Some implications of a copy theory of labeling

Michelle Sheehan

Abstract

A copy theory of labeling makes interesting empirical predictions when combined with a revised version of Kayne's (1994) Linear Correspondence Axiom (LCA) and a category-based version of c-command. Firstly, it forces the pronunciation of the highest copy in a chain, all else being equal, without the need for copy deletion. Secondly, it predicts certain differences between (i) underlying vs. derived head-initial specifiers, and (ii) head-final vs. head-initial specifiers. This serves to extend Uriagereka's (1999) PF-based explanation of Huang's Condition on Extraction Domain (CED) to cover previously problematic or unexplained phenomena: (i) the apparent lack of subject-island effects with head-final specifiers; (ii) the non-islandhood of many derived specifiers (externally merged as complements); (iii) patterns of complement extraposition; and (iv) the Final-over-Final Constraint (cf. Holmberg 2000, Biberauer, Holmberg & Roberts 2007, to appear). All of the above phenomena arise under CoL as side-effects of a revised version of Kayne's (1994) Linear Correspondence Axiom incorporating a head-parameter, whereby asymmetric c-

1

command, defined between categories and unmediated by dominance, maps to precedence as a last resort.

1 Introduction: endocentricity and labeling

Since the work of Chomsky and Jackendoff in the 1970s, it has been widely accepted that natural language syntax is (at least largely) endocentric or headed (cf. Jackendoff 1977).[1] This property is traditionally attributed to the narrow syntactic operation of *projection* which gives rise to *labels*. Thus the label V(P) in (1) shows that the constituent comprising a transitive verb and a direct object has the status of a verb (phrase) rather than that of a determiner (phrase):[2]

(1)   V(P)

   V   D(P)
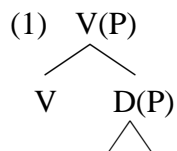
According to standard assumptions, V, the head of the phrase, is privileged in that it passes up (at least a subset of) its features to the phrasal level, unlike its complement, the DP, the features of which remain more deeply embedded, and not visible for selection by a higher category. The resulting constituent V(P) is then able to be selected, probed, moved, elided or co-ordinated and interpreted/manipulated as a V(P) at the LF and PF interfaces.

---

[1] I include this caveat as there are certain areas of natural language syntax which are not so obviously endocentric, such as small clauses (cf. Moro 2000).

[2] The P remains in brackets to indicate that the same holds for both X-bar and Bare Phrase Structure approaches to labeling. We return to these different labeling mechanisms shortly.

Recently, the need for labels has been questioned on conceptual and empirical grounds (cf. Collins 2002, Seely 2006).[3] While Collins accepts the need for a derivational notion of projection (his *locus principle*), he rejects the need for this to result in (representational) labeling. Essentially, his proposal is that the head of a phrase ceases to be such once its features have been satisfied. Moreover, from an empirical perspective, Collins argues that at least some of the effects for which labeling was thought essential can be equally well handled without labels: he argues that c-selection, for example, can be modeled in an empirically superior manner.

The question of whether labeling can be dispensed with entirely and replaced with its derivational equivalent depends on whether all the work done by labels can be assumed by the *locus principle*. Current understanding of at least the following phenomena appeals to labels to greater or lesser degrees: co-ordination, adjunction, ellipsis, phrasal movement, phonological phrasing and LF interpretation (cf. Irurtzun 2007: chapter 1, Hendrick 2007, Hornstein 2009, for discussion). In some cases, it may be a trivial task to reanalyze these phenomena in derivational terms.[4] In other cases, the task is more challenging. Phrasal movement, ellipsis and LF interpretation, for example, seem to resist a derivational explanation as, under our current understanding, they all involve manipulation/interpretation of a phrase after it has ceased to be the derivational *locus*. In the absence of labels, it is difficult to see, for example, how phrasal movement can be both forced and constrained (cf. Adger 2011 for discussion). It therefore seems premature to discard labeling altogether.

---

[3] But cf. also Hornstein (2009) who defends labeling, claiming it to be *the* defining characteristic of natural language, as the operation which enables recursion (once certain assumptions are made).

[4] For example, Collins discusses a potential account of phonological phrasing in a model with multiple points of PF transfer.

Even if labeling cannot be eliminated from syntactic theory, it is nonetheless worth considering to what extent it might be reconceptualized to fall in line with independently motivated narrow syntactic operations. One possibility in this vein is the claim, independently made by Neeleman & Van de Koot (2002, 2010), Hale & Keyser (2002), Boeckx (2008), Citko (2008) and Hornstein (2009) that labels are generated via the operation *copy*.[5] A potential, but to my knowledge unexplored, interpretation of the copy theory of labeling (CoL) is that terminal and non-terminal projection chains (in the sense of Speas 1990: 43) form multi-segment categories, so that complements are formally merged via adjunction, just as specifiers are in Kayne's (1994) approach.[6] In fact, if copying always gives rise to segment rather than category formation then movement chains can also be straightforwardly linearized without the need for deletion (in most cases). The picture which emerges is one in which each lexical item in the derivation projects a single (potentially scattered) category, with each additional instance of copying (segment-formation) serving to alter the c-command relations of the multi-segment category as a whole. This approach to labeling/copying seems to necessitate a (representational) category-based definition of c-command, but it in turn gives rise to a series of novel, empirically-supported predictions. It is the goal of this paper to explore some of these predictions, in particular those concerning movement, linearization, word order asymmetries and the Condition on Extraction Domain (Huang 1982).

---

[5] Hornstein (2009: 55) takes labeling to be "an operation that renders its output as type identical to one of its inputs". This would appear to imply copying of the *cloning* variety discussed below. Note also that Hornstein (2009: 59) explicitly contrasts this view of labeling with that in BPS.

[6] I use the terms 'segment' and 'adjunction' here because of their familiarity from the work of May (1985), Chomsky (1986) and Kayne (1994). In fact, these terms have a different status in the proposal made here as *all* projection results in segment formation. I clarify these matters in due course.

The structure of the paper is as follows. Section 2 discusses one influential theory of labeling, Chomsky's (1995) Bare Phrase Structure (BPS), and notes that, despite appearances, labels are by no means *copies* of terminals in this approach. Section 3 introduces the copy theory of labeling (CoL) and minimally revises Kayne's definitions of dominance and c-command so that they hold over mixed terminal/non-terminal multi-segment categories. The resulting definitions are then shown to be equivalent to those of BPS in many contexts, but divergent with respect to: (i) points of symmetry and (ii) branching specifiers. Section 4 introduces the revised version of the LCA adopted here, whereby asymmetric c-command mediates linearization only as a last resort. Section 5 goes on to consider the predictions of this version of the LCA in conjunction with the CoL, arguing that the two together serve to explain some otherwise puzzling word order gaps and exceptions to the CED (building on Uriagereka 1999 and Biberauer, Holmberg & Roberts 2007, 2008, to appear). Finally, section 6 concludes.

2   Bare Phrase Structure

2.1   Complementation

Chomsky's (1995) Bare Phrase Structure (BPS) superficially appears to be a version of the copy theory of labeling (CoL):

(2)   X
          /\
        X   Y
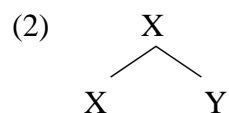
Although Chomsky does not use the term copy to describe BPS labels, he does state that a label *is* the projecting head, and so something akin to copying appears to be involved in label

generation. In fact, Seely's (2006: 185) characterization of the operation Merge in BPS (which, as he notes, is stipulated to require labeling) is as follows:

(3) The algorithm for the Merge of A and B, where A projects

    i.   Create the set {A, B}

    ii.   Make a copy of A

    iii.   Create the two-membered set consisting of {A, {A, B}}

Crucially, though, Chomsky also stipulates that labels, unlike terminal nodes and phrase markers are *not* syntactic terms (Chomsky 1995: 399). As such, labels differ from the heads they replicate in being syntactically inert (cf. Seely 2006: 183, Hornstein 2009: 59 for discussion). Thus, in BPS, labels are certainly not treated as copies of terminals at the narrow syntactic level.[7]

    It is clear, then, that under BPS neither the label X nor the node it labels has the status of a *copy* of the terminal X, despite superficial appearances. The claim that BPS labels are generated via the operation *copy* is therefore misleading as labels are actually a new kind of entity altogether in this approach.

2.2 Segments and categories

Thus far the discussion has focused on complementation. Traditionally, a distinction is made between argument saturation and adjunction, the idea being that merge with an argument leads to the projection of a new category, whereas merge with an adjunct leads to the

---

[7] Note the crucial distinction here between *labels* and *nodes*. As Seely observes, the non-terminal *node* labeled X in (2) represents "the sub-tree of which it is the root" (i.e. the set formed by merge with the label X or {X,{X,Y}}, Chomsky 1995: 398-99).

projection of an additional segment of the same category.[8] Interestingly, then, the labels created by segment formation/adjunction *are* interpreted like copies of a certain kind.[9]

Note also that Kayne's (1994) Linear Correspondence Axiom (LCA) forces the distinction between the two types of projection to become blurred as, in order to comply with the LCA, (branching) subjects/specifiers must be merged as adjuncts (cf. Abels 2010, Abels & Neeleman to appear for discussion):

(4) Linear Correspondence Axiom (Kayne 1994:6)[10]

[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 MS], d(*A*) is a linear ordering of *T*.

Essentially, branching *specifiers* are not asymmetric enough to satisfy the requirements imposed by the LCA, whereas branching *adjuncts* are. This problem applies to BPS as much as to X-bar theory, and to node-based and category-based definitions of c-command (cf. Chomsky 1995:414):[11]

---

[8] Chomsky (2001) calls these two flavours of merge set-Merge and pair-Merge.

[9] Indeed, this is part of the motivation for segments, as type-theoretically adjuncts map onto the identity function.

[10] For Kayne 1994, the LCA operates at all levels of phrase structure, ruling out symmetry at all stages in the derivation. Chomsky (1995), however, reinterprets the LCA as a linearization device which applies only at the mapping to PF. I will adopt the latter stance here.

[11] Branching specifiers are obviously problematic for the LCA under node-based definitions of c-command (Reinhart 1981), as they represents points of structural symmetry.

(5)
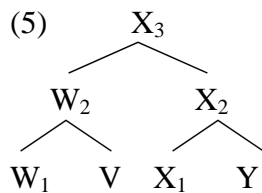
(6) X c-commands Y iff X and Y are categories and X excludes Y and every category that

dominates X dominates Y. (Kayne 1994: 18)

(7) X excludes Y if no segment of X dominates Y. (Chomsky 1986: 9)

According to (6), the tree in (5) creates a problem for the LCA if X_2 and X_3 are distinct

categories because W_2 will asymmetrically c-command the categories X_1 and Y, while X_2

asymmetrically c-commands W_1 and V, giving rise to contradictory precedence pairs. To

avoid this problem, Chomsky (1995) is forced to maintain a category-based definition of c-

command, and to assume, following Kayne , that branching specifiers are actually adjuncts.[12]

In this way, the point of symmetry in (5) disappears as X_3 and X_2 form a two-segment

category which fails to exclude W_1 and V and so cannot c-command them.[13] As such,

Chomsky and Kayne are essentially forced to the position that what were previously thought

to be specifiers are really adjuncts, or rather that the adjunct/specifier distinction does not

exist.[14] As a result a head/terminal category can project either a new category or a segment,

reflecting complementation vs. adjunction respectively, but non-terminal categories can

---

[12] Or to make some additional stipulation such as X-bar invisibility.

[13] This is assuming that Chomsky means to maintain Kayne's exclusion proviso.

[14] This is the case for branching specifiers, atomic specifiers are actually linearizable in either labelling system (cf. Chomsky 1995, Guimarães 2008). If fact it is non-branching adjuncts which create a problem for the LCA (as discussed in section 3.4).

project only to segment (by stipulation). As Abels (2010) notes, this is a rather unprincipled system.

Chomsky uses labels consisting of ordered pairs of the projecting category <X, X> to denote segments. Abstracting away from this arbitrary labeling convention though, where two nodes bear labels denoting segments of the same category, they can essentially be thought of as copies. Interestingly, then, in both X-bar theory and BPS there is a notion of labeling akin to copying which serves to introduce both arguments and adjuncts into the derivation. In fact, under the LCA, the only time labels are not segments, but rather distinct (albeit inert) categories, is where a head merges with a complement. If projection to segment could be extended to cover complementation then the theory of projection and labeling would, therefore, be substantially simplified. Moreover, things would improve further if the copies which occur in movement chains are also segments, and the c-command domain of a given category is sensitive to the position of all of its segments (copies). It will be argued in the next section that such a simplification is possible, and that it actually serves to resolve certain recalcitrant problems facing BPS as well as yielding a number of new empirically supported predictions.[15]

---

[15] The term adjunction is potentially misleading here as the claim is ultimately that complements and specifiers are merged in the same manner, under projection to segment. The question of whether adjuncts should also be merged in this way is a separate issue (cf. Chametzky 1994 for related discussion). Note that for Kayne, it must already be the case that the semantic differences between arguments and adjuncts result from their denotations rather than the manner in which they enter the derivation. It is a small step to assume the same for the complement/adjunct distinction. On the other hand, if adjuncts are not merged via projection to segment, then a configurational distinction remains possible. I return to this matter briefly below.

3 Labels as copies

3.1 The basic proposal

Segment formation can be thought of as copying only under a certain interpretation of the term.[16] There are, in fact, different ways of understanding the copying mechanism. The most obvious way, which we might call *cloning*, takes it to be an operation which copies X to produce a new category, identical to X in all ways, but independent of it. A distinct understanding of copying, which I will call *segmentation*, takes an object X and makes a copy of X to form a multi-part object comprising the totality of these copies. This kind of copying is instantiated in citation networks, which, irrespective of the number of copies, still reference a single article.

It is often assumed that syntactic copies are basically of the *cloning* type, hence the fact that copying necessitates deletion for Nunes (2004). My contention here is that things become simpler if syntactic copying is actually of the *segmentation* kind. If this is the case, then copying always gives rise to segment formation, whether the copy serves the traditional function of labeling or that of movement. The result is a syntax in which each head selected from the lexicon projects exactly one mono or multi-segment category.[17] C-command relations are then calculated for each category over the totality of its segments without the need for deletion (except in exceptional circumstances).

---

[16] Thanks to an anonymous reviewer for urging me to clarify this distinction.

[17] Note, that the idea that a lexical item can itself act as a label is not a new proposal. In BPS, terminal nodes can be both minimal and maximal hence both lexical items and (the equivalent to) XP labels. Assuming that all XP (and X-bar) labels are also copies of lexical items therefore bestows a certain uniformity on the system, which is absent from the BPS system of labeling (in which terminal labels but not non-terminal labels are terms).

10

The question of why/how one category projects rather than another is at once central and orthogonal to the issue of labeling. In what follows, I will not address the deeper issue of how or why headedness emerges, and will simply assume, in the spirit of Chomsky 2006, that any of the following asymmetries are sufficient to trigger projection:

(8) Where two categories A and B merge, let A project if (i) A, but not B, is a mono-segment category (ii) A dominated B prior to merge or (iii) B, but not A, bears uF.

This heterogeneous approach to projection is far from satisfactory, but sufficient for our purposes.[18] It basically means that a category will project (i) where it is a non-branching head and the category it merges with is not, (ii) where the category it merges with has undergone internal merge or (iii) where the category it merges with bears an uninterpretable feature.[19] This approach takes c-selection to be distinct from agree, as c-selection features must not count as uF for (8).[20] The canonical instance of an uninterpretable feature is Case. It appears

---

[18] Ultimately, any theory of phrase structure needs some account of how the locus/label is chosen, so this issue goes beyond the specific proposal made here. The CoL is a theory of how labels are generated and interpreted and it is compatible with any coherent proposal of how the head of a phrase actually emerges, a problem which remains very much open at present (cf. Adger 2011 for recent discussion).

[19] If little v serves both to introduce the external argument and to check accusative Case then a potential problem arises as both v and D will bear uF. This can be resolved, though if the uF (Case) feature on v is checked before the specifier is merged. Issues remain for syntactic ergativity, which I put to one side here.

[20] There are principled reasons to make this distinction. Firstly c-selection must be maximally local, whereas agree need not be and so the former but not the latter must be satisfied at the point of external merge. Secondly, c-selection is a one-way dependency, whereas uninterpretable features like Case/Phi lead to two-way dependencies being established.

to be a fact about Case that (inherent Case aside) it is not valued by external merge, but rather by some non-local agree dependency, optionally coupled with movement. The presence of an uninterpretable Case feature on D will therefore be sufficient to trigger the projection of v where projections of v and D merge, and of V, where V merges with a non-branching pronominal complement. This raises the question what happens in instances of adjunction. If we take adjuncts to be multi-segment categories (non-heads) which lack uninterpretable features, then no projection will be forced by (8) where an adjunct is externally merged. This creates the possibility that true adjunction results not in segment formation after all but rather in a label-free phrase.[21]

This projection mechanism will give rise to structures such as the following, which bear only a superficial resemblance to BPS (with subscripts for expository reasons only):

(9)
$$\begin{array}{c} X_3 \\ \diagup \diagdown \\ Y \quad X_2 \\ \diagup \diagdown \\ X_1 \quad Z \end{array}$$
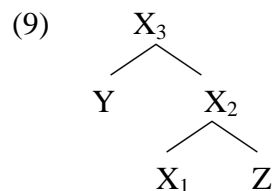
In (9), X_1, X_2 and X_3 are all segments of a single category X, rather than constituting two/three distinct categories (as in BPS). In the remainder of this section, I minimally revise Kayne's definitions of dominance and c-command so that they hold over such categories.

---

[21] As a reviewer points out, there is a sense in which adjuncts do select the phrase they modify, which raises problems for the selection-based approaches to projection in Collins 2002 and Cecchetto & Donati 2010. Under the heterogeneous approach to projection taken here, this problem is avoided, as long as the modified phrase is without uF.

3.2 Reformulating c-command

At first sight, it seems that the CoL is not asymmetric enough in relation to either Reinhart's (1981) or Kayne's (1994) definition of c-command to reflect the asymmetries of natural language. Consider first, Reinhart's definition, which is stated in terms of nodes:

(10) Reinhart's c-command: node A c-commands node B iff the branching node most immediately dominating A also dominates B. (Reinhart 1981: 612)

According to (10), the node X_1 in (9) mutually c-commands the node Z and the node X_2 mutually c-commands Y. If X_1 and X_2 are both segments of the same category then the category X fails to asymmetrically c-command either its specifier or its complement. The result is a highly symmetrical phrase structure which is fundamentally incompatible with the LCA, and there is no obvious way to resolve this.

Now recall Kayne's category-based definition of c-command in (6), repeated here as (11), which relies on the notion of exclusion in (7), repeated as (12):

(11) X c-commands Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y. (Kayne 1994: 18)

(12) X excludes Y if no segment of X dominates Y. (Chomsky 1986: 9)

If all copies of X are segments in (9), then the category X fails to exclude either Y or Z and so cannot c-command them according to (11). Assume the following definition of dominance from Chomsky (1986: 7), which is implicit in Kayne 1994:

13

(13) Y is dominated by X if it is dominated by every segment of X.

Applying (13) to CoL we get the result that no category will ever dominate any other category. This is because every category necessarily contains a terminal node which, by definition, fails to dominate anything. As such, Y and Z asymmetrically c-command X, and mutually c-command each other.[22] In this case, then, the CoL in conjunction with Kayne's original definition of c-command also appears to give rise to a highly symmetrical phrase structure which renders the LCA untenable.

It is, however, possible to minimally reformulate Kayne's category-based definitions of c-command to render them compatible with CoL, yielding interesting empirical predictions. To do this, I reformulate the notion of exclusion as partial category dominance and relativize dominance to dominating (i.e. non-terminal) nodes. Drawing on Chomsky (1986) and Wilder (2008), we can then define complete category dominance as follows:[23]

(14) Complete Dominance: a category X completely dominates a category Y iff $X \neq Y$ and the sum of the paths from every terminal segment of Y to the root includes every non-terminal segment of X exactly once.[24, 25]

---

[22] The same will hold of all further specifiers in a spine, which will also be dominated by no categories.

[23] An anonymous reviewer raises the question whether this definition simply re-introduces the X-bar level. I do not believe that this is the case, however, firstly because terminal nodes which by definition do not dominate can be automatically excluded from generalizations about dominance, and secondly, because, under CoL, the category which projects is the same category as the terminal node rather than a distinct inert category. This has important implications which are explored below.

[24] Where path is defined as a sequence of nodes (N1, N2, …, Nk) such that Y=N$_1$, X=N$_k$ (Wilder 2008).

14

(15) Partial category dominance: a category X partially dominates a category Y iff X ≠ Y, the path from every segment of Y to the root includes a copy of X but X does not completely dominate Y.[26]

Given these definitions, Kayne's version of c-command can be restated in the following way:

(16) 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.[27]

In what follows, I explore the repercussions of the CoL and this definition of c-command.

An anonymous reviewer raises concerns about this category-based definition of c-command, given the difficulty in deriving it from virtual conceptual necessity (cf. also Epstein 1999, Seely 2006). While it is true that category-based definitions of c-command appear complex when described verbally, they are more simple to represent graphically. It is

[25] As an anonymous reviewer points out, this definition runs into trouble if head-adjunction exists. The existence of adjunction to heads has been questioned on independent grounds, though, and the problem disappears if heads move to specifier position, as proposed by Matushansky (2006).

[26] That a category can neither partially nor completely dominate itself, is tacitly assumed by Chomsky (1995: 418). There are a number of welcome repercussions to this assumption. Firstly, the ban on reflexive dominance/c-command prevents the reflexive satisfaction of selection features (Neeleman & Van der Koot 2002: 534). Secondly, it removes all of the points of symmetry inherent to BPS, as illustrated below. Finally, it allows a head to c-command out of its phrase, the repercussion of which are explored in section 5.

[27] Note that under this approach, a category X can both dominate and c-command a category Y.

also the case that they appear to be necessitated by the LCA.[28] Others have taken this to be an argument against the LCA, but in section 5 I provide reasons to maintain the idea that asymmetric c-command mediates linearization to some degree. As long as other approaches to labeling/c-command have no alternative explanations for the asymmetries explained by CoL (see section 5), purely conceptual objections are misplaced. Even on a conceptual level though, if we accept the need for a non-derivational notion of endocentricity, then either we must posit some new operation which gives rise to a distinct kind of syntactically inert object (i.e. a BPS label) alongside a copy operation which in turn necessitates deletion or we are forced to a category-based definition of c-command, where labels/copies are merely segments. At the very least then, the two options seem equally bad/good, and the unexplored alternative merits exploration. Ultimately, it may even be possible to simplify the definition of c-command whilst maintaining the basic insight of CoL: that all copying leads to segmentation.

As such, in the remainder of this paper, I will explore the empirical predictions of the category-based definition of c-command above, taken in conjunction with CoL.[29] Unsurprisingly, CoL overlaps substantially in empirical coverage with BPS in relation to c-command relations. In both systems: (i) non-branching specifiers asymmetrically c-command heads (though in CoL, these are projecting categories); (ii) heads/projecting categories asymmetrically c-command into branching complements; and (ii) specifiers c-command out of their containing phrase. CoL and BPS also differ in a number of important respects,

---

[28]As I show in section 5, simply eliminating complex specifiers in the manner proposed by Uriagereka (1999) appears to be empirically problematic.

[29] Henceforth, I will use the term specifier/adjunct to refer to partially dominated phrases and the term complement to refer to completely dominated phrases, as a descriptive device.

however. Firstly, certain points of symmetry in BPS fail to surface in CoL. More interestingly, branching specifiers behave very differently in the two approaches. I first briefly point out the relevant similarities and differences between the two approaches before exploring their interaction with a revised version of the LCA in sections 4 and 5.

3.3   Complements

CoL and BPS make many similar predictions in relation to basic c-command relations. In (17), for example, Y asymmetrically c-commands X, under both BPS and CoL (in BPS, Y asymmetrically c-commands both X_1 and the two-segment category X_2/X_3, and in CoL it asymmetrically c-commands the single multi-segment category X):

(17) X_3 (18) see_2 (19) see_2

Y X_2 see_1 him see_1 the_2

X_1 Z_2 the_1 man

…Z_1…
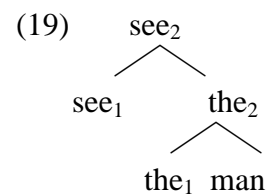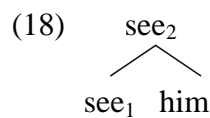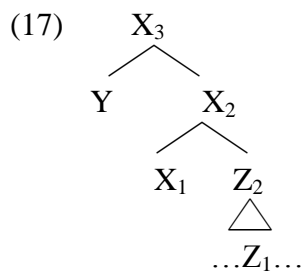
Under CoL the categories X and Y are dominated by the same set of categories (in this case the empty set). However, because X partially dominates Y (and not vice versa), Y asymmetrically c-commands X. The partial dominance caveat therefore does the same work in our definition of c-command as the exclusion proviso: it means that a specifier/adjunct asymmetrically c-commands all other categories in a projection.

X also asymmetrically c-commands a projection of Z in both BPS and CoL in (17).[30] However, whereas in BPS X mutually c-commands Z_2, in CoL, it asymmetrically c-commands the whole projection of Z. An interesting repercussion of this difference emerges where a complement is non-branching. As has often been noted in the literature, the combination of BPS and Kayne's category-based definitions of c-command/dominance creates certain points of symmetry, which (i) are potentially problematic for the LCA, and (ii) reveal a puzzling lack of systematicity in c-command relations. Chomsky (1995) discusses one point of symmetry: that holding between the bottom pair of heads in each spine, the most embedded category being necessarily non-branching.

In (18), under BPS, *him* and *see*$_1$ are dominated by the same set of categories, in this case just *see*$_2$. It follows that the two terminal categories mutually c-command each other. In (19), the problem is displaced to the bottom pair contained in the DP, *man* and *the*$_1$. As such the bottom pair problem is very general and is not obviously mitigated by evidence that V-O sequences in many languages involve movement or null functional heads.[31]

This point of symmetry fails to occur in CoL, because (i) *see*$_1$ and *see*$_2$ are both segments of the same category and (ii) complete category dominance is strongly irreflexive. As such, *see* asymmetrically dominates and therefore asymmetrically c-commands *him*.

---

[30] In CoL, this is because every category which dominates X (in this case no category) also dominates Z, but not vice versa: X dominates Z but not X.

[31] There are of course numerous ways to solve the linearization problem: movement (cf. Chomsky 1995 on clitics, Moro 2000 on small clauses), vacuous projection (Guimarães 2000), null structure, or a PF parameter (cf. Richards 2004). A version of Richards PF-head parameter will be adopted below for independent reasons. Note, however, that the existence of such a parameter is neither derived from nor reliant on mutual c-command at the bottom pair level, if this parameter applies generally to all head-complement pairs.

3.4 Specifiers contained in complements

A second, less-discussed, point of symmetry in BPS is discussed by Barrie (2005) (cf. also Hallman 2004). Barrie points out that a non-branching specifier/adjunct contained in a complement stands in a relationship of mutual c-command with a selecting head.[32] Consider the following examples:

(20) …$that_2$
    $that_1$ is_3
        he is_2
             is_1 eating

(21) $that_2$
    $that_1$ is_3
        the_2 is_2
     the_1 man is_1 eating
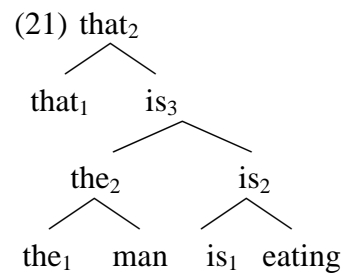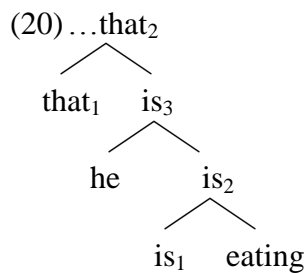
The point of symmetry in (20) arises from the fact that, in BPS, the categories *he* and *that₁* are dominated by exactly the same categories (*that₂* plus any category dominating *that₂*). Interestingly, this point of symmetry disappears if *he* is replaced with a branching category. In (21) *the₂* mutually c-commands *that₁* but *that₁* also asymmetrically c-commands *the₁* (a separate category in BPS) and *man* and as such all the terminals dominated by *the₂* can be ordered with respect to *that₁*. Under CoL, this point of symmetry again disappears. If *that₁* and *that₂* make up a two-segment category, then the multi-segment category *that* completely dominates *he* but not *that* in both (20) and (21). As such, once again a spurious point of symmetry disappears under CoL.

---

[32] Assuming that non-branching specifiers are also merged via adjunction. Interestingly, the problem also disappears if non-branching phrases are merged as specifiers rather than adjuncts, leading to the projection of an additional category.

3.5   Specifiers

In relation to complements, then, CoL resolves two of the points of symmetry which arise under BPS. With respect to specifiers, CoL, again, maintains many of the predictions of BPS, with one subtle but crucial difference.[33] Under CoL, the projecting category of a specifier c-commands into the clausal spine:

---

[33]A further way in which CoL fails to differ from BPS is that, under both kinds of labelling, a specifier c-commands out of its containing phrase. Kayne (1994: 93) gives evidence from negative operators and quantifier phrases in support of this prediction (which also holds in X-bar theory):

(i)   [Nobody's articles] ever get published fast enough.

(ii)   [Everybody's mother] loves him.

Indeed, it appears to be the case that even more deeply embedded negative specifiers c-command into the main clausal spine, as predicted by CoL (cf. also Ruys 1992):

(iv)   [Nobody's friend's articles] ever get published on time.

(v)   [Everybody$_i$'s mother's friend] loves him$_i$.

The idea that surface c-command is relevant to NPI licensing and quantifier-variable binding is, however, challenged by examples of the following kind, from Ruys (2000: 518):

(vi)   [Someone in every city] hates it.

(vii)   [Which picture of which man] pleased him?

Like Kayne, I will assume that DP-internal quantifier raising is possible in such cases. Where no such movement is possible, c-command at LF fails:

(viii)   *[the best friend of everybody's mother] hates him.

As Kayne notes, an apparent problem arises for anaphor binding, for which specifiers appear inactive:

(ix)   *John$_i$'s mother loves himself$_i$.

I assume that examples like (ix) can be ruled out by Relativised Minimality. The nominal determiner category *'s* and its specifier *John* both c-command *himself*, but *John* also asymmetrically c-commands *'s*, meaning that the latter is closer to the anaphor and acts as an intervener. By hypothesis, *'s* does not block the specifier *nobody* as a

(22)

```
            W₃
          /    \
        X₃      W₂
       /  \    /  \
      Y    X₂ W₁   V
          /  \
        X₁    Z
```

$$(22)\quad
\begin{array}{c}
W_3 \\
X_3 \quad W_2 \\
Y \quad X_2 \quad W_1 \quad V \\
X_1 \quad Z
\end{array}$$
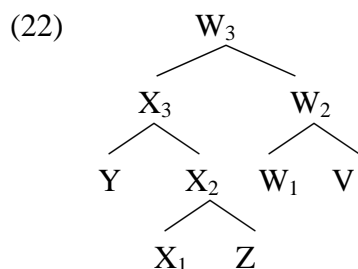
In (22), for example, X asymmetrically c-commands both W and V. This is crucially different from the situation under BPS where X_1 and X_2/X_3 constitute two distinct categories, with only the latter and not the former asymmetrically c-commanding W and V. The repercussions of this small difference are far-reaching, as discussed in section 5.

4   Revising the LCA

In recent years, the issue of a how narrow syntax is externalized has become of increasing interest. In order for languages to be acquirable, there should be some systematic (but potentially parameterized) relation between surface word order and narrow syntactic structure. Kayne's LCA takes the mapping to be universally based on asymmetric c-command (mediated by dominance), the combination of which, as he notes, has the right properties to provide a total linear order for terminal nodes. Kayne (1994, 2004) backs up his

potential negative binder. A reviewer points out that in other cases an intervening D fails to block anaphor binding:

(x)   ?Mary introduced me to herself.

Note, however, that these examples appear to require the anaphor to be focused:

(xi)   Mary introduced me to herself not to her husband/#not my husband.

They may not therefore be true anaphoric uses. The ability of specifiers to c-command out of their containing projection is an apparently unavoidable consequence of a category-based definition of c-command. Whether it is a welcome repercussion is a matter of ongoing debate.

proposal with strong empirical evidence that specifiers generally precede the phrase they are contained in (from leftwards movement, initial subjects and the lack of verb-penultimate orders etc.) and he infers from this that asymmetric c-command systematically maps to *precedence*, rather than *subsequence* at PF. In the case of the order of heads and complements, however, the empirical evidence for a cross-linguistic precedence bias is missing. Heads generally c-command their complements (unless movement has occurred), but VO and OV orders appear to occur with roughly equal frequency (Dryer 2005) and the many asymmetries cited by Kayne fail to concern the order of head and complement, as, Richards (2008) and Abels and Neeleman (to appear) note. While in some non-rigid OV languages there is good evidence that OV order is derived via narrow syntactic movement (cf. Zwart 1994 on Dutch, Haddican 2004 on Basque), this is not so clearly true of all OV languages.[34] In fact, some head-final languages present a formidable challenge for the LCA. Because of its reliance on a category-based definition of c-command, the cannot linearize multiple specifiers, creating problems for the analysis of certain well-studied head-final languages (cf. Sheehan to appear).[35] Moreover, as Richards (2008) notes, roll-up movement of the kind needed to yield head-finality creates massive redundant structure, which strongly

---

[34]Based on observations of this kind, Abels & Neeleman (2009, to appear) propose that the relevant asymmetry reduces to a ban on rightward movement. The Final-over-Final Constraint (Biberauer, Holmberg & Roberts 2007), however, suggests that the asymmetry is deeper than this, as discussed below.

[35] One potential argument in favor of Kayne's version of the LCA is that movement is independently required, and so a movement-based approach to head-finality is economical in that it does not require a head-parameter. This is not necessarily the case, though, as local comp-to-spec movement of the kind proposed by Kayne (1994) appears to be banned elsewhere (cf. Abels (2003) on anti-locality). The addition of null heads to act as the targets of snowballing movement, as per Kayne (1998) is also problematic inasmuch as these heads lack an interpretation at both interfaces.

suggests that head-finality should be marked. This is particularly true because, as Abels and Neeleman (to appear) and Richards (2008) note, the antisymmetry hypothesis alone does not constrain word order possibilities. Given the availability of (and need for) remnant movement, it is clearly possible to derive spec-final or inverse V2 orders in an antisymmetric framework. As such, the LCA can only explain typological asymmetries to the extent that some notion of markedness is associated with movement. But if this is the case then head-finality should also be marked, hence less frequent, contrary to fact. Moreover, if the narrow syntax has an LF bias as argued by Berwick & Chomsky (2011), then encoding what appears to be (at least in some cases) PF variation at the narrow syntactic level is a questionable move: while disallowing a head-parameter arguably simplifies the mapping to PF, it appears to complicate the narrow syntax and the mapping to LF. In what follows, I will therefore adopt a weaker version of the LCA whereby asymmetric c-command determines linear order only as a last resort (cf. Richards 2004, Jurka 2010 for similar ideas).

4.1   Linearizing complements

If (asymmetric) linear order is established only of the mapping to PF, then it is to be expected that it will be parasitic on pre-existing asymmetric narrow-syntactic dependencies Asymmetric c-command is one such relation, of course, but it is not the only one available. Arguably the most local asymmetric dependencies in grammar is c-selection, which is established at the maximally local level between two categories. It is therefore not unexpected that the PF component should use locally-determined c-selection dependencies to provide a linear order between two categories, where one selects the other. This is especially true as, by transitivity, locally determined precedence pairs (head parameters) of this kind will often serve to provide an unambiguous linear order without the redundancy of a system

relying on non-local asymmetric relations which massively overdetermine linear order (cf. Fox & Pesetsky 2005).

Consider the following (simplified) harmonic structures, where categories follow (subscript F) or precede (subscript P) the categories they c-select at PF, as per (23), where c-selection is an asymmetric relation between two categories:

(23) If a category A c-commands and c-selects a category B, then A precedes/follows B at PF[36]

(24)

W_F
W_F X_F
        X_F Y_F
                Y_F Z

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

(25)

W_P
W_P X_P
        X_P Y_P
                Y_P Z
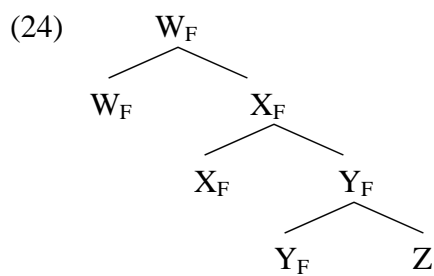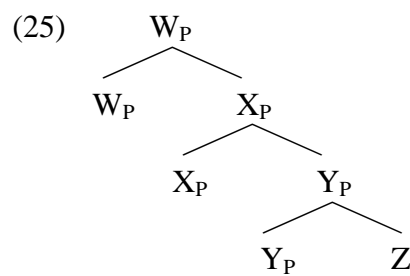
W>X, X>Y, Y>Z,
=W>X>Y>Z

In both cases, transitivity serves to give an unambiguous order of categories based on maximally local asymmetric dependencies, without redundancy. In such cases, c-command, by hypothesis, is irrelevant, and for this reason W can both asymmetrically c-command and follow Z in (24).

4.2 Linearizing specifiers

There are, however, two obvious disadvantages to using c-selection (as opposed to asymmetric c-command) as a linearization algorithm. Firstly, transitivity fails to provide an

---

[36] The proviso that A must c-command B means that the order of selection can be interrupted via movement. Richards (2004) argues that (at least in some languages) this proviso is not required, hence Holmberg's generalization (HG), which basically prevents an object from raising past its selecting category.

unambiguous order of terminals as soon as a structure is disharmonic (with mixed head-initial and head-final orders). Secondly, c-selection obviously fails to order categories which are not c-selected by any other category (i.e. specifiers/adjuncts). In such circumstances a less local asymmetry such as asymmetric c-command will be required. As such, asymmetric c-command becomes a kind of last resort linearization mechanism. I consider the case of disharmonic orders in section 5.1.6, but consider here the case of specifiers. Because specifiers are not c-selected by any category (cf. Svenonius 1994), they cannot be ordered based on the most local kind of narrow syntactic asymmetry. In such cases, asymmetric c-command must determine their linear position with respect to other categories:

(26) v_P v>hit, hit>Paul

      John v_P

          v_P hit_P

             hit_P Paul
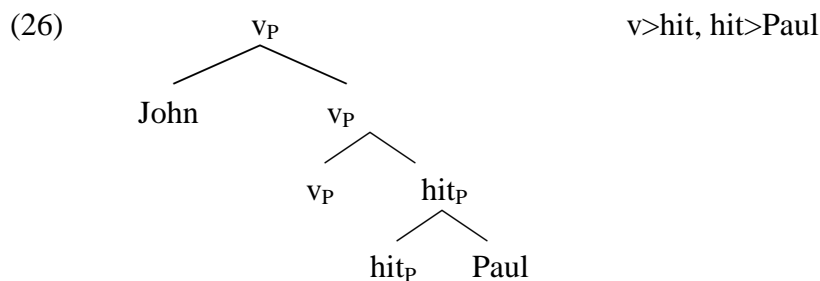
In (26), no order is specified between *John* and any other category based on c-selection, and so asymmetric c-command is required. The revised version of the two-step LCA is as follows:

(27) Revised LCA

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

    (ii) If no order is specified between A and B even transitively by (i), then A precedes B at PF if A asymmetrically c-commands B.

25

Applying (27ii) to (26) yields the following additional precedence pairs: John>v, John >hit, John>Paul.[37] The combination of (27i) and (27ii) thus serve to give the unambiguous linear order John>v>hit>Paul. Note that (27ii), unlike Kayne's LCA, contains no mention of dominance, and this will prove crucial in the discussion to follow. In the following sections, I argue that this revised version of the LCA in conjunction with CoL provides a principled account of many core word order asymmetries as simple effects of the mapping to PF.[38]

5   Exploring the implications

5.1   Specifiers and the Condition on Extraction Domain

As Uriagereka (1999) notes, the inclusion of dominance in the LCA is relevant only for the linearization of branching specifiers from a BPS perspective. He calls this the induction step of Kayne's LCA:

---

[37] It is obviously the case that even this reduced appeal to asymmetric c-command introduces redundancy as *John* can be ordered with respect to *hit* and *Paul* transitively, based on its position with respect to *v*. I leave open the possibility that the application of (ii) stops as soon as an unambiguous linear order has been calculated, avoiding redundancy.

[38] This raises the question how adjuncts fit into this system. If neither category projects, as predicted by the projection algorithm discussed above, then the two categories mutually c-command each other and cannot be ordered based on asymmetric c-command. Now if adjuncts c-select the category which they modify, as a reviewer suggests, then they would be ordered via (i) rather than (ii) of the revised LCA. The idea that c-selection is involved is supported by that fact that modifiers of V-projections (adverbs) are morphologically distinct from modifiers of N-projections (adjectives) in many languages. See Abels & Neeleman (2009) for a discussion of cross-linguistic asymmetries in adjective and adverb ordering without the LCA.

(28) A more explicit version of Kayne's LCA (based on Uriagereka 1999:252)

(a) Base step: If α asymmetrically c-commands ß, then α precedes ß.

(b) Induction step: If α precedes ß and α completely dominates γ then γ precedes ß

Interestingly, the induction step becomes untenable under CoL because the same category X can both c-command and dominate a category Y and this creates a paradox whereby Y must precede itself. Fortunately, as Uriagereka has independently argued, the induction step can readily be eliminated from the LCA, substantially increasing its explanatory power. More specifically, the removal of (b) provides a principled explanation for Huang's (1982) Condition on Extraction Domain. In the following subsections, I argue that the CoL serves to extend Uriagereka's analysis to cover a number of previously problematic exceptions to the CED.

5.1.1 Head-initial externally merged specifiers

Consider first, an externally merged head-initial specifier:[39]

(29)

---
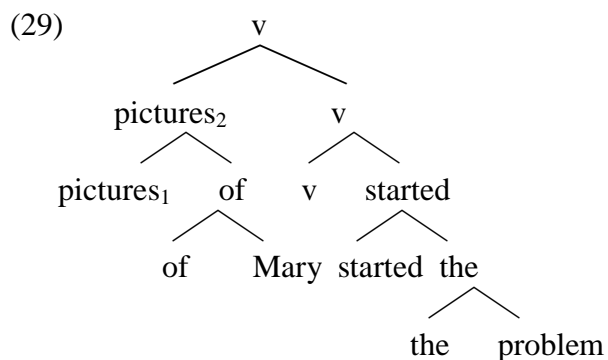
[39] I use simplified structures here for ease of exposition.

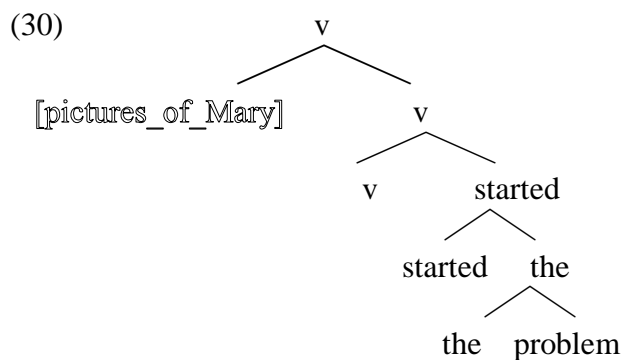Under CoL, the fact that the categories *of* and *Mary* do not enter into any c-command (or c-selection) relation with the categories *v, started, the* or *problem* means that a tree such as (29) cannot be linearized by the revised LCA in (27).[40] Following Uriagereka 1999, the fact that the simple LCA cannot order all the terminal categories in (29) can be resolved by spelling out and atomizing the branching specifier:

(30)

```
                        v
            ┌───────────┴───────────┐
   [pictures_of_Mary]               v
                            ┌────────┴────────┐
                            v              started
                      ┌─────┴─────┐    ┌──────┴──────┐
                   started     the   the        problem
```

According to the conservative version of Multiple Spell-Out (MSO) outlined by Uriagereka, in such cases, *pictures_of_Mary* behaves essentially like "a giant lexical compound", lacking in internal structure but with "syntactic terms [which] are obviously interpretable but are not accessible to movement, ellipsis, and so forth" (Uriagereka 1999:256-7).[41] In our terms this

---

[40] Although *of* and *Mary* are partially dominated by *v*, they are completely dominated by *pictures* which fails to dominate *v*.

[41] As an anonymous reviewer points out, and Uriagereka acknowledges, there are notable differences between Spelled-out chunks of the derivation and real lexical items. Lexical items cannot contain bound variables, for example, whereas complex subjects can. Another difference relates to the possibility of (re)projection. Finally, the fact that they their subconstituents can still bind variables and NPIs also raises some questions, as an anonymous reviewer points out, and suggests a potential mismatch between spell-out to the two interfaces. This idea is implicit in Huang's original formulation of the Condition on Extraction Domain as operative at surface structure only and not at LF. A full discussion of this idea would take us too far afield here, unfortunately.

means that the atomized unit behaves like a single category with respect to linearization. For this reason, as Uriagereka notes, externally merged specifiers of this kind are predicted to be strong islands. This appears to derive Chomsky's (1973) subject condition, which holds of English and many other languages:

(31) *Who did [pictures of t] start the problem?   (two step SO)

(32) Who did you see [pictures of t]?   (one step SO)

Crucially, complex phrases in a complement position, unlike those in a specifier/adjunct position, need not be atomized, as they can be ordered with respect to all other categories by the revised LCA. Assuming that adjuncts are non-complements, the fact that many branching adjuncts are strong islands reduces to the same effect (with some complications to which we return in section 5.1.9):

(33)   a. What kind of films do you like [watching t]?   (complement)

        b. *What kind of films does [watching t] annoy you?   (specifier)

        c. *?What kind of films do you eat popcorn [watching t]?   (adjunct)

Thus, as Uriagereka notes, the rejection of the induction step of the LCA serves to derive a version of Huang's (1982) Condition on Extraction Domain (CED).[42]

---

[42] Huang's CED is actually stated in terms of government and as such the predictions it makes are more nuanced than presented here, and subject to parametric variation:

(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.

29

A problem with the elimination of the induction step in conjunction with BPS (or label-free syntax for that matter) is that it predicts that all branching specifiers will be atomized and hence behave like strong islands, with the only exception being parasitic gaps (cf. Nunes & Uriagereka 2000). As Stepanov (2007) has pointed out, this prediction is too strong, as there are certain notable exceptions to the CED reported in the literature. From a CoL perspective, however, the predictions are more nuanced. Because the projecting category of a branching specifier itself c-commands into the clausal spine, atomization will only be forced where other material in said specifier cannot be ordered with respect to the clausal spine.

5.1.2 Head-final externally merged specifiers

CoL makes the prediction that harmonically head-final projections will not require atomization to be ordered with respect to categories in the clausal spine. Consider the following example:

(34)

```
              Z
          ╱       ╲
      X              Z
    ╱   ╲          ╱   ╲
  W       X      Z      …
        ╱   ╲
      Y      X_F
```
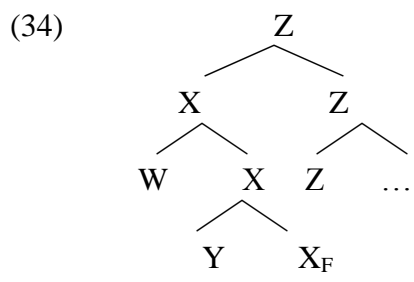
First (i) of the revised LCA specifies that the category X must follow the category that it c-commands and c-selects at PF (Y>X). Next, (ii) applies, giving the following precedence pairs based on asymmetric c-command: W>Y, W>X, W>Z, X>Z. When taken together, these precedence pairs serve to define an unambiguous linear order between all categories: W>Y>X>Z. The fact that Y fails to c-command/c-select or be c-commanded/c-selected by Z

is immaterial in this case, as an order is defined transitively between Y and Z based on independently specified precedence pairs (Y>X, X>Z). The prediction is therefore that harmonically head-final externally merged specifiers will not require atomization and hence, all else being equal, need not be strong islands.

There is suggestive empirical support for this prediction. Stepanov (2007) questions the cross-linguistic plausibility of the subject island constraint and thus the CED more generally, based on a survey of the literature. He reports that several well-studied languages appear to allow subextraction from externally merged specifiers. The overwhelming majority of these languages are agglutinating and head-final (cf. Kural 1997 on Turkish, Lasnik & Saito 1992, Saito & Fukui 1998 on Japanese, Stepanov 2007 on Navajo and Hungarian):[43]

(35) [Op_i [Ahmet-in t_i kırma-sı]-nın   beni

         Ahmet-GEN   break.INF-3SG-GEN   I-ACC

   üz-dü-ğü]   bardak

   sadden-PAST-3SG   glass   [Turkish, Kural (1997:502)]

   The glass that [Ahmet's breaking t] saddened me

---

[43] It must be acknowledged at this point that there are certain other languages which are claimed to permit subextraction from subjects, for which this explanation does not hold (e.g. Spanish, German, Russian and other Slavic languages). Note that in the case of Spanish, it appears that subextraction from subjects is possible only with VOS word order, preverbal subjects resist subextraction as predicted (cf. Gallego & Uriagereka 2007). Following Narita (2009), we might propose that in instances of VOS order it is VP which is atomized rather than the subject, leaving the latter open to subextraction (cf. section 5.1.9. for a similar approach to extraction from low adverbs). Sturgeon, Polinsky, Kravtchenko, Gallo, Medová & Koula (2010) report experimental findings suggesting that subextraction from subjects is degraded in Russian and Czech, and that subextraction is sensitive both to the complement/non-complement distinction as well as information structure and freezing.

(36) [Op$_i$ [Mary-ga   t$_i$ yonda no]-ga   akirakana   yorimo

     Mary-NOM   read C-NOM   is.obvious   than

     John-wa   takusan-no   hon-o   yonda.

     John-TOP   many-GEN   book-ACC   read

     John read more books than [that Mary read t] is obvious.

[Japanese, Stepanov (2007:89)]

According to the analysis advanced here, the crucial factor in all cases is the head-finality of the specifiers in question (cf. also Kayne 1983: fn2). Jurka, Nakao & Omaki (2011) present interesting experimental data from German and Japanese which serves as a virtual minimal pair for comparison.[44] While there is some debate in the literature as to the level of acceptability of subextraction from externally merged subjects in both of these languages (cf. Jurka 2010), there appears to be a contrast in their data between German, in which subextraction from externally merged subjects is significantly degraded when compared to subextraction from objects, and Japanese in which there is no significant subject/object

---

[44] An anonymous reviewer notes that it would also be useful to compare subextraction from head-initial and head-final phrases in VO languages. The relevant languages are difficult to find, however. It is well-known that final complementizers are generally banned in VO languages, so there are presumably no […C]-V-O languages. It might be possible to compare subextraction from head-initial and head-final DPs in VO languages but added problems arise from the fact that DPs are often strong islands in any case. In Chinese, for example, a language with head-final DPs and head-initial VPs, subextraction from DPs is generally banned, in subject and complement position. As Huang (1982: 461) notes, this means that "the effect of the subject condition cannot be directly seen in Chinese, nor is there any evidence suggesting its irrelevance in this language."

distinction.[45] This mirrors the claim in the theoretical literature that Japanese lacks subject island effects (cf. Lasnik & Saito 1992). Further empirical research is required to ascertain whether the predictions hold more robustly across a representative sample of languages. This apparent difference between head-initial and head-final specifiers, then, is predicted by CoL, as a direct result of the linearization mechanism. In the following sections I consider the predictions of CoL with respect to derived specifiers.

5.1.3 Moved specifiers

Recall that externally-merged head-initial projections must be atomic.[46] Where these specifiers move, then, they do so as a single category. Interestingly, CoL and the revised LCA provide a principled account of why moved categories are generally spelled out in their derived position. Under standard assumptions, movement reduces to the operation *copy + internal merge*.[47] If the operation copy in such contexts also gives rise to segment formation then the result is a multi-segment category which can be ordered with respect to other categories without the need for deletion (in most cases).[48] Consider the following example:

(37)   X

---

[45] Though they interpret the data rather differently.

[46] Cf. Sheehan (2010) for a more explicit characterization of the timing of atomization.

[47] A reviewer points out that this is not the case in multi-dominance approaches, which share much in spirit with the current approach: both take movement to involve a single category occupying multiple positions. Note that if labelling and movement can be unified, however, and some labelling mechanism is independently necessary then much of the motivation to model displacement via multi-dominance as opposed to segmentation falls away. Moreover, in section 5.1.7, I provide an empirical reason to favor segmentation over multi-dominance.

[48] This is not to say that deletion is impossible. LF-reconstruction, for example, might involve deletion of certain copies to re-establish earlier c-command relations.

Y_2 X
    X   Z
      Y_1 Z
         Z   …
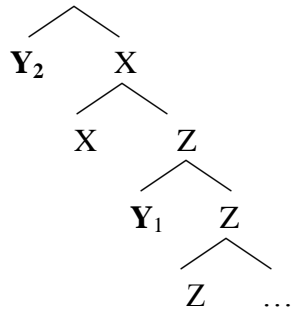
Prior to movement, the category Y asymmetrically c-commands Z and is asymmetrically c-commanded by X. After movement, without the need for deletion of Y_1, the two-segment category Y asymmetrically c-commands both X and Z, according to the definitions from section 3.2, because Y is partially dominated by X and Z is completely dominated by X.[49] Assuming that the mapping to PF is sensitive to the c-command domain of the complete multi-segment category, this provides a much needed explanation for the fact that copy+remerge generally results in leftwards movement. In our terms there is only one category involved and movement, in the form of segment formation, serves to place the entire category in a higher position.

5.1.4   Derived head-final specifiers

Matters are different with head-final phrases which need not be atomized in specifier positions, for the reasons discussed above. A head-final projection involving multiple categories can move from a complement or specifier position, yielding a derived specifier:

(38) W_3 (39) W_3

---

[49] The sum of paths from terminal copies of Y to the root includes the root copy of X twice, and so X does not completely dominate Y, but the path from every segment of Y to the root includes a copy of X. The path from Z to the root includes all non-terminal copies of X exactly once.

Tree (38):

X_6 W_2
Y_2 X_5 W_1 V_2
Z_2 X_F V_1 X_3
Y X_2
Z_1 X_F

Tree (39):

X_6 W_2
Y_2 X_5 W_1 V_3
Z_2 X_F X_3 V_2
Y X_2 V_i …
Z_1 X_F
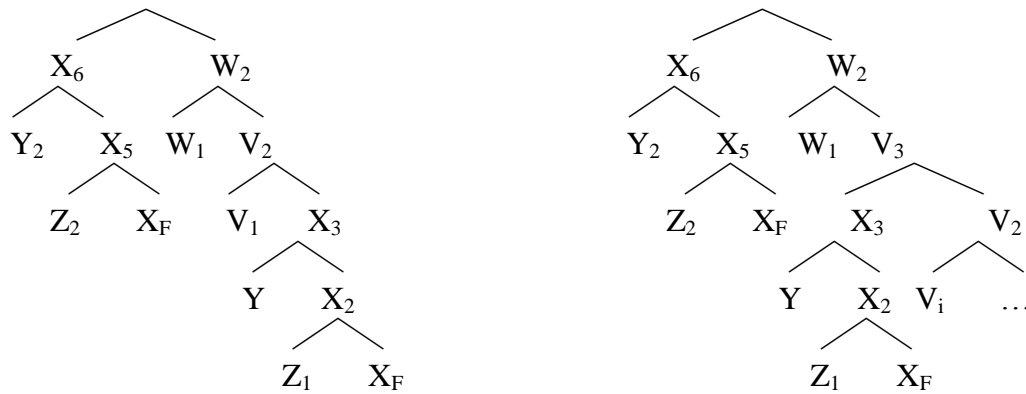
After movement of the head-final projection in (38) and (39), without deletion, X still asymmetrically c-commands Z, which it also c-selects.[50] W partially dominates X and Y and completely dominates V and so X and Y asymmetrically c-command both W and V. Finally, X partially dominates Y and completely dominates Z and so Y asymmetrically c-commands X and Y. According to the revised LCA then, the order of categories is as follows: Y>Z>X>W>V, as expected. This means that, even in the case of derived head-final specifiers, no atomization is necessarily required as an unambiguous linear order of categories can be calculated by the revised LCA. The prediction is that even derived head-final specifiers need not be strong islands, avoiding so-called freezing effects. Müller (2010) provides some German data which appear to support this prediction:

(40)   a. Was$_1$ hat es sich nicht gehört [$_{CP}$ t$_1$ zu beanstanden]?

What has it not been proper   to object.to

b. Was$_1$ hat [$_{CP}$ t$_1$ zu beanstanden] sich nicht gehört?[51]

---

[50] The paths from the two terminal copies of Z to the root of the tree include all non-terminal copies of X exactly once.

[51] I follow Müller's glossing.

What has to object.to   not been proper

Assuming that extraposed clauses in German actually occupy a complement position (cf. Bennis 1987) the examples in (40) can been taken as evidence that subextraction from derived head-final specifiers is also possible (cf. also Jurka et al. 2011: 128, fn 5 who report that this finding is replicated experimentally). It remains to be seen how general this effect is, and I leave this matter to future research.[52]

5.1.5 Derived head-initial specifiers

In the case of (non-atomic) head-initial projections, moved from a complement position, however, matters differ. In (41), the revised LCA yields the following precedence pairs:

(41)

W_3
X_6 W_2
Y_2 X_5 W_{P1} V_2
X_{P4} Z_2 V_1 X_3
Y X_2
X_{P1} Z_1
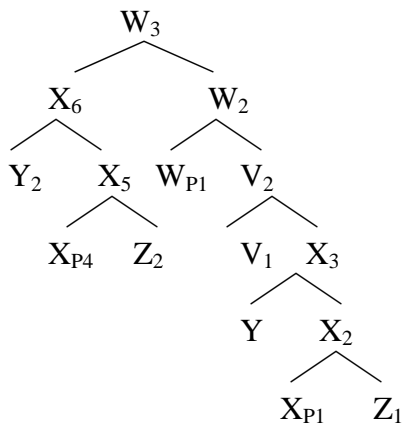
(i) By c-command and c-selection: X>Z, W>V

(ii) By asymmetric c-command: Y>X, Y>Z, Y>W, Y>V, X>W, X>V

The sum of (i) and (ii) does not, however, give a linear position for Z. While Z must follow both X and Y, neither (i) nor (ii) provides any information about the position of Z with

---

[52] Note that there are other factors which affect the availability of extraction in addition to the CED (factivity, bridge vs. non-bridge verbs, specificity, finiteness, and potentially the kinds of movement involved, Abels 2007). The prediction is *not* therefore that *all* head-final derived specifiers will permit subextraction. Rather the prediction is that movement alone will not rule out subextraction in the case of head-final projections.

respect to W and V. Whereas in (38) & (39) where the moved projection is head-final, the position of Z is given by transitivity (Z>X and X>W>V so Z must precede W and V), the same is not true of (41). Where Z follows X, in a head-initial structure, the precedence pairs provide an underspecified linear order, creating a problem for the revised LCA. In the following section, I give evidence that this prediction of CoL is empirically supported by the Final-over-Final Constraint.

5.1.6 The Final-over-Final Constraint

There is growing evidence in favor of the following word order generalization:

(42)   **The Final-over-Final Constraint** (FOFC) (based on Biberauer, Holmberg & Roberts 2007)

   A head-final phrase cannot immediately dominate a head-initial phrase.

The evidence for FOFC comes from an asymmetry in the attestation/grammaticality of disharmonic combinations in a number of different syntactic contexts in a diverse range of languages (cf. Holmberg 2000, Biberauer, Holmberg & Roberts 2007, 2008, Biberauer, Newton & Sheehan 2009, Biberauer, Newton & Sheehan 2009, Biberauer, Sheehan & Newton 2010, Biberauer & Sheehan in press). For example, Holmberg (2000) notes that while Finnish allows both V-O and O-V word orders as well as both V-Aux and Aux-V orders, the combination V-O-Aux is completely ungrammatical:

(43)   a.   Milloin Jussi **olisi**   *kirjoittanut*   ROMAANIN?   [Aux-V-O]

             when   Jussi   would.have   written   novel

      b.   *Milloin Jussi olisi*   *ROMAANIN*   *kirjoittanut?*   [Aux-O-V]

	when	Jussi	would.have		novel		written

c. *Milloin Jussi* <small>ROMAANIN</small> <u>*kirjoittanut*</u> **olisi***?*   [O-V-Aux]

	when	Jussi	novel		written		would.have

When would Jussi have written a novel?

d. *Milloin Jussi <u>*kirjoittanut*</u> <small>ROMAANIN</small>   **olisi***?*   *[V-O-Aux]

	when	Jussi	written		novel		would.have

Biberauer, Holmberg & Roberts (2007, to appear), citing Haddican (2004), show that this same gap holds synchronically in Basque as well as both synchronically and diachronically in all known varieties of Germanic. This underlines the curious thing about this generalization: it appears to be a hierarchical universal, and yet it holds even in languages with variable word order, in which the ungrammatical order (*V-O-Aux) is arguably derived via movement. As such, FOFC must be more than a constraint on the distribution of head-final triggers, as proposed by Biberauer et al. (to appear). Rather any analysis of FOFC must account for why V-O-Aux is ruled out both as a basic word order and also in languages with variable word orders such as Basque and Finnish.[53] In earlier work, I took this as evidence that all head-finality is movement-derived (cf. Sheehan 2010). It is now clear, however, that this is not necessary, as the CoL serves to rule out V-O-Aux as a surface order whether it corresponds to

---

[53] The same gap is also observed in other contexts:

(i)   *[V O] ...C   (Hawkins 1994, Kayne 1994, BHR 2007)

(ii) *[C TP] V   (Biberauer & Sheehan in press)

(iii) *[Q TP] C   (Biberauer, Newton and Sheehan 2010)

(iv) *[Asp V] T   (Julien 2002)

I focus on the *V-O-Aux case here for reasons of space.

a base-generated disharmonic order or a derived order, as long as asymmetric c-command is involved in linearization *as a last resort* (i.e. in step (ii) of the revised LCA).

The absence of V-O-Aux as a basic word order is immediately explained by the combination of the head-parameter and the revised LCA, because asymmetric c-command regulates the order of categories in disharmonic combinations. Consider the two disharmonic combinations, initial-over-final (inverse FOFC) and final over-initial (FOFC-violating).

(44) Aux_P [Aux_P , V_F [O , V_F]]

(45) * Aux [V [V , O] , Aux_F]
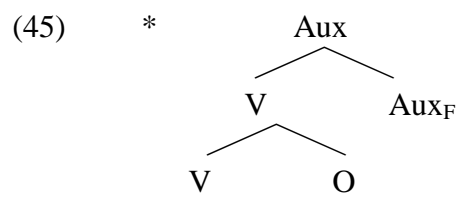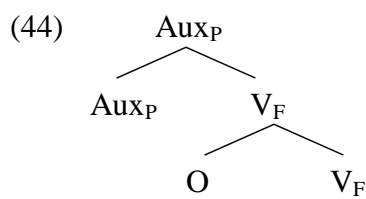
The inverse-FOFC order in (44), which is widely attested, poses no linearization problem. V c-commands and selects O and is specified to follow it, yielding O>V. Aux c-commands and c-selects V and is specified to precede it, yielding Aux>O. The sum of these two precedence pairs fails to give any order between Aux and O, and so (ii) of the revised LCA applies, yielding Aux>O. The result is the unambiguous linear order Aux>O>V.

The FOFC-violating order (45), on the other hand, is different. V c-commands and selects O and is specified to precede it, yielding V>O. Aux c-commands and c-selects V and is specified to follow it, yielding V>Aux. The sum of these two precedence pairs again fails to give any order between Aux and O, and so again (ii) of the revised LCA applies, yielding Aux>O. The linear order of such a structure is therefore specified by the revised LCA as V>Aux>O rather than V>O>Aux. The apparent prediction is that where what is descriptively a head-initial phrase is selected by a head-final phrase, the result will be discontinuous linearization of the lower phrase. As such the absence of V-O-Aux in base-generated orders

falls out as an effect of the mapping to PF, rather than a fact about narrow syntax. This is a welcome result if linear order is a property only of PF.

This explains why VOAux is unattested as a basic word order, but it does not explain why it is ruled out as a derived order, where movement has taken place.[54] The basic order in Finnish is SVO, and there is good evidence that Aux-final order is derived via phrasal movement (cf. Holmberg 2001). What also needs to be ruled out, then, is the linearization of a head-initial VP in a specifier position above the position of the auxiliary, which would give derived V-O-Aux. Interestingly, this is precisely the kind of structure which creates a linearization problem in (41), because the complement of the derived specifier cannot be ordered with respect to the clausal spine. Crucially, the word order combinations O-V-Aux and Aux-O-V, even if they are movement-derived, present no such linearization challenge. CoL thus predicts V-O-Aux to be ruled out as the surface order even with derived word orders.[55]

An interesting question is what happens to resolve this problem in such contexts. The answer appears to be that extraposition can often occur as a repair strategy, and the problematic category is spelled out in final position. Thus while V-O-Aux surface order is not a possible linearization of [[V-O]-Aux], V-Aux-O is:

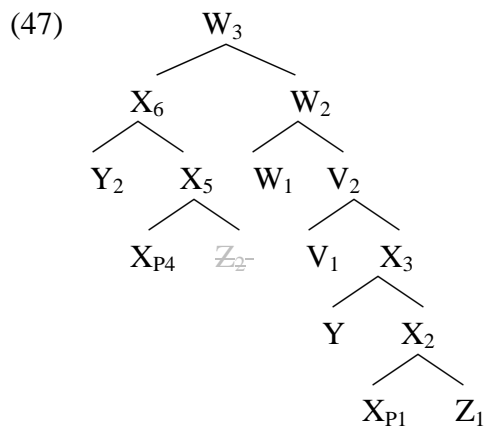(46)   Milloin   Jussi kirjoittanut olisi   romaanin?

---

[54] It also raises the question whether V-Aux-O is attested. Where this order exists as a basic word order it is likely that Aux is analysed as a suffix rather than an auxiliary. Below, I give examples where it occurs as a secondary word order.

[55] Although this provides a non-uniform account of the asymmetry, in both cases the crucial effect comes from the combination of the fact that (a) asymmetric c-command is involved in linearization and (b) terminal nodes are part of the same category as their projections.

When	Jussi	written	would.have	novel

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

If V-Aux order is derived via VP-movement, as Holmberg suggests, then V-Aux-O order can plausibly be seen as the result of the movement of a head-initial VP to the specifier of a higher functional head.[56] The VP is then linearized discontinuously, giving rise to object extraposition. This discontinuous linearization, I assume, stems from the deletion of the derived copy of the problematic complement, after (i) of the LCA has applied:

(47)

$$
\begin{array}{c}
W_3 \\
X_6 \quad W_2 \\
Y_2 \quad X_5 \quad W_1 \quad V_2 \\
X_{P4} \quad \sout{Z_2} \quad V_1 \quad X_3 \\
Y \quad X_2 \\
X_{P1} \quad Z_1
\end{array}
$$

This deletion operation takes place as a last resort in order to permit unambiguous linearization via the revised LCA. After Z_2 has been deleted, the category Z is asymmetrically c-commanded by W and V and so must follow them giving the unambiguous linear order Y>X>W>V>Z. Note that deletion of the higher copy is forced, as deletion of Z_1 would give rise to an underdetermined order of categories, failing to aid linearization.[57] The FOFC asymmetry therefore provides further evidence in favor of the CoL.

---

[56] Actually, Holmberg assumes vP movement where V has raised to v. This makes no difference to the analysis.

[57] If Z_1 were deleted, then Z would still fail to enter into a c-command relation with W and V.

5.1.7 Subextraction from derived head-initial specifiers

The fact that derived specifiers do not require atomization, and rather can be linearized discontinuously potentially explains the fact that they permit subextraction much more readily than externally merged specifiers in English, as noted by Kuno (1973) (cf. also Chomsky 2008 for a recent discussion). Consider the following:

(48)   a. ?Of which celebrity have several pictures appeared recently?

b. ?About which topic do several books seem to have been borrowed?

Failure to pied-pipe the preposition from its post-nominal position leads to degradation:

(49)   a. *?Which celebrity have several pictures of appeared recently?

b. *Which topic do several books about seem to have been borrowed?

Interestingly, though, prepositions can less marginally be stranded in an extraposed position, and subextraction from extraposed complements is often possible for some speakers:

(50)   a. From which types of clauses has extraction taken place?

b. ?Which types of clauses has extraction taken place from?

c. *Which types of clauses has extraction from taken place?

(51)   a. Who has a chance arisen [for you to visit t]?

b. *?Who has [a chance for you to visit t] arisen?[58]

This follows if complement extraposition occurs in exactly those contexts where a head-initial derived specifier is *not* atomized (i.e. where it has raised from a complement position). In such contexts, subextraction will remain possible and extraposition will be forced (cf. Sheehan 2011 for further discussion).

5.1.8 Other kinds of phrasal movement

As noted by an anonymous reviewer, an apparent prediction of the CoL is that movement of head-initial phrases from a complement position will always result in extraposition.[59] This prediction is apparently challenged by simple examples such as the following:

(52) [that we would arrive in good time] was hoped for.

(53) [This/a beautiful picture of Mary] has been published

(54) …and [eat the cake] I shall!

---

[58] One possible interpretation of these data is that PPs which permit extraction/subextraction are not actually modifiers of DP but are rather modifiers of VP (cf. Bach & Horn 1976). This cannot be the case in the following example, though, as subextraction is sensitive to definiteness effects (Davies & Dubinsky 2003, Sheehan 2011):

   (i)   *?Of which celebrity have your new pictures appeared?

   (ii) *?About which topic do all John's books seem to have been borrowed?

In cases where PP is a VP modifier, no definiteness effects obtain:

   (iii) About which topic was your best book written?

[59] Note that the prediction does not extend to head-initial phrase which move from a specifier position, as these will necessarily be atomized.

There are at least three possible explanations for such examples: either (i) the specifiers are not actually raised from a complement position, but rather externally merged in a specifier position; or (ii) these phrases are atomized prior to movement for independent reasons, making them strong islands; or (iii) some PF condition blocks scattered deletion.

The CP in example (52) falls into the first category according to Alrenga (2005), as it is based generated as a topic, rather than being moved from a complement position. Evidence for this position comes from the fact that with CP topics, a DP occupies the thematic position:

(55) a. I wish (*for) that the giants would win the World Series.

b. That the giants would win the World Series was wished *(for)

(example from Alrenga 2005:183, but the observation goes back to Higgins 1973)[60]

Alrenga takes this as evidence that a null DP satisfies the thematic requirements of the verb and the CP is externally merged as a topic (in a specifier position).

Sheehan (2011) proposes that (53) is an example of (b) as specific DPs are generally strong islands even in complement positions in English, suggesting they are always atomized. Note that where the nominal subject is indefinite, it appears to allow optional complement extraposition. In the terms outlined here, this is because non-specific nominals permit subextraction when in a complement position:

(56) A beautiful picture has been published of Mary.

(57) Who did you see a/*the picture of?

Sheehan (2011) provides an account of these facts based on a structural difference between non-specific NumPs and optionally specific DPs, whereby the latter but not the former are atomized.

Other phrases behave differently. VP, for example, when topicalized, does not permit extraposition. I assume that a PF principle which requires the stressed auxiliary to be final rules out discontinuous linearization in such contexts. Although material can follow the stressed auxiliary, it is strongly destressed and separated from the auxiliary by a pause. A full understanding of the distribution of extraposition with other categories is the subject of ongoing research.[61]

5.1.9   Extraction from low adverbials

Thus far it has been argued that the more nuanced predictions of the CoL and the revised LCA are empirically supported inasmuch as head-final specifiers are not always strong islands, and head-initial specifiers trigger FOFC-effects. A final apparent problem for the CED is that, as a reviewer points out, a subclass of post-VP adjuncts remains open to subextraction (cf. Kayne 1983, Manzini 1992, Truswell 2007):[62, 63]

---

[61] As anonymous reviewer points out, even with the categories which permit extraposition/scattered deletion, it is severely constrained. There appear to be PF constraints on the effect. The extraposed material must be stressed, so the presence of a stressed postposed adverb prevents scattered deletion:

   (i)   Which friend did John meet of Mary's?

   (ii)   Which friend did John meet of Mary's *unexpectedly/this morning?

If scattered deletion is a PF process, it is unsurprising that it is sensitive to PF constraints.

[62] The discussion here is limited to English, but subextraction from adjuncts appears also to be possible in other Germanic languages as well as Italian (cf. Cinque 1990) and Mandarin (cf. Huang 1982: 480).

[63] A reviewer raises also example of the following kind from Abels (2007):

   (i)   ?Which football team do you want the manager of to pay a fine/be punished/buy new players?

(58) a. What temperature should I wash my jeans [at]?

b. Which room does Karen teach her class [in]?

c. *Which way does climate change affect the weather [in]?

d. *What extent is Google Earth useful [to]?

(59) a. Which play did you fall asleep watching/*before watching/??after watching?

b. Which man did you return home without talking to/in order to talk to/*because you had talked to?

There are a number of factors affecting the grammaticality of this subextraction. In the case of PP adverbials, subextraction appears to be dependent on the function of PP: measure/locative/rationale/temporal PPs often allow subextraction (58a-b), but manner/extent

---

These examples are not straightforward examples of subextraction from subjects, though, as they clearly dependent on the possibility of a semantic gap in the main clause, thus they have the following meaning (where the preposition is optionally not pronounced):

    (ii) Which football team do you want the manager of to pay a large fine FOR/BECAUSE OF?

As the reviewer also notes, the gap in the main clause with a null preposition is also parasitic on the gap in the ECM subject, so we have a case of symbiotic gaps, as it were:

    (iii) *Which football team do you want John to pay a fine/be punished/buy new players?

Nonetheless, the fact that the subject-internal gap cannot be a PP, and is blocked where no gap in the main clause is permitted, strongly suggest that the subject gap in (i) is parasitic on the main clause gap:

    (iv) *Of which football team do you want the manager to pay a large fine?

    (v) *Which football team do you want the manager of to pay a large fine for speeding/his wife?

    (vi) *Which football team do you want the manager of to be invited?

However these examples are to be analysed then, they do not seem to be cases of simple extraction from externally merged subjects, and I leave them aside here.

PPs (which can be replaced by *how/how much*) seem to disallow it. (58c-d). Where the adjunct is clausal, tense appears to play a crucial role, as Manzini (1992:29) notes. Truswell (2007) proposes that there is also a semantic constraint on subextraction the single event condition, which requires that the adjunct "occupy an event position in the argument structure of the main clause verb". In relation to the revised LCA, though, the fact that subextraction from adjuncts is possible at all remains problematic even if this subextraction is constrained by other semantic/syntactic factors.

A possible account of these is sketched in Sheehan (2010) relying on the idea that in instances of adverbial modification either the adverbial or the main clause can undergo atomization. This follows if neither phrase projects and so the phrases stand in a symmetric relation (cf. Narita to appear for a similar idea). Evidence from the interaction of A and A-bar movement in such contexts supports this proposal. Thus both low adverbials and the phrases they modify are potentially open to A-bar subextraction:

(60) Which exam did you give John some tuition [(in order) to help him pass t]?

(61) What did you give John t [(in order) to help him pass the exam]?

Condition C effects show that where an adverbial has been targeted for wh-extraction, it is adjoined below the indirect object (to $P_{have}P$ in Ramchand's (2008) terms):

(62) How much tuition did she [$_{resP}$ give him$_1$ t] [to help John$_1$ pass the exam]?

(63) *Which exam did she give him$_1$ [$_{PhaveP}$ some tuition] [to help John$_1$ pass t]]?

As predicted, where the adverbial has been the target of A-bar extraction, A-movement from $P_{have}P$ is blocked (in those dialects which otherwise allow it), suggesting $P_{have}P$ is atomized:[64]

(64) *[Which exam]$_i$ was [some tutoring]$_j$ given John [$_{PhaveP}$ t$_j$] [in order to help him pass t$_i$]?

(65) [Which exam]$_i$ was John$_j$ given t$_j$ [$_{PhaveP}$ some tutoring] [in order to help him pass t$_i$]?

Although many complications remain, notably the additional restrictions governing subextraction from adverbials and the differences in adjunction sites, it would appear that there is the potential for an account of these facts consistent with the MSO approach which makes testable predictions.

6   Conclusions

It has been argued that a segment-based copy theory of labeling and movement coupled with a minimally revised version of Kayne's category-based definition of c-command leads to a viable model of phrase structure. The single operation *copy* serves to generate labels and to establish movement operations leading to multi-segment categories over which c-command relations are defined. The result is a system in which no deletion is required (except as a last resort) and in which the PF preference to target the derived position of a phrase follows from the extension condition without stipulation. This model serves to explain possibilities of subextraction from externally and internally merged specifiers as well as a typological word order gap and patterns of extraposition.

---

[64] Thanks to Marcel Den Dikken for pointing out the possibility of this test to me.

One drawback of the approach is the apparently complex definitions of dominance and c-command which it requires. While it is not clear at present how CoL can be aligned with simpler definitions of c-command, it is important to note that it is not crucially tied to any specific definitions. What the CoL allows us to do is vastly simplify the system of labeling and movement, define the notions of specifier and complement in terms of dominance and use them to explain a wide range of word order asymmetries. It is hoped that the definitions necessary to implement this idea can be distilled further in future work, without the loss of empirical coverage.

References

Abels, K. 2003. Successive cyclicity, anti-locality, and adposition stranding, PhD dissertation, University of Conneticut.

Abels, K. 2010. Anti-locality, snowballing movement, and their relation to a theory of word order. Paper presented at *GIST 1*, Ghent, Belgium.

Abels, K. & Neeleman, A. 2009. Universal 20 without the LCA . In *Merging features*, eds. J.M. Brucart, A. Gavarro and J. Solá, 60-80. Oxford: Oxford University Press.

Abels, K. & Neeleman, A. to appear. Linear asymmetries and the LCA, *Syntax*.

Adger, D. 2011. Labels and structures, ms., Queen Mary, London.

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

Bach, E. & Horn, G. M. 1976. Remarks on "Conditions on transformations" , *Linguistic Inquiry*, 7 (2): 265-99.

Barrie, M. 2005. On unifying antisymmetry and bare phrase structure. In *Proceedings of the 35th annual meeting of the North East Linguistic Society*, eds. L. Bateman and C. Ussery, 1-12. University of Connecticut: BookSurge Publishing.

Bennis, H. 1987. *Gaps and Dummies*, Dordrecht: Foris.

Berwick, R. C. & Chomsky, N. 2011. The biolinguistic program: the current state of its evolution and development. In *The biolinguistic enterprise*, eds. A.M. Di Sciullo and C. Boeckx, 19-41. Oxford: Oxford University Press.

Biberauer, T., Holmberg, A. & Roberts, I. 2007. Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC). In *XXXIII Incontro di Grammatica Generativa*, eds. A. Bisetto and F. Barbieri, 86-105.

Biberauer, T., Holmberg, A. & Roberts, I. to appear. A syntactic universal and its consequences, *Linguistic Inquiry*.

Biberauer, T., Newton, G. & Sheehan, M. 2009. Limiting synchronic and diachronic variation and change: the Final-Over-Final Constraint, *Language and Linguistics*, 10 (4): 701-43.

Biberauer, T. & Sheehan, M. in press. Disharmony, Antisymmetry, and the Final-over-Final Constraint. In *Ways of structure building*, eds. M. Uribe-Etxebarria and V. Valmala. Oxford: Oxford University Press.

Biberauer, T., Sheehan, M. & Newton, G. 2010. Impossible changes and impossible borrowings: the Final-over-Final Constraint. In *Continuity and change in grammar*, eds. A. Breitbarth, C. Lucas, S. Watts and D. Willis, 35-60. Amsterdam: John Benjamins.

Boeckx, C. 2008. *Bare syntax.* Oxford: Oxford University Press.

Cecchetto, C. & Donati, C. 2010. On labeling: Principle C and head movement, *Syntax*, 13 (2): 241-278.

Chametzky, R. 1994. Chomsky-adjunction, *Lingua*, 93 (4): 245-64.

Chomsky, N. 1986. *Barriers.* Cambridge, Mass.: MIT Press.

Chomsky, N. 1995. Bare phrase structure. In *Government and Binding Theory and the Minimalist Program*, ed. G. Webelhuth, 383–439. Oxford: Blackwell.

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

Chomsky, N. 2006. Approaching UG from Below. In *Interfaces+ Recursion= Language?*, eds. U. Sauerland and H.M. Gärtner, 1-29. New York: Mouton de Gruyter.

Chomsky, N. 2008. On Phases. In *Foundational Issues in Linguistic Theory*, eds. R. Freidin, C.P. Otero and M.L. Zubizarreta, 133-66. Cambridge, Mass.: MIT Press.

Cinque, G. 1990. *Types of A-bar dependencies.* Cambridge, Mass.: MIT Press.

Citko, B. 2008. Missing labels, *Lingua*, 118: 907–44.

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

Davies, W., D. & Dubinsky, S. 2003. On extraction from NPs, *Natural Language and Linguistic Theory*, 21: 1-37.

Dryer, M. 2005. Order of verb and object. In *The world atlas of language structures*, eds. M. Haspelmath, M. Dryer, D. Gill and B. Comrie, 338-41. Oxford: Oxford University Press.

Epstein, S. D. 1999. Un-principled syntax and the derivation of syntactic relations. In *Working Minimalism*, eds. S.D. Epstein and N. Hornstein, 315-345. Cambridge, Mass.: MIT Press.

Fox, D. & Pesetsky, D. 2005. Cyclic linearization of syntactic structure, *Theoretical Linguistics*, 31 (1-2): 1-45.

Gallego, A. & Uriagereka, J. 2007. Conditions on Sub-extraction. In *Coreference, modality, and focus: studies on the syntax-semantics interface* eds. L. Eguren and O.F. Soriano, 45-70. Amsterdam: John Benjamins.

Guimarães, M. 2000. In defense of vacuous projections in bare phrase structure, *University of Maryland Working Papers in Linguistics*, 9: 90-115.

Guimarães, M. 2008. A note on the strong generative capacity of standard Antisymmetry Theory, *Snippets*, 18: 5-7.

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

Hale, K. L. & Keyser, S. J. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, Mass.: MIT Press.

Hallman, P. 2004. Symmetry in Structure Building, *Syntax*, 7 (1): 79-100.

Hawkins, J. A. 1994. *A performance theory of order and constituency*, Cambridge: Cambridge University Press.

Hendrick, R. 2007. Syntactic labels and their derivations. In *Phrasal and Clausal Architecture: Syntactic derivation and interpretation*, eds. S. Karimi, V. Samiian and W.K. Wilkins, 93-107. Amsterdam: John Benjamins.

Higgins, F. R. 1973. On J. Emonds' Analysis of Extraposition. In *Syntax and Semantics*, ed. J.P. Kimball: 395-420 New York.

Holmberg, A. 2000. Deriving OV Order in Finnish. In *The Derivation of VO and OV*, ed. P. Svenonius, 123-52. Amsterdam: John Benjamins.

Holmberg, A. 2001. The syntax of yes and no in Finnish, *Studia Linguistica*, 55 (2): 140-74.

Hornstein, N. 2009. *A theory of syntax: minimal operations and universal grammar*, Cambridge: Cambridge University Press.

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

Irurtzun, A. 2007. The gammar of focus at the iterfaces, Ph.D dissertation, The University of the Basque Country.

Jackendoff, R. S. 1977. *X-bar syntax: a study of phrase structure*, Cambridge, Mass.: MIT Press.

Julien, M. 2002. *Syntactic heads and word formation.* Oxford: Oxford University Press.

Jurka, J. 2010. The importance of being a complement: CED-effects revisited, Ph.D dissertation, Maryland.

Jurka, J., Nakao, C. & Omaki, A. 2011. It's not the end of the CED as we know it: Revisiting German and Japanese subject islands. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, eds. M.B. Washburn, K. McKinney-Bock, E. Varis, A. Sawyer and B. Tomaszewicz, 124-32. Somerville, MA: Cascadilla Proceedings Project.

Kayne, R. 1983. Connectedness, *Linguistic Inquiry*, 14 (2): 223-49.

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

Kayne, R. 1998. Overt vs. covert movement, *Syntax*, 1 (2): 128-91.

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

Kuno, S. 1973. Constraints on Internal Clauses and Sentential Subjects, *Linguistic Inquiry*, 4 (3): 363-85.

Kural, M. 1997. Postverbal Constituents in Turkish and the Linear Correspondence Axiom, *Linguistic Inquiry*, 28 (3): 498-519.

Lasnik, H. & Saito, M. 1992. *Move alpha.* Cambridge, Mass.: MIT Press.

Manzini, M. R. 1992. *Locality.* Cambridge, Mass.: MIT Press.

Matushansky, O. 2006. Head movement in linguistic theory, *Linguistic Inquiry*, 37 (1): 69-109.

May, R. 1985. *Logical form: its structure and derivation*. Cambridge, Mass.: MIT Press.

Moro, A. 2000. *Dynamic antisymmetry.* Cambridge, Mass.: MIT Press.

Müller, G. 2010. On deriving CED effects from the PIC, *Linguistic Inquiry*, 41 (1): 35-82.

Narita, H. 2009. Cyclic transfer in service of recursive merge. Ms., Harvard University.

Neeleman, A. & Van de Koot, H. 2002. The configurational matrix, *Linguistic Inquiry*, 33: 529-74.

Neeleman, A. & Van de Koot, H. 2010. A local encoding of syntactic dependencies and its consequences for the theory of movement, *Syntax*, 13 (4): 331–72.

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

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

Ramchand, G. 2008. *Verb meaning and the lexicon.* Cambridge: Cambridge University Press.

Reinhart, T. 1981. Definite NP anaphora and c-command domains , *Linguistic Inquiry*, 12 (4): 605-35.

Richards, M. 2004. Object shift and scrambling in North and West Germanic: a case study in symmetrical syntax, Ph.D dissertation, Cambridge University.

Richards, M. 2008. Desymmetrization: parametric variation at the PF-interface, *The Canadian Journal of Linguistics*, 53 (2/3): 275-300.

Ruys, E. 1992. The scope of indefinites. Ph.D dissertation, Utrecht University.

Ruys, E. 2000. Weak crossover as a scope phenomenon, *Linguistic Inquiry*, 31 (5): 513-39.

Saito, M. & Fukui, N. 1998. order in phrase structure and movement, *Linguistic Inquiry*, 29 (3): 439-74.

Seely, D. 2006. Merge, derivational c-command, and subcategorization. In *Minimalist Essays*, ed. C. Boeckx, 182–217. Amsterdam: John Benjamins.

Sheehan, M. 2010. The resuscitation of CED. To appear in *Proceedings of the 40th annual meeting of the North East Linguistic Society*.

Sheehan, M. 2011. Extraposition and Antisymmetry. In *Linguistic Variation Yearbook 10 (2010),* ed. J. Van Craenenbroeck: 203-254. Amsterdam: John Benjamins.

Sheehan, M. to appear. Explaining the Final-over-Final Constraint: formal and functional approaches. To appear in *Theoretical approaches to disharmonic word orders,* ed. T. Biberauer & M. Sheehan. Oxford: Oxford University Press.

Speas, M. 1990. *Phrase structure in natural language*, Dordrecht: Kluwer.

Stepanov, A. 2007. The end of CED? *Syntax*, 10 (1): 80-126.

Sturgeon, A., Polinsky, M., Kravtchenko, E., Gallo, C. G., Medová, L. & Koula, V. 2010. Subject islands in Slavic: the syntactic position matters. Paper presented at *Formal Approaches to Slavic Linguistics 19*.

Svenonius, P. 1994. C-selection as feature checking, *Studia Linguistica*, 48 (2): 133-155.

Truswell, R. 2007. Extraction from adjuncts and the structure of events, *Lingua*, 117: 1355-1377.

Wilder, C. 2008. Shared constituents and linearization. In *Topics in ellipsis*, ed. K. Johnson: 229-58. Cambridge: Cambridge University Press.

Zwart, C. J.-W. 1994. Dutch is Head Initial, *The Linguistic Review*, 11 (3/4): 377-406.

Michelle L. Sheehan

Department of Linguistics

University of Cambridge

Sidgwick Avenue

Cambridge, CB3 9DA

United Kingdom

mls54@cam.ac.uk