On Heads and the Linear Correspondence Axiom

Hans Broekhuis

0. Introduction

In this article, I will discuss Kayne's (1994) claim that the Linear Correspondence Axiom (LCA) blocks the coordination of heads.[1] Assuming without discussion that this is a desirable result, I will investigate the consequences of Chomsky's (1995a,b) claim that traces do not count for the LCA. What this claim amounts to is that any structure that is not in accordance with the LCA can in principle be saved by means of an operation (call it REPAIR) that moves the offending category to a position that is allowed by the LCA. This makes excluding coordination of heads not an easy task. I will argue, however, that this can be done, provided that we assume not that the locus of application of the LCA is PF (as claimed in Chomsky 1995a,b), but that the LCA applies throughout the grammar (which is more in line with but not identical to what is claimed in Kayne 1994), for which independent evidence will be given in section 4. The idea is that REPAIR may apply to every substructure that is created by GT* (=MERGE and MOVE α) if it is not in accordance with the LCA: if this results in a structure that is licit according to the LCA, the derivation may proceed; if not, the derivation is cancelled.

1. The LCA and heads

According to the LCA, asymmetric c–command between nonterminal nodes maps into a linear order of the terminal nodes. As Kayne (1994) has shown, various properties of phrase structure that are stipulated by X–bar–theory can be derived from the LCA. Further, he has noted that the LCA excludes the coordination of heads. Consider the structure in (1).

(1) $[_{ConjP} X [_{ConjP} Conj Y]]$

The structure in (1) is excluded, since the coordinating conjunction Conj and the head Y symmetrically c–command each other, and hence no relative order is specified between the terminals that are dominated by these heads. The con-

---

clusion that coordination of heads is impossible is probably a desirable result. On the assumption that a clitic is a head, it would follow immediately that clitics cannot be coordinated:

(2)   *Je [le   [et   la]] vois souvent.
          I   him   and her   see often

On the other hand, however, Kayne (1994:61) notes that 'if [...] the LCA applies to all representations, then it must be the case that a clitic does not originate as a pure head that is the complement of the verb [as in (3a)], but rather as a subpart of the complement, as in [(3b)]'.

(3)   a   ... V [$_D$ clitic]
      b   ... V [$_{DP}$ [$_D$ clitic] [$_{NP}$ pro]]

If the clitic indeed originates as the determiner of a complex DP, Kayne's account of (2) is undermined. In order to block coordination of clitics, we must therefore reject the analysis in (3b), and follow Chomsky (1995a,b) in assuming that (3a) *is* the correct structure, that is, that clitics do not originate as the head of a complex phrase.[2] According to Chomsky, this in fact accounts for a typical property of clitics, viz., that they move in overt syntax, even (and this is crucial) if no strong feature forces the movement. We return to this in section 3.

2. The LCA and traces

If (3a) is the correct structure, how can the problem with respect to the LCA be solved? Chomsky (1995a,b) suggests that traces are not subject to the LCA. If this is true, structures that are not in accordance with the LCA can in principle be saved by means of an operation (which I will call REPAIR) that moves the offending category to a position that is licit according to the LCA. So, REPAIR may save the structure in (3a) by adjoining the clitic to the verb, as in (4):
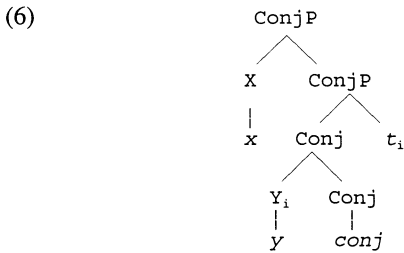
(4)

---

[2] According to Chomsky (1995a,b), the clitic is not even dominated by the nonterminal head D. Since this will generally not confuse our discussion, I will give the phrase markers in their traditional format, kindly requesting the reader to adjust them to Chomsky's format, if needed.

According to Kayne's (1994) definition of c-command in (5):

(5)   X c-commands Y iff (i) X and Y are categories, (ii) X excludes Y (where X excludes Y iff no segment of X dominates Y), and (iii) every category that dominates X dominates Y,

the nonterminal V does not c-command the nonterminal D_i in (4), since it includes it. D_i, on the other hand, does c-command V, since the first category that dominates D_i (=VP) also dominates V. Hence, D_i asymmetrically c-commands V, which gives rise to the order of terminals $<clitic,v>$.[3]

An apparent drawback of this proposal is that REPAIR can also save the structure in (1) by adjoining Y to Conj, as in (6). As a result, Y asymmetrically c-commands Conj, which gives rise to the order of terminals $<y,conj>$.

(6)

```
              ConjP
              /\
           X     ConjP
           |     /\
           x   Conj   t_i
               /\
             Y_i  Conj
              |    |
              y   conj
```

Since X asymmetrically c-commands both Y and Conj, the structure in (6) is in accordance with the LCA, so long as the trace does not count. Consequently, we seemingly can no longer exclude the coordination of heads, unless the structure in (6) is excluded for some other reason. In section 3, it will be shown that this is indeed the case.
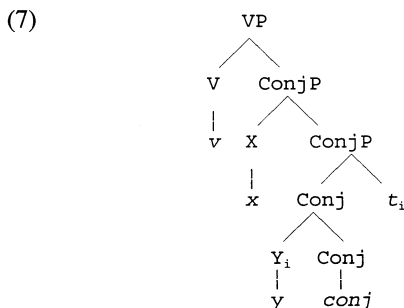

3. Heads adjoined to a maximal projection

According to Kayne (1994:section 3.7), a head cannot function as a specifier, which would still exclude the structure in (6). This restriction follows from the assumption that 'the highest element of a chain of heads must have a specifier, in the sense of having a phrase that asymmetrically c-commands it within its maximal projection (or within the maximal projection of the head it is adjoined to)', which is merely a stipulation.

---

[3] Note that this solution is not compatible with Kayne's assumption that "the LCA applies to all representations". Chomsky (1995b:334) realizes this and claims that the LCA applies at PF only. We will return to the locus of application of the LCA in section 4.

It must be noted, however, that the unacceptability of (6) would also follow from the fact that structures such as (6) are always embedded in a larger structure: (6) occurs as the complement of a head Z or as the specifier of a complex phrase ZP. It can be shown easily that this does not give rise to a linear order. (The fact that this can be done for any phrase that has a head as a specifier or a head adjoined to it, shows that Kayne's restriction is not only stipulative, but in fact entirely superfluous.)

First, consider the structure in (7), in which the structure in (6) is the complement of a verb. In (7), V and X symmetrically c–command each other; ConjP does not dominate X since it includes it, so that both heads are immediately dominated by the category VP. As a result, no order is specified between v and x. Since the order is not total, the structure is illicit.[4]

(7)
```
                    VP
                   /  \
                  V   ConjP
                  |   /   \
                  v  X    ConjP
                     |    /   \
                     x  Conj   t_i
                        /  \
                     Y_i   Conj
                      |     |
                      y    conj
```
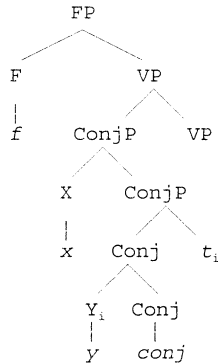
Next, consider the structure in (8), in which the structure in (6) acts as the specifier of VP. The problem that arises in this structure is that X and the first functional head above VP (= F) c–command each other. As a result, no order is specified between f and x, and the structure is illicit. Note that the same result would arise if ConjP is not a specifier but an adjunct to VP (see fn. 4).[5]

---

[4] In Chomsky (1995a,b), the two ConjP-nodes are not considered two segments of the same category, and consequently V asymmetrically c–commands X. The argument in the main text thus crucially presupposes that Kayne's assumption that specifiers involve adjunction is correct.
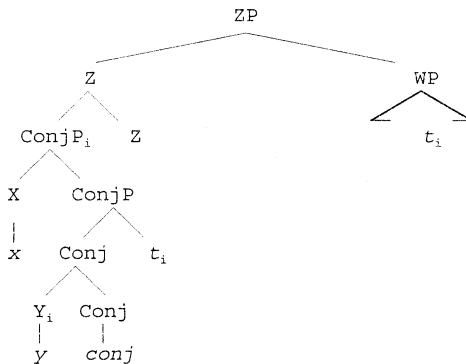
[5] For completeness, note that if each root clause has an abstract beginning node A (cf. Kayne 1994:section 4.3), the same result arises if ConjP is an adjunct or a specifier of the root phrase.

(8)

```
                        FP
                      /    \
                   F          VP
                   |        /   \
                   f    ConjP      VP
                       /    \
                     X      ConjP
                     |     /    \
                     x   Conj      t_i
                        /    \
                    Y_i      Conj
                     |        |
                     y       conj
```

If ConjP were adjoined to a higher head Z, as in (9), the structure would be illicit, too. Since the structure in (9) cannot be base-generated, it must be the result of movement of ConjP from out of WP, the complement of Z (unless ConjP = WP, a case to which we will return in section 6). Since ConjP asymmetrically c–commands everything dominated by WP, the terminals dominated by ConjP (that is, *conj* and *y*) should precede the terminals dominated by WP. However, since WP also asymmetrically c–commands everything dominated by ConjP, the terminals dominated by WP must also precede the terminals dominated by ConjP. Consequently, the order is not antisymmetric.

(9)

```
                                  ZP
                           /              \
                      Z                      WP
                    /   \                   /  \
               ConjP_i     Z              /_____\   t_i
                /    \
              X      ConjP
              |     /    \
              x   Conj      t_i
                 /    \
             Y_i      Conj
              |        |
              y       conj
```

Recall from the discussion of (3) and (4) above that the structure in (7) would also be illicit if the complement of V is not a complex ConjP but a simple head. Of course, this is a desirable result, since this accounts for the fact that object clitics must be moved, even if no strong N–feature is present that would force this movement (as in the case of the Romance languages). Note that something similar is true if we replace the complex ConjP in (8) by a simple head X, as in (10).

(10) $[_{FP} [_F f] [_{VP} [_X x] \text{VP}]]$

In (10), F and X symmetrically c–command each other and consequently no
order between f and x is specified. As in (8), the same result would arise if X is
not a specifier of VP but an adjunct, which implies that adjunction of a head to a
maximal projection is correctly excluded by the LCA (contrary to what is
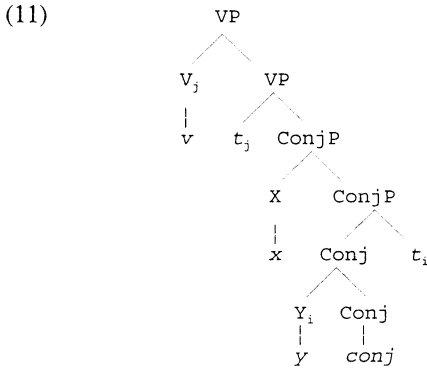claimed in Chomsky 1995a).

Consequently, it follows as a theorem of the LCA that a head must be ad-
joined to some higher head not only if it is generated in complement position, but
also if it is generated in specifier (or adjunct) position. This indicates that the
obligatoriness of movement of clitics can be completely reduced to the LCA, pro-
vided that they are not generated as the head of a functional projection but as a
head in a regular argument position.

The fact that the structures in (7) and (8) do not give rise to a linear order is
still not sufficient to exclude coordination of heads. Since traces do not count for
the LCA, these structures can in principle be saved by applying further move-
ment that destroys the symmetric c–command relations. This can be done in two
ways, viz. (i) by moving the verb V or the functional head F to a position in
which it is no longer c–commanded by X or (ii) by adjoining X to V or F. Con-
sequently, if we want to maintain the conclusion that heads cannot be coordinated,
saving the structures in (7) and (8) in these ways must be excluded. How this can
be done will be discussed in section 4 and 5. We will start with the first option.

4. Reference sets are determined step-wise

In this section, I will argue that the LCA applies to all substructures in the
following way: if a substructure built by means of GT* (=MERGE and MOVE α)
does not give rise to a linear order, the operation REPAIR must apply immediately
(i.e., before the subsequent application of MERGE); if this results in a structure
that is in accordance with to the LCA, the derivation may proceed; if not, the
derivation is cancelled. Note that this proposal is similar in spirit to Chomsky's
(1995b:234ff.) account of the cyclicity effects induced by the presence of a strong
formal feature: 'The intuitive idea is that the strong feature merged at the root
must be eliminated before it becomes part of a larger structure by further oper-
ations' (i.e. before the subsequent application of MERGE). Both can be seen as
instantiations of Chomsky's (1995a:432) 'basic assumption about reference sets
[...] that they are [...] determined step-wise', which is evoked by the need of
avoiding too much computational complexity (Chomsky 1995b:227–8).

Now, consider again the structures in (7) and (8). In principle, REPAIR can
save these structures by adjoining V/F to VP/FP: as a result of this adjunction
operation X no longer c–commands V/F, since VP/FP no longer dominates V/F.
In (11), this is shown for the VP case.

(11)

```
                    VP
                   /  \
              V_j      VP
              |       /  \
              v     t_j   ConjP
                        /   \
                      X      ConjP
                      |     /   \
                      x   Conj    t_i
                         /   \
                      Y_i    Conj
                       |      |
                       y     conj
```

However, this option would violate the part of the 'extended' structure-preserving hypothesis according to which only heads can adjoin to heads, which in its turn can be derived from Chomsky's (1995a:406,(14)) uniformity condition on chains, which requires that a chain be uniform with respect to phrase structure status: since the head of the head-chain formed by movement is a maximal projection, by definition, and the trace is a head, an improper chain is created, and hence the derivation crashes at LF.[6]

Note that this account crucially presupposes that structures that are not in accordance with the LCA cannot be the input of further applications of MERGE. If this were possible, we could wait until a proper landing-site for V/F were available (e.g. the higher functional head T), and apply REPAIR then. In this sense, the present proposal is more or less in line with Kayne's (1994) assumption that the LCA applies throughout the grammar. It is, however, completely incompatible with Chomsky's (1995a,b) assumption that the LCA applies only in the morphology, i.e., in the PF–wing of the grammar.[7]

Recall from Kayne (1994) and Chomsky (1995a,b) that one of the goals of introducing the LCA has been to derive the fact that the phrase markers are strictly binary branching (in overt syntax), without explicitly stating this by means of an X–bar-scheme, or without building this into some operation of the grammar such as GT*. If we want to retain this result, it is clear that Chomsky's assumption about the locus of application of the LCA cannot be maintained. Consider (12).

---

[6] Of course, (11) is also a case of self-attachment, which is ruled out by Chomsky (1995a:407) for reasons that are intrinsically related to his theory of phrase markers. These need not carry over to the more traditional view on phrase markers, adopted in Kayne (1994).

[7] My proposal does not imply, however, that the LCA is relevant for covert syntax. Given the fact that representations must be in accordance with the LCA at the moment of *spell-out* for convergence at PF, and given the hypothesis that LF-movement does not involve categories, i.e. non-terminal nodes, but features (see section 7), the LCA may be void in covert syntax.

(12)   [$_{VP}$ put [$_{NP}$ the book] [$_{PP}$ on the shelf]]

In (12), we are dealing with a ternary branching VP: the verb takes both the NP and the PP as its sister. The structure in (12) would give rise to a contradictory order, since NP asymmetrically c-commands the nonterminal nodes dominated by PP *and* PP asymmetrically c-commands the nonterminal nodes dominated by NP. However, if traces do not count for the LCA, the structure can be saved by moving NP into the specifier of some higher projection, say, SpecAgr$_o$P. Thus, if the application of the LCA is restricted to the PF-wing of the grammar, we cannot block the structure in (12) by means of the LCA only.

If we assume that REPAIR must make the structure in accordance with the LCA before the derivation proceeds, the structure in (12) can only be saved by applying some operation to the structure in (12) as such. One option would be adjunction of NP to VP, as in (13).

(13)   [$_{VP}$ [$_{NP}$ the book]$_i$ [$_{VP}$ put t_i [$_{PP}$ on the shelf]]]

Although the structure in (13) does give rise to a linear order (the PP no longer c-commands the nodes dominated by NP), it is arguably blocked for other reasons. For instance, if we assume that the external argument of *put* must be adjoined to the structure in (13), too, we would be dealing with multiple adjunction to VP, and a contradictory linear order (which cannot be repaired by an additional adjunction operation) would arise again. Or, alternatively, if we assume that (13) is a Larsonian VP-shell, the structure would be blocked by considerations of economy: (13) violates the 'fewest steps' requirement, since we could also have generated the NP as the specifier of the VP-shell immediately.

5. *Checking theory*

The second option (adjoining X to V/F) is blocked by Checking Theory. Since the specifier and complement of VP have [-interpretable] Case features, they must be moved into the checking domain of, say, an Agr-head (or whatever replaces it in Chomsky 1995b). If the ConjP is split before entering the local domain of Agr, checking the Case features is no longer possible.[8] This is clear

---

[8] Various possibilities to formally account for this come to mind. For instance, one may assume that the features of the moved conjunct may get a "free" ride to the AGR-head, so that the Case features of both the moved conjunct and the AGR-head are deleted and erased in overt syntax. The problem would then be caused by the second conjunct; the Case feature of AGR being erased, the Case feature of the second conjunct cannot be checked at LF, and the derivation crashes. Whether a similar approach can be adopted to account for other instances of Ross' Coordinate Structure Constraint, such as the *wh*-example in (14), is not clear to me at this moment.
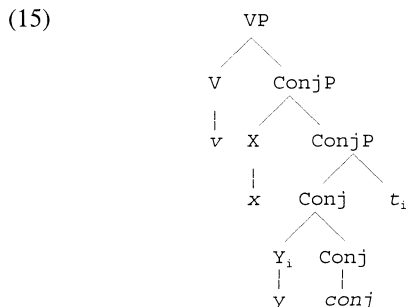
from the following parallel case of *Wh*-movement: if the *wh*-phrase is split, its [+wh]-feature cannot be checked, and the derivation crashes.

(14)   *Which boy$_i$ did you see [$_{ConjP}$ t_i [$_{ConjP}$ and which girl] yesterday

6. LEAST CATEGORY

So far, there is only one problem left. Consider again the structure in (9), in which ConjP is adjoined to a higher head Z. As we have already discussed, this structure is excluded if ConjP originates in a position *dominated* by WP, since this would give rise to a conflicting order. The structure is licit, however, if ConjP *equals* WP, e.g. if (9) is derived from the structure in (7), repeated here for convenience as (15), by adjunction of ConjP to V.

(15)

```
                    VP
                   /\
              V      ConjP
              |      /\
              v    X    ConjP
                   |    /\
                   x  Conj   t_i
                     /\
                 Y_i   Conj
                  |     |
                  y    conj
```

As we have seen in section 3, the problem with respect to (15) is that V and X c-command each other, and hence that the order is not total. We can exclude the operation that creates (9) by assuming an economy condition LEAST CATEGORY that requires that REPAIR (or rather: MOVE in general) select the smallest category possible. In other words, LEAST CATEGORY allows REPAIR to adjoin the first conjunct X to V, but excludes adjunction of the complete ConjP to V. As a result, the derivation crashes for the reason indicated in section 5. Note that LEAST CATEGORY cannot be overridden for convergence at LF, given the basic assumption about reference sets that they are determined step-wise (cf. section 4).[9]

---

[9]   In other words, I assume that the determination of the optimal continuation of a derivation does not take into account the problems that may arise later in the derivation; if the optimal continuation leads to a crashing derivation, the derivation cannot be saved by selecting a suboptimal continuation. Potentially, this opens the possibility to eliminate the notion of reference set from the theory (cf. Broekhuis and Dekkers in prep.).

The economy condition LEAST CATEGORY receives support from various facts in Kayne (1994). Consider the examples (16a,b), which Kayne assigns the base structure in (16c). According to Kayne, the NP *Bill* is assigned Case by the conjunction *with*, but Infl must check the Case of the NP *John*. *John* must therefore asymmetrically c–command (or, which would be equivalent, be in the checking domain of) Infl. This could be obtained by moving either just the NP *John* or the complete ConjP *John with Bill* into SpecIP, as in (16a) and (16b), respectively. LEAST CATEGORY provides an explanation for the impossibility of the latter (cf. Kayne 1994:64 for various parallel cases).

(16) a   John collided with Bill
     b   *John with Bill collided
     c   [$_{IP}$ spec I [... collided [$_{ConjP}$ John [$_{ConjP}$ with Bill]]]]

Similarly, LEAST CATEGORY may provide an account for the judgments on the examples in (17). Kayne explicitly relates the unacceptability of (17a) to the fact that English allows for preposition stranding, which amounts to saying that *wh*-movement as in (17a) is blocked by the fact that *wh*-movement in (17b) involves a smaller category.

(17) a   *We want to know about what you're thinking
     b   We want to know what you're thinking about

Of course, LEAST CATEGORY can be overridden for convergence at PF in these cases, i.e., if overt movement of the *wh*-element is excluded, Pied Piping will apply (cf. Chomsky 1995b:263). This accounts for the fact that languages that do not allow for preposition stranding do have (17a), and for the fact that English allows for (18b).

(18) a   *We know who$_i$ those are [t_i 's articles]
     b   We know [whose articles]$_i$ those are t_i

7. Overriding LEAST CATEGORY

Note that there is an apparent inconsistency in the discussion in section 6. While discussing the structure in (15), it has been claimed that LEAST CATEGORY cannot be overridden for convergence at LF, whereas it has been claimed that LEAST CATEGORY can be overridden for convergence at PF on the basis of the examples in (18). This inconsistency can be solved in two ways. First, we may assume that ConjP in (15), being a complex phrase, has not the phrase structure status of a head. If this is the case, adjunction of ConjP to V would violate the 'extended'

structure-preserving-hypothesis, mentioned earlier (see the discussion below (11)).

Alternatively, we may distinguish the two cases by slightly adapting Chomsky's (1995b:263,(28)) notion of 'generalized' pied-piping, according to which movement of a feature F carries along just enough material for convergence. Assume that all features of a lexical item LI constitute a complex feature bundle and that moving a certain feature F of LI must carry along the complete bundle (which can be reduced in the course of derivation by operations such as checking and subsequent erasure of a [−Interpretable] formal feature, or SPELL-OUT, which strips away the phonological features). In (18), movement of the wh–feature must then carry along the phonological features of *whose*. However, movement of *whose* would involve movement of a non-constituent (viz., a specifier and a head), so that the complete NP *whose articles* must be pied-piped. In (15), on the other hand, movement of X already involves pied-piping of all features of X. Consequently, since X is a constituent, the remainder of ConjP must be stranded. Note that this solution presupposes the assumption in fn. 9.
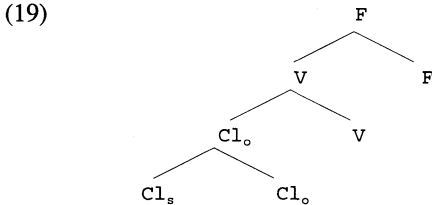
8. *The order of the clitics*

REPAIR predicts the correct order of the clitics, at least in far as the object and the subject are concerned. If GT^* forms $[_{VP}\ V\ Cl_o]$, the LCA requires that the clitic Cl_o be adjoined to V before MERGE applies again. Thus, Cl_o–V is formed. Subsequently, MERGE may form the structure $[_{VP}\ Cl_s\ [_{VP}\ Cl_o$–V]], which is in accordance with the LCA. The next application of MERGE creates $[_{FP}\ F\ [_{VP}\ Cl_s\ [_{VP}\ Cl_o$–V]]]. This structure is again not in accordance with the LCA, but can be repaired by movement of Cl_o–V to F, and adjunction of Cl_s to the cluster Cl_o–V–F thus formed.[10]

Assuming that adjunction is free, three possibilities arise in principle: (i) Cl_s adjoins to F, (ii) Cl_s adjoins to V, or (iii) Cl_s adjoins to Cl_o. The first two options are excluded by the LCA, since these would involve multiple head–adjunction: V is already adjoined to F, and Cl_o is already adjoined to V. This leaves us with the third option, given in (19).[11] The structure in (19) satisfies the LCA, and since

---

[10] As far as the LCA is concerned, we could also first adjoin CL_s to F, and subsequently adjoin the complex $[_V\ CL_o$–V] to CL_s. This would however run into various problems later in the derivation. Given LEAST CATEGORY, the subsequent movement of the verb to the next functional head, say T, must strand CL_s. Since CL_s must move at least as high as AGR_s, movement of the clitic would violate the Head Movement Constraint. Moreover, given the structure of (19) in the main text, the clitics may get a "free" ride to the functional projections against which they must check their features; since the option in this footnote implies that CL_s must move to AGR_s of its own accord, this option is more expensive, hence excluded.

[11] For some reason that is unclear to me, Kayne (1994) seems not to be very happy with this kind of adjunction operation. At present, however, nothing seems to block this option.

Cl_s is included in F, V and Cl_o, it is able to c–command its trace, thus satisfying the antecedent-'government' requirement.

(19)

```
                          F
                        /   \
                     V        F
                   /   \
               Cl o     V
              /    \
          Cl s      Cl o
```

   Note that the analysis given in this section is not incompatible with Kayne's (1994) claim that, in at least some instances, clitics adjoin to separate functional heads, if we allow for excorporation of the adjoined clitics (see Kayne 1994:17, especially his fn. 6, for a discussion of the (im)possibilities of excorporation). The linear order of the subject and object clitics can be preserved by assuming that the antecedent-'government' requirement must be satisfied at LF: Cl_s must in that case end up higher than Cl_o, which contains its trace.

9. Conclusion

In this article, I have demonstrated that coordination of heads can be excluded, even if it is assumed that traces are not relevant for the LCA. Further, I have provided evidence that the LCA applies throughout the grammar and may trigger the application of an adjunction operation REPAIR. This "step-wise" perspective on the saturation of the LCA is in perfect conformity with the "local" interpretation of reference sets defended in Chomsky (1995a and 1995b:chapter 4), and may contribute to the reduction of the problem of computational complexity. Finally, the economy condition LEAST CATEGORY has been introduced (which in fact is already implicitly present in Kayne 1994, and which is even taken a step further in Chomsky 1995b), according to which movement affects the smallest category possible only.

References

Broekhuis, H. and J.Dekkers (in prep). 'Reference Sets'. Ms. University of Amsterdam.
Chomsky, N. (1993) 'A Minimalist Program for Linguistic Theory', in K. Hale and J. Keyser, eds., *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, MIT Press, Cambridge (Mass.)/London.
Chomsky, N. (1995a) 'Bare Phrase Structure', in G. Webelhuth, ed., *Government and Binding Theory and the Minimalist Program*, Blackwell, Oxford/Cambridge (Mass.).
Chomsky, N. (1995b) *The Minimalist Program*, MIT Press, Cambridge (Mass.)/London.
Kayne, R. S. (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge (Mass.)/London.