

J Philip. 'Harmonic Word Order Constraints are Not Created Equal.'
5th Newcastle Postgraduate Conference in Linguistics, 23rd March 2010.

Harmonic Word Order Constraints are Not Created Equal: The Final-Over-Final Constraint as an Epiphenomenon*

Joy Philip
j.philip@ucl.ac.uk

UCL

1 Introduction

We are indebted to the **Final-over-Final Constraint** (FOFC, Holmberg 2000; Biberauer, Holmberg & Roberts 2007, 2008a, 2008b, 2009, 2010; cf. Hawkins 1994; Julien 2002) for highlighting an unusual phenomenon in linearisation: the **absence** of certain **disharmonic** word orders.

FOFC is a descriptive observation generalising over one such absence, whereby **a head-initial phrase cannot be dominated by a categorially non-distinct head-final phrase**:

(1) *The Final-over-Final Constraint*

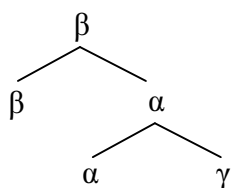
If α is a head-initial phrase and β is a phrase immediately dominating α , then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α , then β can be head-initial or head-final, where:

- (i) α and β are in the same Extended Projection [categorially non-distinct, and α P is a complement to β]¹
- (ii) α P has not been A'-moved to Spec β P. (Biberauer, Holmberg & Roberts 2010:53, ex 1''')

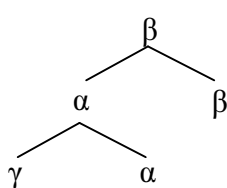
(2)

Harmonic orders

a) Initial-over-initial

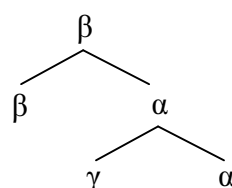


b) Final-over-final

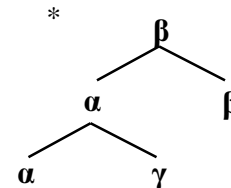


Disharmonic orders

c) Initial-over-final



d) *Final-over-initial



However, there is some confusion as to whether FOFC is an **invariant principle**, or simply a **non-absolute trend**.

In this talk I make the following claims:

- FOFC **only** acts as a **robust principle** in regard to the distribution of **subordinating complementisers**.
- The **distribution of subordinating complementisers** can be derived **independently** of FOFC.
- **Elsewhere**, there is no absolute evidence for FOFC: it is at best a **trend**.

I will propose an **alternative** theory, whereby the presence or absence of disharmony is determined not by syntactic configuration, but by the presence or absence of certain **semantic properties** on a head. I will furthermore suggest that this alternative captures a **wider range of data**.

* Many thanks to my supervisor Ad Neeleman, for stimulating discussion, comments and encouragement, also to the audience at UCL PhD Day, 2nd December 2009. My thanks are also due to Hadja Habi Sali and Hamza Tidjani for Lagwan judgments. Further comments are welcome.

¹ Note that Biberauer *et al*'s definition of Extended Projection differs from Grimshaw's (1991, 2000).

2 Subordinating complementiser distribution

2.1 Complementiser distribution and FOFC

The most **robust** evidence for FOFC is found in certain **left-right asymmetries** in the distribution of **subordinating complementisers** (henceforth simply 'complementiser' or C).

FOFC makes the following predictions:

> **VO** languages allow only **clause-initial Cs**;

OV languages allow both **clause-initial Cs** and **clause-final Cs**.

(*inter alia* Biberauer, Holmberg & Roberts 2007, 2008a, 2008b, 2009, 2010;
Biberauer, Newton & Sheehan 2009)

Where the CP is complement to a verb (and therefore categorially alike), FOFC predicts that a C-initial complement cannot be dominated by a final verb (Sheehan 2008):

> **C-initial** complement clauses must be **postverbal**.

These predictions are **borne out** in the attested distribution of Cs:

(3) *VO languages:* *OV languages:*

V[CVO]	[OVC]V
*V[VOC]	*[COV]V
*[CVO]V	*V[OVC]
*[VOC]V	V[COV]

(cf. Kuno 1974; Grosu and Thompson 1977; Dryer 1980, 1992, 2009; Hawkins 1988, 1990, 1994;
Bayer 1996, 1997, 1999; Kayne 2000; Cinque 2005)

That is, while OV languages allow head-initial Cs, C-initial complement clauses cannot appear in canonical object position in such languages.

OV languages showing the V[COV] pattern include the Indo-Iranian languages Kudmali (or Kurmali), Maithili, Punjabi, Sindhi, Hindi-Urdu, Kashmiri, Pashto, Wakhi, Persian (or Farsi), Tajik and Zazaki, the West Germanic languages Afrikaans, Dutch and German, Latin (Italic), Hittite (Anatolian), Sorbian (Balto-Slavic), the Cushitic languages Iraqw and Somali, Neo-Aramaic (Semitic), the North Dravidian languages Brahui (or Brahvi) and Malto, the South Caucasian languages Georgian and Svan, the Atlantic-Congo languages Tunen and Sare², the Malayo-Polynesian languages Gapapaiwa and Tawala, Anywa (or Anuak, Eastern Sudanic), Djapu (Pama-Nyungan), Mangarrayi (Gunwingguan), Mauka (or Mahou, Central-Southwestern Mande), Pari (Munduruku), Pima Bajo (Southern Uto-Aztecan), Teribe (Chibchan), Tsova-Tush (or Bats, Northeast Caucasian), Turkish (Turkic), Wappo (Yuki-Wappo) and Yaqui (Taracahitic).

(Dryer 1980, 2009; Bayer 2001; Cinque 2005; Davison 2007; Sheehan 2008;
Biberauer, Holmberg & Roberts 2010)

² Dryer (2009) also includes the Atlantic-Congo language Supyire in his list. However, an earlier example shows that in Supyire the rightwards extraposed clause is doubled by an overt pronoun in (preverbal) object position. Since the extraposed clause itself is not the complement of the verb, this example is irrelevant. Since Dryer does not give examples from all the other languages he lists, it is possible that some of these should also be excluded on the same grounds.

- (4) a) Er hatte gewusst [*dass* er nicht lange leb-en würde]. *German*
 he had known COMP he not long live-INF would³
 'He had known that he would not live long.'
- b) * Er hatte [*dass* er nicht lange leb-en würde] gewusst
 he had COMP he not long live-INF would known
(Hawkins 1994:302, ex 5.43)
- (5) a) An zan mi danat [*ke* an mard sangi partab kard]. *Persian*
 that woman CONT knows COMP that man rock threw
 'The woman knows that the man threw a rock.'
- b) * An zan mi [*ke* an mard sangi partab kard] danat
 that woman CONT COMP that man rock threw knows
(Dryer 1980:130, ex 15-16)
- (6) a) aapo hunen hia [*ke* hu hamut tutu?uli]. *Yaqui*
 he thus say COMP this woman pretty
 'He says that this woman is pretty.'
- b) * aapo hunen [*ke* hu hamut tutu?uli] hia
 he thus COMP this woman pretty say (p131, citing Lindenfeld 1973)
- (7) a) Adam ban=a söyle-di-ø [*ki* Aysekitab=i oku-du-ø]. *Turkish*
 man me=DAT tell-PST-3SG COMP Aysebook=ACC.DEF read-PST-3SG
 'A man told me that Ayse read the book.' (ex 20)
- b) * Adam ban=a [*ki* Aysekitab=i oku-du-ø] söyle-di-ø
 man me=DAT COMP Aysebook=ACC.DEF read-PST-3SG tell-PST-3SG (ex 21)

In Turkish, extraposition only occurs where there is a complementiser:

- (8) Adam ban=a [Ayse=nin kitab=i oku-duğ-u=nu] söyle-di-ø. *Turkish*
 man me=DAT Ayse=GEN book=ACC.DEF read-NOM-3SG.POSS=ACC tell-PST-3SG
 'The man told me that Ayse read the book/of Ayse's reading the book.' (Dryer 1980:131, ex 19)
- (9) Herkes [ben=i Ankara=ya git-ti] san-ıyor.
 everyone me=ACC.DEF Ankara=DAT go-PST consider-PROG
 'Everyone considers me to have gone to Ankara.' (Özsoy 2001:217, ex 5a)

There is evidence that these **C-initial** complements are **base-generated in postverbal** position, since at least in Dutch, German, Hindi-Urdu, Persian and Turkish they are **not islands for extraction**.⁴

(see Bennis 1987; Karimi 2001; Aghaei 2006; Biberauer, Newton & Sheehan 2009 and references cited in these works)

³ Abbreviations in glosses are as follows: ACC accusative; APPL applicative; ASP aspect; ASS assertive; AUX auxiliary; CL classifier; COMP complementiser; CONT continuous; DAT dative; DECL declarative; DEF definite; DEM demonstrative; DEP dependent; DU dual; FUT future; GEN genitive; IMMED immediate; IMP imperative; IMPF imperfective; INCOMPL incomplete; INDEF indefinite; INF infinitive; LOC locative; M masculine; n- n-word; N nominaliser; NEG negative; NOM nominative; OBJ object; PASS passive; PERF perfective; PL plural; POSS possessive; POT potential; PRES present; PROG progressive; PST past; Q interrogative; QUOT quotative; SBJ subject; SG singular; TOP topic

⁴ On the other hand, where extraposed clauses in these languages are doubled by a pronoun, and are therefore adjuncts, extraction is impossible (Karimi 2001; Aghaei 2006; Biberauer, Newton & Sheehan 2009 and references cited in these works).

- (10) [Un ketab-a=ro]_i mæn mi-dun-æm [ke Kimea t_i xær-id-e]. *Persian*
 that book-PL=ACC I IMPF-know-1SG COMP Kimea buy-PERF-3SG
 'As for those books, I know that Kimea has bought (them).' (Karimi 2001, ex 69)

Although it is therefore possible for OV languages to have base-generated postverbal complement, this option is **not** available where there is a **final C: *V[OVC]**

This is particularly striking in certain OV languages allowing **both initial and final Cs**: mainly Indo-Aryan languages with close geographical or historical contact with Dravidian, such as Assamese, Bengali (or Bangla), Oriya, Marathi, Gujarati and Nepali, and also the Turkic language Uzbek and the Atlantic-Congo language Vata (or Dida)⁵. (Bayer 1996, 1997, 1999, 2000, 2001; Cinque 2005; Davison 2007)

C-initial clauses are obligatorily **postverbal**; **C-final** clauses are obligatorily **preverbal**:

- (11) a) chele-ta [or baba aS-be (bole)] Sune-che. *Bengali*
 boy-CL his father come-FUT.3 COMP hear-PST.3
 b) chele-ta Sune-che [or baba aS-be (*?bole)].
 boy-CL hear-PST.3 his father come-FUT.3 COMP
 'The boy has heard that his father will come.' (Bayer 1996:255, ex 9)

- (12) a) chele-ta [(?je) or baba aS-be] Sune-che
 boy-CL COMP his father come-FUT.3 hear-PST.3
 b) chele-ta Sune-che [(je) or baba aS-be].
 boy-CL hear-PST.3 COMP his father come-FUT.3
 'The boy has heard that his father will come.' (ex 11)

Evidence suggests that the C-initial clause is **base-generated** in postverbal position:

- The extraposed clause is **c-commanded** by the **indirect object**:

- (13) ami [prottek-Ta chele-ke]_i bole-chi [je Ek-jon ta-ke_i durga pujo-Y notun
 I each-CL boy-OBJ say-PST.1 COMP one-CL he-OBJ Durga Puja-LOC new
 jama kapoR de-be].
 clothes give-FUT.3
 'I told [each boy]_i that someone will give him_i new clothes at the festival of Durga Puja.'
 (Bayer 2000:2, ex 5)

- The extraposed clause is **not** an **island** for extraction:

- (14) kriSno mEleria-te_i bhab-che [je ram t_i mara gE-che]. *Bengali*
 Krishna malaria-LOC think-PERF.3COMP Ram die go-PERF.3
 'Krishna thinks that Ram died of malaria.' (Simpson & Bhattacharya 2000, ex 13)

⁵ It is debatable whether Vata genuinely has an initial complementiser. Koopman (1984) argues that only the final subordinator *kā* is a true complementiser.

2.2 *Complementiser distribution and Head Proximity*

(3) *VO languages:* *OV languages:*

V[CVO]	[OVC]V
*V[VOC]	*[COV]V
*[CVO]V	*V[OVC]
*[VOC]V	V[COV]

The data in (3) can be summed up by the following two observations:

- **Final Cs** are only permitted in **OV** languages.
- Subordinating **Cs** must be **base-generated** such that they **intervene linearly** between their selecting verb and its complement clause.

Here I propose that both observations can be captured by the **interaction** of three **independently-motivated harmonic word order constraints**:

> **Head Proximity**

(adapted from Rijkhoff 1984, 1986, 1990, 1992; cf. Lehmann 1973; Surface Recursion Restriction, Emonds 1976, 1985; Head-Final Filter, Williams 1982; Early Immediate Constituents, Hawkins 1990, 1994)

(15) *Principle of Head Proximity*

The head of a [subordinate] domain prefers to be contiguous with the head of its superordinate domain. (Rijkhoff 1986:5)

For example, Head Proximity accounts for the cross-linguistic tendency to avoid the placement of dependents of a noun between this noun and its selecting verb.

(cf. Greenberg 1963; Lehmann 1973; Hawkins 1983)

In (3), **C**, as head of the dependent clause, must be positioned such that it is **adjacent** to its **selecting head**, the matrix verb.

> **Head Uniformity** (cf. Natural Serialisation Principle, Bartsch and Venneman 1972, Vennemann 1974; Cross-Categorical Harmony, Hawkins 1980, 1983; Head Parameter, *inter alia* Chomsky 1981; Branching Direction Theory, Dryer 1992, 2009; Principle of Cross-Domain Harmony, Dik 1997)

This refers to the preference for heads in a given language/domain to be uniformly head-initial/-final (cf. harmonic orders in (2)a) and b)).

As regards Cs, there should therefore be a preference for **initial Cs** in **VO** languages and for **final Cs** in **OV** languages.

> **CP-Final** (cf. Sentential NP Position Hierarchy, Dryer 1980; Language Independent Preferred Order of Constituents, Dik 1997)

Dryer (1980) showed that there is a cross-linguistic preference for **clausal arguments** to appear in **sentence-final** position.

Where these three constraints compete, **Head Proximity** takes precedence:

(16) *Harmonic Word Order Ranking*

Head Proximity > Head Uniformity, CP-Final

In **VO** languages, the constraints do not compete, resulting in a **single optimal order**:

(17)

	<i>Head Proximity</i>	<i>Head Uniformity</i>	<i>CP-Final</i>
V[CVO]			
*[VOC]V		*!	*!
*[CVO]V	*!		*
*V[VOC]	*!	*	

In **OV** languages, no single order obeys all three constraints.

In order to obey the **dominant** constraint, **Head Proximity**, either the **Head Parameter** or **Final-over-elsewhere** must be **violated**, resulting in **two possible orders**:

(18)

	<i>Head Proximity</i>	<i>Head Uniformity</i>	<i>CP-Final</i>
[OVC]V			*
V[COV]		*	
*V[OVC]	*!		
*[COV]V	*!	*	*

3 FOFC elsewhere: trend rather than principle

We have seen that the **distribution of Cs** (in (3)) can be derived **independently** of FOFC, by the interaction of **pre-existing** word order constraints.

However, this result will be irrelevant if FOFC is required elsewhere.

For **any category other than C**, there is evidence suggesting that FOFC **does not hold** as an **absolute** principle: all four logically possible orders in (2), both harmonic and disharmonic, appear to be attested:

(Note that Dryer's (1992) data only uses **auxiliaries** that are '**specifically verbal**' and **negative auxiliaries** that '**exhibit verbal properties**')

- (19) a) Initial-over-initial: [P [N PossP]] = 134 languages (40%)
 b) Final-over-final: [[PossP N] P] = 177 languages (53%)
 c) Initial-over-final: [P [PossP N]] = 14 languages (4%) (7% of N-final lgs)
 d) Final-over-initial: [[N PossP] P] = 11 languages (3%) (8% of N-initial lgs)

(Hawkins 2010:1, using data from Hawkins 1983)

- (20) a) Initial-over-initial: [Aux [V O]] = 28 genera (39%)
 b) Final-over-final: [[O V] Aux] = 36 genera (51%)
 c) Initial-over-final: [Aux [O V]] = 3 genera (4%) (8% of OV genera)
 d) Final-over-initial: [[V O] Aux]] = 4 genera (6%) (13% of VO genera)⁶

(Data taken from Dryer 1992:100, table 28)

⁶ Dryer documents only the relative order of verb and object and of verb and auxiliary (i.e. not the relative order of object and auxiliary). However, since Julien (2002:235) states that the order V T/Asp O (where T/Asp is a free-standing morpheme) does not occur, it is safe to conclude that for the four genera exhibiting VO and VAux, the auxiliary follows the object.

- (21) a) Initial-over-initial: [Neg [V O]] = 13 genera (52%)
 b) Final-over-final: [[O V] Neg] = 8 genera (32%)
 c) Initial-over-final: [Neg [O V]] = 3 genera (12%) (27% of OV genera)
 d) Final-over-initial: [[V O] Neg]] = 1 genus (4%) (7% of VO genera)
 (Data taken from Dryer 1992:101, table 29)

- (22) a) Initial-over-initial: [Q [V O]] = 28 genera (30%)
 b) Final-over-final: [[O V] Q] = 32 genera (35%)
 c) Initial-over-final: [Q [O V]] = 13 genera (14%) (29% of OV genera)
 d) Final-over-initial: [[V O] Q]] = 19 genera (21%) (40% of VO genera)⁷
 (Data taken from Dryer 1992:102, table 30)

> Biberauer, Holmberg & Roberts (2007, 2008a, 2008b, 2009, 2010) claim that FOFC applies only to **inflected** auxiliaries (as opposed to uninflected T/Asp markers, which may not be heads).

- (23) yə=ca dəyo lə. *Bwe Karen*
 1SG=see picture ASP
 'I'm looking at a picture.'
 (Dryer 2008:20, ex 24, citing Henderson 1997:39)

- (24) cə=də [mɪ jə=khó phi má nə]?
 3=say COMP 3=FUT take what
 'What did he say he would take?'
 (p21, ex 29, citing Henderson 1997:187)

However, the evidence suggests that the alleged 'agreement inflection' is in fact a **proclitic pronoun**, since it appears in **complementary distribution with full NPs**, in **canonical subject position**:

- (25) jɛ ní dðkhí tə-dó *Bwe Karen*
 trap catch barking.deer one-CL
 'the trap catches a barking deer'
 (Dryer 2008:14, ex 2a, citing Henderson 1997:258)

- (26) yə=bðdá [mɪ yə=cɛ ɓe-nu lémè thó].
 1SG-think COMP 1SG-book CL-that lost PERF
 'I thought that my book was lost'
 (p21, ex 26, citing Henderson 1997:379)

Moreover, Matthew Dryer (p.c.) notes two Central Sudanic languages in his database with **verbal auxiliaries** with the order **VOAux**: Mbay and Ngambay.

> As regards **negative** and **interrogative** markers, Biberauer, Holmberg & Roberts (2010:§5) claim that they are **syncategorematic**: they are **not involved in c-selection** and are therefore **outside the extended projection**, hence **outside the scope of FOFC**.

However, in Ma'di (Central Sudanic), the marker of **negation** also **encodes tense**, a **verbal** feature, suggesting that it is indeed part of the **extended projection** of the **verb**:

- (27) m'-āwí dʒótī kū. *Ma'di*
 1SG-open door NEG.NONPST
 'I won't/am not opening the door/don't open doors.'
 (Blackings & Fabb 2003:14, ex 8)

⁷ Bailey *et al* (2010:13) and Biberauer, Holmberg and Roberts (2010:81) point out that some final interrogative particles may in fact be initial disjunctive elements with an elided second conjunct. The numbers for orders b) and d) in (22) may therefore be misleadingly high.

- (28) m'-āwí dʒótī kūrù.
 1SG-open door NEG.PST
 'I did not open the door.' (ex 7)

Notice that in the above examples, the lexical verb is unmarked for tense. However, in the absence of negation, the same past/nonpast distinction is marked on the lexical verb:

- (29) ká gbándà òǎ.
 3SG cassava NONPST.eat
 'He is eating/eats cassava.' (Blackings & Fabb 2003:13, ex 1)

Even where a **final particle** is **uninflected**, there may be evidence not only that the particle is a **head dominating the verb**, and that this head is involved in **c-selection**:

Lagwan (Chadic) exhibits the order TVONeg:

- (30) Sà-dî gír kàskú dīyásín sá.
 FUT-3F go market tomorrow NEG
 'She won't go to the market tomorrow.' *Lagwan*

There is evidence for the projection of **NegP**, since the negative marker is required to **license n-words**:

- (31) Bìl=á shímá á ló *(sá).
 man=LNK.M n-M 3MSG come NEG
 'Nobody came.' *Lagwan*

There is evidence that **final Neg dominates** both **initial V and T**, since there is **no true negative imperative**; a surrogate (including T) is required (cf. Zanuttini 1994, 1996). This shows that **Neg c-selects T**:

(See also (30) above, where Neg follows (and therefore dominates) a TP-adjoined adverbial)

- (32) Slà á !
 push.IMP up
 'Get up!' *Lagwan*

- (33) a) * Slà á sá !
 push.IMP up NEG
 b) Sà-gì slá á sá !
 FUT-2SG push up NEG
 'Don't get up!'

Finally, there is evidence from the 'Why not?' test (Merchant 2001) that the negative marker in Lagwan is indeed a **head**, since it cannot adjoin to a phrasal category:

- (34) a) * Àgé ghwàni sá ?
 because what NEG
 b) * Sá àgé ghwàni ?
 NEG because what *Lagwan*

4 (Dis)harmony and semantics

We have seen evidence suggesting then that disharmony is possible for any category except C, which always obeys the Harmonic Word Order Ranking in (16).

So, why do other categories not always obey this Harmonic Word Order Ranking? That is, **why** does **disharmony** arise?

I propose the following explanation:

- **Linearisation rules** pertaining to the **semantics** of a head can require it to appear in a **prominent position**, either initial or final.
- Where such rules **conflict** with, and **override**, the Harmonic Word Order Ranking in (16), **disharmony** arises.

For example, disharmony is relatively common for negative markers. The expression of negation is highly significant semantically. Cross-linguistically, there is a tendency to place negative markers in one of two prominent positions: initially, so that negation will be expressed as soon as possible (Jespersen 1917:4, 1924:297); or finally, the position reserved for new or significant information (Mazzon 2004:97). Where the choice of prominent position differs from the headedness of the verb, disharmony arises (cf. (21)c) and d)).

However, **subordinating C** is **impervious** to such **linearisation rules**:

- **Subordinating C** doesn't contribute to the compositional semantics of its extended projection, but rather serves to **mark a relationship** between two extended projections; it is **semantically vacuous**.
- Disharmonic orders arise when linearisation rules require a head with specified semantics to appear in a certain (prominent) position; however, since C doesn't have relevant semantics, the effects of **Head Proximity cannot be overridden** by such linearisation rules.

This makes two predictions:

- > For any **C** that is not a subordinator, but rather **contributes to the compositional semantics**, **disharmony** should be possible.

While subordinating C is consistently harmonic (see section 2), interrogative markers and other discourse C-particles are frequent violators of both FOFC and its inverse (cf. (22), Julien 2002; Paul 2009, Biberauer, Holmberg & Roberts 2010):

(35) Ni yao kan zhe-ben shu ma? *Mandarin Chinese*
you want read this-CL bookQ
'Do you want to read this book?'

(Biberauer, Holmberg & Roberts 2010:53, ex 77a, citing Aldridge 2009)

- > Any other **semantically vacuous** head, serving only to **mark a relationship** should also always **obey Head Proximity** in its base-generated position.

Relative clause markers and linkers in the noun phrase, that serve only to mark the relationship between the head noun and its dependent, **always intervene linearly** between noun and dependent:

- (36) a) Initial-over-initial: [N [REL TP]] = 56 languages (95%)
b) Final-over-final: [[TP REL] N] = 3 languages (5%)
c) *Initial-over-final: [N [TP REL]] = 0 languages (0%)
d) *Final-over-initial: [[REL TP] N]] = 0 languages (0%)

(Data taken from De Vries 2002:386-384, table 2)

- (37) a) Initial-over-initial: [N [LNK XP]] = 49 languages (66%)
b) Final-over-final: [[XP LNK] N] = 25 languages (34%)

c) *Initial-over-final: [N [XP LNK]] = 0 languages (0%)

d) *Final-over-initial: [[LNK XP] N]] = 0 languages (0%)

(see Appendix B for languages and classification)

Here we find a complete **absence** of **both disharmonic orders** (cf. (2)). Note that **neither absence** falls under the descriptive/explanatory **scope of FOFC**.

Moreover, it may be necessary to **violate FOFC** in order to **obey Head Proximity** (note that linkers always form a constituent with the dependent of the noun, Philip 2009):

- (38) [zuotian chi yurou de] ren *Mandarin Chinese*
yesterday eat fish LNK person
'The people who ate fish yesterday' (Paul 2009:4, ex 8a)

5 Summary and conclusion

- **Subordinating complementisers** (and other **semantically vacuous subordinators**) always obey **optimal** word order principles, governed by the **interaction** between **Head Proximity**, **Head Uniformity** and **CP-Final**, with **Head Proximity** taking **precedence**. This allows us to capture a wider range of absent disharmonic word orders than FOFC.
- This is because **linearisation rules** pertaining to the **semantics** of a head **cannot apply** to **heads lacking semantics**.
- **Elsewhere**, such **linearisation rules** are permitted to **override** harmonic word order principles, giving rise to **disharmony**, including FOFC violations.

Appendix A: Superficial counterexamples

The literature cites various (superficial) counterexamples to the observations in (3).

These fall under one or both of two categories:

- The complement clause is (embedded under a) **nominal**, therefore not a direct complement of V.
- The postverbal complement clause is not base-generated there as a complement of the verb, but appears in an **adjoined** position (either by movement, or it is an adjunct associated with a null nominal).

> Harar Oromo (or Galla, Cushitic), allegedly [COV]V:

(Cinque 2005, citing Julien 2001; Dryer 2007, 2009)

- (39) inní [akka deem-u] good'-ám-é. *Harar Oromo*
he that go-DEP order-PASS-PST
'He was ordered to go.' (Owens 1985:145, ex 49)

Owens (1985) refers to the embedded clause headed by *akka* as a '**noun clause**'. Where an *akka*-clause expresses purpose, it is optionally marked with the **dative case**:

- (40) [akka na árka-aní-f]d'uf-an. *Harar Oromo*
that me see-PL-DAT came-PL
'They came to see me.' (Owens 1985:146, ex 54)

Elsewhere, the lack of case-marking is expected, since ordinary direct objects appear in the **morphologically unmarked absolutive case**.

Moreover, the **semantics** of *akka* suggest that it is not a subordinating complementiser:

'according as, just as, like, how, manner, way, (in order) to/that, (the fact) that'

(Hodson & Walker 1922; Owens 1985)

> Lakhota (or Lakota, Siouan), allegedly V[OVC] (Dryer 1980):

- (41) [Tohá slolyáya he] [wakpála ektá ohĩhpaye kǐ]. *Lakhota*
when you.know Q creek to fall the
'When did you find out that he fell in the creek?' (Rood 1973:73, ex 8)

kǐ, which Rood (1973) terms a **nominaliser**, is identical to the **definite determiner**, which also appears in final position.

It appears in **complementary distribution** with an **indefinite** marker:

- (42) a) [Tuwá omákiyi kte kǐ] slolwáye. *Lakhota*
who help.me POT the I.know
'I know who will help me.'
- b) [Tuwá omákiyi kte čha] slolwáye.
who help.me POT INDEF I.know
'I know somebody will help me.' (Rood 1973:72, ex 6)

For verbs such as 'say', 'want', 'tell', 'be able', 'be supposed to' a bare clause appears as complement. *kǐ* (definite) or *čha* (indefinite) are only required with certain verbs. This suggests that the latter set of verbs select **nominal**, rather than clausal, complements.

In any case, the **extraposition** in (41) is **optional**.

Moreover, the position of the complement following the interrogative marker *he* suggests that it is **adjoined externally to the matrix clause**, rather than base-generated as a complement.

> Dhivehi (or Maldivian, Indo-Iranian), allegedly V[OVC] (Cinque 2005):

The complement clause is marked by **dative** or **locative case**:

- (43) ahannaš lafā kurevenī [hama jessēne kam-aš]. *Dhivehi*
me.DAT guess do.IN.PREFOC even touch.FUTthing-DAT
'I guess (the event) that it will even out.' (Cain & Gair 2000:48, ex 163)

> Ngiti (Central Sudanic), allegedly V[OVC] (Cinque 2005):

The complement clause is embedded under a **noun**:

- (44) k=ùni [ma m-í-rà dhu]. *Ngiti*
3SG=know.PERF.PRES I 1-AUX-come.N thing
'He knows (the fact) that I am coming.' (Kutsch Lojenga 1994:395)

> Khoekhoe (or Nama, Khoe), allegedly V[OVC] (Dryer 2009):

The complement clause is marked by **objective case**:

- (45) tsĩ // *ĩ-p-à-kxm̄ ke kè mĩ-pa [!úũ-kxm̄ ta !xáís=à]. *Khoekhoe*
and 3-3MSG-OBJ-1M.DU.SBJ DECL PST say-APPL go-1M.DU.SBJ COMP=OBJ
'And we told him that we were going.' (Hagman 1977:138, gloss Güldemann 2006, ex 20)

Appendix B: Languages with linkers in the noun phrase:

<i>Classification</i>	<i>No. of lgs. in sample</i>	<i>Language</i>	<i>Position of linker</i>
Afro-Asiatic	10		
Chadic	(9)		
Biu-Mandara	(8)		
A	(1)	Gude	Postnominal
B	(7)		
Kotoko-Yedina			
Kotoko	(5)	Afade	Postnominal
		Goulfey	Postnominal
		Lagwan	Postnominal
		Mpade	Postnominal
		Mser	Postnominal
Zina	(2)	Mazera	Postnominal
		Zina	Postnominal
West Chadic	(1)	Nyam	Postnominal
East Cushitic	(1)	Dasenech	Prenominal
Austronesian	7		
Malayo-Polynesian	(6)		
Central Philippine	(1)	Tagalog	Both
Oceanic	(5)		
Central-Eastern Oceanic	(2)		
Central Pacific	(1)	Rotuman	Postnominal
Micronesian	(1)	Kiribati	Postnominal
Meso-Melanesian	(1)	Bali-Vitu	Postnominal
Polynesian	(1)	Samoan	Postnominal
Southern Oceanic	(1)	Malo	Postnominal
Nuclear Malayo-Polynesian	(1)	Palauan	Both
Tsouic	(1)	Tsou	Prenominal
Basque	1	Basque	Prenominal
Creole languages	2		
Dutch Creole	(1)	Berbice Dutch Creole	Postnominal
English Creole	(1)	Tok Pisin	Postnominal
Indo-European	18		
Albanian	(1)	Albanian	Postnominal
Indo-Iranian	(11)		
Indo-Aryan	(2)	Hindi	Prenominal
		Urdu	Both
Western Iranian	(9)		
Northwestern Iranian	(7)		
Caspian	(2)	Gilaki	Both
		Mazandarani	Both
Kurdish	(5)	Balochi	Postnominal
		Hawrami	Postnominal
		Kurmanji	Postnominal
		Sorani	Postnominal
		Zazaki	Postnominal
Southwestern Iranian	(2)	Persian	Postnominal
		Tajik	Postnominal
Italic	(3)		
Romance			
East Romance	(1)	Romanian	Postnominal
Italo-Western	(2)		
Italo-Dalmation	(1)	Italian	Postnominal
Western	(1)	French	Postnominal

West Germanic	(3)		
Anglo-Frisian	(1)	English	Postnominal
High German	(1)	German	Postnominal
Low Franconian	(1)	Dutch	Postnominal
Japonic	1	Japanese	Prenominal
Korean	1	Korean	Prenominal
Kwadi-Khoe	1		
Khoe	(1)	Khoekhoe	Prenominal
Mayan	1		
Cholan-Tzeltalan	(1)	Tzeltal	Prenominal
Niger-Congo	10		
Atlantic-Congo	(9)		
Benue-Congo	(8)		
Bantoid			
Central Bantu			
Zone D	(1)	Kilega	Postnominal
Zone E	(2)		
Kikuyu-Kamba		Gikuyu	Postnominal
		Kiitharaka	Postnominal
Zone G	(1)	Swahili	Postnominal
Zone J	(2)		
Haya-Jita	(1)	Haya	Postnominal
Konzo	(1)	Kinande	Postnominal
Zone N	(1)	Chichewa	Postnominal
Zone S	(1)	Shona	Postnominal
Senegal-Guinea	(1)	Wolof	Postnominal
Western Mande	(1)	Bambara	Prenominal
Nilo-Saharan	4		
Central Sudanic	(1)	Lendu	Prenominal
East Sudanic	(2)		
Western Nilotic			
Southern Luo			
Luo-Acholi		Dholuo	Postnominal
		Lango	Postnominal
Songhay	(1)	Koyra Chiini	Prenominal
Penutian	1	Tsimshian	Postnominal
Sino-Tibetan	8		
Sinitic	(3)		
Chinese		Cantonese	Prenominal
		Mandarin	Prenominal
		Taiwanese	Prenominal
Tibeto-Burman	(5)		
Himalayish	(2)		
Mahakiranti	(1)	Newari	Prenominal
Tibeto-Kanauri	(1)	Byansi	Prenominal
Lolo-Burmese	(2)		
Burmish	(1)	Burmese	Prenominal
Loloish	(1)	Lahu	Prenominal
Northeast Tibeto-Burman	(1)	Bai	Prenominal
Tai-Kadai	1		
Tai		Thai	Postnominal
Trans-New Guinea	1		
Madang		Amele	Prenominal

References

- Aghaei, Behrad. 2006. 'The Syntax of *ke*-clause and Clausal Extraposition in Modern Persian.' Ph.D. dissertation, University of Texas.
- Aldridge, E. 2009. 'Neg-to-Q: The Historical Development of one Clause-Final Particle. To appear in Theresa Biberauer (ed.), *Cambridge Occasional Papers in Linguistics* (Special edition on particles).
- Bailey, Laura, Theresa Biberauer, Anders Holmberg & Michelle Sheehan. 2010. 'The Particle Problem.' Handout from workshop on The Final-Over-Final Constraint: Processing, Typology and Acquisition, Cambridge, 29th January 2010. (Available from *Research – Newcastle University*, [<http://research.ncl.ac.uk/linearization/Particle%20problem%20handout%20final.pdf>], accessed 16/02/10)
- Bartsch, Renate & Theo Venneman. 1972. *Semantic Structures: A Study in the Relation Between Semantics and Syntax*. Frankfurt am Main: Athenäum.
- Bayer, Josef. 1996. *Directionality and Logical Form: On the Scope of Focussing Particles and Wh-in-situ*. Dordrecht: Kluwer.
- Bayer, Josef. 1997. 'CP-Extraposition as Argument Shift.' Beerman *et al* (eds.), 37-58.
- Bayer, Josef. 1999. 'Final Complementizers in Hybrid Languages.' *Journal of Linguistics* 35, 233-271.
- Bayer, Josef. 2000. 'Symmetry and Asymmetry in South Asian Languages.' Handout from Workshop on Antisymmetry Theory, Palazzone della Scuola Normale Superior, Cortona, 15-17 May 2000. (Available from *Universität Konstanz – Fachbereich Sprachwissenschaft*, [<http://ling.uni-konstanz.de/pages/home/bayer/pdf/symmetry-handout%5B1%5D.pdf>], accessed 07/10/2009)
- Bayer, Josef. 2001. 'Two Grammars in One: Sentential Complements and Complementizers in Bengali and other South-Asian Languages.' Peri Bhaskarorao & Karumuri 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.
- Bennis, Hans. 1987. *Gaps and Dummies*. Dordrecht: Foris.
- Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2007. 'Disharmonic Word-Order systems and the Final-Over-Final-Constraint (FOFC).' Antonietta Bisetto & Francesco Barbieri (eds.), *Proceedings of XXXIII Incontro di Grammatica Generativa*, 86-105.
- Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2008a. 'Linearizing Disharmonic Word Orders: The Final-over-Final Constraint.' K.-A. Kim (ed.), *Perspectives on Linguistics in the 21st Century*. Proceedings of the ICLK conference, Seoul, January 2007, 301 - 318.
- Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2008b. 'Structure and Linearization in Disharmonic Word Orders.' Charles B. Chang & Hannah J. Haynie (eds.), *The Proceedings of 26th Western Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 96-104.
- Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2009. 'Linearization and the Architecture of Grammar: A View from the Final-Over-Final Constraint.' To appear in V. Moscati & E. Servidio (eds.), *Proceedings of the Incontro di Grammatica Generativa XXXV*. (Available from *Research – Newcastle University*, [http://research.ncl.ac.uk/linearization/BHR_IGG35_Proceedings_paper.pdf], accessed 11/01/10)
- Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2010. 'A Syntactic Universal and its Consequences.' Ms, University of Cambridge, Stellenbosch University & Newcastle University.
- Biberauer, Theresa, Glenda Newton & Michelle Sheehan. 2009. 'Limiting Synchronic and Diachronic Variation and Change: The Final-Over-Final Constraint.' *Language and Linguistics* 10, 701-743.
- Blackings, Mairi & Niger Fabb. 2003. *A Grammar of Ma'di*. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Cain, Bruce D. & James W. Gair. 2000. *Dhivehi (Maldivian)*. Munich: Lincom Europa.
- Cinque, Guglielmo. 2005. 'A Note on Verb/Object Order and Head/Relative Clause Order.' Laura Brugè (ed.), *University of Venice Working Papers in Linguistics* 15, 49-104.
- Davison, Alice. 2007. 'Word order, Parameters, and the Extended COMP Projection.' Josef Bayer, Tanmoy Bhattacharya & M. T. H. Babu (eds.), *Linguistic Theory and South Asian Languages*. Amsterdam/Philadelphia: John Benjamins, 175-198.
- Dik, Simon C. 1997. *The Theory of Functional Grammar Part 1: The Structure of the Clause*. 2nd edition. Berlin: Mouton de Gruyter.
- Dryer, Matthew S. 1992. 'The Greenbergian Word Order Correlations.' *Language* 68, 81-138.

- Dryer, Matthew S. 2007. 'Word Order.' Timothy Shopen (ed.), *Clause Structure*. Language Typology and Syntactic Description, vol. 1. Second Edition. Cambridge: Cambridge University Press, 61-131.
- Dryer, Matthew S. 2008. 'Word Order in Tibeto-Burman Languages.' *Linguistics of the Tibeto-Burman Area* 31, 1-88.
- Dryer, Matthew S. 2009. 'The Branching Direction Theory Revisited.' Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of Language Today*. Berlin: Springer, 208-230.
- Dryer, Matthew Synge. 1980. 'The Positional Tendencies of Sentential Noun Phrases in Universal Grammar.' *Canadian Journal of Linguistics* 25, 123-195.
- Emonds, Joseph. 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.
- Emonds, Joseph. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris
- Greenberg, Joseph H. 1963. 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements.' Joseph H. Greenberg (ed.), *Universals of Language*. London: MIT Press, 73-113.
- Grimshaw, Jane. 1991. 'Extended Projection.' Ms, Brandeis University.
- Grimshaw, Jane. 2000. 'Locality and Extended Projection.' Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical Specification and Insertion*. Amsterdam/Philadelphia: John Benjamins, 115-133.
- Grosu, Alexander & Sandra Thompson. 1977. 'Constraints on the Distribution of NP Clauses.' *Language* 53, 104-151.
- Güldemann, Tom. 2006. 'Structural Isoglosses between Khoekhoe and Tuu: The Cape as a Linguistic Area.' Yaron Matras, April McMahon & Nigel Vincent (eds.), *Linguistic Areas: Convergence in Historical and Typological Perspective*. Hampshire: Palgrave Macmillan, 99-134.
- Hagman, Roy S. 1977. *Nama Hottentot Grammar*. Bloomington: Indiana University.
- Hawkins, John A. 1980. 'On Implicational and Distributional Universals of Word Order.' *Journal of Linguistics* 16, 193-235.
- Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.
- Hawkins, John A. 1988. 'On Explaining some Left-Right Asymmetries in Syntactic and Morphological Universals.' Michael Hammond, Edith A. Moravcsik, and Jessica R. Wirth (ed.), *Studies in Syntactic Typology*. Amsterdam: John Benjamins, 321-57.
- Hawkins, John A. 1990. 'A Parsing Theory of Word Order Universals.' *Linguistic Inquiry* 21, 223-261.
- Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. 2010. 'FOFC from a Processing Typology Perspective.' Handout from workshop on The Final-Over-Final Constraint: Processing, Typology and Acquisition, Cambridge, 29th January 2010. (Available from *Research – Newcastle University*, [http://research.ncl.ac.uk/linearization/Hawkins_FOFC_handout.pdf], accessed 16/02/10)
- Henderson, Eugénie J. A. 1997. *Bwe Karen Dictionary: With Texts and English-Karen Word List*. Volume 2. London: School of Oriental and African Studies, University of London.
- Hodson, Arnold W. & Craven H. Walker. 1922. *A Grammar of the Galla or Oromo Language*. London: Society for Promoting Christian Knowledge.
- Holmberg, Anders. 2000. 'Deriving OV Order in Finnish.' Peter Svenonius (ed.), *The Derivation of VO and OV*. New York: Oxford University Press, 123-152.
- Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: A.F. Høst.
- Jespersen, Otto. 1924. *The Philosophy of Grammar*. London: George Allen & Urwin.
- Julien, Marit. 2001. 'Word Order Type and Syntactic Structure.' *Linguistic Variation Yearbook* 1, 17-59.
- Julien, Marit. 2002. *Syntactic Heads and Word Formation: A Study of Verbal Inflection*. New York: Oxford University Press.
- Karimi, Simin. 2001. 'Persian Complex DPs: How Mysterious are they?' *Canadian Journal of Linguistics* 46, 63-96.
- Kayne, Richard S. 2000. *Parameters and Universals*. New York: Oxford University Press.
- Koopman, Hilda. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.
- Kuno, Susumo. 1974. 'The Position of Relative Clauses and Conjunctions.' *Linguistic Inquiry* 5, 117-36.
- Kutsch Lojenga, Constance. 1994. *Ngiti: A Central Sudanic Language of Zaire*. Cologne: Rüdiger Köppe.
- Lehmann, W. P. 1973. 'A Structural Principle of Language and its Implications.' *Language* 49, 47-66.

- Lindenfeld, Jacquelin. 1973. *Yaqui Syntax*. University of California Publications in Linguistics, No. 76. Berkeley: University of California Press.
- Mazzon, Gabriella. 2004. *A History of English Negation*. Harlow: Pearson Longman.
- Merchant, J. 2001. 'Why no(t).' Ms, University of Chicago.
- Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia): Including a Text and a Glossary*. Hamburg: Helmut Buske.
- Özsoy, A. Sumru. 2001. 'On "Small" Clauses, Other Bare Verbal Complements and Feature Checking in Turkish.' Eser Erguvanlı Taylan (ed.), *The Verb in Turkish*. Amsterdam: John Benjamins, 213–237.
- Paul, Waltraud. 2009. 'Consistent Disharmony: Sentence-Final Particles in Chinese.' To appear in Theresa Biberauer (ed.), *Cambridge Occasional Papers in Linguistics* (Special edition on particles).
- Philip, Joy. 2009. 'Restrictions on the Morphosyntactic Marking of Dependencies: What Linkers Tell us.' Ms, UCL. (Available from *LingBuzz*, [<http://ling.auf.net/lingBuzz/000878>], accessed 04/06/09)
- Rijkhoff, Jan. 1984. 'The Principle of Head Proximity.' MA dissertation, University of Amsterdam.
- Rijkhoff, Jan. 1986. 'Word Order Universals Revisited: The Principle of Head Proximity.' *Working Papers in Functional Grammar* 14.
- Rijkhoff, Jan. 1990. 'Explaining Word Order in the Noun Phrase.' *Linguistics* 28:5–42.
- Rijkhoff, Jan. 1992. 'The Noun Phrase: A Typological Study of its Form and Structure.' PhD dissertation, Universiteit van Amsterdam.
- Rood, David. 1973. 'Aspects of Subordination in Lakhota and Wichita.' *You Take the High Node and I'll Take the Low Node*. Chicago Linguistic Society: Chicago, 71–78.
- Sheehan, Michelle. 2008. 'FOFC and Phasal Complements of V.' Handout from FOFC seminar series, Cambridge, 15th February 2008. Available from [http://research.ncl.ac.uk/linearization/MS_FOFC_and_phases_Cambridge.pdf], *Research – Newcastle University* (accessed 17/09/09).
- Simpson, Andrew & Tanmoy Bhattacharya. 2000. 'Wh-CP Pied Piping in Bangla.' Masako Hirotani, Andries Coetzee, Nancy Hall & Ji-yung Kim (eds.), *NELS* 30, Amherst: University of Massachusetts, GLSA, 583–596.
- Vennemann, Theo. 1974. 'Topics, Subjects, and Word Order: From SXV to SVX via TVX.' J. Anderson & C. Jones (eds.), *Historical Linguistics: Proceedings of the First International Congress of Historical Linguistics, Edinburgh, September 1973* (vol. II). Amsterdam: North-Holland, 339–376.
- Vries, Mark de. 2002. *The Syntax of Relativization*. Utrecht: LOT.
- Williams, Edwin. 1982. 'Another Argument that Passive is Transformational.' *Linguistic Inquiry* 13, 160–163.
- Zanuttini, Raffaella. 1994. 'Speculations on Negative Imperatives.' *Rivista di Linguistica* 6, 67–89.
- Zanuttini, Raffaella. 1996. 'On the Relevance of Tense for Sentential Negation.' Adriana Belletti & Luigi Rizzi (eds.), *Parameters and Functional Heads. Essays in Comparative Syntax*. Oxford: Oxford University Press, 181–207.