Proceedings XXXV Incontro di Grammatica Generativa
Università di Siena

Labelling, Multiple Spell-Out and the Final-over-Final Constraint*

MICHELLE SHEEHAN
Newcastle University
Michelle.sheehan@ncl.ac.uk

This paper explores the relationship between labelling and Kayne's Linear Correspondence Axiom, arguing that the latter requires some version of the former. More specifically, a copy theory of labelling, compatible with inclusiveness, is proposed which renders phrase structure 'linearizable' via the LCA. This labelling mechanism, coupled with Uriagereka's (1999) Multiple Spell-Out, makes a number of interesting predictions about the properties of left-branches, which are shown to have empirical support. Most importantly, an apparently correct distinction is predicted to hold between underlying specifiers and derived specifiers, and more generally between head-final vs. head initial specifiers. These differences, it is proposed, serve to explain a number of seemingly unrelated facts: (i) the restrictions on extraction from derived specifiers in English, (ii) patterns of complement extraposition, (iii) the lack of CED effects in harmonically head-final languages, and (iv) the Final-over-Final Constraint (cf. Holmberg 2000).

1. Kayne's Linear Correspondence Axiom

Kayne's (1994) Linear Correspondence Axiom (LCA) proposes a direct correlation between hierarchical structure and linear order, often thought of in the following way:

(1) Linear Correspondence Axiom (informal version)
Asymmetric c-command maps to precedence.

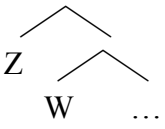
This correlation is straightforward as long as a structure is uniformly right-branching (abstracting away from the bottom pair) as in (2a), however, structures such as those in (2b) present a non-trivial challenge for (1), because no total order is specified directly by the asymmetric c-command domains of terminal nodes:

* Thanks to other members of the AHRC-funded project 'Structure and Linearization in Disharmonic Word Orders' (http://research.ncl.ac.uk/linearization/) for comments and stimulating discussions: Laura Bailey, Theresa Biberauer, Ángel Gallego (albeit briefly), Glenda Newton, Ian Roberts, Sten Vikner and especially Anders Holmberg. The ideas in this paper have also benefited from comments and critique from the audiences at LAGB (Essex), IGG (Siena), GLOW (Nantes) and at Newcastle, Cambridge and York Universities. All errors and omissions are my own.

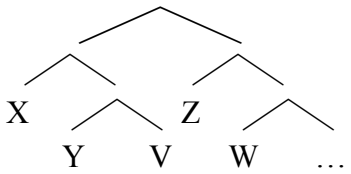
BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

(2)

a.

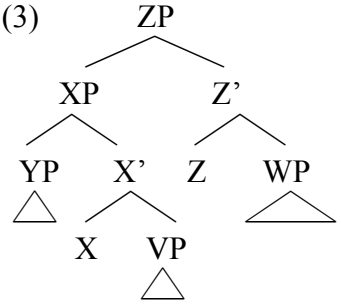


b.

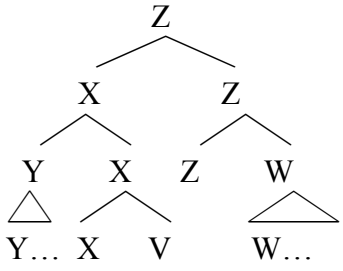


Taking a sisterhood definition of c-command, in (2b), neither X nor Y asymmetrically c-commands Z nor W, nor vice versa. Given that natural language is widely assumed to allow structures such as those in (2b) (i.e. branching specifiers), the question is how the LCA can hold in such cases. In Kayne (1994) this problem is avoided by assuming that the structures created by merge are more complex than those in (2), containing labels which themselves create additional precedence pairs via asymmetric c-command. In Kayne's terms, projection is conceived of as X-bar theory (3), but the same effect holds of Chomsky's (1995) Bare Phrase Structure (4):

(3)



(4)



These supplementary precedence pairs in turn create a new problem for linearization: contradictory ordering commands. For example, taking a simple sisterhood definition of c-command, Z' asymmetrically c-commands YP and X in (3), but XP also asymmetrically c-commands Z and WP. Kayne (1994) resolves this problem by distinguishing between categories and segments in his definition of c-command:[1]

(5) Definition of c-command (original version, Kayne 1994:18)
X c-commands Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y.[2]

The elimination of Z' as a c-commander removes contradictory precedence pairs but once again leads to an underspecified order: no asymmetric c-command holds in either direction between Z and VP for example. This problem is only apparent, though, as Kayne makes use of dominance in his formal version of the LCA:

(6) Linear Correspondence Axiom (Kayne 1994:6)
[For a given phrase marker *P*, where *d* is the non-terminal to terminal dominance relation, *T* the set of terminals and *A* the set of ordered pairs $<X_j, Y_j>$ such that for each j, X_j asymmetrically c-commands Y_j], d(*A*) is a linear ordering of *T*.

[1] Chomsky (1995) proposes that this idea can be maintained from a BPS perspective. For discussion cf. Chomsky 1995:242, 437 fn. 33, 2001:40.

[2] Note that another consequence of this version of the category/segment distinction is that specifiers actually c-command out of their phrase. I abstract away from this prediction here as it will ultimately prove irrelevant.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

In these terms, the linearization of the terminals dominated by VP depends on the precedence pairs established by XP. This ensures that a whole complex specifier like XP (including YP and VP) will precede Z (and WP).[3]

To summarize, the LCA can only linearize structures such as (2b) with the help of a number of ancillary assumptions: (i) some theory of projection/labelling, (ii) a segment/category distinction, (iii) a notion of dominance as well as asymmetric c-command in the LCA. From a Minimalist perspective one is urged to question both whether these assumptions are absolutely necessary and if so, whether they can be justified on independent grounds.

In the remainder of this article I will claim that in order to maintain the insight of the LCA (i) seems to be required, whereas (ii) & (iii) do not. It will be further argued that the elimination of (iii), in particular, renders the LCA both more restrictive and more explanatory. Section 2 will discuss labelling and explore but ultimately reject a possible way of maintaining the LCA in the absence of labelling. Section 3 will propose that the most Minimalist way to think of labelling is in terms of copying. This makes some interesting predictions about word order possibilities which are explored in the remainder of section 3. Section 4 briefly considers some rather obvious challenges facing this approach and suggests some possible solutions. Finally, section 5 concludes and outlines avenues for future research.

2. LCA without Labelling

In recent years, there has been much debate as to the status of labels in Minimalist theory. It has been noted that labelling, as described by Chomsky (1995), involves adding material not present in the numeration during the course of the derivation, and as such violates inclusiveness (cf. Collins 2002, Seely 2006). From an empirical perspective, it has been claimed that effects previously attributed to labelling can be derived by other means. For example, Collins (2002) argues that labelling is not needed for c-selection, and that a label-free theory of selection actually seems more explanatory in some cases. Nonetheless, it seems that labels are theoretically problematic but in some cases empirically necessary (cf. Boeckx 2008, Citko 2008).[4] A relevant question is therefore whether labelling is a crucial component of the LCA. In this following section I will consider an attractive way to eliminate the need for labelling from the LCA, based on Uriagereka's Multiple Spell-Out, before showing that it is empirically problematic.

2.1. Multiple Spell-Out

Uriagereka (1999) proposes an approach to linearization which would appear to eliminate the need both for labelling and dominance in the LCA. As we have seen, the label-free product of merge is not asymmetric enough to enable linearization via the LCA as soon as a complex specifier is present (cf. (2b) above repeated here as

[3] I am simplifying here for reasons of space. Kayne makes a distinction between terminals and categories, as is standard in X-bar theory, so dominance is required in any case to relate the two. Under BPS, however, the role of dominance becomes more suspect.

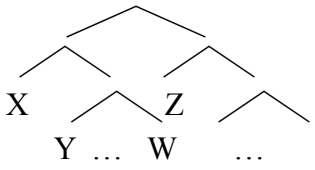
[4] This is implicitly recognized by Hornstein (2009), when he proposes that labelling or rather endocentricity is *the* defining characteristic of natural language (NL). In his terms, labelling is *the* central first-factor property of NL, merge and other locality principles are third factor, and thus more general properties of biological systems.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

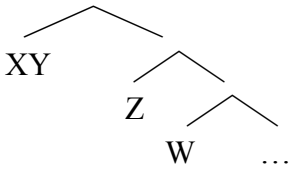
(8)).[5] What Uriagereka proposes is that a strict LCA, as in (7) coupled with the need for total order essentially rules out structures like (8) and forces them to look like (9):

(7) *Strict LCA*: If α asymmetrically c-commands ß, then α precedes ß.

(8)



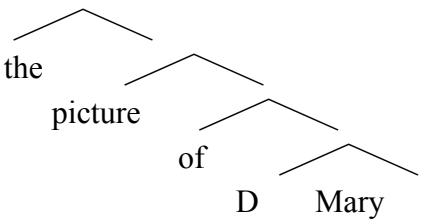
(9)



To derive this, he proposes that the number of applications of Spell-Out (SO) in the course of a derivation is subject to last resort. Because of the existence of Multiple Spell-Out (MSO) and (7), all complex phrases will undergo SO before they can be merged in a specifier position. Without this additional application of SO, complex specifiers could not be ordered with respect to the clausal spine via (7):

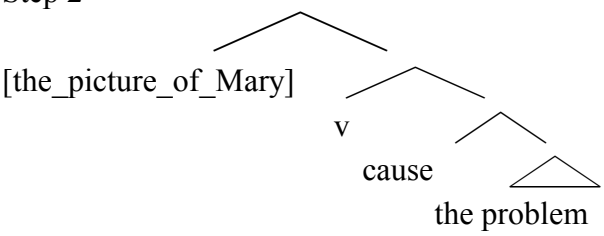
Figure 1: Multiple Spell-Out

Step 1



SO ⇒[the_picture_of_Mary],

Step 2



For this reason, all underlying specifiers are predicted to disallow subextraction, as they behave like complex lexical items. This appears to derive Huang's (1982) Condition on Extraction Domains (CED):[6]

(10) ***Who** did [**a picture of** t] cause the problem? (two applications of SO)
(11) **Who** have you seen **a picture of**? (one application of SO)

MSO and the strict LCA are therefore highly Minimalist as they allow us to eliminate the need for labels and dominance and also serve to derive the CED.

2.2. Empirical problems with Multiple Spell-Out

Despite the obvious appeal of MSO, it is unfortunately too restrictive. The immediate challenges can be summarized as follows:

[5] A potential objection arises from the fact that without labels it is actually not clear why it is specifiers and not head+complements which are targeted for SO. I assume that SO of head+complement is ruled out by the presence of selection features on the projecting head.

[6] Huang's CED is actually stated in terms of government and as such the predictions it makes are more nuanced than presented here.

(i) Condition on Extraction Domains (Huang 1982:505)
A phrase A may be extracted out of a domain B only if B is properly governed.

The strong islandhood of most adjuncts plausibly has the same explanation, as Uriagereka (1999) proposes.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

(i) The existence of remnant movement,[7]
(ii) The special behaviour of derived specifiers,
(iii) The non-universality of subject islands,
(iv) Object extraction in head-final languages.

Nunes and Uriagereka (2000) claim that extraction from derived subjects is impossible, using the following example:

(12) *Who was [a picture of t] taken (by Bill)?

They propose to explain the ungrammaticality of (12) by making SO a subcomponent of move (giving us 'freezing'). In their system, this is not stipulated, but rather it is forced by chain uniformity, based on the premise that any complex specifier will (eventually) require an additional application of SO to be linearized. This assumption, coupled with the chain uniformity condition conspires to rule out extraction from an eventual derived specifier at any point in the derivation as follows:

(13) *Who was [a picture of ~~who~~] taken [a picture of ~~who~~](subject unlinearisable)
(14) *Who was <a picture of who> taken [a picture of ~~who~~] (non-uniform chain)
(15) *Who was <a picture of who> taken <a picture of who>(no extraction)

In (13) the derived specifier cannot be linearized according to the strict LCA, for by now familiar reasons. In (14), extraction should technically be permitted from the lower, un-linearized copy, but the derivation crashes because the two copies of *a picture of who* form a non-uniform chain. Finally, in (15), the chain is uniform, but both copies behave like complex lexical items, banning extraction. Nunes & Uriagereka (2001) note that their account predicts that extraction from moved constituents should generally be banned. In actual fact, their analysis also rules out any kind of remnant movement, which would be an equivalent configuration to (13-15): extraction taking place from a complex phrase which will eventually move to a specifier position. Given the strong empirical evidence for remnant movement (Müller 1996), this is the first way in which label-free MSO is too restrictive.

A further problem comes from the fact that the data regarding derived subjects are more complex than implied above. Ross (1967) notes that in some cases, pied-piping seems to rescue extraction from derived specifiers. As Chomsky (2008) notes, this is true only of *derived* subjects, pied-piping does not substantially improve extraction from *underlying* subjects:[8]

(16) Of whom has a picture been taken?
(17) *Of whom did a picture cause the problem.

It would appear that MSO, as it stands, cannot account for this distinction.

A further empirical challenge comes from the fact that not all languages have a subject/object asymmetry with regard to subextraction (cf. Lasnik and Saito 1992, Mahajan 1992, Stepanov 2007). In Japanese, for example, extraction from a clause marked with nominative case is no worse than extraction from a clause marked with

[7] Note that the existence of remnant movement of head-initial phrases will also prove problematic for the alternative proposed later in the paper. I return to this point below.
[8] Additional, highly interesting complications are discussed by Sauerland & Elbourne (2002).

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

accusative case, according to Lasnik & Saito. Thus extraction is marked but not impossible in both cases:

(18) ??dono hon-o Mary-ga [John-ga *t* katta koto]-o
which book-acc Mary-nom John-nom bought fact-acc
mondai-un siteru no
problem-to marking
'Lit. Which book is it that John is calling the fact that John bought it into question?' object

(19) ??dono hon-o Mary-ga [John-ga t katta koto]-ga
which book-acc Mary-nom John-nom bought fact-nom
mondai-da to omotteru no
problem-cop COMP think
'Lit. Which book is it that Mary thinks that the fact that John bought it is a problem?' subject [Japanese, Lasnik and Saito (1992:42)]

This is also unexpected given MSO, which predicts *all* left branches in *all* languages to behave like strong islands.

Finally, given the crucial role of the LCA in regulating MSO, it is critical that the latter can accommodate LCA-based approaches to head-finality. The latter necessarily involve either snowballing, or roll-up movement (cf. Kayne 1994, Koopman and Szabolcsi 2000 *inter alia*). As MSO appears to completely block remnant movement, it follows that head-finality must be derived along the lines proposed by Kayne (1994): very local comp-to-spec movement. Unfortunately, a problem also arises in relation to this kind of approach: left-branching structures in OV languages are expected to be strong islands, contrary to fact (e.g. German, Hindi):

(20) saare phal ravi-ne socaa ki siitaa-ne [$_{XP}$ t khaaye] the
all fruits Ravi-ERG thought that Sita-ERG eat.PERF.PL be.PAST.PL
'All the fruit, Ravi thought that Sita had eaten.' [Hindi, Mahajan (2003:233)]

(20) presumably involves movement of XP (VP or a larger constituent) to a position asymmetrically c-commanding the auxiliary.[9] If SO were a subcomponent of move, then this would result in freezing of the object inside this XP, contrary to fact. In fact, as is often noted, OV languages are often more permissive than VO languages with respect to object scrambling (cf. Kayne 2004). It would appear, then, that MSO is too restrictive on several counts. In the next section I show that these empirical challenges can be resolved in an explanatory way by a copy theory of labelling.

3. Why the LCA needs labelling after all

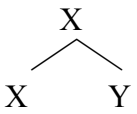
3.1. Labelling as copying

While the operation 'copy' is involved in the labelling operation as described in Chomsky (1995), it is clear that the *label* of a complex phrase does not have the status of a *copy* in any real sense. Rather, as discussed by Seely (2006), the label X in (21) is shorthand for the set formed by merge (i.e. {Z,{X,Y}), because "each node is understood to stand for the sub-tree of which it is the root" (Chomsky 1995: 398-99):

[9] This is less obvious in German because of verb cluster effects.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

(21)



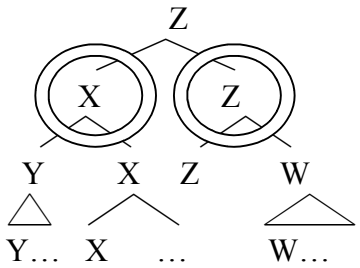
In relation to the LCA, the asymmetric c-command domain of a label X will necessarily define the linear position not just of X but of *all the terminals dominated by X* (in the set denoted by X, i.e. X & Y). As a result, if we adopt the Chomskyan notion of labelling then we are forced also to adopt dominance as part of the LCA.

If, on the other hand, we take the 'copy theory of labelling' seriously and take X to be simply a copy of the terminal X then we get a very different result. In this case, the asymmetric c-command domain of X will have no effect on the linear position of Y. Let us consider what the implications of this would be. Assume a simple projection algorithm such as the following (abstracting away from adjunction here):

(22) If X selects and merges with Y then copy X to dominate {X, Y}.

This will lead to the same contradictory ordering commands discussed above in relation to X-bar theory and BPS:

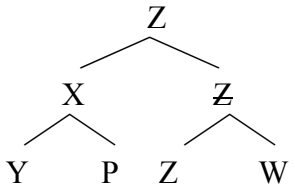
(23)



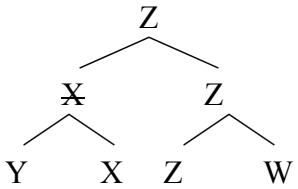
In (23), a copy of Z asymmetrically c-commands a copy of X and vice versa. Given that all three copies of Z are non-distinct, we cannot have recourse to the notion of segment/category to distinguish them so this problem must be otherwise resolved.

In influential work, Nunes (1995, 2004) has proposed that where contradictory ordering commands arise as the result of XP-movement, they are resolved via 'deletion' at the mapping to PF. Let us assume that the same thing applies to labels: problematic labels/copies are deleted at the PF interface in order to enable linearization. In (23), the PF component must delete one of the circled labels in order to enable a total ordering between Z & X. Following Nunes, we can also assume that economy will favour the deletion of lower copies over higher copies, all else being equal.[10] Thus deletion of the highest label of X will be less economical than the deletion of the intermediate copy of Z and (24) will be preferable over (25):

(24)



(25)



[10] This holds if selection is mediated by uninterpretable features in the same way that movement is. If this is the case then different copies of Z will presumably have different feature specifications in the same way that different copies of moved XPs will have different feature specifications.

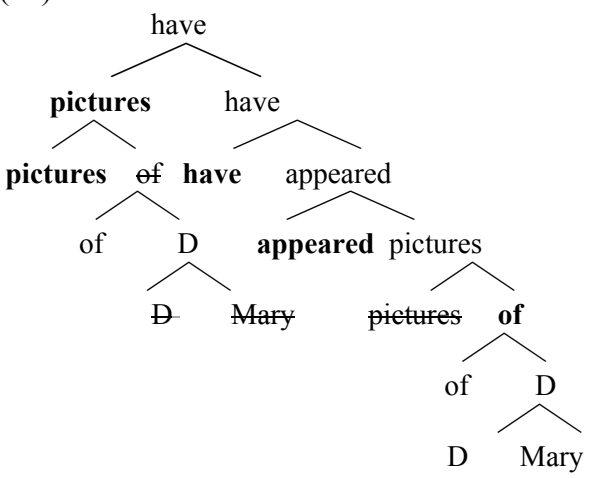
BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

Note that there is no advantage here to deleting the other copies of Z/X as they do not give rise to ordering paradoxes. Assuming deletion to be a last resort mechanism, these copies should remain at PF. Up to now it might seem that 'labelling as copying' is a notational variant of BPS. In actual fact, though, labelling as copying makes clearly distinct predictions regarding patterns of linearization and CED, which I explore in the following sections.

3.2. Derived subjects

The copy theory of labelling makes it possible to linearize derived specifiers in one fell swoop in relation to the clausal spine. This is true because the asymmetric c-command domain of *pictures* in (26) does not affect the linear position of *of Mary*. In a sense, when the phrase *pictures of Mary* moves, the projecting head *pictures* moves 'higher' but its complement does not. Following Bošković (2001), I assume that scattered deletion is available as a last resort at PF, where full copy deletion is blocked. While I assume, following Nunes, that PF privileges derived copies, I propose that the need for total linear order overrides this preference. In its base-generated position, *of Mary* can be linearized with respect to the clausal spine (i.e. *have* & *appeared*), whereas in its derived position it cannot.[11] This situation gives rise to obligatory 'complement extraposition', whereby the complement contained in a moved XP is stranded in its base (complement) position via scattered deletion. In this way, a total order is obtained without the need for an additional application of SO:

(26)



Note that complement stranding can only apply to complements contained in a phrase merged in a complement position, underlying specifiers will always require an additional application of SO. More generally, the prediction is that complement extraposition via scattered deletion will be preferred over an additional application of SO.

A complication arises from the fact that (27) is available as an alternative to (26):

(27) Pictures of Mary have appeared (recently).

Sheehan (2009) argues that this is linked to specificity, which induces strong islandhood with representational nominals. Complex NPs which are strong islands in a complement position will also be strong islands in a specifier position. 'Indefinites' are ambiguous between a specific and non-specific reading (Diesing 1992). Sheehan

[11] Note that neither of the two copies of *of Mary* is any 'higher' than the other in terms of c-command.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

proposes that these readings correlate with different underlying syntactic structures (DP vs. NumP). NumPs, but not DPs, give rise to obligatory complement stranding:

(28) a. *The pictures have been published of the fire.
b. The pictures of the fire have been published.

(29) a. Many pictures have been published of the fire.
b. Many pictures of the fire have been published.

Now recall the fact that pied-piping served to rescue extraction from derived subjects.

(30) Of which person have many pictures appeared?

This can be seen to result from a structure involving scattered deletion, hence where NumP is not a strong island (cf. Sheehan 2009 for details).

Note also, that some speakers accept stranding of *of* after appeared, and all speakers accept this where a heavier preposition is used:

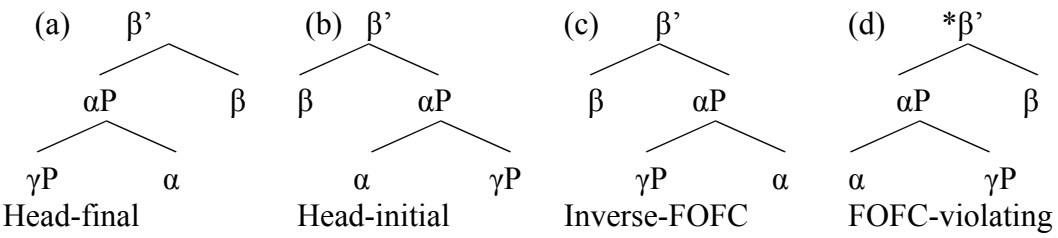
(31) (??)Which celebrity have some nice pictures appeared of?
(32) Which subject has a decent book appeared about?

This is predicted by the account proposed here. Stranding affects all moved XPs which are not inherently islands (i.e. strong islands also when in complement position). For this reason, all else being equal, extraction from stranded complements is expected to be possible.[12]

3.3. The Final-over-Final Constraint

Holmberg (2000:124) points out an interesting asymmetry between the two possible disharmonic (mixed head-initial/final) word order combinations in natural languages: (33c) is fairly common, whereas (33d) is unattested in many syntactic domains:

(33) Harmonic and disharmonic combinations



(34) The Final-over-Final Constraint (FOFC) [First Version]:
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.

[12] Note that adjunct extraposition behaves differently from complement extraposition in being permitted form underlying specifiers and in many other respects. I leave this matter to one side here for reasons of space.

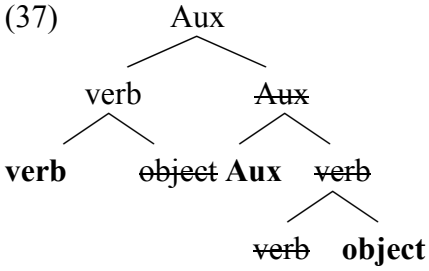
BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

Evidence in favour of FOFC comes from a number of typological gaps as well as from the ungrammaticality of certain word order combinations in certain mixed languages such as Basque and Finnish. For a full discussion see Biberauer, Holmberg and Roberts (2007, 2008), Biberauer, Newton and Sheehan (to appear-b, to appear-b). A simple way to think of FOFC is (35) which can be restated as in (36):

(35) Head-finality must begin at the bottom of the tree.
(36) Roll-up movement must begin at the bottom of the tree.

If all head-final phrases are the result of roll-up movement then FOFC receives a principled explanation from the copy theory of labelling coupled with Uriagereka's strict LCA. As long as there is no SO trigger, we expect roll-up movement to be subject to obligatory complement stranding: if a non-spelled-out head-initial phrase moves to the specifier of a higher head, it will always strand its complement:

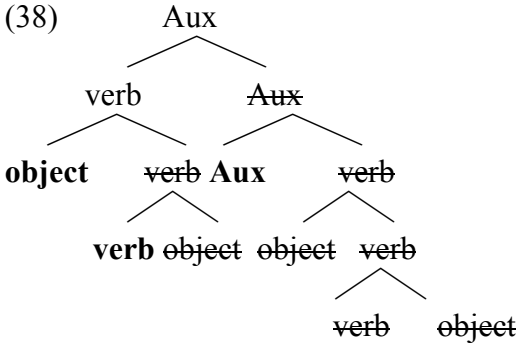
(37)



As Biberauer, Holmberg and Roberts (2008) note V-O-Aux is unattested in any diachronic/synchronic variety of Germanic, but V-Aux-Obj is well-attested in Old Norse and Old English (cf. Hróarsdóttir 2008 & Pintzuk 2005).

Note that harmonically head-final specifiers are correctly predicted not to display FOFC or CED effects. Head-finality is derived by comp-to-spec movement and specifiers, unlike complements, always asymmetrically c-command a phrase head at PF. For this reason they so can be ordered parasitically based on the position of said head:

(38)



In (38), the derived copy of *object* can be targeted for SO as it asymmetrically c-commands *verb*, and a copy of *verb* asymmetrically c-commands *Aux,* giving the total linear ordering object>verb>Aux. The crucial difference between (37) & (38) is that in (37) the derived copy of *object* has no order with respect to *Aux*, whereas that in (38) does (parasitic on that of *verb*). This fact accounts for the FOFC asymmetry as well as the lack of CED effects in Japanese and Turkish.

4. Apparent counterexamples

An apparent problem with the proposal is that it appears to predict that any XP which moves will necessarily strand its complement. Firstly, it is important to emphasise that this prediction relates only to first-merged complements. XPs which are

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

externally merged in a specifier position will necessarily be spelled out prior to insertion as per Uriagereka's MSO. Thus underlying specifiers are predicted never to permit complement extraposition, as appears to be correct, once we differentiate between adjunct and complement extraposition (cf. Coopmans and Roovers 1986):

(39) a. **More pictures of celebrities** would provoke a scandal.
b. ***More pictures** would provoke a scandal **of celebrities**.

More problematic are non-nominal categories which appear to move without giving rise to complement stranding in English: PP, VP, CP. In many cases, these very same phrases do *not* appear to be strong islands in complement position:

(40) a. [That pigs eat apples] is believed by most people.
b. *That is believed (by most people) pigs eat apples.
c. What do most people believe [that pigs eat t]?

There are at least two potential solutions to this problem. Either (i) (40a) does not involve movement, but rather base-generation of the complex CP in a specifier position, or (ii) the moved CPs in (40a) is a covert DP. I will not pursue this matter here for reasons of space, but it is interesting to note that both (i) and (ii) have actually been proposed in relation to sentences like (40a) on independent grounds (cf. Alrenga 2005 for an overview of the debate).[13]

5. Conclusions

It has been argued that the LCA necessitates some system of labelling in order to allow for complex specifiers. A simple copy theory of labelling which satisfies inclusiveness has been proposed and some of its implications explored. Crucially, it has been shown that the copy theory of labelling is not a notational variant of BPS but that it actually makes some interesting predictions regarding linearization, at least some of which seem to hold. For example, complement extraposition and the Final-over-Final Constraint are explained, as are patterns of extraction from derived subjects. Ultimately, the result is a reduction in the number of word orders which can be generated via relatively free movement combined with a strict LCA.

References

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

Biberauer, T., A. Holmberg, & I. Roberts (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, T., A. Holmberg and I. Roberts (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.

[13] Note that Scandinavian remnant VP-fronting remains problematic for the approach proposed here. In fact the prediction is roughly that OV languages should allow remnant VP movement whereas VO languages should not. I leave this interesting matter to future research.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

Biberauer, T., G. Newton, and M. Sheehan (to appear-a) Impossible changes and impossible borrowings: the Final-over-Final Constraint. In A. Breitbarth, C. Lucas, S. Watts and D. Willis (Eds.), *Continuity and Change in Grammar*. Amsterdam: John Benjamins.

Biberauer, T., G. Newton, and M. Sheehan (to appear-b) Limiting synchronic and diachronic variation and change: the Final-Over-Final Constraint. *Language and Linguistics*.

Boeckx, C. (2008) *Bare syntax*. Oxford: OUP.

Bošković, Ž. (2001) *On the nature of the syntax-phonology interface: cliticization and related phenomena*. Amsterdam: Elsevier.

Chomsky, N. (1995) Bare phrase structure. In G. Webelhuth (Ed.) *Government and binding theory and the minimalist program: principles and parameters in syntactic theory*, 383–439. Cambridge, MA: Blackwell.

Chomsky, N. (2001) Derivation by phase. In M. Kenstowicz (Ed.) *Ken Hale: a life in language*, 28-53. Cambridge, MA: MIT Press.

Chomsky, N. (2008) On Phases. In R. Freidin, C. P. Otero and M. L. Zubizarreta (Eds.) *Foundational Issues in Linguistic Theory*, 133-166. Cambridge, MA: MIT Press.

Citko, B. (2008) Missing labels. *Lingua* 118, 907–944.

Collins, C. (2002) Eliminating labels. In S. D. Epstein and T. D. Seely, (Eds.) *Derivation and explanation in the Minimalist Program*, 42-64. Malden, MA: Blackwell.

Coopmans, P., and I. Roovers (1986) Reconsidering some syntactic properties of PP-extraposition. In P. Coopmans, I. Bordelois and B. D. Smith (Eds.) *Going romance: Formal parameters of generative grammar II*, Dordrecht: ICG Printing.

Diesing, M. (1992) *Indefinites*. Cambridge, Mass.: MIT Press.

Epstein, S.D., E. Groat, R. Kawashima, and H. Kitahara (1998) *A Derivational Approach to Syntactic Relations*. Oxford: OUP.

Holmberg, A. (2000) Deriving OV Order in Finnish. In P. Svenonius (Ed.) *The Derivation of VO and OV*, 123-152. Amsterdam; Philadelphia: John Benjamins.

Hornstein, N. (2009) *A theory of syntax*. Cambridge: CUP.

Hróarsdóttir, T. (2008) Types of DPs in OV order. *Studia Linguistica* 62, 261-286.

Huang, C.-T. J. (1982) Logical relations in Chinese and the theory of grammar. MIT, Ph.D Thesis.

Kayne, R. (2004) Antisymmetry and Japanese. In L. Jenkins (Ed.) *Variation and universals in biolinguistics*, 3-37. Amsterdam: Elsevier.

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

Koopman, H. J., and A. Szabolcsi (2000) *Verbal complexes*. Cambridge, MA: MIT Press.

Lasnik, H., and M. Saito (1992) *Move alpha: conditions on its application and output*. Cambridge, MA: MIT Press.

Mahajan, A. (1992) Specificity Condition and the CED. *LI* 23, 510-516.

Mahajan, A. (2003) Word order and (remnant) VO movement. In S. Karimi (Ed.) *Word order and scrambling*, 217-237. Maulden, MA: Blackwell.

Müller, G. (1996) A constraint on remnant movement. *NLLT* 14, 355-407.

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

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

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)

Nunes, J., and J. Uriagereka (2000) Cyclicity and extraction domains. *Syntax* 3, 20-43.
Pintzuk, S. (2005) Arguments against a universal base: evidence from Old English. *English Language and Linguistics* 9, 115-138.
Richards, M. (2004) Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax. Cambridge University, Ph.D Thesis.
Ross, J. R. (1967) Constraints on variables in syntax. Massachusetts Institute of Technology, Ph.D Thesis.
Sauerland, U., and P. Elbourne (2002) Total Reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33, 283–319.
Seely, D. (2006) Merge, derivational c-command, and subcategorization. In C. Boeckx (Ed.) *Minimalist Essays*, 182–217. Amsterdam/Philadelphia: John Benjamins.
Sheehan, M. (2009) The Final-over-Final Constraint as an effect of complement stranding. *Newcastle Working Papers in Linguistics* 15:104-125.
Stepanov, A. (2007) The End of CED? Minimalism and Extraction Domains. *Syntax* 10, 80-126.
Uriagereka, J. (1999) Multiple Spell-out. In N. Hornstein and S. D. Epstein (Eds.) *Working minimalism*, 251-282. Cambridge, MA: MIT Press.

BDD-A22699 © 2009 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.15 (2026-09-14 15:13:07 UTC)