

The Antisymmetry of Syntax

Kayne (1994)

January 15, 2009

1 Core Ideas, Motivations and Relevance

The syntax of natural languages exhibits both a horizontal and a vertical dimension. The former corresponds to precedence relations of terminals in a sentence and the latter to hierarchical relations of nodes. What is the relationship between these dimensions? For a long time it was – more or less implicitly – assumed that this relationship is rather loose, or accidental. For example, both in English and German the IP layer dominates the VP. Yet in English elements hosted by I^0 precede VP-internal material in the surface string, while standard theories of German have it that I^0 follows VP-internal material in German.

Contrary to this picture Kayne (1994) claimed that hierarchical structure and the linear precedence relation of elements in a sentence are linked: He proposed that surface order reflects hierarchical order and, conversely, that hierarchical relations map onto precedence of elements in a sentence. This idea is captured in the Linear Correspondence Axiom.

To illustrate the shift of perspective that takes place with antisymmetry, let's consider a simple question: Why do internal arguments follow predicates in some languages (such as English), while they precede predicates in their canonical order in others (like German)? Early Principles and Parameters theory attributed this difference to the head parameter. Antisymmetry gives a different answer: in a language like German internal arguments obligatorily undergo a covert licensing movement to a position that is higher in the tree than the verb('s surface) position. Another question concerns the mixed-headedness of languages like German. Why is the V^0 - and the I^0 -domain head-final while the C^0 -domain is head-initial? Under the head-parameter approach a different value (either left or right) for the instantiation of each X-bar projection (both functional and lexical) must be assumed. Under antisymmetry the answer is simple: there is no mixed-headedness: Wherever a head superficially follows its complement, this order must be the result of movement of the complement into a position that c-commands the head¹.

The main features of Antisymmetry are:

- restrictive theory of syntactic structures UG provides (in line with the historical tradition of restrictive theories pertaining to structure, such as X-bar (Chomsky (1970), Jackendoff (1977), binary branching (Kayne (1984)), etc.)
- Question: why does the language faculty display the particular ordering that it does? Answer: the relationship between hierarchical structure and linear order is inflexible. The universal order is S-H-C, and the hierarchical order of syntactic structures is mapped to precedence by asymmetric

¹In and of itself the term “parameter“ does not seem to have an explanatory value but shows a direction to a solution to the question why complements follow their heads in some languages and precede them in others. One of the core ideas of late GB is that parameters are associated with functional heads, i.e. that differences between languages reduce to morphology. I find it striking that, on the one hand, syntacticians readily accepted replacing, for example, something that must have been the “*Wh*-parameter” with *Wh*-features to account for the difference between *Wh*-in-situ languages and *Wh*-movement languages. On the other hand, there does not seem to be an equal willingness to abolish the head parameter (at least, amongst most syntacticians dealing with German).

One might argue that the head parameter is a simpler concept than postulating an additional functional layer as a landing site for arguments to the verb, plus having an additional instance of (covert) movement. However, the latter approach is consistent with the overall idea of late GB, while the former isn't (Thanks to Neven for an inspiring discussion on this topic on the train.).

c-command (roughly: "the higher in the tree a node sits, the earlier it is spelled out and conversely: the deeper embedded it is, the later its spell-out").

- specifiers are instances of adjunction
- there is no right-adjunction and -movement (by this theory, the former violates the ECP)
- X-bar theory is derived, not a primitive of UG; moreover, there is no head-parameter. With the rise of Minimalism there have been a number of attempts to reconcile Antisymmetry with Chomsky's Bare Phrase Structure theory. Getting in touch with the former prior to the latter might be a good exercise.
- influential in late GB/Minimalism. The consequences, possibilities and (often new) problems the theory entails have been explored in many languages aside from English (for German, West Flemish, and Dutch see Zwart (1993), Hinterhölzl (2006); for Swiss German and Greek, see Leu (2008); for French, see Kayne and Pollock (2008); for Italian, see the 1994-monograph by Kayne). In more recent versions of antisymmetry (especially those of Kayne, but also Leu (2008) and others) a number of silent elements have been proposed. Essentially, they resemble familiar silent elements, such as PRO/pro, operators, zero morphemes, ellipses, etc., i.e. they have a semantic (and sometimes morpho-syntactic function), but are phonologically realized in other languages (in fact, this is the prime motivation for their postulation, each of which carries a heavy empirical burden). Moreover, there are lots of movement processes in antisymmetric derivations, often with remnant movement². The good news is these derivations are cross-linguistically parallel. Pollock's famous split-INFL idea from the late 1980s foreshadowed the idea that has become one of the core antisymmetric tenets, namely that there is a Universal Base.
- According to Kayne, the LCA might pave the way to answering the following elementary question: "why are there so many functional heads?" (p. 29) There are functional heads which seem to serve no purpose in certain languages, which make functional heads look dubious: "It is evident from languages like Chinese that overt agreement can be completely dispensed with." The LCA "provides at least a partial answer to this question. Assuming that phrases of various kinds must move out of their base position at some point in the derivation, the answer is that functional heads make landing sites available. Spec-head configurations are used for licensing for a principled and simple reason: there is no other possibility. Given that double adjunction to the same projection is prohibited, there must, for every moved phrase, be a distinct head to whose projection it can adjoin as a specifier." (p. 29-30)

There is a continuum between syntacticians who straightforwardly reject Antisymmetry on the one end, and die-hard Kaynians on the other. Some Minimalists seek a moderate path in between (say, by proposing that the LCA applies at the PF-branch only). In what follows I will stick to the 1994-version of antisymmetry for expository purpose.

2 The Linear Correspondence Axiom (LCA)

A linear order has three properties. It is

transitive $xLy \ \& \ yLz \rightarrow xLz$.

total for all distinct x, y , either xLy or yLx .

antisymmetric $\text{not}(xLy \ \& \ yLx)$.

The dominance relation, by contrast, is not total: there can be two nodes in a given phrase marker such that neither dominates the other. For example, a head does not dominate its complement and vice versa. Yet there is a further commonality between a linear order and a dominance relation apart

²Remnant movement is the movement of a constituent containing the trace of a second constituent that has been moved out of the first constituent before: $[_{XP} X \ t_{YP}] \dots YP \dots t_{XP}$

from antisymmetry and transitivity which Kayne calls *locally total*: Consider a given nonterminal X in a phrase marker, and then the set of nonterminals that dominate X. For any Y that dominates X and Z that dominates X it holds that either Y dominates Z or Z dominates Y. "Although the dominance relation itself is not total, it becomes total when restricted to the set of nodes dominating a given node." (p. 4) Thus in this local sense dominance is total just as linear order is. As a result there is no reason to dissimilate hierarchical relations from precedence relations - both have (locally) the same properties. The relation of c-command must be slightly modified so as to satisfy antisymmetry:

ansymmetric c-command X asymmetrically c-commands Y iff X c-commands Y and Y does not c-command X.

This relation is also locally total, because there is only binary branching: for every X that asymmetrically c-commands Y, any Z that asymmetrically c-commands Y must also asymmetrically c-command X or X must asymmetrically c-command Y.

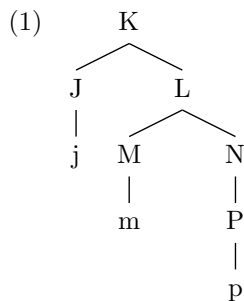
- D is the dominance relation between nonterminals
- d is the dominance relation from nonterminals to terminals
- for a given nonterminal X, $d(X)$ is the set of terminals that X dominates (d 's "image"). This also holds for ordered pairs $\langle X, Y \rangle$. The image of $d\langle X, Y \rangle$ is based on $d(X)$ and $d(Y)$, i.e. the Cartesian product of the set of terminals we get through the mapping of X and the set of terminals we get through the mapping of Y.
- if we say that A is the maximal such set, it contains all pairs of non-terminals (with the first element asymmetrically c-commanding the second)

For a given phrase marker with T the set of terminals and A as just given:

The Linear Correspondence Axiom :

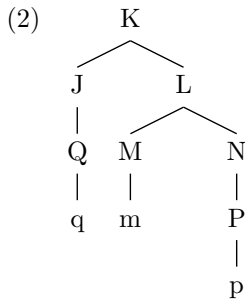
$d(A)$ is a linear ordering of T

3 A Bunch of Examples



In this tree A is $\langle J, M \rangle$, $\langle J, N \rangle$, $\langle J, P \rangle$, $\langle M, P \rangle$. $d(A)$ is: $\langle j, m \rangle$, $\langle j, p \rangle$, $\langle m, p \rangle$. These three tuples determine the linear order of the elements in the set $\{j, m, p\}$ (which, by themselves, are unordered) since the tuples guarantee (1) transitivity (2) totality and (3) antisymmetry. Specifically, j must precede m and p , and m must precede p .

Specifiers and adjuncts don't fit the picture. An example:



A contains both $\langle J, M \rangle$ and $\langle L, Q \rangle$. If mapped by d , this yields the ordered pairs $\langle q, m \rangle$ and $\langle m, q \rangle$. Say J is a specifier of IP , say the NP *the girl*. Let's say, moreover, that M corresponds to I^0 which is occupied by *will* and L to I' . This means that the LCA runs into a contradiction: it yields the pairs $\langle \text{the girl}, \text{will} \rangle$ and $\langle \text{will}, \text{the girl} \rangle$. Thus no proper linearization is possible.

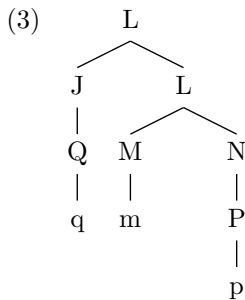
As a solution to this problem Kayne adopts Chomsky's (1986 a) notions of *segment* and *exclusion* and defines specifiers in terms of adjunction, i.e. every specifier is technically an adjunct:

X excludes Y if no segment of X dominates Y .

If we have a classical adjunction structure (i.e. adjunction to a maximal projection), we say that this structure has two segments - and upper and a lower segment. Both segments together form the category. Kayne now defines c -command as follows

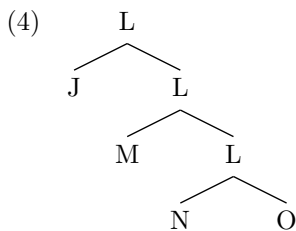
X c -command Y iff X and Y are categories and X excludes Y and every category that dominates X dominates Y .

Now consider the tree in (3) which is like (2) except that K is replaced by L , indicating that we are dealing with an adjunction structure (J is adjoined to L):



J now c -commands M since it excludes L , and the first category that dominates J also dominates M .

How about multiple adjuncts/specs? Again, consider the next tree:



In this tree J and M c -command each other since the first category that dominates M also dominates J (and vice versa). By the LCA both $\langle J, M \rangle$ and $\langle M, J \rangle$ are in A which is a linear contradiction: M cannot precede and follow J at the same time. It follows that a structure like the one above must be excluded from the grammar: multiple adjuncts are not permissible.

These considerations carry over to head-adjunction: only one head can (left-)adjoin to one other head. Clitics are the prime phenomenon that is considered. "(F)rom the perspective of the LCA, sequences of clitics must not be analyzed as successive adjunctions to the same head but instead should be analyzed as involving either adjunctions to distinct functional heads (e.g., one clitic to Tense, one to Agr) or adjunctions of one clitic to another, or some combination thereof." (p. 21)

4 Some Sample Analyses

A general empirical consequence of Kayne's approach is the following: There are languages that are systematically head-final. Japanese, Korean and, I believe, Chinese are examples. In these languages sentential complementizers follow the embedded sentence. By the LCA, the universal structure is S-H-C. If true, how does the complementizer end up in the final position? Answer: By moving IP to Spec,CP - this way, the embedded sentence can precede C^0 as the former asymmetrically c-commands the latter. Among other advantages Kayne mentions, one appeal of this simple large-tree displacement approach is that it provides "a way of accounting for the fact discussed by Bach (1971, p. 161) that interrogative *wh*-movement is generally absent from SOV (i.e., from consistently head-final) languages." (p. 54) If IP occupies Spec,CP in the overt syntax, *Wh*-elements cannot end up there. How does covert *Wh*-movement take place in these languages? Kayne is not explicit about this but makes general remarks about reconstruction in footnote 22 of Chapter 5. If we make the standard assumption that phrases in non-A positions undergo reconstruction, what happens is that IP reconstructs and thus frees Spec,CP for covert *Wh*-movement³.

Let's now go into some data.

4.1 Possessives

First observation: postnominal possessives in English cannot be complements to N, since a true complement can intervene between N and the possessor:

- (5) I have two pictures of John's
- (6) I have two pictures of Mary of John's

By the LCA *of John's* cannot be a right-adjunct to the NP, since rightward adjunction is not admitted. Which option is left? On the basis of Hungarian NPs and data from Szabolcsi (1981, 1983), Kayne (p. 86) proposes that possessors occupy a category sandwiched between DP and NP. This category is AgrP since the "dative possessor necessarily cooccurs with a post nominal AGR morpheme identical to the one that arguably licenses, or contributes to licensing, the nominative possessor" (Kayne (1993:4); if the possessor stays low, i.e. if it is preceded by a definite article, it bears nominative in Hungarian. If the DP is indefinite, the possessor must raise, picking up dative Case). By assumption, Hungarian and English don't differ in their base. It is thus tempting to assume that the Hungarian surface form indicates the base form of English possessors:

- (7) D^0 [_{AgrP} John [_{Agr} 's [two pictures]]]

Thus English has a null-D above the possessor, a head which is phonetically realized in Hungarian if the article is definite (something that corresponds to *the John's two pictures*). According to Szabolcsi "Hungarian also has a (phonetically unrealized) indefinite D^0 that can precede the possessor phrase" (p. 86). In that case the possessor leaves the DP via Spec,DP where it must pick up dative Case because indefinite D cannot case-license the possessor. This option is excluded in English since the possessor *John's* is not a constituent. In English, by contrast, the possessor receives its case not by movement of the possessor, but by insertion of *of* and raising of the embedded NP to Spec,DP:

- (8) a. [_{D⁰} of] [_{AgrP} John [_{Agr} 's [two pictures]]]
- b. [two pictures]_i [_{D⁰} of] [_{AgrP} John [_{Agr} 's [t_i]]]

The generalization that follows from this analysis is that the complement of Agr cannot be a DP, but must an NP or a QP (Kayne assumes that, unlike definite articles, indefinite articles don't occupy D). This correctly predicts the ungrammaticality of 9 and the grammaticality of sentences equivalent to 5 in which there is a quantifier like *all*, *every*, *some* and *any* instead of the numeral.

- (9) *? I found the pictures of John's/his.

³I don't know if a ban on "multiple occupation" within a single derivation has ever been suggested or if there is some other problem with such an analysis; of course, the order of operations is crucial: reconstruction > *Wh*-movement).

Sentences like 10 are excluded since DPs are not recursive (the definite article in cannot select for a structure like 4.2), or, at least, since “*the* cannot have a DP complement“ (p. 86):

- (10) *? I found the two pictures of John’s/his.

Open questions remain: In Hungarian, movement of the possessor is obligatory if covert D is indefinite. In English, however, the embedded NP/QP need not move as in *I like John’s every idea*. How does this apparent optionality of movement in English come about? And how does the English possessor receive Case in this construction (since there is no *of*-insertion)?

4.2 Relative Clauses in English

After what we’ve said about 9 and 10, why 11 is grammatical?

- (11) I found the (two) pictures of John’s/his that you lent me.

Answer: *(two) pictures of John’s/his that you lent me* is a constituent distinct from *the*. Specifically, Kayne proposes that the D-head occupied by *the* selects for a CP. The head of this CP is *that* and *(two) pictures of John* moves to Spec,CP from below *me*:

- (12) I found the [_{CP} [(two) pictures of John’s/his]_i that you lent me t_i]

This N-raising analysis of relative clauses goes back to Vergnaud (1974). This analysis also solves the problem that arises with alternative analyses of relative clauses that assume that the relative clause right-adjoins to NP/DP. The problem is that by the LCA right-adjunction is simply out. Such analyses also run into difficulties concerning Binding theory. Kayne gives the following example:

- (13) John bought the picture of himself that Bill saw.

By the right-adjunction theory, the anaphor cannot be bound by *Bill* since it is not c-commanded by *Bill*. The ambiguous interpretation (either with *John* or *Bill* as antecedent) cannot be derived. With the raising-analysis, by contrast, both readings become available.

How about relative clauses with *Wh*-elements? If promotion of the noun is the correct analysis, *person* and *who* compete for the same position in 14 (namely Spec,CP):

- (14) the person who Bill saw

The Romance languages hold the key to a solution to this problem. In Italian (15a) and French (16a) relative pronouns are impossible in constructions equivalent to 14. Relative pronouns must be accompanied by prepositions to make these sentences grammatical (15b and 16b):

- (15) a. *la persona cui Bill ha visto
 the person who Bill has seen
 b. la persona con cui Bill ha parlato
 the person with whom Bill has spoken

- (16) a. *la personne qui Bill a vue
 b. a personne avec qui Bill a parlé

The interpretation Kayne gives to this contrast is that the prepositions provide the extra space needed by the heads of the relative clauses. In 15a and 16a *cui* and *qui* allow no constituent next to them, which is why the sentences are out. If the preposition is present, however, the noun can raise to Spec,PP. Thus the structure for 16b is 17, and an English equivalent is right below. In the English equivalent the base version (of the CP-internal material) is grammatical as an echo question, which Kayne takes as supporting evidence for his analysis:

- (17) la [_{CP}[_{PP} personne_i [avec qui t_i]] [C⁰ ...
 (18) the [_{CP}[_{PP} hammer_i [with which t_i]] [he broke it t_{PP}]]

In Romance a landing site for the head of the relative is made available by a preposition. If absent, the sentence becomes ungrammatical as 15a and 16a show. “(A) well formed ‘headed’ restrictive relative clause structure requires that *personne* reach in the overt syntax a position governed by D^0 .” (p. 90) Returning to the problem posed by 11, Kayne proposes that it must be analyzed as

(19) the [CP [DP $person_i$ who [t_i]]] [C^0 ...

In English, the head of the relative must obligatorily raise, namely to Spec,DP.⁴ The question that arises is why French and Italian don’t allow for this type of movement while English does.

4.3 Some heavy Stuff

Heavy NP-shift is usually analyzed as rightward adjunction of the “heavy” constituent to the VP. This option is ruled out under the LCA. As a solution, Kayne proposed that what looks like rightward movement is actually leftward movement. Consider 20:

(20) John gave to Bill all his old linguistics books.

Kayne proposes that the sentence has the following structure:

(21) John gave [[to Bill]_i [X^0 [[all his old linguistics books] [Y^0 [t_i]] ...

He claims that the rule that moves the PP *to Bill* leftward resembles scrambling (familiar from German). Y^0 is a small clause hosting both internal arguments of the verb. The rule now moves the indirect object leftward into a functional projection XP. In German, the verb ends up in the lowest position while it ends up in a position above XP in English (Kayne remains agnostic as to whether the verb got there by movement or not). One advantage of this analysis is that it accounts for the fact that Dutch and German “by and large lack heavy NP shift to postverbal position,” because “direct objects in Dutch and German must raise leftward to a position higher than the surface position of the verb.” (p. 73)

5 Conclusion

“To a significant extent, the LCA-based theory of syntax proposed here allows us to have the all too infrequent pleasure of seeing the theory choose the analysis.” (p. 132)

References

- Richard S. Kayne. *The Antisymmetry of Syntax*. MIT Press, 1994.
- Noam Chomsky. Remarks on nominalization. In Jacobs and P. Rosenbaum, editors, *Reading in English Transformational Grammar*. Waltham: Ginn, 1970.
- Ray Jackendoff. *X-bar-Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monograph 2. Cambridge, MA: MIT Press, 1977.
- Richard S. Kayne. *Connectedness and Binary Branching*. Foris Publications, Dordrecht, 1984.
- Jan-Wouter Zwart. *Dutch Syntax: A Minimalist Approach*. PhD thesis, Rijkuniversiteit Groningen, 1993.
- Roland Hinterhölzl. *Scrambling, Remnant Movement, and Restructuring in West Germanic*. Oxford: University Press, 2006.

⁴In footnote 15 of page 90 Kayne makes an interesting proposal for German constructions like the following:

- Den Mann, den haben wir nicht gesehen.

Such constructions pose the problem, that there seem to be a constituent plus another element that precede the verb, violating the verb second constraint.

According to Kayne, [den Mann_i [den t_i]] is *one* constituent, just as [hammer_i [with which t_i]] is. He claims that this construction is a counterpart for clitic left dislocation familiar from Romance.

Thomas Leu. *The Internal Syntax of Determiners*. PhD thesis, 2008.

Richard S. Kayne and Jean-Yves Pollock. Towards an analysis of french hyper-complex inversion. *Published in a festschrift*, 2008.