EXPLAINING THE FINAL-OVER-FINAL CONSTRAINT: FORMAL AND FUNCTIONAL APPROACHES[*]

Michelle Sheehan, University of Durham

1. Abstract

Biberauer, Holmberg and Roberts (2007, 2008, 2010) posit the Final-over-Final Constraint, to capture a striking asymmetry in the attestation and grammaticality of disharmonic word orders. As an empirical generalization in its most abstract form, FOFC rules out the possibility of a head-final phrase immediately dominating a head-initial phrase. In this paper, two different explanations of this asymmetry are compared and contrasted: a 'functional' account derived from Hawkins' (1994) Performance-Grammar Correspondence Hypothesis (PGCH) and a 'formal' account based on the syntax-PF interface, which attempts to derive FOFC from a version of the Linear Correspondence Axiom (LCA). The functional approach predicts FOFC to be a statistical effect, stemming from principles of efficient processing

---

1

which can affect grammaticalization paths. While both accounts capture the basic asymmetry and (interestingly) overlap substantially in their predictions, they are by no means notational variants, and the focus of this paper is to test their empirical predictions. The first important difference concerns the status of apparent counterexamples to FOFC. It is proposed that while there is suggestive initial evidence that FOFC is merely a statistical effect, many of the apparent surface violations actually turn out to be illusory and the small number which remain cannot be considered a strong argument in favor of the PGCH. A second difference between the approaches concerns the correlation with harmony. It is proposed that instances where the relevant disharmonic gap does not correlate with a typological preference for 'harmony' create a serious problem for the PGCH-based analysis. The general conclusion is that harmonic tendencies and the FOFC gap cannot receive a unified explanation. This appears to disfavor an account of FOFC based on the PGCH, though it does not rule out a functional account per se. In fact, it is argued that the LCA might itself be partially 'functional' if the fact that it maps asymmetric c-command to precedence rather than subsequence is motivated by third-factor processing pressures. In these terms, FOFC is the arbitrary side effect of a partly functionally determined linearization algorithm which uses narrow syntactic asymmetries to yield linearity.

2.   Introduction: harmony and disharmony in natural language

Typologists have long noted a preference for 'harmony' in the linear alignment of certain heads and modifying dependents (cf. Greenberg 1963, Vennemann 1974, Hawkins 1983, this volume, Dryer 1992). In X-bar Theory, this trend was attributed to a parameter governing the

2

order of syntactic head X and complement/sister (cf. Huang 1994), which can be informally stated as follows:[1]

(1)   The sister of X precedes/follows X

Unfortunately, the neatness of the head parameter is empirically challenged by the fact that many languages are not fully harmonic, as has often been noted. In fact, as Jackendoff (1977: 84-85) notes, even English has some degree of disharmony.[2] A starker example is German, which has what appear to be head-final VPs in non V2 contexts, head-initial CPs and DPs and both head-initial and head-final PPs, as well as circumpositions. This is further complicated by the fact that certain matrix and embedded clauses in German display the V2 property, giving rise to surface VO order.[3] As such, it seems clear that the word order of a given language cannot be straightforwardly settled by a single parameter like that in (1).[4] A similar point emerges from the fact that word order change proceeds on a category-sensitive

---

[1] As Dryer (1992: 88, fn 6) notes, the head-parameter does not regulate the order of heads and adjuncts, and distinct mechanisms are often assumed to regulate the linear positioning of the latter (cf. Ernst 2003, Haider 2004).

[2] Dryer (1992: 109, fn17) notes, for example, that there are few languages which consistently branch in the same direction, though he includes non-complements in his correlation pairs such as the order of subject and verb. As such, it cannot be inferred from this that few languages are consistently head-initial or head-final. Nonetheless, the existence of languages with mixed word orders does indicate that there is more to word order than the head-parameter.

[3] Of course it is possible to posit an underlyingly head-initial system and derive head-final orders via movement, as Zwart (1997) proposes for Dutch.

[4] This point was noted by Koopman (1984) and Travis (1984), who proposed an interaction with theta-marking and Case-marking parameters to account for these more complex word order systems.

3

basis rather than in 'one fell swoop' (cf. Li and Thompson 1974, Biberauer, Newton and Sheehan 2009). The implication is that there can be no across-the-board 'head-parameter', but rather the basic order of a head and its complement must be relativized to subsets of categories (cf. Huang 1994), and regulated either by a series of head parameters or in some other way (cf. Kayne 1994, Cinque 2005 for movement-based theories of word order).

In some cases, there is even evidence that two heads of the same category are specified differently for directionality within a single language. Thus Bengali has both phrase-initial and phrase-final complementizers with synchronically equivalent subordinating functions, as Bayer (1999, 2001) shows. The 'quotative' complementizer *bole* derived from the Bengali verb meaning 'to say' always follows the phrase it selects, whereas the complementizer *je*, derived from a relative pronoun, precedes it:[5]

(2)   a.   chele-Ta   Sune-che [**je** [or   baba   aS -be]]

        boy -CF   hear-PTS3   C   his   father come-FUT3

        'The boy has heard that his father will come.'

b.   chele-Ta [[or   baba   aS -be]   **bole**] Sune-che

        boy -CF   his   father come-FUT3   C   hear-PTS3

        'The boy has heard that his father will come.'   [Bengali, Bayer 1999]

Such cases indicate that surface disharmony is a very real phenomenon in natural language, and one that any theory of word order must allow for. Even if the head parameter exists at

---

[5] In fact, many Indo-Aryan languages display a similar dual complementizer system (cf. Davison 2007).

some abstract level, at the very least there must be other forces interacting with it to give rise to disharmony.

## 3.   Degrees of disharmony and the Final-over-Final Constraint

The implication of all this disharmony is that word order at least *can* simply be acquired on a case-by-case basis (relative to (subsets of) categories or even individual lexical items). Indeed, evidence from acquisition suggests that children acquire the word order of their native language very early, even in languages such as German with complex disharmony and V2 effects (Clahsen and Muysken 1986). But if word order is simply acquired on a category-by-category basis then the fact that certain pairs of categories display a strong cross-linguistic preference for harmony is mysterious. In a sense, in weakening the head parameter to render it empirically adequate, we lose the GB explanation for harmony altogether. The apparent prediction is that any harmonic/disharmonic combination should be equally syntactically viable, as Biberauer, Holmberg and Roberts (2007) note.[6] In actual fact, though, it appears that certain disharmonic orders are fairly common, whereas others are unattested in many syntactic domains. Examples (3a-b) represent harmonic structures, well-known to occur most frequently in natural languages. Holmberg (2000) claims, however, that between the two possible disharmonic combinations (3c) is fairly common, whereas (3d) is unattested or extremely rare.

---

[6] Of course, harmony may well lie beyond the remit of syntax, deriving from patterns of diachronic change/acquisition (cf. Bybee 1988, Roberts 2007); or functional pressures (Hawkins 1994, 2004, this volume). We return to Hawkins's approach shortly, as well as a potential explanation in terms of economy.

5

(3)   Harmonic and disharmonic combinations

1.

(a)   β'   (b) β'   (c) β'   (d) *   β'

     αP   β   β   αP   β   αP   αP   β

γP   α   α   γP   γP   α   α   γP
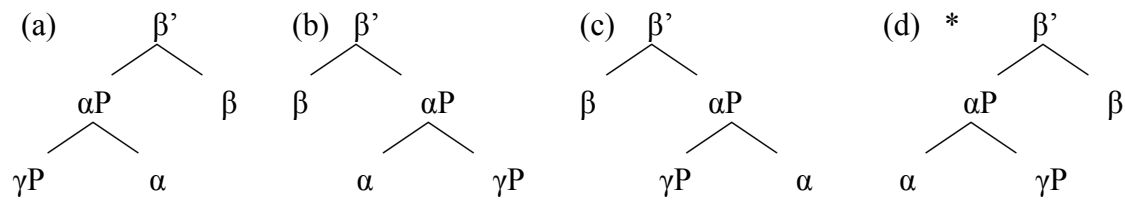
Biberauer, Holmberg and Roberts (BHR) (2007, 2008, 2010) term this effect the *Final-over-Final Constraint* (henceforth FOFC):[7]

(4)   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.

I will term the disharmonic order in (3c) the 'inverse-FOFC order' and that in (3d) the 'FOFC-violating order' for ease of reference, but no theoretical implications should be read into these terms. Section 3 briefly introduces some of the empirical support for FOFC, drawing on work by BHR (2007, 2008, 2010), Biberauer, Newton and Sheehan (BNS) (2009), Biberauer, Sheehan and Newton (BSN) (2010) and Biberauer and Sheehan (2011). Section 4 introduces two different accounts of FOFC, one 'formal', the other 'functional'. The discussion puts aside certain apparent counterexamples to FOFC until section 5, when

---

[7] The version of FOFC given here is the first version offered by BHR and it subsequently undergoes revisions, to accommodate apparent counterexamples. As these counterexamples will be discussed in section 5, I use the maximally general version of the constraint here. A full discussion of BHR's formal account of FOFC is beyond the scope of this paper, but I touch upon it at various points below. .

6

they are used to contrast the predictions of the two accounts. Section 5 shows that while some evidence arguably goes in favor of an account stated in terms of Hawkins' (1994) Performance Grammar Correspondence Hypothesis (PGCH), other evidence raises some serious problems for an explanation along these lines. The implication is that the PF-interface account is a more promising line of explanation for FOFC, despite remaining challenges. Section 6 addresses the problematic status of particles in relation to the two accounts, and proposes that the PF-interface account can provide a new perspective on this problem. Finally, section 7 reconsiders the 'formal' status of the PF-interface account, arguing that it might actually be considered a partially functional account, before section 8 concludes.

4. Empirical evidence for FOFC

4.1. Inflected auxiliary placement and the verb phrase

It is well-known from the typological literature that inflected auxiliaries are verb patterners, meaning that they are more commonly preverbal in VO languages and post-verbal in OV languages (Dryer 1992).[8] It is, however, less well-known that the two potential disharmonic orders between auxiliaries, verb and object are not equally well attested. BHR (2007, 2008, 2010) show that, surprisingly, all possible combinations of verb, object and inflected

---

[8] Auxiliary and modal verbs are usually taken to be base-generated in or to move to some functional head between vP and CP (e.g. *I, Agr, Asp* or *T,* cf. Roberts 1985, Pollock 1989). Ultimately the categorial status of auxiliaries, while undeniably important, does not affect FOFC as an empirical generalization in its most basic form. This is because FOFC holds transitively through the clause and thus, in a VO language, the possibility of any higher head-final phrase is ruled out (cf. BHR 2010). As such, as long as an auxiliary c-commands VP, its position with respect to VP is potentially evidence for/against FOFC.

auxiliary (even discontinuous ones) are attested in diachronic and synchronic stages of Germanic with one exception, *V-O-Aux:

(5) a. **O-V-Aux** German and dialects of German, Dutch and its dialects, Afrikaans; Old English, Old Norse

  b. **O-Aux-V** or so-called *verb-raising/VR* structures: Swiss German dialects, Dutch and its dialects, Afrikaans; Old English, Old Norse

  c. **Aux-O-V** or so-called *verb-projection raising/VPR* structures which involves a head-initial TP and a head-final VP: Swiss German dialects, Dutch dialects, spoken Afrikaans; Middle Dutch, Old High German, Old English, Old Norse

  d. **V-Aux-O**: required for CP-complements in German, Dutch, Afrikaans and their dialects; possible with PP-complements in Dutch and Afrikaans and, to a lesser extent, German; possible with DPs in Old English and Old Norse

  e. **Aux-V-O:** English, Mainland Scandinavian, Icelandic; Old English

  f. ***V-O-Aux:** unattested[9]   [summary based on BHR (2007: 97)]

Note crucially that the effects in (5) go beyond a simple preference for harmony (contra Hawkins, this volume). The inverse-FOFC order, **Aux-O-V** in (5c), is actually very common in Germanic and beyond, notably in Niger Congo and Cushitic languages (cf. Koopman

---

[9] An NWP reviewer points out that sentences involving A-bar VP-fronting are a superficial exception to this otherwise robust gap:

  (i)   I asked him to pay the bill, and pay the bill he did.

Given that these structures involve non-local A-bar movement of VP (past the subject) they fall outside the constraint as described in (4) because VP is not dominated by AuxP in its derived position (cf. BHR 2008, 2010).

1984, on Vata, Creissels 2005, on Mande, and Mous 1993, on Iraqw).[10] The **V-O-Aux** order, on the other hand is systematically banned in languages with inflected auxiliaries, even those with variable word orders which permit all other permutations of verb, object and auxiliary (cf. BHR 2007, 2008, 2010, citing Holmberg 2000 on Finnish and Haddican 2004 on Basque). In a model of word order in which directionality is free, there is nothing to rule out the base generation of the order **V-O-Aux** as per (3d). BHR thus posit (4), initially as a descriptive constraint, to capture this gap. Hawkins (this volume) points out that uninflected tense markers (particles) do not adhere to FOFC (as BHR also note). I return to this complication in section 6.

4.2. Polarity question particles and complementizers

---

[10] Aux O V is relatively common as a surface order, though this does not mean that all surface strings have the same underlying syntax. In Germanic, Aux O V arises as a result of V2, which by hypothesis involves movement of the finite auxiliary to C or verb projection raising in embedded clauses (Den Besten 1981). The fact that the basic word order in Niger Congo is S Aux DO V IO strongly suggests that OV is derived by A-movement of the object, in an otherwise head-initial clause. This is because all constituents except DO occur in a head-initial order (cf. Kandybowicz & Baker 2003 on Nupe). A similar argument might be made for Iraqw, which actually has a number of different object positions with different case and agreement properties (cf. Mous 1993). As the PGCH is concerned with surface strings and how they serve to construct hierarchical structure, the varying derivations of disharmonic orders are of limited relevance, though, as dominance plays a role, syntactic structure is not completely irrelevant. From a formal syntactic perspective, however, any analysis will have to rule out a number of possible ways of deriving the FOFC-violating order. Sheehan (to appear a) discusses this issue and argues that the PF-interface account of FOFC given here extends to rule out instances of V-O-Aux derived by A-movement.

BSN (2010) argue that the presence of an initial polarity question head has blocked the development/borrowing of a final complementizer in a number of Indo-Aryan languages. Drawing on work by Bayer (1999, 2001) and Davison (2007), they show that South Asian languages show great variation in the placement of complementizers and question particles (henceforth Pol(arity) heads) (cf. Masica 1989, Bayer 1999, 2001, Davison 2007). Sanskrit had a final complementizer *iti,* lost in Modern Indo-Aryan. Conversely, all Mainland Modern Indic languages have initial complementizers. The variation in the area concerns the borrowing/development of a final complementizer from either a quotative or demonstrative source. While some Indo-Aryan languages have developed a final C (possibly under influence from Dravidian or other contact languages); others have not (Davison 2007).

Interestingly, the split appears to be syntactically determined: all languages with an initial Pol head have failed to develop/borrow a final C (Marlow 1997, Davison 2007). As an illustration, compare Hindi-Urdu and Marathi. Hindi-Urdu has an (optional) initial Pol head *kyaa* (homophonous with the word meaning 'what') and lacks any kind of final complementizer from either a verbal or demonstrative root:

(6) a. kyaa   aap   wahaaN   aaeeNgii?

     POL   you   there   go.FUT.2PL

     'Are you going there?'

   b. *usee [[vee aa   rahee haiN] **yah/ kah-kar**]   maaluum   hai

      3S-DAT 3PL come   PROG   are   this/   say-PRT   known   is

      'He/she knows [that they are coming].'   [Hindi-Urdu, Davison (2007: 182)]

Marathi, on the other hand, has a final Pol head *kaa(y)* which can co-occur with either a final or initial C:

10

(7) a. [[to kal parat   aalaa   **kaa(y)]**   **mhaaNun/asa**] raam malaa

　　　he yesterday back come.PST.3ms   POL   Quot/such   Ram I.DAT

　　witSaarat   hotaa

　　ask-PROG.   be.PST.3MS

　　'Ram was asking me [whether/if he came back yesterday].'

　 b. raam   maalaa witSaarat   hotaa   [**ki**   to   kal   parat

　　　Ram   I.DAT ask.PROG   be.PST.3MS   that   he   yesterday   back

　　aalaa   kaa(y)

　　come.pst.3ms Pol

　'Ram was asking me [whether/if he came back yesterday].'

　　　　　　　[Marathi, Davison (2007: 184), attributed to R. Pandharipande]

Assuming that C is higher than Pol (as argued by Laka 1994, Rizzi 2001 and Holmberg 2003)

and that polarity question particles are Pol heads in Indo-Aryan, BSN propose that this gap is

again an effect of FOFC. Interestingly, data from the World Atlas of Language Structures

suggest that this gap holds more generally, also outside Indo-Aryan (Dryer 2005a/2011a,

2005b/2011b), as BSN also observe.

Table X.1 Typological positioning of Polarity heads and complementizers

Type	Position of Pol	Position of C	Number of Language: genera: families[11]
A	Initial	Initial Only	72: 35: 13 (78)[12]
B	Final	Final Only	45: 33: 20 (46)
C	Final	Initial	74: 40: 16 (82)
D	Initial	Final	4: 3: 3 (4)

Once again, we see that the two harmonic possibilities as well as the inverse FOFC order are common, whereas the FOFC-violating order (D) is virtually unattested.[13]

4.3. Clausal complements

Biberauer and Sheehan (2011) discuss the placement of CP complements in OV languages, arguing that this area of grammar is also regulated by FOFC. While there are many head-initial languages in which a head-initial CP follows the verb, giving a harmonic head-initial sequence, there seem to be no languages in which a head-initial CP precedes the verb, giving

---

[11] Where genera denote groups of languages whose "relatedness is fairly obvious without systematic comparative analysis" with a time depth no greater than 3500-4000 years (Dryer 2005e: 584). For example, "[t]he standard subfamilies of Indo-European (e.g. Germanic, Slavic, Celtic) are fairly clearly examples of genera…".

[12] The numbers in brackets are the updated versions from Dryer (2011a, b). The breakdown into genera and language families is not provided in the online resource.

[13] The four counterexamples are Tacana and Ese Ejja (Tacanan), Gavião (Tupi) and Resígaro (Arawakan). These languages appear to nominalise embedded clauses (cf. Ottaviano 1980 on Tacana, Moore 1989 on Gavião and Allin 1976 on Resígaro). If nominalization triggers atomization as Sheehan (2011b) suggests then these counter examples might be classified as particles and analysed in the manner proposed in section 7.

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

rise to the FOFC-violating order ***C-TP-V** (cf. Hawkins 1994, this volume, Dryer 2009).

This is true of familiar OV Germanic languages and of Persian but also of unrelated

languages such as Mangarrayi, Iraqw, Neo-Aramaic, Sorbian, Anywa and Päri (cf. Dryer

2009 for further evidence).[14] That this is a FOFC-effect is most obvious in languages like

Turkish in which the canonical position for embedded clauses, both nominalized (8a) and

'direct' (8b), is preverbal, but where C-initial embedded clauses are obligatorily extraposed

(8c):

(8) a. (Ben)   siz-in   Ankara-ya   git- tiğ- iniz]   -i

        I   you-GEN   Ankara-ya   go-NOM-POSS.2PL -ACC

        duy- du-   m

        head-PAST-1SG

        'I heard that you went to Ankara.'   [Turkish, Özsoy (2001:216)]

     b. Biz   [sen- Ø   Ankara-ya   git   -ti   -n]

        we   you-NOM   Ankara-DAT go   -PAST-2SG

        san   -dı   -k

        consider-PAST-1PL

        'We consider you to have gone to Ankara.'   [Turkish, Özsoy (2001:217)]

     c. Anla   -dı   -m   [CP **ki**   onun   bir

---

14 Dryer (2009) finds only one languages with initial Cs in which the unmarked position for embedded clauses

is preverbal (Harar Oromo). To this, we must add Akkadian at a certain stage of development (cf. Deutscher

2000). These languages require further investigation.

13

understand	-PAST-1SG	that	3SG.GEN	one
derdi	var			
problem.POSS.3SG	exists			

'I realized that he had a problem.' [Turkish, Haig (2001:201)]

Hawkins (this volume: PP) claims that the inverse-FOFC combination V-TP-C is also virtually unattested, calling into question the relevance of FOFC as an empirical generalization in this domain. This objection is unfair, though, as FOFC independently rules out V-TP-C as a basic word order, as Biberauer and Sheehan (2011) show. As BHR note, FOFC rules out the possibility of a head-final CP in a VO language, by transitivity. VO order rules out the possibility of a final T (*[[V-O]-T]), and T-VP order in turn rules out the possibility of a final C (*[[T-VP]-C]). If we assume that, all else being equal, CP and DP arguments will surface in the same position in a given language, it follows that wherever a language has a final complementizer, it will also be an OV language and so the clause in question will surface preverbally. Of course the assumption that CP and DP arguments should, all else being equal, surface in the same position might be queried, given empirical evidence from languages like Dutch, German, Persian and Hindi, and Stowell's (1981) influential Case Resistance Principle (CRP). Nonetheless, Biberauer and Sheehan claim that if obligatory 'extraposition' of CP complements in these languages is actually a FOFC effect, then we can maintain the simpler idea that CP/DP complements are base generated in the same position, and that extraposition is a FOFC-compliance strategy. In an OV language, it follows that there is no motivation to obligatorily extrapose a head-final CP. The data from languages such as Turkish, above, strongly support the claim that extraposition in such contexts is tied to directionality (cf. Hawkins 1994 for a processing account of the same

14

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

empirical phenomenon). Note also that languages with greater word order flexibility do display surface V-TP-C order (cf. Uriagereka 1999b on Basque):

(9) Nork   esan   du   [ardoa   bidali   dio-la]?

who   said   have   wine   sent   have-that

Who has he/she said has sent (*the) wine?   [Basque, Uriagereka (1999b: 409)]

Hawkins (this volume) raises a further, more challenging objection relating to apparent absence of the order ***C-TP-N**. As he notes, according to the specific formulation of FOFC in BHR (2010), which relativizes FOFC to heads within an extended projection, the ***C-TP-N** gap falls beyond the remit of the generalization. Given that the order ***C-TP-N** appears to be unattested and also appears to be an instance of a FOFC gap, it is arguably problematic that it falls outside the empirical scope of the generalization. The two accounts of FOFC presented here differ from that in BHR (2010) in several respects and predict clausal complements of N to fall within their remit, and thus sidestep one of Hawkins' objections.[15] An additional problem remains, though, as the order ***N-TP-C** is, according to Hawkins, also unattested, and this inverse-FOFC order is not ruled out by transitivity, in fact, in many cases the order of categories in DP operates completely independently of those in VP (cf. section 5 below). As such, the lack of ***C-TP-V** can be taken as straightforward empirical evidence for FOFC, but the lack of ***C-TP-N** is more controversial.

---

[15] BHR (2008) add a category proviso to their formal account of FOFC to allow for the fact that head-initial DPs and PPs are possible in OV languages. In BHR (2010) this is reformulated as a restriction to extended projections, though this requires a CP complement and its selecting verb to be in the same extended projection. The rationale behind such a move is questioned in section 5 below.

15

4.4.   Summary

In this section, I have considered three syntactic contexts in which something akin to FOFC appears to hold. The reader is referred to BHR (2007, 2008, 2010), BNS (2009) and BSN (2010), Sheehan (to appear) for further discussion. The status of FOFC as a true empirical generalization is challenged by Hawkins (this volume) on a number of counts. One the one hand, Hawkins claims, FOFC is too strong, as it seems to face certain potential counterexamples. On the other hand, he claims, FOFC is too weak, because the other disharmonic order is also not widely attested. The data in sections 3.1-3.4 partly confirm Hawkins' suspicions and partly allay them. On the one hand, it is true that there are a small number of counterexamples to FOFC. There are four languages from four genera, for example, which display the surface word order ***Pol-TP-C,** though there may be an independent explanation for their behaviour. It is not true, however, that the other disharmonic combination is equally rare in the cases discussed, as there are 70 languages from 16 genera with the combination **C-TP-Pol**. In fact, in the above discussion we have seen that the inverse-FOFC order is in fact widely attested in the first two contexts (**Aux-O-V, C-TP-Pol**). In the third context, where both disharmonic orders are virtually unattested (***V-TP-C, *C-TP-V**), FOFC provides an independent explanation for the lack of the first order at least. Of course, this is not to say that these data are sufficient to prove that FOFC holds universally across all categories in all languages. Clearly substantial careful cross-linguistic investigation is required to ascertain whether this is the case, and there are already clear classes of FOFC-exceptions, to be discussed in sections 5.2 and 6. Nonetheless, the fact that FOFC holds in the three domains discussed above is sufficient to make it an interesting potential generalization, worthy of theoretically informed investigation.

16

Ultimately, FOFC remains open to falsification, based on a careful consideration of potential counterexamples. As a methodological point, though, it seems fair to begin with the hypothesis that it is a hard universal and investigate it as such, using the standard methodology of generative linguistics. One might equally begin with the even stronger hypothesis that the other disharmonic order is equally disprefferred, as Hawkins (this volume) suggests, and indeed this was the working hypothesis under the head-parameter. The data discussed in this section, however, suggest that there is more at stake than a preference for harmony. In many ways, as Hawkins points out, more problematic than the few counterexamples to FOFC are the many unattested orders, which appear similar to the FOFC gap, but which do not involve complementation, and so would fall outside the remit of BHR's general approach.[16] Sheehan (to appear b) investigates cases where FOFC appears to hold in instances of adjunction, and claims that many of them fall under the PF-interface account of FOFC provided here, as discussed in Sheehan (to appear b). A full consideration of these facts is beyond the remit of the present paper, however.

The fact that Hawkins (this volume) rejects FOFC as an empirical generalization is, in a sense, surprising, as an elegant account of (a statistical version of) the asymmetry emerges from one version of the Performance-Grammar Correspondence Hypothesis (PGCH)

---

[16] Theresa Biberauer (p.c.) reminds me that the distinction between complements and adjuncts is murky territory. She claims that from a Cinquean perspective, adjuncts form part of the clausal/nominal spine and so are expected to be subject to a version of FOFC stated in terms of complementation. To me this appears to be true only inasmuch as adjuncts are themselves construed as functional heads (as suggested by Abney 1987 for adjectives). If adverbs are specifiers of functional heads then they should not be affected by roll-up movement (cf. Sheehan in prep for discussion). Cinque's roll-up approach to adverb and adjective ordering specifically requires the presence of null functional heads, which give rise to FOFC-violations (cf. Cinque 2005).

(Hawkins 1994).[17] In the remainder of this paper I describe the 'functional' account of FOFC emerging from the PGCH and contrast it with a 'formal' approach based on a version of Kayne's (1994) Linear Correspondence Axiom (cf. Sheehan 2010a). Interestingly, the two approaches make many similar predictions, but they also differ in certain respects, as discussed in section 5. While both approaches face different challenges, those facing the PGCH, it is argued, are more fundamental in nature.

5.   Explaining the Final-over-Final Constraint

5.1.   Performance-Grammar Correspondence Hypothesis (PGCH)

It is often claimed that word order universals lie beyond the remit of generative grammar and stem ultimately from parsing/processing principles or patterns of diachronic change (cf. Newmeyer 2005, Whitman 2008, Abels and Neeleman 2009).[18] One highly articulated processing-based theory of word order is that proposed by Hawkins (1994, 2004, this volume), which provides a potential explanation for the tendency for cross-categorial

---

[17] I focus here on Hawkins (1994) rather than Hawkins (2004) because the earlier version of the theory makes clear predictions in relation to FOFC. Hawkins (this volume) implies that this is no accident as he now rejects the empirical basis of the FOFC asymmetry. In personal communication, he notes that his 2004 approach could explain the asymmetry only if there were some preference for mother attachment overt daughter attachment. Interestingly, this is reminiscent of the core idea behind the Linear Correspondence Axiom whereby there is a requirement for higher categories to precede lower categories. One might say, then, that Hawkins' (2004) approach leads us irrevocably back to the LCA as an explanation of FOFC. See Walkden (2009) for further discussion of Hawkins (2004) in relation to FOFC.

[18] In section 7, I propose a partial endorsement of this position.

harmony discussed in section 1, as well many other on-line word order tendencies (heavy NP

shift and other kinds of 'rightward' displacement). As mentioned several times already,

Hawkins (this volume) rejects FOFC as an empirical generalization, and claims that all that is

observed cross-linguistically is a tendency towards harmony in the order of grammatical

categories and their respective complements. Interestingly, though, the PGCH, when applied

to disharmonic structures, appears to predict a statistical version of FOFC without further

stipulation. In this section, I explore this 'functional' account before comparing it with a

'formal' alternative.

Hawkins (1994) proposes a theory of Early Immediate Constituents (EIC), which

favors harmony over disharmony for processing/parsing reasons:

(10)   Early Immediate Constituents (EIC) (Hawkins 1994: 77)

The human parser prefers linear orders that maximize the I[mmediate]C[onstituents]-

to-non-I[mmediate]C[onstituents] ratios of Constituent Recognition Domains.

Immediate constituents (ICs) are the constituents required to identify a certain grammatical

category. For example a transitive VP consists of two ICs: the category V, which is

constructed by the verb and the category NP, constructed by the determiner or noun, whose

presence indicates that it is a transitive VP. *Constituent Recognition Domain* and *IC-to-non-*

IC ratio are defined in the following ways by Hawkins:

(11)   Constituent Recognition Domain (CRD) (Hawkins 1994: 58-59)

The CRD for a phrasal mother node M consists of the set of terminal and non-

terminal nodes that must be parsed in order to recognize M and all ICs of M,

proceeding from the terminal node in the parse string that constructs the first IC on

19

the left, to the terminal node that constructs the last IC on the right, and including all intervening terminal nodes and the non-terminal nodes they construct.

(12)   IC-to-non-IC ratio =   Number of ICs in domain

                                      Number of non-IC nodes in domain

Obviously, as Hawkins acknowledges, the number of non-IC nodes in a given structure will vary depending on a number of independent syntactic assumptions (i.e. binary vs. ternary branching, heads and movements posited, and more pertinently nowadays X-bar theory vs. Bare Phrase Structure, vP-internal subject hypothesis etc.). Hawkins ignores null functional structure and allows ternary branching meaning that his trees look very different from standard trees in the current literature. Aware of such issues, he also provides a simplified IC-to-word ratio, which factors out syntactic assumptions and gives results broadly similar to the IC-to-non-IC ratio:[19]

(13)   IC-to-word ratio =   Number of ICs in domain

                                      Number of words in domain

Finally, he proposes that the ratio for a CRD is the average of the scores for all IC-to-word ratios calculated left-to-right.

---

[19] Theresa Biberauer (p.c.) asks me to clarify what is meant by a word. As far as I can see, Hawkins gives no definition, and rather uses the term in an informal sense. This may, of course prove problematic, especially if Julien (2002) is right, and words have no syntactic status. One might plausibly reformulate his proposal, nonetheless, substituting 'morpheme' for 'word', as an anonymous reviewer suggests.

20

(14)   Calculating left-to-right IC-to-non-IC ratios

The L-to-R IC-to-non-IC ratio for a non-optimal CRD is measured by first counting the ICs

in the domain from left to right (starting from 1), and then counting the non-ICs (or words

alone) in the domain (again starting from 1). The first IC is then divided by the total number

of non-ICs that it dominates (e.g. 1/2); the second IC is divided by the highest total for the

non-ICs that it dominates (e.g. if this IC dominates the third through seventh non-IC in the

domain, then 2/7 is the ratio for the second IC); and so on for all subsequent ICs. The ratio

for each IC is expressed as a percentage, and these percentages are then aggregated to achieve

a score for the whole CRD.   (adapted from Hawkins 1994: 82)

According to Hawkins, these processing principles predict that harmonic head-initial and

head-final constructions should be most common cross-linguistically as these kinds of

structures are optimal in terms of processing, with CRD and ratios as small as possible.

Consider, by way of illustration, the harmonic orders for VPs containing an adpositional

complement and a verb (Hawkins 1994: 96-97, this volume):

(15)   **[vp   [v go [pp [p to]** [NP school]]   IC-to word ratio = 1/1, 2/2, average= 100%[20]

                   1   2

(16)   **[vp [pp** [NP school] **[p to]] [v go]]**   IC-to-word ratio = 1/1, 2/2 average =100%

                         1   2

---

[20] I use Hawkins' labels here for consistency.

In both (15) and (16) the NP complement of P is not included in the CRD of VP as NP is an IC of PP but not of VP. In (15), the IC-to-word ratio of V (the first IC) is 1/1, as the word *go* serves to construct it (though at this point it remains unclear whether the V is transitive/intransitive/ditransitive etc.). The IC-to-word ratio of PP is 2/2, as the second IC (PP) dominates only the second word contained in the CRD of VP, namely *to*. The average IC-to word ratio is therefore 100% as the number of words is exactly equal to the number of ICs constructed. In (16) a similar effect holds, except that this time, the head-final language is constructed 'bottom up'. Once again, the NP is not included in the CRD of VP. As such, the CRD begins with *to*, which constructs the first IC of VP, namely PP. Because NP is outside the CRD of VP, PP dominates only one word in the CRD (namely *to*), giving it an optimal 1/1 IC-to-word ratio. The word *go*, similarly, serves to construct the second IC of VP, and V also dominates the second word in the CRD (namely *go*). Once again the average of these two IC-to-word ratios gives a perfect 100% as two adjacent words serve to construct the two ICs of VP. As such, Hawkins' approach means that harmonically head-final and harmonically head-initial languages are equally optimal in processing terms. Assuming that frequency correlates with processing efficiency, Hawkins' EIC theory predicts that harmonic structures will be most frequent in the world's languages, as appears to be the case.

Hawkins (1994: 255) also discusses the two disharmonic combinations, where NP complements of P are necessarily included in the CRD of VP. He gives the following IC-to-word ratios, assuming that V and P are single words and that NP comprises a determiner and a noun:

(17)   [$_{VP}$ [$_V$ go [$_{PP}$ [$_{DP}$ the shops] [$_P$ to]]]

            1   2   3   4   IC-to-word-ratio = 1/1, 2/4 = 75%

(18)   [$_{VP}$ **[$_{PP}$ [$_P$ to]** [$_{DP}$ the shops]] **[$_V$ go**]]

1   2   3   4   IC-to-word ratio = **1/3,** 2/4 = 42%

In (17), the first word *go* serves to construct the first IC of VP, V, giving an IC-to-word ratio of 100%. The second IC of VP (PP) is constructed by *to*. Now PP dominates the second through fourth words in the CRD, and so the IC-to-word ratio of PP is 2/4. The aggregate IC-to-word ratio is thus 75% for this word order. In (18) on the other hand, the first IC which is constructed is PP. In this case, PP dominates the first through third words in the CRD of VP (i.e. *to, the* and *shops*). According to the definition in (14), then, the first IC-to-word ratio in (18) is 1/3. The IC-to-word ratio of the second IC, namely V, is 2/4 because the IC V dominates the fourth word in the CRD (i.e. *go*). As such, (18) has a substantially lower efficiency rate of 42%. Moreover, the greater the number of words in the intervening DP constituent, the larger the difference in efficiency between the two disharmonic word orders will be. In effect, the NP complement of P is parsed twice in (18), once in the construction of the first IC (PP), and again in the construction of the second IC (V). Crucially, this makes the prediction that structures/orders like (18) will be more difficult to process, and hence less frequent than those in (17). As (18) is the FOFC-violating order, this version of the PGCH thus appears to derive a statistical FOFC from independently justifiable principles of efficient processing.[21]

In these terms, FOFC reduces to the fact that (i) CRDs are constructed left-to-right and (ii) higher heads are privileged in constructing more independent ICs. Another way to

---

[21] It should be made clear that in the final cut, DP and PP complements to V are beyond the remit of FOFC for BHR (2008, 2010) because of a category or extended projection proviso. Nonetheless, as discussed below, there is a sense in which a kind of FOFC-asymmetry is observed with PP complements of V, suggesting that PPs should perhaps not be ruled out of FOFC.

think of this is that where a phrase appears sandwiched between two heads, the most economical way to parse it is if it forms a constituent with the second head. The prediction, then, is that this parsing preference will be reflected in the linear orders of the word's languages.[22]

Note crucially that the PGCH does not necessarily rule out completely the possibility of FOFC-violating orders. Rather it predicts that they will be infrequent and certainly less frequent than harmonic or inverse-FOFC orders. IC-to-word ratios are also influenced by the relative 'heaviness' of the constituents involved and the implication is that average relative weights might affect grammaticalization trends (cf. Hawkins, this volume, for discussion).[23] As such, it is predicted that FOFC effects may be (i) directly sensitive to heaviness or (ii) category-sensitive, where different categories have different tendencies towards heaviness. For example, the FOFC-violating order should surface more frequently with the lightest categories like DP and less frequently with heavy categories like CP. We will see below that these predictions seem to hold, though the numbers of exceptions are very small in all cases.

---

[22] Interestingly, a similar processing effect appears to apply in phonology if we assume that vowels/nuclei are equivalent to syntactic heads. The maximal onset principle prefers consonants to be parsed with a following rather than preceding nucleus (as an onset rather than a coda, subject to partially language-specific phonotactic constraints) (Blevins 1995: 230):

*/…VC-V…/   /…V-CV…/

A full consideration of these facts is beyond the scope of this paper. When discussing potential FOFC-effects in morphology, BHR (2010) note that Hawkins' theory stops at the word level. As such, if a unified processing account of both onset maximization and FOFC is possible, it will look different from the PGCH in its current form. An anonymous reviewer claims, though, that there is nothing *per se* to prevent an extension of Hawkins' approach to the morphemic and even phonemic level. This is an important area for future research.

[23] As Theresa Biberauer (p.c.) points out, some light categories such as R-pronouns very frequently extrapose whereas heavier PPs do not, raising some problems for the heaviness constraint.

24

However, the PGCH also makes other predictions, which are not so well supported. Crucially it predicts that where two categories display a typological preference for harmony they will also display a FOFC effect and vice versa.[24] This is because the same principle which gives rise to the preference for harmony (i.e. EIC) also gives rise to FOFC. While there might be additional (historical/sociolinguistic) factors which skew the typological sample away from harmony, where this happens, these same factors should also serve to rule *in* a FOFC-violating order. The prediction of EIC is therefore biconditional:

Cross-linguistic preference for harmony between X and Y iff FOFC holds between X and Y

I will argue below that there is suggestive evidence against such a biconditional relation, hence against an account of FOFC based on the PGCH.

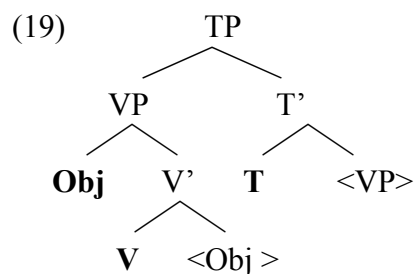## 5.2.   A Phonological Form-interface account

There is strong empirical evidence that specifiers uniformly precede the head/complement in their containing phrase (e.g. from the preference for leftwards movement and initial subjects and the lack of verb-penultimate orders, cf. Kayne 1994, 2004). This is taken as evidence for the Linear Correspondence Axiom (LCA), a linearization mechanism which maps asymmetric c-command to linear precedence.[25] In relation to the order of head and

---

[24] I am not concerned here with conceptual objections to Hawkins' general approach, but rather with teasing out its empirical predictions. For a critique of the PGCH cf. Mobbs (2008).

[25] As an anonymous reviewer notes, strictly speaking the LCA at its most basic is the proposal that asymmetric c-command maps to order (Kayne 1994: 6), the claim that the relevant order is precedence rather than

complement, however, there is less empirical evidence for any such asymmetry. Both VO and OV are equally frequent, for example (Dryer 2005c/2011c) and the asymmetries cited by Kayne do not relate to the order of head and complement. The lack of verb-penultimate, for example, is evidence only of the lack of final specifiers (Richards 2008: 280). Based on observations of this kind, Abels and Neeleman (2009) propose that the empirical asymmetries usually taken as evidence for the LCA actually reduce to a can on rightward movement. FOFC is of particular interest in this respect as it provides crucial evidence of a cross-linguistic asymmetry in the order of heads. The implication is that if the linear order of heads displays the same kind of asymmetry observed with specifiers, then word order asymmetries cannot be reduced to a ban on rightward movement. FOFC can thus be taken as crucial empirical evidence in favour of the LCA.

Previous formal accounts of FOFC have capitalized on this fact and relied upon the idea that head-finality is derived via very local comp-to-spec movement, following the ideas in Kayne (1994) (cf. Holmberg 2000, BHR 2007, 2008, 2010, BNS 2009, Sheehan 2009a, b). In these terms, a harmonically head-final TP is derived in the following way: the object of the verb first moves locally to spec VP, then the whole VP, complement of T, 'rolls-up' to spec TP:

(19)
```
                    TP
            ┌───────┴───────┐
           VP               T'
        ┌───┴───┐       ┌───┴───┐
       Obj      V'      T      <VP>
             ┌───┴───┐
             V    <Obj >
```

---

subsequence is separate and is justified independently by Kayne (ibid: 33), partly on an empirical basis. I return to this issue in section 8.

If it is assumed that internal merge is costless, in some sense, then head-finality arises in those cases where a head simply merges twice with its complement. According to Kayne's (1994) category-based definition of c-command, the complex specifier VP asymmetrically c-commands the head T and so must precede it, giving a head-final order. BHR (2010) provide an account of FOFC in these terms by claiming that movement in such cases is triggered by a caret feature ^ which is optionally associated with c-selection features. In their terms FOFC stems from certain restrictions in the distribution of ^, which serve to ban local comp-to-spec movement of a head-initial phrase. Sheehan (2009a, b) provides an alternative account whereby this kind of movement is permitted at the narrow syntactic level, but gives rise to discontinuous linearization at PF (as developed below).

There are several potential objections to these kinds of accounts. Firstly, Kayne's same category-based definition of c-command which ensures that a specifier (VP) asymmetrically c-commands a head (T) also means that multiple specifiers of T are ruled out (cf. Kayne 1994). If the LCA holds at the narrow syntactic level, then it follows that any phrase containing an externally merged specifier cannot be the target of roll-up movement (cf. Julien 2002). This creates considerable problems for FOFC, as vP, assuming that it introduces the external argument, will necessarily be head-initial. This in turn implies that there will be no phrase-final auxiliaries or complementizers in any languages (contrary to fact). There is an escape from this problem in the form of Chomsky's (1995) reconceptualization of the LCA as, essentially, a linearization algorithm. If this is the status of the LCA then the ban on multiple specifiers holds only at the mapping to PF and not at the narrow syntactic level (cf. Moro 2000). From this perspective, a strong prediction is that in languages with roll-up movement of VP to spec vP the subject should vacate vP by the point of Spell-Out, moving to some (possibly null) head which is (i) not the target of roll-up

27

movement, and (ii) higher than the highest head-final phrase. Interestingly, evidence suggests that this does *not* happen in some well-studied head-final languages, as Julien (2002: 130-136) notes. Thus in Japanese and Turkish evidence from the licensing of NPIs and scope interactions have been taken as evidence that the subject remains inside vP.[26] Given that heads above vP are final in both languages (e.g. negation, question particles, certain complementizers), these facts raise a serious empirical challenge for accounts of FOFC relying on comp-to-spec movement. Altering the definition of c-command so that it allows multiple specifiers to be linearized is a potential solution to this problem, though this is no simple task, as the basic spec-head-comp order itself relies crucially on a category-based definition of c-command, and multiple specifiers of a single category by definition do not differ in relative 'height'.

A second objection to roll-up movement concerns Abels' (2003) anti-locality condition, which blocks movement which cannot result in feature-checking. His argument is that the head-complement relationship is the closest syntactic dependency possible and so comp-to-spec movement cannot possibly be motivated by feature-checking. Strong empirical support for the ban on comp-to-spec movement comes from evidence that extraction of the complement of a phase head is generally banned. Of course, it is possible to make the ^ feature triggering comp-to-spec movement distinct from other features and immune to anti-locality, but this in turn weakens the main conceptual argument for deriving head-finality via

---

[26] In Turkish, for example, NPIs are licensed in subject5 position (Julien 2002: 132, citing Kural 1997: 502):

(i) Kimse   uyu-ma-dı

     anyone   sleep-NEG-PAST

     'No one slept.'

movement, namely that the mechanisms required to do so are independently needed by the grammar.[27]

Thirdly, as Richards (2008) notes, roll-up movement of this kind creates massive redundant structure, which strongly suggests that head-finality should be marked. This is particularly true because, as Abels and Neeleman (2009) and Richards (2008) note, the antisymmetry hypothesis is not as restrictive as it first appears. Given the availability of (and need for) remnant movement, it is actually eminently possible to derive spec-final or inverse V2 orders. As such, the LCA can only explain typological asymmetries if some notion of markedness is associated with movement. If this is the case then spec-final orders cannot be categorically ruled out by the LCA, but can be predicted to be rare. The implication of this move, though, implies that head-finality will also be marked, and as noted above, there is no evidence that, for example, OV order is rarer than VO order. In fact, the only evidence that head-finality is 'marked' comes from FOFC. It follows that if FOFC can be explained without the need for comp-to-spec movement then the only evidence that head-finality is movement-derived also disappears.

The traditional alternative to (18) is to assume that the order of head and complement is regulated by a parameter, applying either in the narrow syntax or at the mapping to Phonological Form (PF). If we take the latter option then head-initial and head-final phrases have potentially identical syntax but different PF-forms. Moreover, as Richards (2004, 2008) has shown, approaches assuming a PF head-parameter are actually more successful in

---

[27] An alternative to this very local movement is what Aboh (2004) terms 'snowballing', where null functional heads lacking phonetic/semantic content are present to attract the complement of a given phrase past its selecting head (cf. Kayne 1998). While these kinds of approaches get around the anti-local problem, they introduce further problems, notably the positing of functional heads illegible at both the PF and LF interface.

accounting for certain VO/OV asymmetries than analyses which assume all head-finality to be derived via narrow syntactic movement.[28] All in all, then, it is worth considering to what extent it is possible to state previous analyses of FOFC without the need for local comp-to-spec movement. In this section, I argue that this is eminently possible. More specifically, I propose a simple restatement of Sheehan's (2009a, b) analysis of FOFC in head-parameter-based terms, whereby the LCA regulates word order only where it remains underspecified by the settings of c-selection-based PF-parameters. This analysis, it will emerge, makes many similar predictions to Hawkins' account, but some differences nonetheless emerge, as discussed in section 5.

Assuming that language acquires linear order only at the PF-interface, it follows that in a minimalist system, PF will rely on independently necessary syntactic asymmetries in order to impose this order. A number of potential candidates exist. Most locally, c-selection creates an immediate asymmetry between heads and complements.[29] Non-locally other asymmetric relations hold, all of which seem to be parasitic on asymmetric c-command

---

[28] Richards takes the LCA to rely on c-command rather than asymmetric c-command. For this reason, a phrasal complement and its selecting head will always stand in a relation of mutual c-command, replicating the bottom pair problem and forcing the need for a PF head-parameter. As a result it is true that Richards derives the need for a head-parameter from the version of the LCA he posits, and the same is not true of the account put forth here, as an anonymous reviewer asks me to clarify. However, while Richard's approach is undeniably elegant, it remains unclear why a symmetric relation would be the crucial determiner of linear order, when syntax is otherwise full of asymmetries.

[29] Theresa Biberauer (p.c.) suggests that some specifiers are also selected. If we limit our attention to c-selection, then evidence suggests that this is not the case (cf. Svenonius 1994). Of course externally-merged subjects are thematically selected, and this too appears to be a very local asymmetrical relation, suggesting that in-situ subjects might also be ordered by a PF-parameter. This provides the potential for a novel account of VS orders in Romance, which will be explored elsewhere.

(probe-goal, copy-trace). Kayne (1994) proposes that only the latter, non-local asymmetry is used to determine word order, but the weight of empirical evidence suggests that both kinds of dependencies are used. More specifically, local c-selectional dependencies regulate head-complement orders and then non-local c-command relations order remaining categories. The result is a system in which something akin to the LCA regulates the order of disharmonic orders and specifiers.

Assume that where two categories stand in a relation of c-selection, a PF parameter of the following kind exists:

(20)   If a category A c-selects a category B, then A precedes/follows B at PF

Following Richards (2004), if phase-internal movement must respect this ordering command at least in some languages, then (20) serves to derive Holmberg's generalization.[30] I further propose, following Sheehan (to appear a), that a head and its projections form a single category, so that projection is effectively reconceptualized as copying, and all arguments/adjuncts are formally merged via adjunction (cf. Sheehan to appear a). The key implication of this proposal is that there is a one-to-one mapping between terminal nodes and categories so that a terminal, its projected 'labels' and any copies of it generated by movement constitute a single multi-segment category.[31] Combined with a minimally revised

---

[30] Richards (2004) shows that this allows a simple account of Holmberg's generalization, stating that an object can shift out of VP in a VO language, only if V also leaves VP. The head movement constraint is arguably a more general case of Holmberg's generalization, in these terms.

[31] A reviewer objects to this idea as "methodologically and ontologically non-minimalist". I disagree. Sheehan (to appear) argues that representational headedness is, at our current level of understanding, uneliminatable from narrow syntax. Given this fact, labels, or some notational variant thereof are required, and copying is the most

category-based definition of c-command, based on Kayne (1994), this notion of copying effectively derives FOFC.

The following definitions reformulate Kayne's definitions of c-command/inclusion/exclusion in a manner compatible with the copy theory of labelling:

(21) **Complete Dominance**: A category X completely dominates a category Y iff X≠Y and the shortest path from every copy of Y to the root of the tree includes all non-terminal copies of X.

(22) **Partial category dominance**: a category X partially dominates a category Y iff X≠Y, the shortest path from every copy of Y to the root of the tree includes a copy of X, but X does not completely dominate Y.

(23) **C-command**: A c-commands B iff A and B are categories, A≠B, A does not partially dominate B and any category which completely dominates A also completely dominates B.[32]

The outcome of these definitions are that a specifier asymmetrically c-commands a head and a head asymmetrically c-commands its complement. The one-to-one correlation between categories and terminals simplifies the linearization process substantially. Firstly, where an

---

minimalist method available to generate them. The unification of 'labeling' and 'movement' has considerably minimalist ontological implications and in conjunction with a reformulation of Kayne's category-based definition of c-command, makes well-supported empirical predictions (cf. Sheehan to appear a for discussion).

[32] Note that under these definitions a category can both dominate and c-command another.

atomic category moves, no deletion is necessary in order for the resultant two-segment

category to be linearized:

(24)

```
        Y
       / \
      X   Y
         / \
        Y   W
           / \
          W   X
```
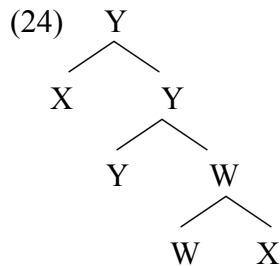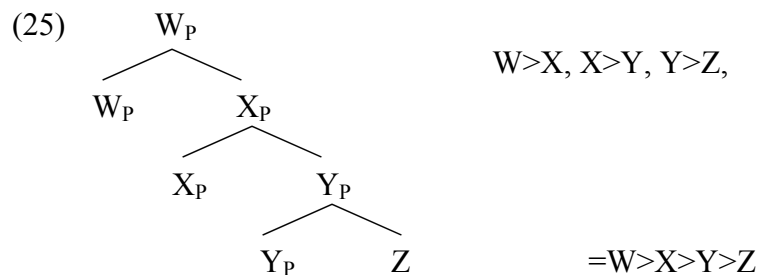
In (24), the category X will simply c-command Y and W without the need for deletion,

because both Y and W partially dominate X. This serves to derive the hitherto poorly

understood fact that moved phrases are spelled out in their derived position at PF.[33] It has the

further implication that problematic non-terminal (X-bar) categories can be dispensed with

and c-selection reduced to a category-category relation.

Putting aside specifiers/adjuncts and the added complications they introduce here, for

reasons of space, let us consider how this kind of system works where a structure is

harmonically head-initial.[34] In such a structure, all categories are specified to precede the

category which they select (at PF) as indicated by subscript P. Crucially, the copy theory of

labelling clarifies the fact that it is only the order of the selecting and selected *categories*

which is specified by this parameter. Just because W must precede X, for example, this

---

[33] Nunes (1995, 2004) provides an alternative account of this relying on the idea that lower copies have more

unchecked features than higher copies. It is not clear whether this account can be maintained in a system without

spec-head agreement, however (cf. Fernández-Salgueiro 2008).

[34] Theresa Biberauer raises many interesting questions concerning specifiers and adjuncts which I take up

elsewhere (cf. Sheehan to appear b).

implies nothing about the order between W and Y or W and Z. [35] Nonetheless, the sum of all

the PF-parameters in a harmonic system does serve to provide a single unambiguous order of

categories by transitivity:

(25)

$$
\begin{array}{c}
W_P \\
\diagup\ \ \diagdown \\
W_P \qquad X_P \\
\diagup\ \ \diagdown \\
X_P \qquad Y_P \\
\diagup\ \ \diagdown \\
Y_P \qquad Z
\end{array}
$$

W>X, X>Y, Y>Z,

=W>X>Y>Z

In such cases, W c-commands Y and Z and also precedes them but this information is not, by

hypothesis, required to linearize the categories in (25). If X>Y and Y>Z then it follows, by

transitivity, that X>Z, irrespective of any direct c-command relation between X and Z (cf.

Fox and Pesetsky 2005). As such, we can assume that only very local c-selection based PF-

parameters are used to linearize harmonic structures such as (25).[36] This issue becomes more

salient when we consider harmonically head-final orders such as that in (26), where all

categories are specified to follow their selected complement at PF, and indicated by subscript

F:

---

[35] Recall that if projection is copying then there is no head/category distinction. All the copies of a given

head/terminal form a single category.

[36] Note that if this holds then there is a sense in which harmonic structures are actually more economical to

linearize than disharmonic structure as well as more efficient to parse.

34

(26)

W_F

W_F X_F

X_F Y_F

Y_F Z

X>W, Y>X, Z>Y

=Z>Y>X>W
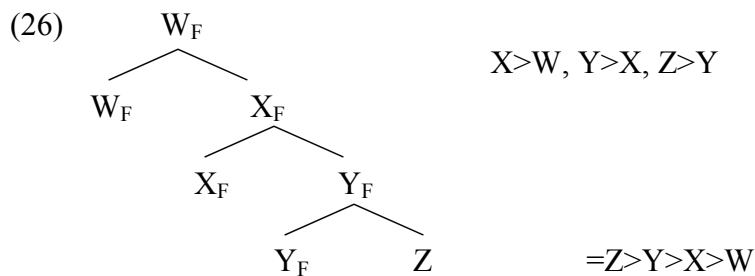
In (26), once again, we obtain a total linear order of categories by the sum of locally defined PF-parameters. In such cases, despite the fact that W still (asymmetrically) c-commands Y and Z, it fails to precede them. As such, the LCA fails to hold in such cases.[37] In our terms, this is because it is never invoked. The fact that W c-commands Y and Z is irrelevant as local PF-parameters based on c-selectional asymmetries are sufficient to order all the categories in (26).

Local c-selection asymmetries between categories will not, however, always be sufficient to impose an unambiguous linear order on a hierarchical structure. In disharmonic contexts, for example, linear order will be underdetermined by the sum of PF-parameters.[38] In such cases, I propose that non-local syntactic asymmetries between categories, in the form of asymmetric c-command are also used. In short, a revised LCA is proposed whereby linear

---

[37] Note that this problem arises also for Richards (2008), who is also forced to accept that the LCA is overridden in such circumstances.

[38] This provides the rationale for using both local and non-local asymmetries to determine linear order: local asymmetries are not sufficient to establish an unambiguous order between all categories. Of course it would be possible to use only non-local asymmetries, as Kayne suggests, bypassing more local asymmetries. I return to this objection shortly.

order between categories is first determined as far as possible based on local c-selection

relations and only as a last resort by non-local asymmetric c-command relations:

(27)   Revised LCA

If a category A c-selects a category B, then A precedes/follows B at PF

If no order is specified between A and B by the sum of all precedence pairs defined by

(i), then A precedes B at PF if A asymmetrically c-commands B.[39]

In these terms, asymmetric c-command is a last resort linearization aid, necessary only where

no linear order between two categories is determined by the sum of all locally determined PF-

parameters.

Before continuing, it is necessary to consider certain objections raised by an

anonymous reviewer. The main objection runs as follows: by Occam's razor, an linearization

algorithm which relies only on asymmetric c-command (strong LCA) should be favored over

one which relies on both c-selection and asymmetric c-command (revised LCA). Note, first

of all, that this objection is valid only if both approaches are empirically equivalent. As

mentioned above, strong LCA-based approaches to head-finality face certain empirical and

conceptual challenges which render them independently problematic, making an appeal to

Occam's razor irrelevant. Even if both approaches were directly comparable, though, the

above characterization of the difference between the approaches is misleading. Head-finality

exists, in descriptive terms, in that languages like Japanese have essentially the opposite

---

[39] This raises the question why (ii) is not parameterized as (i) isIn section 8 I propose that it is at the level of UG,

but that third factor pressures so strongly prefer precedence that it has the appearance of a principle (cf.

Biberauer, Holmberg, Roberts & Sheehan 2010).

36

linear order of spine categories to English-type languages. As noted above, even deriving these orders under the strong LCA gives rise to certain amounts of redundant structure. It might be that the apparent simplicity of the strong LCA option is somewhat illusory.

As such, while conceptual considerations of 'methodological and ontological minimalism' must be taken seriously, explanatory adequacy must also be achieved. One could just as easily observe, for example, that agree has the capacity to value features at a distance and so it is methodologically and ontologically unminimalist for thematic roles to be determined locally. Nonetheless, it is still generally accepted that thematic roles cannot be assigned/valued via agree, based on empirical evidence. Likewise, agree is able to establish non-local dependencies between categories, precluding the conceptual need for movement, nonetheless it is generally accepted that movement exists in addition to agree. In short, theories are also constrained by empirical considerations. In this case, the analysis proposed represents an attempt to salvage the crucial insights of the LCA from certain serious empirical and conceptual challenges. To the extent that it succeeds, the apparent departure from the strong minimalist thesis (SMT) may be justified. Alternatively, it may turn out that there is some more principled reason why the linearization algorithm favours local over non-local asymmetries.

Now let us consider how this version of the LCA serves to derive FOFC. Let us assume that individual heads/categories can be freely specified to precede or follow the categories they select, as appears to be empirically necessary (see sections 1-2 above). It follows that in such cases part (ii) of the revised LCA will be required to aid the linearization of such heads.[40] Consider first a structure where W_P selects X_F (the inverse-FOFC order):[41]

---

[40] The linear position of specifiers will always be regulated by asymmetric c-command, hence the robust tendency for them to surface on the left, and hence for movement to be leftwards (cf. Kayne 1994, Cinque 2005)

(28)
```
        W
       / \
     W_P   X
          / \
         Y   X_F
```
W>X, Y>X
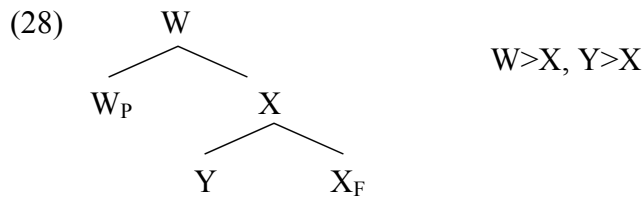
The sum of locally determined PF-paraneters here leaves the order of terminals

underspecified for (28). Both W and Y must precede X but the order between W and Y

themselves is simply not specified. In such cases, linearization appeals to less local

asymmetries between categories. As such, it is necessary also to consider the c-command

relations between W and Y, as per (27 (ii)). In (28), W asymmetrically c-commands Y and so

W must precede Y (according to (27 (ii))), giving the total linear order W>Y>X. The inverse-

FOFC order is therefore straightforwardly linearizable by the revised LCA in (27).

Crucially, the same is not true of the 'FOFC-violating' order in (29):

(29)
```
        W
       / \
      X   W_F
     / \
   X_P   Y
```
X>W, X>Y
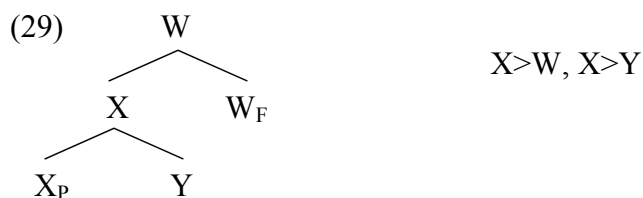
In (29), the sum of locally determined PF-parameters once again leaves the linear order of

categories underspecified. This time it is known that both W and Y must follow X, but again

no order is specified between W and Y themselves. In this case, too, W asymmetrically c-

commands Y and so via (ii) of the revised LCA the additional precedence pair W>Y is added.

As such, the linear order established for (29) is X>W>Y, rather than the FOFC-violating

---

[41] From now on PF-parameter settings will be represented in tree diagrams. No theoretical implications should

be read into this move, which is for ease of reference only.

38

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

order *X>Y>W. Crucially, the surface order X>Y>W cannot be derived from such a structure by the revised LCA unless the constituent [X Y] is spelled out separately (an option which I return to in section 8).[42] This raises the question whether orders of the X>W>Y kind are indeed attested. In the case of verb, object and auxiliary, V-Aux-O is possible in Old English, Finnish and Basque:

(30)   þæt   ænig mon atellan   mæge ealne þone demm

        that   any man relate   can all   the   misery

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

                [Old English, Pintzuk 2005: 13 (coorosiu,Or_2:8.52.6.1011)]

In Old English, V-Aux-O order is attested where O is *not* a negative object, pronominal object or particle, as Pintzuk (2005) notes. Interestingly, the latter often raise out of VP in Old English, making it plausible that (30) involves an in-situ object as per (29).

6.   **Contrasting the approaches**

Many of the predictions made by the two approaches are very similar. Both attach a certain cost to disharmony and both predict that the inverse-FOFC order will be more frequent than the FOFC-violating order. In what follows I outline both the similarities and differences

---

[42] In actual fact, as Sheehan (to appear b) notes, FOFC also appears to hold where a disharmonic word order is derived via movement rather than via non-harmonic head parameters. In such cases, she argues, the copy theory of labeling predicts that FOFC will hold for similar reasons to those outlined here (cf, also Sheehan to appear a, section 7)..

39

between the two accounts. Data which are equally problematic for both accounts is discussed in section 6.

6.1. The core cases

Both approaches account for the core cases of FOFC described in section 3. ***V-O-Aux** is predicted by Hawkins (1994) to be rarer than **Aux-O-V** for exactly the same reasons that ***P-DP-V** is predicted to be rarer than **V-DP-P.** Consider the following example:
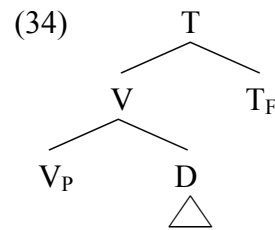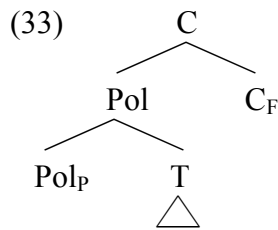
(31) $[_{TP} [_{VP} [_{V}$ eaten] $[_{NP}$ the cake]]$[_{T}$ has]] 1/3, 2/4 = 42%

The first IC of the TP (namely VP) will be constructed by *eaten*. VP dominates the first through third word in the CRD and so its IC-to-word ratio will be 1/3. The second IC is T, which is constructed by *has*. T dominates the fourth word in the CRD, so its ratio is 2/4. Once again, the inverse-FOFC order has a higher efficiency rating:

(32) $[_{TP} [_{T}$ has] $[_{VP} [_{NP}$ the cake] $[_{V}$ eaten]]] 1/1, 2/4 = 75%

In this case the first IC of TP, namely T, is constructed by *has*. As T dominates only the first word in the CRD, it has a maximally efficient IC-to-word ratio 1/1. The second IC, namely VP is constructed by *eaten*. VP dominates the second through fourth words in the CRD and so has a ratio of 2/4. The case of polarity question markers and complementizers works in the same way, as long as Pol serves to construct PolP and C constructs CP.

The PF-interface account explains the first two cases in exactly the manner outlined in section 4.2:

40

(33)
```
          C
        /   \
      Pol     C_F
     /   \
  Pol_P   T
          △
```

(34)
```
          T
        /   \
       V      T_F
     /   \
    V_P    D
           △
```

In both (33) and (34), PF-parameters alone fail to provide an unambiguous linear order of

categories. In (33), no order is specified between T and C, and in (34) the same is true

between D and T. Crucially, in both cases the FOFC-violating orders are predicted never to

surface.[43]

6.2.   **Exceptions to FOFC**

Thus far we have focused on the empirical support in favor of FOFC. It must be noted,

however that BHR (2007, 2008, 2010) also discuss certain apparent counterexamples to the

generalization in its most abstract form, and actually present a rather different version of

---

[43] An anonymous reviewer asks whether Pol-C-T and C-V-TP are attested as predicted. A potential example of

Pol-C-TP occurs in Mecayapan Nahuatl, assuning that the conditional marker occupies the same position as

question particles, which often have the same form:

(i)   <u>Si iga </u>quijliisquej iga quena, huel monaamictij.

       <u>If that </u>he gets permission, he can marry.'   [Mecayapan Nahuatl Gutiérrez-Morales (2008: 181)]

Si iga alternates with a fused form *siga*. Note that *iga* always surfaces in a clause-initial position, but this is as

predicted if all clauses contain an optionally covert polarity head, triggering 'extraposition' of TP to the right of

C. C-V-TP appears to be unattested, but full CP extraposition in such contexts is common (cf. Biberauer &

Sheehan to appear on the connection between CP extraposition and FOFC.

41

FOFC, in order to accommodate these counterexamples. They propose, for example, that FOFC holds only within extended projections, and that particles, being acategorial, are not subject to the generalization (cf. BHR 2010). Hawkins (this volume) raises some objections to limiting the scope of FOFC in this way, as it appears to rule out of the scope of the generalization some apparent FOFC-effects. The two approaches to FOFC explored here actually make slightly different predictions about which [[head-complement]-head] surface orders will be permitted. These overlap only partially with the kinds of exceptions allowed by BHR's account. This section explores which surface FOFC-violations are attested and which of the two approaches fairs better in accounting for these exceptions. [44]

The first apparent exception comes from DP complements of V, which surface in a preverbal position:

(35)   Johann hat   das/ein Buch   ausgeliehen

        John   has   the/a book   borrowed

        'John has borrowed the/a book.'

In actual fact, there are relatively few OV languages with clear determiners which allow a FOFC-violating configuration between DP and V. Most OV languages either (i) lack determiners distinct from demonstratives (which are plausibly specifiers) or (ii) have final determiners (cf. Dryer 1992:104). Moreover, OV languages with initial articles are no less common (taking into account areal and genetic factors) than VO languages with final articles. For this reason, there is no evidence for a FOFC asymmetry in this domain. A relevant question, then, is whether either approach can account for this fact.

---

[44] It is beyond the scope of this paper to provide a critical discussion of BHR's account.

Interestingly, the PGCH appears to have a ready explanation. Recall that the PGCH is sensitive to the relative heaviness of phrases. It follows, therefore, that DPs, tending to be the lightest phrases, might allow FOFC-violations more often than other categories from a cross-linguistic perspective, unlike heavier phrases such as PP and CP.[45] The lack of a skewing between the two disharmonic combinations of article, NP and V is therefore not an immediate problem for Hawkins' approach. In fact, given the status of articles in Hawkins' theory, it is not clear that there should actually be any difference between the inverse FOFC-order and the FOFC-violating order in terms of EIC. Hawkins rejects Abney's (1987) DP-hypothesis and assumes that NPs comprise Det and N, with either being sufficient to construct an NP. As such, N-initial and Det-initial nominals are equally efficient in processing terms. It follows that in the absence of other material, the order of constituents in the NP will have no effect on processing efficiency, and thus is expected to be irrelevant. A potential problem for Hawkins, however, is Dryer's (1992: 103) claim that articles are nonetheless verb patterners, as this means that we have a context in which there is a preference for harmony but no evidence of a FOFC-effect. Hawkins (1994, this volume) however, offers an alternative account of the distribution of articles, which potentially overcomes this problem.

The PF-interface account also has a potential account of the lack of FOFC effects between V and DP. Firstly, note that specific DPs arguably fall outside the purview of FOFC,

---

[45] Theresa Biberauer (p.c.) asks how heaviness is determined. Heaviness equates to number of words for the purposes of online reordering processes. According to the PGCH, though, grammars can develop categorical rules stemming from heaviness tendencies. Thus the fact that PP always contains DP makes PP heavier than DP in general terms, even though both can have different heaviness values in different contexts. As such, an extraposition rule might be grammaticalised to apply only to PP and not DP irrespective of their relative online heaviness.

if Kayne (2004: 10) is correct and, in all OV languages, specific DPs scramble outside VP.[46] In these terms, then, we might even say that the existence of obligatory scrambling predicts that no FOFC asymmetry will be observed between specific DP and V. Whatever triggers scrambling, then, plausibly serves to render specific DPs immune to FOFC. Non-specific nominals, however, can clearly remain insider VP in OV scrambling languages, meaning that they nonetheless constitute an apparent FOFC-violation. If the indefinite article were a projecting head in such cases then the following discontinuous word order would be predicted to surface, contrary to fact:

(36)   *Johann hat   ein   ausgeliehen   Buch
        John   has   a   borrowed   book

Interestingly, these non-specific representational nominals permit subextraction, whereas specific scrambled DPs do not (Müller 1996: 401-402, citing Webelhuth 1992, Fanselow 1991):[47]

(37)   Worüber   hat   Antje ein/*das   Buch   gelesen?
        What.about   has   Antje a/the   book   read
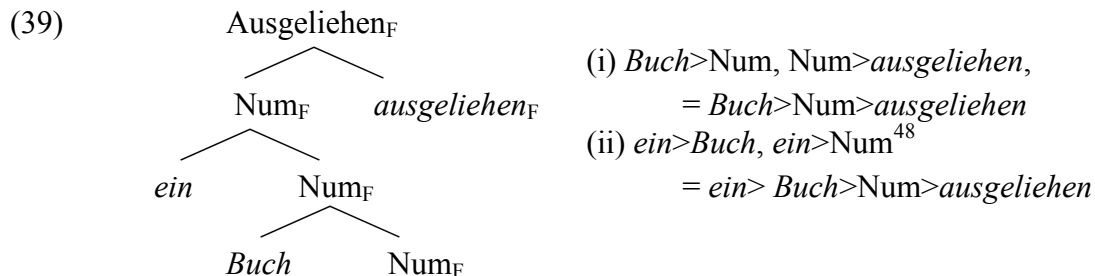   'What has Antje read a/*the book about?'   [German (Müller 1996: 402)]

44

Even in languages without scrambling, however, subextraction from specific representational DPs is severly degraded (cf. Fiengo and Higginbotham 1981, Davies and Dubinsky 2003):

(38)   Which topic did you borrow a/*the/*my book about?

As such, the effect in (37) does not reduce to a complement/specifier distinction as per Huang's (1982) Condition on Extraction Domain (contra Mahajan 1992). Sheehan (2010) proposes, following an idea in Bowers (1987), that non-specific indefinites are not full DPs, but rather truncated NumPs. As such *ein* is not a projecting head in (35) hence the ungrammaticality of (36). Rather the structure of (35) is a projection of a head-final (at PF) Num as follows:

(39)

$$\text{Ausgeliehen}_F$$

$$\text{Num}_F \qquad \textit{ausgeliehen}_F$$

$$\textit{ein} \qquad \text{Num}_F$$

$$\textit{Buch} \qquad \text{Num}_F$$

(i) *Buch*>Num, Num>*ausgeliehen*,
    = *Buch*>Num>*ausgeliehen*
(ii) *ein*>*Buch*, *ein*>Num[48]
    = *ein*> *Buch*>Num>*ausgeliehen*

Evidence in favor of this structure comes from instances of 'extraposition' from indefinite non-specific NumPs in German:[49]

---

[48] If order is computed online we might assume that once an unambiguous order of all categories has been calculated, no further asymmetric c-command relations are used. If this is the case then either *ein*>*Buch* or *ein*>Num would be sufficient here. This becomes more pertinent in relation to (40) below, where *ein* asymmetrically c-commands four categories.

45

(40) a. Er   hat   ein/*das   Buch   ausgeliehen   über Syntax.

He   has   a/the   book   borrowed   about syntax
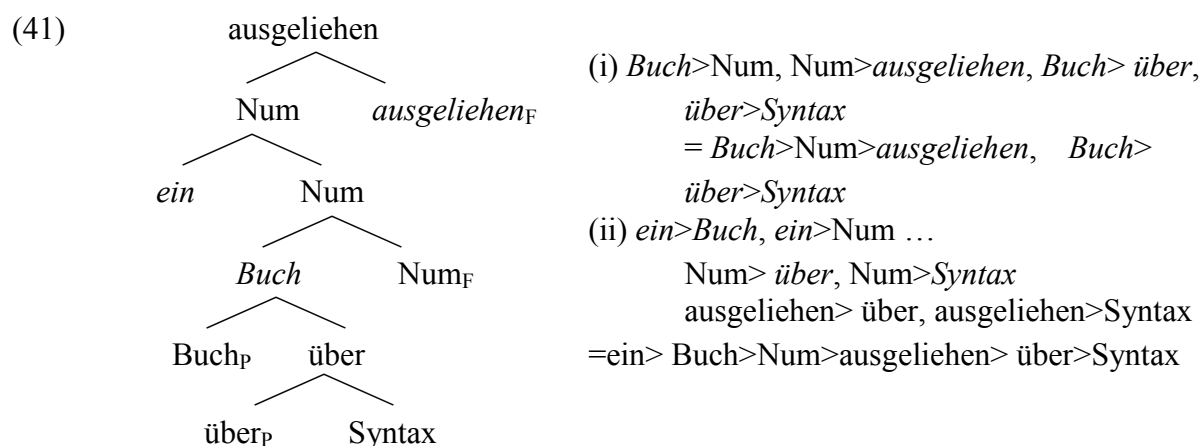
'He has borrowed a book about syntax.'

b. *Er   hat   ein   Buch doch   ausgeliehen   über Syntax.

He   has   a   book already   borrowed   about syntax

'He has already borrowed a (specific) book about syntax.'

As (40) shows, 'extraposition' is only possible from non-specific unscrambled nominals. This is as predicted if *Buch* projects to a head-initial (at PF) phrase, giving rise to discontinuous linearization as described in section 4.2:

(41)

```
              ausgeliehen
              /         \
          Num          ausgeliehen_F
         /    \
      ein     Num
             /    \
         Buch     Num_F
         /   \
     Buch_P  über
            /    \
        über_P  Syntax
```

(i) *Buch*>Num, Num>*ausgeliehen*, *Buch*> *über*,
     über>*Syntax*
     = *Buch*>Num>*ausgeliehen*,   *Buch*>
     über>*Syntax*
(ii) *ein*>*Buch*, *ein*>Num …
     Num> *über*, Num>*Syntax*
     ausgeliehen> über, ausgeliehen>Syntax
=*ein*> Buch>Num>ausgeliehen> über>Syntax

In these terms, the claim is that NumP is not immune to FOFC, but rather that a FOFC-effect does obtain where N has the relevant kind of complement.

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

BHR (2007, 2008) also discuss another kind of exception to FOFC involving prepositional phrases in OV languages. Once again, German provides a relevant example:

(42) Sie ist [VP [PP nach Berlin]   gefahren]

she is   to   Berlin   driven

'She went to Berlin.'

As Hawkins (this volume) notes, it appears, nonetheless, that despite a small number of robust counterexamples, there is cross-linguistic evidence of a FOFC effect between PP complements and V. Firstly, the raw data from Dryer (2005c/2011c, 2005d/2011d) indicate that there are more inverse FOFC languages than FOFC-violating languages, and that these languages occur in more macro-areas and represent a more genetically diverse group than the small number of FOFC-violating languages:

Table X.2 Languages with disharmony between the ordering of VP and PP

	FOFC-violating P-NP and OV	Inverse-FOFC NP-P and VO
N of language	12[50]	38
N of macro areas	3	5
Language families	5	16
Genera	8	22

---

[50] I have added German and Dutch to this category although they are categorized as having mixed order of verb and object on WALS because of their V2 property.

47

These data seem to be consistent with the spirit of the PGCH, whereby FOFC is a statistical rather than a categorical effect, as there appear to be a number of languages which allow the FOFC-violating combination. Note, however, that WALS contains data about the directionality of PP and VP but it does not consider the actual placement of head-initial and head-final PPs in otherwise OV and VO languages. An examination of the problematic languages reveals that many of the 12 languages, despite being OV, actually require PPs to be obligatorily postverbal, meaning that the surface FOFC-violating construction *P>DP>V does not actually occur (as Hawkins also notes).

Of the 12 languages, only the Indo-European languages (German, Dutch, Persian, Tajik, Kurdish and Sorbian) and Tigré (Semitic) allow structures of the kind P>DP>V. The other languages all either lack true adpositions or are languages in which PPs appear in a post-verbal position. Mangarrayi (Australian, Mangarrayi, Northern Territories) lacks true prepositions according to Merlan (1989: 26), instead it has 'prepositional-like phrases consisting of an adverb followed by a noun appropriately case-marked to complement the combined meaning of the adverb and verb in the clause'. Moreover, the order of the adverb and its complement is not fixed, both adv>comp and comp>adv are possible. As such it is not clear that Mangarrayi is a true exception to FOFC.[51] In Päri (Nilo-Saharan, Nilotic, Sudan), Tobelo (West Papuan, North Halmaheran, Indonesia), Iraqw (Afro-Asiatic, Southern Cushitic, Tanzania) and Neo-Aramaic (Afro-Asiatic, Semitic, Israel) PP complements to V seem to surface in a post-verbal position, as the following examples illustrate:

(43)   á-lw'ʌʌr'   kí   kwàc   [Päri, Anderson (1988: 303)]

---

[51]As such, it seems that Mangarrayi is wrongly classified on WALS.

1S-fear   PREP   leopard

'I am afraid of leopards.'

(44)   lăbulmunne   [ta-Bagdàd]   [Neo-Aramaic, Khan (1999: 338)]

        Take-me   to-Bagdad

        'Take me to Bagdad!'

(45)   i-na   ta'<a'>ín   [ay dí-r   konkomo] [Iraqw, Mous (1993: 100)]

        S-PAST   run<HAB>3SGM   to   place.F-CON   cock

        'He ran to the cock'

In Päri, this is obligatory (Anderson 1988: 303), as it is in all dialects of Neo-Aramaic (Geoffrey Khan, personal communication). There is insufficient information to say the same for Iraqw, though all examples suggest it to be the case. Matters are slightly less clear in Tobelo.[52] Following Kandybowicz and Baker's (2003) analysis of Nupe, we might take the

---

[52] The vast majority of the examples show PPs to be extraposed:

(ii)   ngohi-o   to-modeke   de   o-Matias   [Tobelo, Holton (2003:30)]

      I-also   1-agree   PREP   NM-Matias

      'I also agreed with Matias.'

However, Holton (2003:55) explicitly states that oblique arguments can occur 'either before or after the verb'. Unfortunately, he uses the term 'oblique arguments' to refer to both PPs and DPs marked with a locative or directional case suffix. The only example with PP-V order which he gives involves an instrumental PP which may be a topic in the CP layer:

(iii)   de   ma-kakatama   n-a-lye-ino   [Tobelo, Holton (2003:55)]

      P   NM-tongs   2-3-roll-ALL

49

position of PP as evidence that these languages are basically head-initial with DP-V order derived via A-movement. This means that the number of languages displaying FOFC-violations of this type is extremely small so that P>DP>V is far rarer than the inverse-FOFC order [V-[DP-P]. If this were not a FOFC effect, but merely a preference for harmony then there would be no explanation for the fact that postpositional phrases in VO languages are rarely preposed (cf. Dryer 1992: 92).

These data can be accommodated by the PGCH, if they are taken to indicate a statistical FOFC-effect. PP is a relatively 'light' IC, so FOFC violations between V and PP are predicted to be typologically rare but not impossible. In these terms, Germanic/Iranian/Tigré are not optimal for processing purposes, but not ruled out by UG. Moreover, optional PP extraposition is available to varying degrees in the languages in question, meaning that the problematic surface order P>NP>V may be avoided to a certain extent in usage. This raises the question, though, pointed out by Theresa Biberauer (p.c.) why German has not lost preverbal PPs over time. One might also question why there are so few counterexamples to FOFC in this domain, given that, for Hawkins, PP is a fairly light category.

The small number of well-studied exceptions are more problematic for the PF-interface account.[53] There is no obvious reason why PPs should fail to be subject to FOFC

'Roll it up with the tongs.'

In the absence of further data is it impossible to say for certain whether Tobelo PPs adhere to FOFC or not.

[53]It is worth reiterating, though, that in the hard sciences analyses are rarely required to explain 100% of a dataset (cf. Johnson 2007).

only in this particular subset of Indo-European and Semitic languages.[54] In relation to the

Indo-European counterexamples, one possible approach might be to exploit the idea that P is

parameterized as to whether or not it is a phase-head (Abels 2003). In languages allowing

P>DP>V, PPs would constitute separate spell-out domains or phases and would therefore

behave like atoms for the purposes of linearization. In this respect, it is initially suggestive

that all Germanic VO languages allow preposition stranding, whereas most Germanic OV

languages disallow it (as does Persian).[55] This suggests that in standard German, Dutch and

the relevant Iranian languages PPs are phases and preposition stranding is ruled out by anti-

locality (Abels 2003). As expected, non-local subextraction from PPs in German is possible,

as Abels shows:

(46)　　[Über　welches Thema]$_i$ hast du mich noch mal [$_{PP}$ nach einem Buch t$_i$]　gefragt?

　　　　Overt　which　subject have you me again　　　　after a　　book　asked

　　　　'Which topic did you ask me about a book on again?'

This is not a satisfactory solution to the problem, though, as FOFC effects are observed

across phasal barriers and this is as expected if spell-out leaves the edge of the phase active in

the derivation. Note that CP complements of V are presumably phasal but nonetheless display

a FOFC asymmetry (as discussed above). At present then, the existence of a non-categorical

---

[54] Baker (2010) argues that Amharic, another Semitic language which also allows the FOFC-violating order,

may not be a true counterexample to FOFC as the adpositions in question are not actually prepositions but rather

clitics which are ordered via independent PF-rules.

[55] Though, as Theresa Biberauer (p.c.) notes, colloquial Afrikaans and some dialects of Dutch and German allow

preposition stranding, so this correlation is not perfect.

FOFC-effect between PP and VP remains problematic for the PF-interface account.[56]

Ultimately, the behaviour of PPs in these languages should be attributed to some independent property of their grammar.[57]Unfortunately, this property remains elusive at present.

As such, it must be conceded that the fact that a statistical FOFC-asymmetry appears to be attested between PP and is suggestive evidence against the PF-interface approach and in favor of the PGCH-based analysis of FOFC.[58] This is because the PGCH-based account predicts the effect to be statistical rather than absolute and formal approaches simply cannot account for this kind of non-categorical effect. In the following section, however, I raise a potentially more serious problem with the PGCH-based approach.

6.3.   The relation to harmony

According to the PGCH, harmony and FOFC are two effects of the same processing principle (EIC). The clear prediction is that if two categories show a preference for harmony then they should also show a FOFC skewing and vice versa. The PF-interface account makes no such

---

[56] BHR (2010) and a caveat to their formulation of FOFC so that it holds only within extended projections. This is also unsatisfactory, though, as only a small number of closely related languages permit the order [[P-DP]-V].

[57] Potentially related phenomena, though, include the availability of circumpositions in Iranian/Germanic, which are in themselves problematic for FOFC, and the positioning of head-initial adjuncts. Note also that unlike many languages, which have few or even a single preposition, Iranian and Germanic are notable for their rich inventory of Ps. A full consideration of these structures is beyond the scope of the present paper.

[58] Theresa Biberauer raises a further potential FOFC-exception in Germanic verb clusters, which

(iv)   ... dat hy die boek moes koop het   [2, 3, 1]

   that he the book must.past buy have = "... that he should have bought the book"
These raise problems for both the accounts proposed here.

prediction. The preference for harmony may be neutralized for some independent reason, but FOFC will always hold as long as the two categories in question are linearized in the same application of Spell-Out. This is because on the PF-interface account, the effect results from basic facts about asymmetric c-command and linearization.

Some categories display both a preference for harmony and a FOFC effect. This is the case for Aux and VP, V and PP, C and VP and V and CP (cf. Dryer 1992). As such, these pairs of categories are well-behaved by the standards of EIC. More problematic for the PGCH is the fact that according to Dryer (1992) articles are also verb patterners, but not subject to FOFC effects cross-linguistically (see previous discussion).

Polar question particles, on the other hand, seem to display the opposite situation. From a functional perspective, question particles (Pol) arguably serve the function of identifying a clause as a polar question, and C serves to construct an embedded clause. For this reason, it is surprising that according to the data in table 1 from Dryer (2005a/2011a 2005b/2011b) repeated here, the disharmonic inverse-FOFC combination initial C and final Pol is more common than the harmonic combinations in terms of languages and genera and almost as common in terms of language families.

Table X.1 Typological positioning of Polarity heads and complementizers

Type	Position of Pol	Position of C	Number of Languages (genera: families)
A	Initial	Initial Only	72: 35: 13 (78)
B	Final	Final Only	45: 33: 20 (46)
C	Final	Initial	74: 40: 16 (82)
D	Initial	Final	4: 3: 3 (4)

There is thus no evidence of a preference for harmony between C and Pol. Crucially, though, FOFC appears to hold in this domain (as discussed in section 3.2). It therefore appears to be the case that one and the same category displays a FOFC effect but no preference for harmony. This is highly problematic for the PGCH, which attributes both FOFC and the preference for harmony to the same principle, namely EIC. Of course, it remains an open question why there is no harmonic preference between Pol and C.[59] Under both approaches this fact can be accommodated as an effect of grammaticalization (cf. Bailey 2010). A difference emerges in relation to whether the approaches can deal with a FOFC effect in such contexts. From a formal perspective, even if two heads are shown not to display a cross-linguistic tendency towards harmony, they are still expected to be subject to FOFC, as FOFC is an unavoidable fact about linearization. From a processing perspective, on the other hand, once two heads fail to display a cross-linguistic tendency towards harmony, they fail to be subject to EIC and as such any adherence to FOFC cannot be attributed to the processing principle. In such cases, then, FOFC effects simply cannot be explained by EIC.

7.   The problem of Particles

The final type of exception to FOFC, also discussed by BHR (2007, 2008, 2010) comes from particles. It is well-known from the typological literature that tense/aspect particles, unlike inflected auxiliaries, are not verb patterners (cf. Dryer 1992: 114-115, Hawkins, this volume). Interestingly, they also fail to be subject to FOFC:

---

[59] Ultimately this might also shed light on the reason why final Q is so frequent in VO languages (cf. Dryer 2005a.2011a, 2005c/2011c).

(47) Wo-men   daoda shung- ding   **le**   [Mandarin Chinese]

  1SG-PL   reach   mountain-top PART

  'We have reached the top of the mountain.'   [BHR, citing Soh and Gao, 2004]

Mandarin Chinese also has a plethora of final discourse particles, limited to matrix clauses, which appear to violate FOFC (cf. Paul 2008). BHR (2007, 2008) note that these kinds of structures tend to cluster in a given language with languages allowing many different C particles in the same VO…PART structure.

The first thing to observe is that there is no clear independent diagnostic for particlehood and this potentially challenges the validity of FOFC as a generalization (cf. Biberauer and Sheehan to appear). Particles are often claimed to be 'deficient' in some sense. In empirical terms, this equates to at least the following properties:

  i.   lack of inflection/ morphologically invariance (cf. Greenberg 1963, Dryer 1992, Cinque 1999);

  ii. phonological deficiency (but cf. Cardinaletti in press)

  iii. immobility;

  iv. lack of category (BHR 2010);

  v.   inability to c-select (BHR 2010).

Note that while (i) may be a necessary condition for particlehood, it is arguably not sufficient. Lack of inflection is the norm for many categories and indeed in isolating languages such as Mandarin, it is the norm across all categories. Diagnostics (ii) and (iii), likewise, are necessary but not sufficient conditions for particlehood, as they are more general properties of functional heads (cf. Roberts and Roussou 2003: section 5.3.2). Diagnostics (iv) and (v)

55

are theoretically interesting but empirically ill-defined. It is not clear what the empirical diagnostic for lack of categoryhood is, for example. In conjunction with (v), though, (iv) suggests that particles should be fairly free in their distribution, able to merge with a number of different complements and be invisible to selection. Indeed there are 'particles' with this behaviour, so-called focus-related particles such as *only* are a case in point as they can be associated with any kind of phrase, suggesting that they do not have a place in the functional sequence (cf. Cinque 1999 for this conclusion). The particles which fail to be subject to FOFC in many cases fail to have these properties, though. Final high particles in Mandarin and Cantonese occur in a rigid sequence for example. Even this diagnostic, then is not sufficient to determine the class of heads which are immune to FOFC.

As a result, it is even impossible even to add a stipulation to FOFC stating that it does not apply to particles, as at present the only diagnostic for particlehood is insensitivity to FOFC. This problem is compounded by the fact that some categories appear to be subject to FOFC in one direction but not the other. Final question particles, for example are frequently attested in VO languages and also fail to harmonize with the order of verb and object (Dryer 2005a/2011a, 2005c/2011c). However, this does not render question particles immune to FOFC per se. As discussed in section 3.2., evidence suggests that FOFC rules out the order ***Pol-TP-C**, suggesting that it is not possible to simply classify question markers as particles and rule them out of the FOFC generalization altogether.

This problem, in a sense, extends to the PGCH. On the one hand, the PGCH allows that a pair of categories might fail to display a cross-linguistic tendency for harmony, and in such cases makes the prediction that these categories will also fail to be subject to FOFC. On the other hand it offers no real explanation as to why certain categories fall outside its remit

56

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

in this way.[60] In the absence of an independent definition/diagnostic of particlehood, the

account again risks circularity.

The PF-interface account potentially avoids this problem by simply avoiding any

reference to the term 'particle'. More specifically, it makes a strong prediction regarding

surface orders which appear to be FOFC-violations: the phrase preceding a final particle must

be an atomic strong island, unless it is harmonically head-final. This stems from the fact that

as long as the complement of a 'particle' is a hierarchical head-initial phrase (at PF), the

particle will fail to appear in final position. If the complement of the 'particle' is already

spelled-out, though, and hence atomized in the sense of Uriagereka (1999a), then its internal

hierarchy is obliterated and it is rendered immune to FOFC. Interestingly, there is some initial

evidence that the prediction holds with certain final particles in Mandarin Chinese (a VO

language). Lin (2006) observes that the presence of certain final particles indeed correlates

with the strong islandhood of the preceding phrase. More specifically, he observes that the

presence of the final aspectual particle *le* blocks extraction of manner adverbials from VP:

(48) a. Zhangsan   zenmeyang   xiu   che __?

         Zhangsan   how   repair car

         'How does Zhangsan repair the car?'

      b. *Zhangsan   zenmeyang   xiu   che   le?

         Zhangsan   how   repair car   SFP

         'How did Zhangsan repair the car?'   [Mandarin, Lin (2006: 4)]

---

[60] An anonymous reviewer suggests that particles might be non-projecting elements which therefore fail to head

a projection and hence do not reconstruct a category. Again, although this is an interesting theoretical

characterization, it provides no independent diagnostic for particlehood.

57

Adopting a Kaynean position, Lin attributes this to Huang's (1982) Condition on Extraction Domain, whereby specifiers, as they are ungoverned, are strong islands, unlike complements, which are governed and hence permit subextraction. The idea is that (39b), unlike (39a), involves movement of vP to a specifier position above Asp.

In relation to FOFC, this explanation of the facts is problematic, though, as it overgenerates, and would allow us to derive any number of surface FOFC-violating orders with final inflected heads ***V-O-Aux, *Pol-TP-C** etc. In fact, as Sheehan (to appear b) observes, even languages which appear to derive a head-final order via this kind of movement still appear to adhere to FOFC. Thus Finnish allows both VO and OV orders (depending on information structure) and both Aux-V and V-Aux orders. Moreover, Holmberg (2001) argues at length that Finnish V-Aux orders are derived via VP-movement, and this is consistent with the fact that they are optional. Nonetheless, ***V-O-Aux** order is still ungrammatical in Finnish, suggesting that the FOFC generalization applies even where phrasal movement is involved. This is especially suggested by the grammaticality of (49), which is arguably the linear order resulting from movement of a head-initial VP:

(49)   *Milloin Jussi *kirjoittanut* romaanin   olisi?   *[V-O-Aux]

        when   Jussi written   novel   would.have

                                                        [Finnish, Holmberg (2000)]

(50)   Milloin   Jussi   kirjoittanut   olisi   romaanin?

        When   Jussi   written   would.have   novel

        'When would Jussi have written a novel?'   [Finnish, Anders Holmberg, p.c.]

So what is the explanation for Lin's observation, then and how does it solve the particle problem? It must be the case that the islandhood of vP in (48b) is not caused by movement to a specifier position, but rather that the surface violation of FOFC is possible because the phrase which is selected happens to be atomic. Let us further capitalize on one of the necessary properties of particles and assume that only heads which lack uninterpretable features other than a c-selection specification can select for an atomized complement. This is because the presence of uninterpretable features on a given category require that head to probe some other category in its c-command domain. If the complement of a head is atomized it follows that that the categories inside that domain are closed off to probing. In these terms, we can define 'particle' as a functional head without uninterpretable features (other than c-selection features) which merges with an atomized phrase. This, at least, yields falsifiable predictions about where (and how) surface FOFC violations will arise. The implication is that FOFC does hold absolutely as a condition on the linearization of structures: all superficial exceptions involve a reduced structure in which the head-initial phrase has been atomized for independent reasons.[61]

8.   Formal or functional?

---

[61] Theresa Biberauer (p.c.) raises some empirical problems for this prediction which warrant close attention, notably from the presence of final particles in clauses with wh-movement in Singaporean English. Given that wh-movement in Singaporean English is optional. One possibility is that wh-movement is to a Focus position below the position occupied by Pol. The fact that extraction from circumpositions in Dutch and German is also possible remains problematic.

Up to this point, the PF-interface account has been characterized as a 'formal' account of the FOFC-asymmetry essentially because it takes the constraint to be categorical rather than statistical, and to derive from I-language rather than E-language principles. This formal/functional divide is, however, misleading in the context of the minimalist program which attempts to declutter the biological component of UG by assigning some core properties of language to second and third factors (cf. Chomksy 2005). As such, the PF-interface account is formal only in as far as the LCA is a formal principle of first factor UG (the genetic endowment). In recent work, Biberauer, Holmberg, Roberts and Sheehan (2010) question this assumption, and argue rather that the LCA might actually be an emergent parameter of language arising from the inexorable need for linear order in externalisation. Note that, in these terms, parameters are defined as points of underspecification in UG (Biberauer and Richards 2006, Richards 2009, Roberts and Holmberg 2010). As such, then, there would actually be a parameter setting in both part (i) and (ii) of the revised LCA:

(51)   Revised LCA

(i) If a category A c-selects a category B, then A **precedes/follows** B at PF

(ii) If no order is specified between A and B by the sum of all precedence pairs defined by (i), then A **precedes/follows** B at PF if A asymmetrically c-commands B.

The idea is that linearization is parasitic on narrow syntactic asymmetries (which are part of the biological endowment) because this is the simplest way to translate a hierarchical structure into linear order. However, the linearization algorithm is free to map the relevant asymmetries to either precedence or subsequence, the two logical possibilities arising from linearity. This raises the question why asymmetries of the FOFC-kind exist at all. The answer provided by Biberauer, Holmberg, Roberts and Sheehan (2010) is that spoken languages

60

unanimously set the parameter in (ii) to 'precedence' rather than 'subsequence' so that (ii) essentially constitutes a 'no-choice parameter'. Interestingly, the reason why non-local asymmetries in narrow syntax preferentially map onto precedence might plausibly stem from processing preferences of the kind which Hawkins has also discussed. First, note that filler-gap dependencies are easier to parse if the filler precedes the gap (Hawkins 2001, Abels and Neeleman 2009). This is reflected by experimental evidence which suggests that subjects look for a gap once they have heard/read a filler (cf. Crain and Fodor 1985, Wagers and Phillips 2009 for a recent overview). The hypothesis is that this exerts a strong third factor pressure favoring the precedence setting of (ii) above. Why then, doesn't the same thing happen with (i)? Arguably this is because there is no such parsing pressure affecting (i). As Hawkins argues, head-final and head-initial orders are equally optimal in parsing terms.

Crucially, then, as Biberauer, Holmberg, Roberts and Sheehan (2010) note, once (ii) has been set to precedence it may have further more arbitrary implications for word order which do not afford any direct processing advantage: hence FOFC. In these terms, the PF-interface account of FOFC, while not functional, has a functional basis of sorts. The linear order of natural languages is derived from narrow syntactic asymmetries which exist in all natural languages, but the basic asymmetry which ultimately gives rise to FOFC (as a side-effect) actually stems from third factor processing pressures.

9. Conclusions and unresolved issues

In conclusion, there is some evidence that something akin to FOFC seems to hold of natural language, despite a small number of counterexamples. However, it seems to be the case that FOFC is independent of any preference for cross-categorial harmony. This renders the PGCH problematic as it predicts harmony and FOFC to be inextricably linked. As such, the PF-

interface account remains the more plausible account at present, subject to further investigation (notably of the problematic PP structures in Germanic/Iranian). Characterising the PF-interface account as a purely formal' account might, however, be misleading as the LCA plausibly reduces to an emergent 'no-choice parameter', which means that where non-local dependencies are concerned higher categories preferably precede lower categories for processing reasons. The implication is that Hawkins' approach, while it cannot explain FOFC in directly functional terms, might explain the forces affecting parameter-setting which give rise to FOFC as a side-effect.

10. References

Abels, Klaus (2003). *Successive Cyclicity, Anti-locality, and adposition stranding*. Ph.D dissertation, University of Connecticut.

Abels, Klaus and Neeleman, Ad (2009). Universal 20 without the LCA. In J. M. Brucart, A. Gavarró and J. Solà, (eds.), *Merging features*. Oxford: Oxford University Press, 60-80.

Abney, Steven P. (1987). *The English noun phrase in its sentential aspect*. Ph.D dissertation, Massachusetts Institute of Technology.

Aboh, Enoch (2004). *The morphosyntax of complement-head sequences.* Oxford: Oxford University Press.

Allin, Trevor R. (1976). *A grammar of Resígaro*. Horsleys Green, Buckinghamshire: Summer Institute of Linguistics.

Anderson, Torben (1988). 'Ergativity in Pari, a Nilotic OVS language', *Lingua* 75: 289-324.

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

Bailey, Laura (2010). 'Sentential word order and the syntax of question particles', *Newcastle Working Papers in Linguistics* 16: 23-43.

Baker, Mark (2010). 'On the nature of apparent mixed headedness in Amharic', Ms. Rutgers.

Bayer, Josef (1999). 'Final complementizers in hybrid languages', *Journal of Linguistics* 35: 233-271.

Bayer, Josef (2001). 'Two grammars in one: sentential complements and complementizers in Bengali and other South Asian languages', in P. Bhashkararao, and K. V. Subbarao (eds.), *The Yearbook of South Asian languages: Tokyo symposium on South Asian languages- contact convergence and typology*. New Delhi: Sage Publications, 11-36.

den Besten, Hans (1981). 'On the interaction of Root Transformations and lexical deletive rules', *Groninger Arbeiten zur germanistischen Linguistik* 20: 1-78.

Biberauer, Theresa, Holmberg, Anders and Roberts, Ian (2007).'Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC)', in A. Bisetto, and F. Barbieri (eds.), *The proceedings of XXXIII Incontro di Grammatica Generativa*, 86-105.

Biberauer, Theresa, Holmberg, Anders and Roberts, Ian (2008). 'Structure and linearization in disharmonic word orders', in C. B. Chang and H. J. Haynie (eds.), *The proceedings of the 26th Western Coast Conference on Formal Linguistics*. Somerville, MA.: Cascadilla Proceedings Project, 96-104.

Biberauer, Theresa, Holmberg, Anders and Roberts, Ian (2010). 'A syntactic universal and its consequences', to appear in *Linguistyic Inquiry.*

Biberauer, Theresa, Holmberg, Anders, Roberts, Ian G., and Sheehan, Michelle (2010). 'Reconciling formalism and functionalism: A Minimalist perspective', paper presented at *Linguistics Association of Great Britain*. Leeds. September 2010.

Biberauer, Theresa, Newton, Glenda and Sheehan, Michelle (2009). 'Limiting synchronic and diachronic variation and change: the Final-Over-Final Constraint', *Language and Linguistics* 10: 701-743.

Biberauer, Theresa, Sheehan, Michelle and Newton, Glenda (2010). 'Impossible changes and impossible borrowings', in A., Breitbarth, C., Lucas, S.Watts and D. Willis (eds.), *Continuity and change in grammar.* Amsterdam: John Benjamins, 35-60.

Biberauer, Theresa, and Richards, Marc (2006). 'True optionality: When the grammar doesn't mind', in C. Boeckx (ed.), *Minimalist Essays.* Amsterdam: John Benjamins, 35-67.

Biberauer, Theresa and Sheehan, Michelle (2011). 'Disharmony, Antisymmetry, and the Final-over-Final Constraint', to appear in M. Uribe-Etxebarria, and V. Valmala, (eds.), *Ways of structure building.* Oxford: Oxford University Press, PP.

Biberauer, Theresa and Sheehan, Michelle (eds.) (to appear). *The Linguistic Review: special edition on particles.*

Blevins, Juliette (1995).'The syllable in phonological theory',in J. Goldsmith (ed.), *Handbook of phonological theoy.* Oxford: Blackwell, 206-244.

Bowers, John. (1987). 'Extended X-bar theory, the ECP and the left-branch condition', *Proceedings of the West Coast Conference on Formal Linguistics*, 47-62.

Bybee, Joan L. (1988). 'The diachronic dimension in explanation'. In J. A. Hawkins (ed.), *Explaining language universals.* Oxford: Blackwell, 350-379.

Chomsky, Noam. (1995). 'Bare phrase structure', in G. Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program.* Oxford: Blackwell, 383–439.

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

Cinque, Guglielmo (1999). *Adverbs and functional heads: a cross-linguistic perspective.* New York: Oxford University Press.

64

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

Cinque, Guglielmo (2005). 'Deriving Greenberg's Universal 20 and Its Exceptions',
Linguistic Inquiry 36: 315–332.

Clahsen, Harald , and Muysken, Pieter (1986). 'The availability of universal grammar to
adult and child learners - a study of the acquisition of German word order', *Second
Language Research* 9: 93-119.

Crain, Stephen and Fodor. Janet D. (1985). 'How can grammars help parsers?', in D. Dowty,
L. Karttunen and A. Zwicky (eds.), *Natural language parsing*. Cambridge:
Cambridge University Press, 94–128.

Creissels, Denis (2005). 'SOVX constituent order and constituent order alternations in West
African languages', *Proceedings of the thirty-first annual meeting of the Berkeley
Linguistics Society.* Berkeley: Berkeley Linguistics Society, 1-15.

Davies, William, D. and Dubinsky, Stanley (2003). 'On extraction from NPs', *Natural
Language and Linguistic Theory* 21: 1-37.

Davison, Alice (2007). 'Word order, parameters, and the Extended COMP projection', in J.
Bayer, T. Bhattacharya, T. and M. T. H. Babu (eds.), *Linguistic theory and South
Asian languages*. Amsterdam/Philadelphia: John Benjamins, 175-198.

Deutscher, Guy (2000). *Syntactic change in Akkadian : the evolution of sentential
complementation*. Oxford: Oxford University Press.

Diesing, Molly (1992). *Indefinites*. Cambridge, MA.: MIT Press.

Dryer, Matthew S. (1992). 'The Greenbergian word order correlations', *Language* 68: 81-
138.

Dryer, Matthew S. (2005a/2011a). 'Position of Polar Question Particles', in M. Haspelmath,,
M. S. Dryer, D. Gill and B. Comrie (eds.), *The world atlas of language structures*.
Oxford: Oxford University Press, 374-377. [reissued as Dryer, Matthew S. (2011a)
'Position of Polar Question Particles', in M. Dryer, M. Haspelmath, D. Gill and B.

Comrie (eds.), *The world atlas of language structures online.* Munich: Max Planck digital library, chapter 92, available online at http://wals.info/chapter/92. Accessed 2011-05-09]

Dryer, Matthew S. (2005b/2011b). 'Order of adverbial subordinator and clause', in M. Haspelmath, M. Dryer, D. Gill and B. Comrie (eds.), *The world atlas of language structures.* Oxford: Oxford University Press, 382-385. [reissued as Dryer, Matthew S. (2011b) 'Position of Polar Question Particles', in M. Dryer, M. Haspelmath, D. Gill and B. Comrie (eds.), *The world atlas of language structures online.* Munich: Max Planck digital library, chapter 94, available online at http://wals.info/chapter/94. Accessed 2011-05-09]

Dryer, Matthew S. (2005c/2011c). 'Order of verb and object', in Haspelmath, M., Dryer, M., Gill, D. and Comrie, B. (eds.), *The world atlas of language structures* 338-341. Oxford: Oxford University Press. [reprinted as Dryer, Matthew S. (2011c). Order of object and verb. In M. Dryer, M. Haspelmath, D. Gill and B. Comrie (eds.), *The world atlas of language structures online.* Munich: Max Planck digital library, chapter 83, available online at http://wals.info/chapter/83. Accessed 2011-05-09]

Dryer, Matthew S. (2005d/2011d). 'Order of Adposition and Noun Phrase', in M. Haspelmath,, M. S. Dryer, D. Gill and B. Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press, 346-349. [reissued as Dryer, Matthew S. (2011a) 'Order of Adposition and Noun Phrase', in M. Dryer, M. Haspelmath, D. Gill and B. Comrie (eds.), *The world atlas of language structures online.* Munich: Max Planck digital library, chapter 85, available online at http://wals.info/chapter/85. Accessed 2011-05-09]

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

Dryer, Matthew S. (2005e). 'Genealogical language list', in M. Haspelmath,, M. S. Dryer, D. Gill and B. Comrie (eds.), *The world atlas of language structures*. Oxford: Oxford University Press, 584-644.

Dryer, Matthew S. (2009). 'The branching direction theory revisited;, in S. Scalise, E. Magni and A. Bisetto (eds.), *Universals of language today.* Berlin: Springer, 185-207.

Ernst, Thomas (2003). 'Adjuncts and word order asymmetries', in A.-M, Di Sciullo (ed.), *Asymmetry in Grammar*. Amsterdam: John Benjamins, 187-208.

Fanselow, Gisbert (1991). *Minimale Syntax*, unpublished habilitation thesis, Universitat Passau.

Fernández-Salgueiro, Gerardo (2008). 'The case-F valuation parameter in Romance', in T. Biberauer (ed.), *The limits of syntactic variation*. Amsterdam: John Benjamins, 295-310

Fiengo, Robert, and Higginbotham, James (1981). 'Opacity in NP', *Linguistic Analysis* 7: 395-422.

Fox, Danny and Pesetsky, David (2005).' Cyclic linearization of syntactic structure', *Theoretical linguistics* 31: 1-45.

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

Gutiérrez-Morales, Salomé (2008). *Borrowing and Grammaticalization in Sierra Popoluca: The Influence of Nahuatl and Spanish*. Ph.D dissertation, University of California at Santa Barbara.

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

Haider, Hubert (2004). 'Pre- and postverbal adverbials in OV and VO', *Lingua* 114: 779-807.

Haig, Geoffrey (2001). 'Linguistic diffusion in present-day Easter Anatolia: from top to bottom', in A. Y. Aikhenvald and R. M. W. Dixon (eds.), *Areal diffusion and genetic inheritance : problems in comparative linguistics*. Oxford: Oxford University Press, 195-224.

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

Hawkins, John A. (1994). *A performance theory of order and constituency*. Cambridge, New York: Cambridge University Press.

Hawkins, John A. (2001). 'Why are categories adjacent?', *Journal of Linguistics* 37: 1-34.

Hawkins, John A. (2004). *Efficiency and complexity in grammars*. Oxford, New York: Oxford University Press.

Holmberg, Anders (2000). 'Deriving OV Order in Finnish', in P. Svenonius (ed.), *The derivation of VO and OV*. New York: Oxford University Press, 123-152.

Holmberg, Anders (2001). 'The syntax of yes and no in Finnish', *Studia Linguistica* 55: 140-174.

Holmberg, Anders (2003). 'Yes/no questions and the relation between tense and polarity in English and Finnish', *Linguistic Variation Yearbook* 3: 45-70.

Holton, Gary (2003). *Tobelo*. Munchen: Lincom Europa.

Huang, C.-T. James (1982). *Logical relations in Chinese and the theory of grammar.* Ph.D dissertation. Massachusetts Institute of Technology.

Huang, C-T. James (1994). 'More on Chinese word order and parametric theory', in B. Lust, M. Suñer, J. Whitman, G. Hermon, J. Kornfilt and S. Flynn (eds.), *Syntactic theory and first language acquisition: cross-linguistic perspectives.* Hillsdale, NJ: L. Erlbaum Associates, 15-35.

Jackendoff, Ray (1977). *X-bar syntax: a study of phrase structure.* Cambridge, MA.: MIT Press.

68

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

Johnson, Kent (2007). 'The legacy of methodological dualism', *Mind and language* 22: 366-401.

Julien, Marit (2002). *Syntactic heads and word formation*: Oxford studies in comparative syntax. Oxford ; New York: Oxford University Press.

Kandybowicz, Jason and Baker, Mark C. (2003). 'On Directionality and the Structure of the Verb Phrase: Evidence from Nupe', *Syntax* 6: 115-155.

Kayne, Richard S. (1998). 'Overt vs. covert movement', *Syntax* 1: 128-191.

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

Kayne, Richard S. (2004). 'Antisymmetry and Japanese', in L. Jenkins (ed.), *Variation and universals in biolinguistics*. Amsterdam: Elsevier, 3-37.

Kayne, Richard S. (2008). 'Antisymmetry and the lexicon', *Linguistic Variation Yearbook* 8: 1-31.

Khan, Geoffrey (1999). *A grammar of Neo-Aramaic : the dialect of the Jews of Arbel*. Leiden: Brill.

Kiparsky, Paul (2008). 'Universals constrain change; change results in typological generalizations', in J. Good (ed.), *Language Universals and Language Change*. Oxford: Oxford University Press, 23-53.

Koopman, Hilda (1984). *The syntax of verbs: from verb movement rules in the Kru languages to universal grammar*. Dordrecht, Cinnaminson, N.J.: Foris Publications.

Kural, Murat (1997). 'Postverbal Constituents in Turkish and the Linear Correspondence Axiom', *Linguistic inquiry* 28: 498-519.

Laka, Itziar (1994). *On the syntax of negation*. New York, London: Garland.

Li, Charles N., and Thompson, Sandra A. (1974). 'An Explanation of Word Order Change SVO→SOV', *Foundations of Language* 12: 201-214.

Lin, T.-H. Jonah (2006). 'Complement-to-Specifier movement in Mandarin Chinese', Ms. National Tsing Hua University.

Mahajan, Anoop (1992). 'Specificity Condition and the CED', *Linguistic Inquiry* 23: 510-516.

Marlow, Patrick E. (1997). *Origin and development of the Indo-Aryan quotatives and complementizers: an areal approach*. Ph.D dissertation, University of Illinois.

Masica, Colin P. (1989). *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Merlan, Francesca (1989). *Mangarrayi*. London: Routledge.

Mobbs, Iain (2008). 'Functionalism', the design of the language faculty, and (disharmonic) typology', ms. Cambridge University.

Moore, Denny (1989). 'Gavião Nominalizations as Relative Clause and Sentential Complement Equivalents', *International Journal of American Linguistics* 55:309-325.

Moro, Andrea (2000). *Dynamic antisymmetry*. Cambridge, Mass.: MIT Press.

Mous, Maarten (1993). *A grammar of Iraqw*. Hamburg: H. Buske.

Müller, Gereon (1996). 'A constraint on remnant movement', *Natural Language & Linguistic Theory* 14: 355-407.

Newmeyer, Frederick J. (2003). 'Grammar Is Grammar and Usage Is Usage', *Language* 79: 682-707.

Newmeyer, Frederick J. (2005). *Possible and probable languages: a generative perspective on linguistic typology*. Oxford /New York: Oxford University Press.

Nunes, Jairo (1995). *The copy theory of movement and linearization of chains in the Minimalist Program*. Ph.D dissertation, University of Maryland.

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. Cambridge, MA..:
    MIT Press.

Ottaviano, Ida de (1980). *Textos tacana*. Riberalta, Bolivia: Instituto Linguístico de Verano :
    En colaboración con el Ministerio de Educación y Cultura.

Özsoy, A. Sumru (2001). 'On 'small' clauses, other 'bare verbal complements and feature
    checking in Turkish', in E. E. Taylan (ed.), *The verb in Turkish*. Amsterdam: John
    Benjamins, 213-237.

Paul, Waltraud (2008). 'Consistent disharmony: sentence final particles in Chinese'. Paper
    presented at *Workshop on particles*, University of Cambridge, 2008.

Pintzuk, Susan (2005). 'Arguments against a universal base: evidence from Old English',
    English Language and Linguistics 9:115-138.

Pollock, Jean-Yves (1989). 'Verb movement, universal grammar and the structure of IP',
    Linguistic Inquiry 20: 365-424.

 Richards, Marc (2004). *Object shift and scrambling in North and West Germanic: A case
    study in symmetrical syntax*. Ph.D dissertation, Cambridge University.

Richards, Marc (2008). 'Desymmetrization: parametric variation at the PF-interface', *The
    Canadian Journal of Linguistics* 53: 275-300.

Richards, Marc (2009). 'Two Kinds of Variation in a Minimalist System'. Ms: University of
    Leipzig.

Rizzi, Luigi (2001). 'On the position 'Int(errogative)' in the left periphery of the clause', in
    G. Cinque and G. Salvi (eds.), *Current studies in Italian syntax. essays offered to
    Lorenzo Renzi.* Amsterdam: Elsevier, 287–296.

Roberts, Ian (1985). 'Agreement Parameters and the Development of English Modal
    Auxiliaries', *Natural Language and Linguistic Theory* 3: 21-58.

Roberts, Ian (2007). *Diachronic syntax*. Oxford: Oxford University Press.

Roberts, Ian and Anders Holmberg. (2010). 'Introduction: Parameters in Minimalist Theory', in T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Roberts, Ian, and Roussou, Anna (2003). *Syntactic change*. Cambridge: Cambridge University Press.

Sheehan, Michelle (2009a). 'The Final-over-Final Constraint as an effect of Complement Stranding', *Newcastle Working Papers in Linguistics* 15: 104-125.

Sheehan, Michelle (2009b).' Labelling, Multiple Spell-Out and the Final-over-Final Constraint', in V. Moscati and E. Servidio (eds.), *Proceedings of XXXV Incontro di Grammatica Generativa.* Siena: STiL - Studies in Linguistics, 231-43.

Sheehan, Michelle (2010a). 'Some implications of a copy theory of labelling for the Linear Correspondence Axiom', to appear in *Syntax.*

Sheehan, Michelle (2010b). 'Extraposition and Antisymmetry', *Linguistic Variation Yearbook* 10: 203-254.

Sheehan, Michelle (in press). The resuscitation of CED. To appear in *Proceedings of NELS 40.*

Sheehan, Michelle (to appear) Adjunct placement and the Final-Over-Final Constraint. To appear in T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan, The Final-Over-Final Constraint. Cambridge, MA.: MIT Press.

Soh, Hooi Ling and Gao, Meijia (2004). 'Perfective aspect and transition in Mandarin Chinese: an analysis of double le sentences', in P. Denis, E. McCready, A. Palmer, and B. Reese (eds.), *Proceedings of the 2004 Texas Linguistics Society Conference: issues at the semantics-pragmatics interface*. Somerville, MA.: Cascadilla Press, 107–122.

Prefinal draft, to appear in *Theoretical approaches to disharmonic word orders*. OUP

Stowell, Tim (1981). *Origins of phrase structure.* Ph.D dissertation, Massachusetts Institute
    of Technology.

Svenonius, Peter (1994). 'C-selection as feature checking'*Studia Linguistica* 48: 133-155.

Travis, Lisa (1984). *Parameters and Effects of Word Order Variation*. Ph.D dissertation,
    Massachusetts Institute of Technology.

Uriagereka, Juan (1999a). 'Multiple Spell-Out', in N. Hornstein and S. D Epstein (eds.),
    Working minimalism. Cambridge, MA.: MIT Press, 251-282.

Uriagereka, Juan (1999b). 'Minimal Restrictions on Basque Movements', *Natural Language
    & Linguistic Theory* 17: 403-444.

Vennemann, Theo (1974). 'Topics, subjects, and word order: from SXV to SVX via TVX', in
    J. M. Anderson (ed.), *Historical Linguistics: Proceedings of the First International
    Congress of Historical Linguistics*, *Edinburgh, 2nd-7th September 1973* (vol. II).
    Amsterdam: North-Holland, 339–376.

Wagers, Matthew and Colin Phillips. (2009). 'Multiple dependencies and the role of the
    grammar in real-time comprehension', *Journal of Linguistics* 45: 395–433

Walkden, George (2009). 'Deriving the Final-over-Final Constraint from third factor
    considerations', *Cambridge Occasional Papers in Linguistics* 5: 67-72.

Webelhuth, Gert (1992). *Principles and Parameters of Syntactic Saturation*. New
    York/Oxford: Oxford University Press.

Whitman, John (2008). 'The Classification of constituent order generalizations and
    diachronic explanation', in J. Good (ed.), *Linguistic Universals and Language
    Change*. Oxford : Oxford University Press, 233-253.

Zwart, Jan Wouter (1997). *Morphosyntax of verb movement: a minimalist approach to the
    syntax of Dutch*: Studies in natural language and linguistic theory 39. Dordrecht;
    London: Kluwer Academic Publishers.