no free deletion in narrow syntax and that β is not a Probe or a Goal, it will have to have an LF interpretation. Moreover, any configurational aspect of this interpretation, e.g. in particular scope, will be determined by its direct c-command relations with other elements in the configuration. Let us call categories with the properties of β as just described syncategorematic elements (slightly extending and adapting the usual use of that term).

Looked at in these terms, interrogative particles are fairly good candidates for syncategorematic elements: they appear in the clausal spine, but are not verbal; they neither select the clause, nor are they selected by any higher C-element; and their scopal properties follow fairly directly from their position, i.e. they have clausal scope. To the extent that they qualify as syncategorematic, they are able to introduce FOFC violations simply in virtue of introducing an independent occurrence of the freely available ^. One proviso should be added in the case of interrogative particles: indirect questions are of course selected by higher predicates (e.g. *wonder* vs *believe*).

We thus predict that FOFC-violating interrogative particles can only appear in root clauses. As far as we are aware, this is correct.

We conclude then that sentence-final Q-particles in VO languages are syncategorematic. Recall now our observation in §2.1.1.4 that interrogative markers were among the elements which disobeyed the general suffixing preference in morphology, and therefore arguably violate FOFC at the word-level. If these elements are inherently syncategorematic, we can now see why this is. We conjecture that the same is true for the other types of particles, primarily clause-typing or mood particles, discussed in §3. Note that both types of particles can be thought of as "high" mood markers, and in that respect fall in with one of the inflectional classes which Myler observed not to obey FOFC. It seems that the notion of syncategorematic category is relevant for both syntax and morphology, just as we would expect under the "Single Engine Hypothesis" (cf. Marantz 1997).

5.1.3 Further examples of syncategorematic elements

In the previous section, we gave a kind of negative definition of what a syncategorematic element is, in relation to the general notion of Extended Projection. Let us summarise these properties:

(117) Syncategorematic elements:
- i. are not c-selected;
- ii. do not c-select;
- iii. (therefore) occupy no fixed position in the clausal hierarchy;
- iv. have surface scope determined by their position;
- v. may violate consistent word-order patterns of the language;
- vi. may violate FOFC.

Many of the most recalcitrant cases of FOFC violations discussed in §3 appear to be analysable as syncategorematic elements. Let us now look at some of these, and one or two others, more closely.

5.1.3.1 Negation. Negation is not readily selected: negative complementisers can be selected (e.g. Latin *ne*, archaic English *lest*), but these are negative subordinators, not

pure negators.[48] Negation appears to lack a c-selection feature. Very frequently, the same negation word negates clauses as well as other constituents (English *not*, French *pas*, etc). Further, negation appears in a range of different positions in the clause across languages: very low in German and Swedish (see (53)), in C in Celtic (see Acquiviva 1995, McCloskey 1996 on Irish; Rouveret 1994, Willis 2006 on Welsh), intermediate in most Romance languages (but see also Zanuttini 1997 and Poletto 2008; see also DeVos & van der Auwera 2009 on Bantu). Moreover, negation does not fit readily into the clausal hierarchy, as observed by Cinque (1999: 126). In other words, negation is everywhere.

Furthermore, the scope properties of overt negation are predictable from surface structure in the sense that the ability to license NPIs was formerly seen as an S-structure constraint. In current terms, it is usually said that these relations do not undergo reconstruction. On the other hand, the semantic scope of clausal negation may not always correspond to the surface position of the negative morpheme owing to the possibility of a null "true" negator, in relation to which the overt negator functions in a concord relation (see Zeijlstra 2004, 2008, Biberauer & Zeijlstra 2009, Biberauer & Roberts 2009). We will leave these complexities aside here, and simply note that the surface position of negation plays a major role in determining its scope.

Concerning word order, then, the prediction is that Neg should be able to be initial, even in head-final languages, and final in head-initial languages. There are many examples of initial negation in otherwise head-final projections: Hindi (and most other Indo-Aryan languages, with the exception of Gujarati (Masica 1991:389) and Marathi (Nayudu 2008, Dhongde & Wali 2009,), German (see (53)) and Latin. As noted in §3, there are also numerous examples of final negation in otherwise head-initial languages, which yield what looks like FOFC-violations (worth noting here, though, is the very strong areal effects in play as far as the distribution of this phenomenon is concerned – cf. Reesink 2002 and Dryer 2009b): Abun (Papuan; Reesink 2002: 251), Jukun (Niger-Congo; Dryer 2009b:338), and Bagirmi, in (123):

[48] In Modern English, *that* can be negative when selected by the appropriate verb, as pointed out by Progovac (1994:8-9, 68, 99, 134)):
(i) I deny that he saw anyone.
(ii) *I deny anything.

(118) Bagirmi (Nilo-Saharan; Dryer 2009b:317)

deb-ge tol kobio **li**

person-PL kill lion not

'The people didn't kill the lion'

We conclude that negation is another syncategorematic element. Note also that negative morphemes also feature among Myler's exceptions to the suffixing preference.

5.2.3.2. Coordinating conjunctions. If a predicate selects a nominal complement, it will allow a conjoined nominal complement and so on; and coordinating conjunctions typically coordinate almost any categories. We predict therefore that we will commonly find head-initial coordinate structures in head-final languages. This seems to be true. In fact, Zwart (2009a), in his cross-linguistic survey of coordination, concludes that "/A & B/ is the universal type of monosyndetic noun phrase coordination". Some head-final languages show head-final coordination, showing that Conj^ does exist (assuming that coordination has the structure [$_{\text{ConjP}}$ A [$_{\text{Conj'}}$ Conj B]] (see Munn 1993, Kayne 1994, Zwart, 2009a). We are unaware of any head-initial languages with final coordination, however.

Clearly there is much more to say about coordination, but the core cases appear to act as syncategorematic elements in the sense being developed here.

5.2.3.3. Serial verbs in SOV languages. Serial Verb Constructions (SVCs) are found in both VO and OV languages (although they may be commoner in the former). (119) is an example of an SVC in an SVO language (Ewe; Collins 1997):

(119) me fo kad-gb- gba

I hit lamp break

"I hit the lamp and broke it"

(120) is an example of an SVC in Gojri, an Indo-Aryan SOV language (Bukhari, 2009):

(120) kaloo-nə sentro chillii khayo
kaloo-ERG orange peel ate
'Kaloo peeled the orange and ate it'

We might expect the word order in a head-final language to be as in (121), with the second VP moving to the specifier of the first VP. This is not the case:

(121) *kaloo-nə sentro khayo chilli
Kaloo-ERG orange ate peel

Gojri also has complex predicates made up of a main verb and a light verb:

(122) kaloo-nə TuGRo kha choRyo
kaloo-ERG food eat left
"Kaloo ate the food completely."

In this case, the word order is the expected one: The main verb VP *kha* "eat" moves into the spec of the higher light verb *choR* "left". The difference between the two cases lies in the fact that the light verb selects the complement VP in (122), while the first verb in the SVC does not select the second VP. Head initial order arises because the first verb can act in a disharmonious fashion, being essentially syncategorematic here.[49]

We thus observe a small class of elements which cross-linguistically tend to have the properties listed in (117), which we have referred to as syncategorematic. These elements are able be interspersed among the elements of a single Extended Projection, without breaking that Projection up. They are thereby able to give rise to anomalous orders, including FOFC violations. The elements include negation, root interrogative markers, coordination and the first verb of an SVC. This class of elements ties in closely with those which Myler observed to violate the suffixing preference in morphology, something we can now account for (assuming that c-selection relations also hold in morphology). In addition to negation, discourse

[49] The argument holds most clearly if verb serialisation is essentially complementation (as in Collins 1997). The analysis of Gojri in Bukhari (2009) is that the first verb in the serial verb construction heads a VP which is left-adjoined to the vP headed by the second verb.

markers and mood markers (which may be similar to interrogatives), Myler observed that causative prefixes don't obey the suffixing preference. The first verb in an SVC very often has a causative reading, and so there may be a connection here.[50] Finally, Myler also observed that agreement markers are in this category. Since Chomsky (1995:4.10) it has generally been assumed that agreement features cannot on their own form a syntactic category. More generally, we observe that these elements fall outside of functional hierarchies (since these constitute Extended Projections).[51]

5.2 Conclusion

In this section we have shown how the system of linearization described in §§4.3 and 4.4, which is aimed at deriving FOFC, can account for the attested cases of disharmonic word order and for the unattested, FOFC-violating ones. This led to the development of a notion of syncategorematic elements, which gave us a way to account for many of the recalcitrant counterexamples to FOFC. Once again the notion of Extended Projection plays a central role.

This almost completes our account. There is one remaining issue to be dealt with, concerning on the one hand the relation of A-movement to FOFC and linearization, and on the other hand the exact nature of the Movement Proviso.

6. A-movement and FOFC

In this section, we return to the question of the relationship of A-movement to FOFC. The purpose of this section is to show that there is evidence that A-moved categories are in fact subject to FOFC, once the Category Proviso is taken care of. This will lead to a reformulation of the notion of Extended Projection, which will in turn allow us to eliminate the A'-Movement Proviso.

First, then, what is the evidence that A-moved categories are subject to FOFC? The standard cases of A-movement (passive, raising, unaccusative) involve DP-movement within the Extended Projection of V. As such, they are exempted from FOFC by the Category Proviso and the Satellite Proviso. So we need to look at DP-

[50] It should, however, be noted that causatives are proposed by Cinque (2004:74ff.) to be part of the functional hierarchy. These elements therefore require closer investigation.

[51] Although causatives again pose a difficulty; see the previous note.

movement within DP, or, if possible vP- or VP-movement within the clause. DP-movement within DP seems not to give rise to FOFC violations, as we observed in §3:

(123) a. the emperor's destruction of the city
b. the city's destruction by the emperor
c. the destruction of the city by the emperor

Arguably in both (123a) and (123b), the head of the whole DP is realised as *'s* (see Abney 1987), and the initial DP is A-moved to the highest SpecDP. So we have the substructure in (124):

(124) $[_{DP1}\ DP_2\ [_{D1}$'s] … (DP_2) ..]

Since DP_2 is head-initial, this violates FOFC for $\alpha=D_2$ and $\beta=D_1$, but the examples are grammatical. As we pointed out in §3, this in fact follows from the Satellite Proviso. Therefore, to the extent that the Satellite Proviso falls under the notion of Extended Projection, we can conclude that DP_2 is not in the same Extended Projection as DP_1.[52]

Turning now to the clausal Extended Projection, do we find cases of VP-to-SpecvP movement or vP-to-SpecTP movement which can in any sense be seen as A-movement?

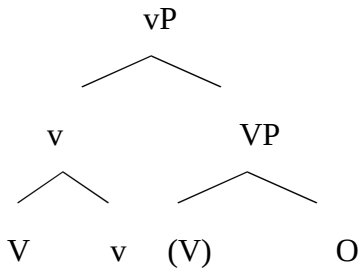
In fact we do, in an analysis of West Germanic clause structure which has recently been put forward. This analysis involves VP-movement of a head-initial VP to SpecvP (Biberauer 2003, Biberauer & Richards 2005, Biberauer & Roberts 2005, Richards & Biberauer 2006), where this movement is taken as a pied-piping option for satisfying a movement trigger associated with v's probing features for the object DP. The same in fact applies to subject-movement, which pied-pipes the vP. To see how this works, consider the following simple German subordinate clause:

[52] This is fairly clear for (123a), where DP_2 is the external argument, first-merged in a specifier (presumably SpecnP). But it is not so clear for (123b), where DP_2 is the complement of *destruction*. However, we can treat the complement of N as marked with inherent (genitive) Case by N. We may then be able to consider an oblique-marked nominal (which might be a PP or a KP) as categorially distinct from N, and hence not able to form an Extended Projection with it. This is consistent with our earlier account of Finnish postpositions.

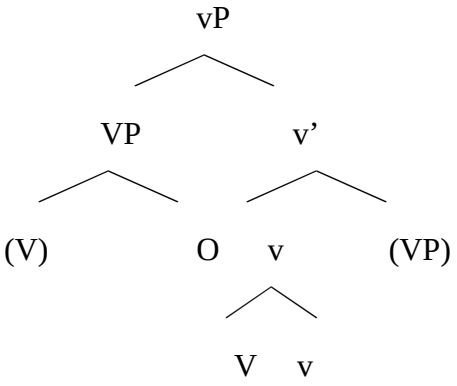
(125) ... dass Johann **das Buch gelesen hat**
that John the book read has
'… that John has read the book'

The order observed in (125) is obtained by means of the operations shown in (126):

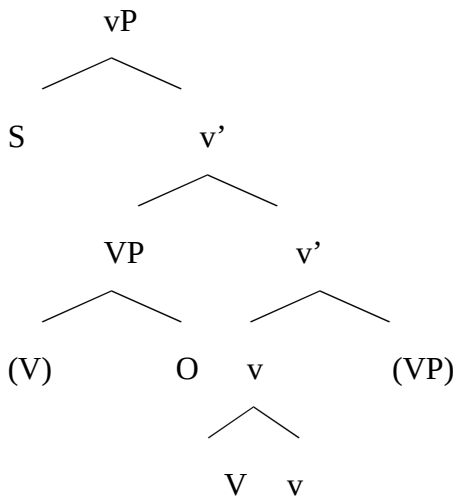
(126) (i) V-to-*v* raising:



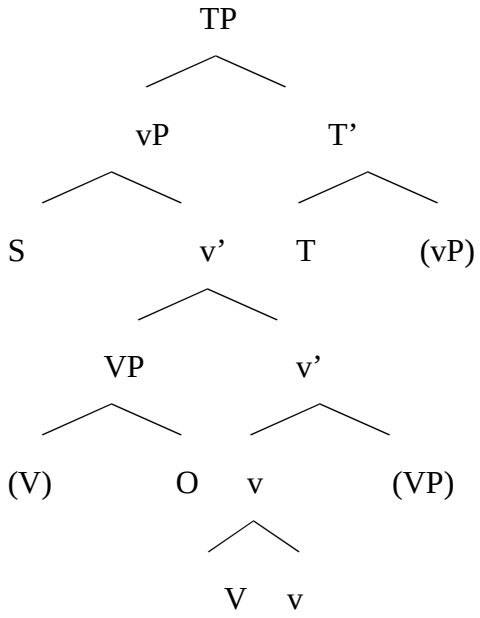
(ii) VP-to-(inner)Spec*v*P movement:



(iii) merger of the subject in the topmost SpecvP:



(iv) vP-movement to SpecTP:



This gives the linear order SOVAux of (125). In terms of this analysis, the structure of (monoclausal) German subordinate clauses may be viewed as differing minimally from that of their counterparts in other Germanic languages. Compared to Modern English, for example, German differs only in two respects. First, v, like T, is associated with [uφ^], while in English only v has [uφ]. Second, the mechanism by means of which ^ is satisfied involves pied-piping of the [iφ]-bearing DPs inside VP and vP.

So, according to this analysis, in German, as the typical representative of OV West Germanic, we have v^. In terms of (90), this is $v_{[u\varphi^]}$, i.e. ^ is associated with v's probing features, and so quite clearly distinct from the ^ associated with L-movement

trigger which is involved in the account of FOFC given in §4. V has no movement trigger, but that is irrelevant to the derivation under consideration. However, under these conditions, the only reason the ^ associated with v's φ-features does not give rise to VOAux order is because V moves to v. We cannot take this to be a consequence of the need for symmetry-breaking movement of V in the absence of object-movement, since the object very often (arguably always) has internal structure; where it doesn't, it is a weak pronoun which moves further to the left in German than other complements, so this movement of a weak-pronoun object may be the symmetry-breaking movement. Neither can the V-to-v be "FOFC-compliance strategy", since the fronted VP is moved by the probe-associated occurrence of ^, and so there is no RM-violation here. It seems to be a mere accident produced by V-movement, then, that we do not find VOAux in OV West Germanic. This is a highly unsatisfactory state of affairs.[53]

The propensity for VOAux in West Germanic is thus only avoided by apparently fortuitous V-to-v movement and is not accounted for by our analysis as it stands. It obviously desirable to bring these cases under our analysis. What is needed is a better understanding of the connection between VP-movement as A-movement pied-piping and VP-movement for linearization. Where V lacks ^, VP-movement *qua* L-movement violates FOFC and falls under our analysis, but VP-movement *qua* A-movement violates FOFC and does not fall under our analysis.

We can relate the A-movement pied-piping case to the linearization case; in fact, we can render them virtually indistinguishable, if we look more closely at V-to-v movement. The pied-piped VP in (126) is part of the same EP as v and the rest of the clausal spine. When V-to-v movement takes place, the two heads then form a complex which bears ^ (since v, as a Probe, bears ^). Effectively, the system cannot distinguish $[_{v^\wedge}$ V v^] from $[_{v^\wedge}$ V^ v^]. More generally, then, (2), repeated once again, violates FOFC, while (127), where α incorporates into β, does not:

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

(127) $[_{\beta P} [_{\alpha P}$ (α) γP] α+β]

[53] In fact, in terms of the theory of head-movement in Roberts (forthcoming), if v has ^ then V must have ^ in order to count as a defective goal, so head-movement is guaranteed. It is nonetheless guaranteed, as it were, by accident, and so the problem raised in the text still holds.

In other words, head-movement of V-to-v is required where v triggers A-movement of a complement (in principle this holds independently of pied-piping of VP, although the latter would only induce a FOFC violation were V-to-v movement not to take place).

Let us now consider the various cases of A-movement in relation to FOFC. Clearly, non-pied-piped A-moved categories generally fall under either the Category Proviso (DP-movement in the clause) or the Satellite Proviso (DP-movement in DP); the above observations regarding the interaction with head-movement take care of the residue of pied-piped Goals. The other possible case of A-movement within the clausal projection that has been reported in the recent literature is VP-fronting in Niuean and a number of other VSO/VOS languages, as analysed by Massam & Smallwood (1997), Massam (2000, 2005), Rackowski & Travis (2000), and many of the papers in Carnie, Dooley & Harley (2005).[54] Massam (2000, 2005) shows that there is a general operation which fronts non-verbal predicates which are clearly larger than heads, e.g. relative clauses, and that what has been called noun-incorporation in Niuean (e.g. by Baker 1988) cannot be movement of N into V (*pace* Baker) since there are clear cases where a constituent larger than N undergoes this operation. She proposes instead that putative noun-incorporation is really the absence of object-shift to a VP-external position. In that case, the fact that the apparently incorporated noun moves with the verb shows that what is moved is VP rather than V. VOS order is thus derived by VP-fronting, and VSO by object-shift to a VP-external position combined with remnant VP-fronting. The landing-site of VP-fronting is taken to be SpecTP. Massam (2000) argues that this is motivated by essentially the same property as that which causes the subject to raise to SpecTP in languages like English, French and Mainland Scandinavian: the operations "can be seen as two reflections of a single EPP predication feature" (Massam 2000: 111). This type of analysis has been applied to a number of languages which display both VOS and VSO orders (mainly, but not exclusively, Polynesian and Mayan languages). Since the VP is head-initial in these languages, and it is fronted within a single Extended Projection, and it is fronted by A-movement, it should give rise to a FOFC violation. However, there are reasons to think that the fronted category is not a VP in the usual sense: Massam (2005: 227)

[54] Actually, Massam (2005) suggests that these orders do not derive from VP-movement to SpectTP, but rather to a higher position. So the parallel with subject-movement becomes less clear.

says that "the Niuean verb is not a morphosyntactic verb, but is more participial in nature, bearing no features for finiteness or tense"; see also Massam (2005: 240), where it is allowed that Niuean verbs may be "participial or even nominal in nature". This observation follows on from a well established tradition in Polynesian linguistics advocating the absence of a (lexical) noun-verb distinction (cf. Massam 2005: 230-231 for brief overview discussion). If these fronted predicative categories are categorially distinct from V, and especially if they are moved to a position in the left periphery above SpecTP and therefore are cases of A'-movement, there is no FOFC violation here.[55]

It seems, then, that all cases of A-movement are exempted from FOFC in one way or another, and we can see how this comes about, either as a consequence of the nature of Extended Projections (since this notion incorporates both the Category Proviso and the Satellite Proviso), or as a result of head-movement. This differs from A'-movement, where, as we saw in §3, fronted VPs (or vPs) in the clausal projection are exempt from FOFC for no obvious reason. We will return to this point in the next section.

7. Conclusions

In this paper, we have tried to provide empirical motivation for FOFC, as an exceptionless syntactic universal. In §2, we adduced a wide range of data, from word order in the clause and in the nominal, and from morphology, diachrony and contact, in favour of an initial formulation of FOFC (see once again (2)). In §3, we looked at a range of counterexamples, which led to a slight complication of the basic generalisation. In §4, we put forward an analysis which derived FOFC from the factors in (106), repeated here:

(106) (i) the LCA (as in (83));
(ii) ^ as the general movement-triggering diacritic;
(iii) the notion of Extended Projection defined in (76') (incorporating the modification to clause (i) given in the previous section).

[55] This may be why VOS orders are much rarer than their mirror image SOV (see Cinque 2010 for discussion of this observation). FOFC blocks the simplest derivation of this order, and so some special operation of category-neutralisation or A'-movement must take place, thereby making the order relatively uncommon.

Assuming that our data is correct, and that we have not missed some significant and intractable set of counterexamples, and assuming that our analysis is coherent, we can now ask a deeper question. In recent work, Chomsky (2005, 2007, 2008) has proposed that much of what is universal to language may derive from "third-factor" constraints, principles relating to optimisation, computational efficiency, and elegant design: "(a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints ... including principles of efficient computation" (Chomsky 2005:6).

Let us look at (106) in this light, taking (i-iii) in reverse order. We could question the postulation of Extended Projection in (106iii), especially given that the definition of this notion is rather complex. Let us therefore look at (76′), in its final form:

(76′) The Extended Projection of a lexical head L (EP(L)) is the sequence of categories EP = $\{\alpha_1 \ldots \alpha_i \ldots \alpha_n\}$ such that:

i) α_i is in the spine defined by L;

ii) for each pair of heads $<H_i, H_{i+1}>$ in EP, H_i c-selects H_{i+1};

iii) each pair of heads $<H_i, \ldots, H_{n>i}>$ in EP is RM-compliant in formal features.

There is no problem in invoking c-selection in clause (ii); to the extent that it relies on syntactic categories and is distinct from s-selection, this is a UG-internal mechanism, but clearly a necessary one. Moreover, our conception of this is completely standard, originating essentially in Chomsky (1965, Chapter 2). Relativised Minimality is clearly a very deep and wide-ranging constraint; it arguably follows from 3rd-factor considerations involving minimisation of search space (see Mobbs 2008).

This leaves the notion of spine and the complication introduced in the previous section for A-movement. We can perhaps understand these in terms of the other basic element of syntactic locality, alongside RM, namely the phase. We can redefine the notion of Extended Projection in terms of phase-linking, as follows:

(128) A phase PH_1 may link to a distinct phase PH_2 just where:

(i) PH_1 is c-selected by an element in PH_2;

(ii) PH_1 and PH_2 are categorially non-distinct.

Following (127), a vP phase can link to an immediately superordinate CP phase, assuming T c-selects v, and C and v are categorially non-distinct (as we have already been doing). This operation can be iterated through a series of vPs and CPs, but will be blocked by a DP. Similarly, satellite elements like DPs, PPs and APs, both arguments and adjuncts, will be excluded from phase-linking with the clausal phases.[56] Our concept of Extended Projection is, then, effectively built on the three core locality relations in the theory: RM, phases, and c-selection. The notion may be partially UG-internal, but at least we are not adding to UG in postulating it.

The approach in (127) may also help us understand why A'-movement is exempt from FOFC. A'-movement always places elements in phase edges (by definition, according to Chomsky (2008)). It is natural to propose that phase edges are excluded from phase-linking. This may be related to the fact that phase edges act as intermediate landing sites for long-distance movement; such positions are, in a sense, outside of the basic locality domains. Since, as we have argued, FOFC holds as a consequence of a locality condition on the formation of Extended Projections, it is natural that it would not hold of the one class of positions exempted from general locality conditions.

A further point concerns the connection of Extended Projections to categories. We tentatively suggest that they in fact play a role in characterising syntactic categories (here again we follow Grimshaw 1991, 2001, 2005). The traditional view of categories is that lexical items are exhaustively assigned to one of a small set of predefined categories, e.g.:

(128)
a. cat → N
b. enjoy → V
c. between → P
d. red → Adj
e. quickly → Adv
f. if → C
g. must → M

[56] It is worth pointing out in this connection that our notion of Extended Projection captures the fact that phases need to "match up", in that CPs take vP complements, DPs take nP complements, etc. Thus the notion may be necessary in order to properly understand why phases interact the way they do.

h. the → D

This view is largely inherited from traditional grammar, and there is no particular reason to think it is correct. Indeed, both bare phrase structure (which treats category membership as just another feature of a lexical item, and not as creating an environment for lexical insertion) and Distributed Morphology (which treats lexical roots as lacking categorical specifications) point away from this view.

We suggest that the notion of syntactic category is diffuse, and that it relates to the different locality domains and to the different types of movement. At the highest level, we have groups of linked phases, the Extended Projection, which probably reflects a basic N, V, A, possibly P, distinction, and corresponds to the domain of wh-movement, and hence plays a major role in defining islands. The next level is the level of the phase itself, which encodes argument structure or inflectional structure, and corresponds to the domain of A-movement. The lowest level is the level of direct c-selection/complementation, corresponding to whatever features can be c-selected, which drives the hierarchies of functional elements of the kind revealed by cartographic work, and corresponds to the domain of linearization movement, as we have seen. Postulating a notion of Extended Projection, then, amounts to part of the theory of syntactic categories. As far as we know, this is part of UG: nouns and verbs do not seem to exist outside the grammar of natural languages.

The movement trigger ^ in (106ii) is part of UG. It is also part of narrow syntax, by virtue of its participation in the definition of EP and the fact that RM is sensitive to it. So narrow syntax makes a contribution to linearization, albeit an indirect one, in that linearization follows from the nature of ^ as a leftward movement trigger. Postulating the movement trigger ^ as part of narrow syntax is not in conflict with the Strong Minimalist Thesis, since this feature is required anyway in order to account for the existence of movement.

That leaves us with a final question: is the LCA a UG principle? We take it, to the extent that it cannot obviously be reduced to third-factor properties, that it is a UG-internal interface-mapping principle. It is worth observing that, on this view, asymmetric c-command in narrow syntax has a dual role: it determines linear order (via the LCA) and it determines LF relations such as scope, anaphora, etc. This seems to be a highly economical architecture for the language faculty (broadly construed, in

the sense of Hauser, Chomsky & Fitch 2002; cf. also Hauser 2008, 2009 on "promiscuous interfaces" as a defining property of "humaniqueness").

In FOFC, then, we seem to have a generalisation which gives an empirical indication regarding the nature of UG, and, indirectly, the way in which the SMT should be interpreted.

References

Abels, K. (2003). *Successive Cyclicity, Anti-locality and Adposition Stranding*. Ph.D. Thesis. University of Connecticut, Storrs.

Aboh, E. (2004). *The Morphosyntax of Complement-Head Sequences. Clause structure and word order patterns in Kwa*. Oxford: OUP.

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

Abney, S. (1987). *The English Noun Phrase in its Sentential Aspect.* Doctoral dissertation, MIT, Cambridge, MA.

Abondolo, D. (1998) *The Uralic Languages*. London: Routledge.

Ackema, P. & A. Neeleman. (2004). *Beyond Morphology: Interface Conditions on Word Formation.* Oxford: OUP.

Adger, D., D. Harbour & L. Watkins (2009). *Mirrors and Microparameters. Phrase Structure Beyond Free Word Order*. Cambridge: CUP.

Aldridge, E. (2004). *Ergativity and Word Order in Austronesian Languages.* Ph.D. dissertation,. Cornell University

Aldridge, E. (2009). Neg-to-Q: the historical development of one clause-final particle. To appear in T. Biberauer (ed.). *Cambridge Occasional Papers in Linguistics (Special edition on particles)*.

Alrenga, P. (2005). A sentential subject asymmetry in English and its implications for complement selection. *Syntax* 8(3): 175-207.

Arsenijević, B. (2009). Clausal complementation as relativization. *Lingua*, 119 (1):39-50.

Baker, M. (1988). *Incorporation; A Theory of Grammatical-Function Changing*. Chicago: Chicago University Press.

Baker, M. (2005). On Verb-Initial and Verb-Final Word Orders in Lokaa. *Journal of African Languages and Linguistics* 26: 125-164.

Baker, M. (2008). The macroparameter in a microparametric world. In: T. Biberauer (ed.). *The Limits of Syntactic Variation.* Amsterdam: Benjamins, 351-375.

Bayer, J. (1995). On the origin of sentential arguments in German and Bengali. In: Haider, H., S. Olsen & S. Vikner (eds.) *Studies in Comparative Germanic Syntax.* Dordrecht, Kluwer, 47-76.

Bayer, J. (2001). Two grammars in one: sentential complements and complementisers in Bengali and other South Asian languages. In: P. Bhashkararao & K. Subbarao (eds.). *The Yearbook of South Asian Languages: Tokyo Symposium on South Asian Languages - Contact Convergence and Typology*. New Delhi: Sage Publications, 11-36.

Barbiers, S. (2000). 'The right-periphery in SOV-languages: English and Dutch.' In: P. Svenonius (ed.) *The derivation of VO and OV*. John Benjamins, Amsterdam, pp. 181-220.

Behagel, O. (1932). *Deutsche Syntax. Eine geschichtliche Darstellung*. Volume 4: *Wortstellung. Periodenbau*. Heidelberg: Winter.

Benmamoun, E. (2000). *The Feature Structure of Functional Categories: A Comparative Study of Arabic Dialects*. Oxford: Oxford University Press.

Benveniste, E. (1968). Mutations of linguistic categories. In *Directions for historical linguistics*, ed. Y. Malkiel and W.P. Lehmann, Austin and London: University of Texas Press¸ 83-94.

Berman, R. (1997) Modern Hebrew. In: R. Hetzron (ed) *The Semitic Languages*. London: Routledge, 312-333.

Bernstein, J. (2001). The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain. In M. Baltin & C. Collins (eds). *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, 536-561.

Berwick, R. & N. Chomsky (2008). The Biolinguistic Program: the current state of its evolution and development. To appear in: A-M. di Sciullo & C. Aguero-Bautista (eds.). *Biolinguistic Investigations*. Cambridge, MA: MIT Press.

den Besten, H. (1983). "On the interaction of root transformations and lexical deletive rules," in W. Abraham (ed) *On the Formal Syntax of the Westgermania*. Amsterdam: John Benjamins, 47-131.

den Besten, H. (1986). Decidability in the Syntax of Verbs of (Not Necessarily) West Germanic Languages. *Groninger Arbeiten zur Germanistischen Linguistik* 28: 232-256. Reprinted as: H. den Besten (1989). *Studies in West Germanic Syntax*. Amsterdam: Rodopi: 111-135.

Biberauer, T. (2003). *Verb Second (V2) in Afrikaans: a minimalist analysis of word-order variation.* Ph.D. dissertation: University of Cambridge.

Biberauer, T. (2009a). Jespersen off course? The case of contemporary Afrikaans negation. In: E. van Gelderen (ed.). *Linguistic Cycles.* Amsterdam: Benjamins, 91-130.

Biberauer, T. (2009b). P-doubling in Afrikaans: Facts and comparisons. Paper presented at the LAGB Annual Meeting (Edinburgh).

Biberauer, T. (2010). Semi pro-drop languages, expletives and expletive pro reconsidered. In: T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 153-199.

Biberauer, T. & S. Cyrino (2009). Negative developments in Afrikaans and Brazilian Portuguese. Paper presented at CGG 19 (Vitoria).

Biberauer, T., A. Holmberg & I. Roberts (2007). Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC). In Bisetto, A. & F. Barbieri (eds). *Proceedings of XXXIII Incontro di Grammatica Generativa*.
[Available on-line at:
http://amsacta.cib.unibo.it/archive/00002397/01/PROCEEDINGS_IGG33.pdf]

Biberauer, T, A Holmberg & I. Roberts (2008). Structure and Linearization in Disharmonic Word Orders. In: C. Chang & H. Haynie (eds.). *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press.

Biberauer, T., A. Holmberg & I. Roberts (2009). Linearization and the architecture of grammar: a view from the Final-over-Final Constraint. In: V. Moscati & E. Servidio (eds.). *StiL - Studies in Linguistics (Proceedings of XXXV Incontro di Grammatica Generativa): 78-91*.

Biberauer, T., G. Newton & M. Sheehan (2009). Limiting synchronic and diachronic variation and change: the Final-Over-Final Constraint. *Language and Linguistics* 10(4): 699-741.

Biberauer, T., M. Sheehan & G. Newton (2010). Impossible changes and impossible borrowings. To appear in: A. Breitbarth, C. Lucas, S. Watts & D. Willis (eds). *Continuity and Change in Grammar*. Amsterdam: Benjamins.

Biberauer, T. & M. Richards (2006). True Optionality: when the grammar doesn't mind. In: C. Boeckx (ed.). *Minimalist Essays*. Amsterdam: John Benjamins, 35-67.

Biberauer, T. & I. Roberts (2005). Changing EPP-parameters in the history of English: accounting for variation and change. *English Language and Linguistics* 9(1): 5 – 46.

Biberauer, T. & I. Roberts (2008). Phases, Word Order and Types of Clausal Complements in West Germanic Languages. Paper presented at CGSW 23 (Edinburgh).

Biberauer, T. & I. Roberts (2009). Negative words and related expressions: a new perspective on some familiar puzzles. To appear in: R. Ingham & P. L'Arrivée (eds.). *Cycles of Grammaticalization. Comparative views on the history of negation*. Cambridge: CUP.

Biberauer, T. & M. Sheehan (2010). Disharmony, antisymmetry, and the Final -over- Final Constraint. To appear in: V. Elguea & M. Extebarria (eds.). *Ways of Structure Building*. Cambridge: CUP.

Biberauer, T. & H. Zeijlstra (2009). Negative Concord in Afrikaans: filling the typological gap. To appear in: S. Lima, J. Mullin & B. Smith (eds.). *Proceedings of NELS 39*.

Boeckx, C. (2008). Approaching parameters from below. To appear in: A-M. Di Sciullo & C. Boeckx (eds.). *Biolinguistic Approaches to Language Evolution*. Oxford: Oxford University Press.

[Available on-line at: http://ling.auf.net/lingBuzz/000641].

Brody, M. (1995). *Lexico-Logical Form*. Cambridge, Mass: MIT Press.

Bukhari, N. (2009). *The syntax of serial verbs in Gojri*. Ph.D. dissertation: Newcastle University.

Burzio. L. (1986). *Italian Syntax: A Government Binding Approach*. Dordrecht: Kluwer.

Carnie, A., H. Harley & S. Dooley (2005). *Verb First.* Amsterdam: John Benjamins.

Cecchetto, C. (2009). Backwards 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. Bibearuer & I. Roberts (eds.). *Challenges to Linearization*. Berlin: Mouton.

Cecchetto, C. & C. Donati (forthcoming).On Labelling: Principle C and head movement. To appear in *Syntax*.

Cheng, L. & Sybesma, R. (2003). Forked modality. In: L. Cornips & P. Fikkert (ed.). *Linguistics in the Netherlands* 2003, 13-23.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Chomsky, N. (1995). Categories and Transformations. In: N. Chomsky. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: the framework. In: R. Martin, D. Michaels & J. Uriagereka (eds). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass: MIT Press, 89-156.

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

Chomsky, N. (2004). Beyond explanatory adequacy. In: A. Belletti (ed.). *Structures and beyond: the cartography of syntactic structures.* Oxford: Oxford University Press, 104-13.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry* 36: 1-22.

Chomsky, N. (2007). Approaching UG from below. In: U. Sauerland & H-M. Gärtner (eds.). *Interfaces + recursion = language?* Berlin: Mouton de Gruyter, 1-29.

Chomsky, N. (2008). On Phases. In R. Friedin, C. Otero & M.L. Zubizarreta (eds). *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. Cambridge, Mass.: MIT Press, 133-165.

Cinque, G. (1991). *Types of A'-Dependencies.* Cambridge, Mass: MIT Press.

Cinque, G. (1999). *Adverbs and functional heads: a cross-linguistic perspective.* New York: OUP.

Cinque, G. (2004) *Restructuring and Functional Heads*. Oxford: Oxford University Press.

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

Cinque, G. (2010). Deriving Word Order Types. A Change of Perspective. Ms. University of Venice.

Clackson, J. & G. Horrocks (2007). *The Blackwell History of the Latin Language.* Oxford: Blackwell.

Collins, C. (1997). Argument Sharing in Serial Verb Constructions. *Linguistic Inquiry* 28: 461-497.

Davison, A. (2007). Word order, parameters, and the Extended COMP projection. In: J. Bayer, T. Bhattacharya & M.T. Hany Babu (eds.). *Linguistic Theory and South Asian Languages.* Amsterdam: Benjamins, 175-198.

Deal, A. 2008. The origin and content of expletives: evidence from "selection". *Syntax* 12(4):285-323.

Devine, A. M. & L.D. Stephens (2006). *Latin Word Order: Structured Meaning and Information.* Oxford: Oxford University Press.

DeVos, M. & J. van der Auwera (2009). *Jespersen Cycles in Bantu: double and triple negation.* Unpublished ms.: Royal Museum for Central Africa and University of Antwerp.

Dhongde, R. V. & K. Wali. (2009). *Marathi*. Amsterdam: Benjamins

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

Dryer, Matthew. (2008a). Order of Object and Verb. In: M. Haspelmath, M. Dryer, D. Gil & B. Comrie (eds.). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 83. [Available online at http://wals.info/feature/83. Accessed on May 15th 2009.]

Dryer, Matthew. (2008b). Order of Adverbial Subordinator and Clause. In: M. Haspelmath, M. Dryer, D. Gil & B. Comrie (eds.). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 94. [Available online at http://wals.info/feature/94. Accessed on May 15th 2009.]

Dryer (2008c) Prefixing and Suffixing in inflectional Morphology. In: M. Haspelmath, M. Dryer, D. Gil & B. Comrie (eds.). *The World Atlas of Language Structures Online*. Munich: Max Planck

Digital Library, chapter 26. [Available online at http://wals.info/feature/94. Accessed on May 16th 2009.]

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

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, 307-362.

Duffield, N. (2001). On Certain Head-final Effects in Vietnamese. In: K. Megerdoomian & L.A. Bar-el (eds.). *Proceedings of WCCFL 20.* Somerville, MA: Cascadilla Press, 150-163.

Emonds, J. (1976). *A Transformational Approach to English Syntax*. New York: Academic Press.

Enfield, N. (2002). *Linguistic Epidemiology. Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge.

Ernout, A. & F. Thomas (1993) *Syntaxe Latine*. Paris: Klincksieck.

Evers, A. 1975. *The transformational cycle of Dutch and German*. Ph.D. dissertation: University of Utrecht.

Ferraresi, G. (1991). *Die Stellung ds gotischen Verbs im Lichte eines Vergleichs mit dem Althochdeutechen*. MA dissertation: University of Venice.

Ferraresi, G. (1997). *Word Order and Phrase Structure in Gothic*. Ph.D. Dissertation: University of Stuttgart.

Fox, D. & D. Pesetsky (2005). Cyclic Linearization of Syntactic Structure. In: K. Kiss (ed.). *Theoretical Linguistics* 31:1-46.

Fukui, N. & M. Saito (1998). Order in Phrase Structure and Movement. *Linguistic Inquiry* 29(3):439-474.

Fuss, E. & C. Trips (2002). Variation and change in Old and Middle English. On the validity of the Double base Hypothesis. *Journal of Comparative Germanic Linguistics* 4: 171-224.

Givon, T. (2001). *Syntax. An Introduction.* Volume II. Amsterdam: Benjamins.

Gildersleeve, B.L. & G. Lodge (1997). *Latin Grammar*. London: Bristol Classical Press.

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

Grimshaw, J. (1981). Complement Selection and the Lexicon. *Linguistic Inquiry* 10(2): 279-326.

Grimshaw, J. (1991). *Extended Projection*. Unpublished ms: Rutgers.

Grimshaw, J. (2001). Extended projection and locality. In: P. Coopmans, M. Everaert & J. Grimshaw (eds.). *Lexical Specification and Insertion.* Amsterdam: Benjamins, 115-133.

Grimshaw, J. (2005). *Words and Structure.* Stanford: CSLI Publications.

Haddican, W. (2002). *Aspects of DO word order across creoles*. Paper presented at CUNY/SUNY/NYU mini-conference, 20 April 2002.

Haddican, W. (2004). Sentence polarity and word order in Basque. *The Linguistic Review* 21: 87 – 124.

Haegeman, L. & H. van Riemsdijk (1986). Verb projection raising, scope and the typology of rules affecting verbs. *Linguistic Inquiry* 7: 417-66.

Hale, K., L.V. Jeanne & P. Platero (1977). Three cases of overgeneration. In: P. Culicover & T. Wasow (eds). *Formal Syntax*. New York: Academic Press, pp. 379-416.

Harbour, D., D. Adger & S. Béjar (2008). *Phi Theory: Phi-Features across Modules and Interfaces.* Oxford: Oxford University Press.

Harley, H. (2009). Compounding in Distributed Morphology. In: R. Lieber & P. Stekauer (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.

Harris, M. (1978). *The Evolution of French Syntax: A Comparative Approach*. London: Longman.

Hauser, M. (2008). *The Seeds of Humanity*. Tanner Lectures, Yale University.

Hauser, M. (2009). The possibility of impossible cultures. *Nature* 460: 190-196.

Hauser, M., N. Chomsky & T. Fitch (2002). The language faculty: What is it, who has it, and how did it evolve? *Science*,298: 1569-1579.

Hawkins, J. (1983). *Word order universals.* New York: Academic Press.

Hawkins, J. (1990a). A parsing theory of word order universals. *Linguistic Inquiry* 21: 223-261.

Hawkins, J. (ed.). (1990b). *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.

Hawkins, J. (2010). *The FOFC from a Processing Typology Perspective.* Talk given at the Cambridge Institute of Language Research Workshop on FOFC, Cambridge.

Hawkins, J. & G. Gilligan (1988). Left-Right Asymmetries in morphological universals. In: J. Hawkins & H. Holmback (eds.). *Lingua* (special issue: *Papers in Universal Grammar*) 74: 219-259.

Hoeksema, J. (1991). Complex Predicates and Liberation in Dutch and English. *Linguistics and Philosophy* 14-4, 1991, 661-710.

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

Holmberg, A. (2003). *Yes/no* questions and the relation between tense and polarity in English and Finnish. *Linguistic Variation Yearbook*. Amsterdam: Benjamins, 45-70.

Hornstein, N. (1999). Movement and control. *Linguistic Inquiry* 30: 69-96

Hróarsdóttir, Th. (1999). *Verb Phrase Syntax in the History of Icelandic*. Ph.D. Dissertation: University of Tromsø.

Hróarsdottir, T. (2000). *Word Order Change in Icelandic: From OV to VO*. Amsterdam: Benjamins.

Huang, C-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. Dissertation: MIT.

Huang, C-T. J. (2007). *The macro-history of Chinese syntax and the theory of language change*. Talk given at the University of Chicago.

Johnson, Kyle. 2002. *Towards an etiology of adjunct islands*. Ms, University of Massachusetts at Amherst.

Jones, M. (1988). Auxiliary verbs in Sardinian. *Transcations of the Philological Society* 86(2): 173-203.

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

Julien, M. (2005). *Nominal phrases from a Scandinavian perspective*. Amsterdam: Benjamins.

Kandybowicz, J. (2008a). On Edge Features and Perfect Extraction. In: C. Chang & H. Haynie (eds.) *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Somerville, Mass: Cascadilla Proceedings Project, 77-86.

Kandybowicz, J. (2008b). *The Grammar of repetition: Nupe grammar at the syntax-phonology interface.* Amsterdam: Benjamins.

Kandybowicz, J. (2009). Externalization and Emergence: On the Status of Parameters in the Minimalist Program. *Biolinguistics* 3.1: 94-99.

Kaye, A. & J. Rosenhouse (1997). Arabic dialects and Maltese. In: R. Hetzron (ed) *The Semitic Languages*. London: Routledge, pp. 263-311.

Kayne, R. (1983). Connectedness. *Linguistic Inquiry* 14(2): 223-250.

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

Kayne, R. (2000). *Parameters and Universals.* Oxford: OUP.

Kayne, R. (2008). *Antisymmetry and the Lexicon*. Ms. New York University.

Kayne, R. (2009). Why Isn't *This* a Complementizer? Ms. New York University.

Keenan, E. (1985). Passive in the world's languages. In: T. Shopen (ed.). *Language Typology and Syntactic Description, Volume I: Clause Structure*. Cambridge: Cambridge University Press, 243-281.

Van Kemenade, A. (1987). *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.

Kiparsky, P. (1982). Lexical Phonology and Morphology. In: I-S. Yang (ed.). *Linguistics in the Morning Calm*. Seoul: Hanshin, 3-91.

Kiparsky, P. (1996). The shift to head-initial VP in Germanic. In: H. Thráinsson, S. Epstein & S. Peter (eds.). *Studies in Comparative Germanic Syntax* II. Dordrecht: Kluwer, 140-179.

Kiparsky, P. (2008). Universals constrain change; change results in typological generalizations. In: J. Good (ed.). *Linguistic Universals and Language Change*. Oxford: OUP, 23-53.

Koopman, H. (1984). *The syntax of verbs: From verb movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.

Koopman, H. & D. Sportiche (2008). The *que/qui* alternation: new analytical directions. Ms, UCLA. [Available at http://ling.auf.net/lingBuzz/000638].

Koptjevskaja-Tamm, M. (1988). *A typology of action nominal constructions*. Ph.D. dissertation: Stockholm University.

Koptjevskaja-Tamm, M. (1993). *Nominalizations.* London/New York: Croom Helm.

Koster, J. (1978). Why Subject Sentences Don't Exist. In: S. J. Keyser (ed) *Recent Transformational Studies in European Languages*. Cambridge, Mass: MIT Press.

Kühner, R. & C. Stegmann (1955) *Ausführliche Grammatik der Lateinischen Sprache*, Leverkusen: Gottschalksche Verlagsbuchhandlung.

Laka, I. (1994) *On the Syntax of Negation*, New York: Garland.

Ledgeway, A. (forthcoming). Syntactic and morphosyntactic typology and change in Latin and Romance. In: A. Ledgeway, M. Maiden & J.C. Smith (eds), *The Cambridge History of the Romance languages*. Cambridge: Cambridge University Press.

Lightfoot, D. (1979). *Principles of Diachronic Syntax.* Cambridge: Cambridge University Press.

Longobardi, G. (1978). *Problemi di sintassi gotica. Aspetti teorici e descrittivi.* MA dissertation: University of Pisa.

Longobardi, G. (1996). The Syntax of N-Raising: a minimalist theory. *OTS Working Papers*. Utrecht: Research Institute for Language and Speech.

Manzini, M.-R. (2009). T*he Structure and Interpretation of (Romance) Complementizers.* Unpublished ms: University of Florence.

Marantz, A. (1997). No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In: A Dimitriadis, L Siegel, C. Surek-Clark & A Williams (eds.). *Proceedings of the 21st Annual Penn Linguistics Colloquium*, University of Pennsylvania Working Papers in Linguistics.

Masica, C. (1991). *The Indo-Aryan Languages.* Cambridge: Cambridge University Press.

Massam, D. (2000). VSO and VOS: aspects of Niuean word order. In: A. Carnie & E. Guilfoyle (eds.). *The syntax of verb-initial languages.* New York: Oxford University Press, 97-116.

Massam, D. 2005. Lexical categories, lack of inflection, and predicate fronting in Niuean. In: A. Carnie, S. Dooley and H. Harley (eds). *Verb First. On the syntax of verb-initial languages.* Amsterdam: Benjamins, 227-242.

Massam, D. and C. Smallwood. 1997. Essential features of predication in English and Niuean. *Proceedings of NELS 27*: 236-272.

May, R. (1985). *Logical Form: Its Structure and Derivation.* Cambridge, Mass: MIT Press.

McCloskey, J. (1996). On the Scope of Verb-Movement in Irish. *Natural Language and Linguistic Theory*. 14:47-104.

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

McCloskey, J. (2002). Resumption, Successive Cyclicity, and the Locality of Operations. In: S.D. Epstein & T.D. Seeley (eds.). *Derivation and Explanation*. Oxford: Blackwell, 184–226.

Mobbs, I. (2008). *'Functionalism', the design of the language faculty, and (disharmonic) typology*. M.Phil. dissertation: University of Cambridge.

Moro, A. (2000). *Dynamic Antisymmetry*. Cambridge, Mass: MIT Press.

Müller, G. & W. Sternefeld (1993). Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24(3): 461-507.

Munaro, N. & C. Poletto (2004). Sentential particles and clausal typing in the Veneto dialects. In: B. Shaer, W. Frey & C. Maiernborn (eds.). *ZAS Papers in Linguistics* 35(2): 375-397.

Munn, A. (1993). *Topics in the Syntax and Semantics of Coordinate Structures.* Ph.D. dissertation: University of Maryland.

Myler, N. (2009). *Form, Function and Explanation at the Syntax-Morphology Interface: Agreement, Agglutination and Post-Syntactic Operations.* M.Phil. dissertation: University of Cambridge.

Nayudu, A. (2008). *The sentential structure of Marathi.* Ph.D. dissertation: Durham University.

Neeleman, A. & H. van de Koot (2002). The Configurational Matrix. *Linguistic Inquiry* 33(4): 529-574.

Newmeyer, F. (2005). *Possible and Probable Languages.* Oxford: Oxford University Press.

Newton, G. (2007). *Complementizers and C-particles*. Paper presented at the FOFC Seminar (Cambridge)

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

[Available on-line at: http://www.stanford.edu/~tann/word_order.pdf]

Nunes, J. & J. Uriagereka (2000). Cyclicity and extraction domains. *Syntax* 3(1): 20–43.

Paul, I. (2000). *Malagasy Clause Structure.* Ph.D. dissertation: McGill.

Paul, W. (2009). Consistent disharmony: sentence-final particles in Chinese. To appear in T. Biberauer (ed.). *Cambridge Occasional Papers in Linguistics (Special edition on particles).*

Pearson, M. (2001). *The Clause Structure of Malagasy. A Minimalist approach*. Ph.D. dissertation: UCLA.

Pesetsky, D. (1982). *Paths and Categories.* Ph.D. dissertation: MIT.

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

Pinkster, H. (1990). *Latin Syntax and Semantics.* London: Routledge.

Pintzuk, S. (1991). *Phrase structure in competition: variation and change in Old English word order*. Ph.D. dissertation: University of Pennsylvania. Published as: S. Pintzuk (1999). *Phrase structures in competition. Variation and change in Old English word order*. New York: Garland.

Pintzuk, S. (2002). Verb-Object order in Old English: variation in grammatical competition. In: D Lightfoot (ed.). *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, 276-299.

Poletto, C. (2008). *I quantificatori e la negazione nei dialetti calabresi del progetto ASIt*. Talk given at the 3rd Italian Dialect Syntax Meeting, Pescara.

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

Progovac, L. (1994) *Negative and Positive Polarity: A Binding Approach*. Cambridge: Cambridge University Press.

Rackowski, A. & L. Travis (2000). V-initial languages: X or XP movement and adverbial placement. In: A. Carnie & E. Guilfoyle (eds). *The Syntax of Verb-initial Languages.* New York: Oxford University Press, 117-141.

Reesink, H. (2002). Clause-final negation: structure and interpretation. *Functions of Language* 9(2): 239-268.

Richards, M. (2004) *Object Shift in North and West Germanic: optionality, scrambling and base-generated OV*. Ph.D. dissertation: University of Cambridge.

Richards, M. (2009). Two Kinds of Variation in a Minimalist System. Ms University of Leipzig. [Available on-line at: http://www.uni-leipzig.de/~richards/papers.htm]

Richards, M & T. Biberauer (2005). Explaining *Expl*. In: M den Dikken & C Tortora (eds.). *The Function of Function Words and Functional Categories*. Amsterdam: Benjamins, 115-154.

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

Rizzi, L. (1997). On the fine structure of the left periphery. In: L. Haegeman (ed.). *Elements of grammar*. Dordrecht: Kluwer, 281-337.

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.* Talk given at the 31st GLOW Colloquium, Newcastle.

Rizzi, L. & I. Roberts (1989). Complex Inversion in French. *Probus* 1: 1-30.

Roberts, I. (2007). *Diachronic Syntax.* Oxford: Oxford University Press.

Roberts, I. (forthcoming). *Clitics, head movement and incorporation.* To appear with MIT Press.

Roberts, I. & A. Holmberg (2009). Introduction. In: T. Biberauer, A. Holmberg, I. Roberts & M. Sheehan. *Parametric Variation. Null subjects in minimalist theory*. Cambridge: Cambridge University Press.

Roberts, I. & A. Roussou (2003). *Syntactic Change. A Minimalist Approach to Grammaticalisation.* Cambridge: Cambridge University Press.

Rosenbaum, P. (1967). *The Grammar of English Predicate Complement Constructions.* Cambridge, Mass: MIT Press.

Rouveret, A. (1994) *Syntaxe du gallois: principes généraux et typologie*. Paris: CNRS Editions.

Salvi, G. (2004) *La formazione di struttura di frase romanza*. Tübingen: Niemeyer.

Schmid, T. (2005). *Infinitival Syntax: Infinitivus pro Participio as a Repair Strategy*. Amsterdam: John Benjamins.

Sheehan, M. (2008a). *FOFC and phasal complements of V/v*. Handout from talk given at the University of Newcastle (January 2008).

Sheehan, M. (2009). The Final-over-Final Constraint as an effect of Complement Stranding. *Newcastle Working Papers in Linguistics* 15: 104-125.

Sheehan, M. (forthcoming). Formal and functional approaches to disharmonic word orders. To appear in *Newcastle Working Papers in Linguistics* 16.

Shlonsky, U. (2004). The form of Semitic noun phrases. *Lingua* 114: 1465-1526.

Simpson, A. (2004). The EPP, fossilized movement and reanalysis. In: E. Fuss & C. Trips (eds.). *Diachronic Clues to Synchronic Grammar.* Amsterdam: Benjamins, 161-218.

Simpson, A. & Z. Wu (2002). IP-raising, tone sandhi and the creation of particles: Evidence for cyclic spell-out. *Journal of East Asian Linguistics* 11(1):1-45.

Sportiche, D. (1988). A Theory of Floating Quantifiers and Its Implications for Constituent Structure. *Linguistic Inquiry* 19: 425-51.

Svenonius, P. (2003). Limits on P: Filling in holes vs. falling in holes. *Nordlyd* 31.2:431-445.

Svenonius, P. (2008). Review of *Non-projecting Words* by Ida Toivonen. *Language* 84 (3): 666-670.

Svenonius, P. (forthcoming). Spatial P in English. In: G. Cinque & L. Rizzi (eds.). *The Cartography of Syntactic Structures*. Oxford: Oxford University Press.

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

Toivonen, I. (2003). *Non-projecting words*. Dordrecht: Kluwer.

Tokizaki, H. & Y. Kuwana (2009). Polar question particles and the Final-over-Final Constraint. To appear in T. Biberauer (ed.). *Cambridge Occasional Papers in Linguistics (Special edition on particles).*

Uriagereka, J. (1999). Minimal Restrictions on Basque Movements. *Natural Language & Linguistic Theory* 17(2): 403-444.

Vincent, N. (1988). Latin. In: M. Harris & N. Vincent (eds). *The Romance Languages*. London: Routledge, 26-78.

Vilkuna, M. (1995). Discourse Configurationality in Finnish. In: K. Kiss (ed.). *Discourse Configurational Languages*. Oxford: Oxford University Press, 244-268.

De Vos, M. (2006). *Quirky verb-second in Afrikaans: complex predicates and head movement*. In: J. Hartmann & L. Molnarfi (eds.). *Comparative Studies in Germanic Syntax: From A(frikaans) to Z(urich German)*. Amsterdam: Benjamins, 89-114.

Walkden, G. (2009). *The Final-Over-Final Constraint and West Germanic Verb Clusters*. Talk given at the LFG 'speed session' (Cambridge).

Wang, Y-F., A. Katz, & C-H. Chen (2003). Thinking as saying: s*huo* ('say') in Taiwan Mandarin conversation and BBS talk. *Language Sciences* 25(5): 457-488.

Whitman, J. (2008). The classification of constituent order generalizations and diachronic explanation. In: J. Good (ed.). *Linguistic universals and language change*. Oxford: OUP, 233-252.

Williams, E. (1981). Argument Structure and Morphology. *Linguistic Review* 1: 81-114.

Willis, D. (2006). Against N-raising and NP-raising analyses of Welsh noun phrases. *Lingua* 116: 1807-39.

Wurmbrand, S. (2006). Licensing Case. *Journal of Germanic Linguistics* 18, 3: 175-236.

Yeung, K. (2006). On the status of the complementizer *waa6* in Cantonese. *Taiwan Journal of Linguistics* 4(1): 1-48.

Zanuttini, R. (1997). *Negation and Clausal Structure.* Oxford: Oxford University Press.

Zeijlstra, H. (2004). *Sentential Negation and Negative Concord*. Ph.D. Dissertation: University of Amsterdam.

Zeijlstra, H. (2008). Negative Concord is Syntactic Agreement. Unpublished ms: Amsterdam University.

Zwart, J-W. (1997). *Morphosyntax of Verb Movement*. Dordrecht: Kluwer.

Zwart, J-W. (2005). *Co-ordination in head-final languages*. Unpublished ms: University of Groningen.
[Available on-line at: http://ling.auf.net/lingBuzz/000161]

Zwart, J-W. (2009a). Relevance of typology to Minimalist inquiry. *Lingua* 119: 1589-1606.

Zwart, J-W. (2009b). Structure and order: asymmetric merge. To appear in: C. Boeckx (ed.). *The Oxford Handbook of Linguistic Minimalism.* Oxford: OUP.
[Available on-line at: http://ling.auf.net/lingBuzz/000833]

Zwart, J-W. (2009c). *The FOFC asymmetry: a layered derivation perspective.* Paper presented at the Theoretical Approaches to Disharmonic Word Orders Conference, Newcastle University