University of Pennsylvania Working Papers in Linguistics

Volume 1
Issue 1 *Working Papers*

Article 8

1-1-1994

Notes on the Antisymmetry of Syntax

Bernard Rohrbacher
University of Pennsylvania

Notes on the Antisymmetry of Syntax

Notes on the Antisymmetry of Syntax

Bernhard Rohrbacher

1. Overview

In what proved to be probably the most influential Principles-and-Parameters manuscript of the last year, Kayne (1993) has proposed 1) a Linear Correspondence Axiom which together with a particular definition of (asymmetric) c-command is supposed to allow only SVO and OVS as underlying word orders and 2) an abstract beginning node asymmetrically c-commanding all other nodes which is supposed to further exclude OVS so that one arrives at the conclusion that SVO constitutes the universal underlying word order. Below, I argue against this conclusion on both theoretical and empirical grounds. While the Linear Correspondence Axiom has desirable effects on clause structure (cf. section 3), neither it nor the assumption of an abstract beginning node has any effects on word order.[1] In particular, Kayne's system actually allows not only SVO and OVS, but also SOV and VOS (cf. section 4). Moreover, it will not do to simply stipulate SVO as the universal underlying word order since word order in German, a language traditionally analyzed as being underlyingly SOV, cannot be adequately treated in the universal SVO approach, especially when it is compared with word order in Yiddish, a closely related SVO language (cf. section 5). The next section introduces the theoretical machinery of Kayne (1993). It should be read even by those who are already familiar with Kayne's paper, since the exposition of the linear ordering concept given in section 2 will help the reader to understand the central theoretical arguments in section 4.

2. Definitions

The Linear Correspondence Axiom, the centerpiece of Kayne's theory, is based on three different concepts: (Asymmetric) c-command, the image under dominance of an ordered pair of non-terminal nodes and linear ordering. Let us briefly look at each of these concepts in turn.

Kayne's definition of c-command in (1b) differs from the one in (1a) familiar from Chomsky (1986) in that it refers to categories (i.e. the sum of all

---

1 Gereon Müller (and quite possibly others) independently reached the same conclusion. This paper grew out of a discussion at the 1993 Summer school on Diachronic and Theoretical Syntax in Melbu/Norway and was presented at the University of Pennsylvania in Philadelphia. I would like to thank the participants and in particular Tilman Becker for various suggestions. The paper also benefitted from comments by two anonymous reviewers. Work on this paper was supported by NSF Grant # SBR-8920230.

segments of a node) and exclusion (where X excludes Y if no segment of X dominates Y).

(1) <u>C-Command</u>

a. α c-commands β if α does not dominate β and every γ that dominates α
dominates β. (Chomsky 1986:8)

b. X c-commands Y iff X and Y are categories and X excludes Y and every
category that dominates X dominates Y. (Kayne 1993:9)

The difference becomes relevant once we take into consideration the role that asymmetric c-command (cf. (2)) will play in the Linear Correspondence Axiom.

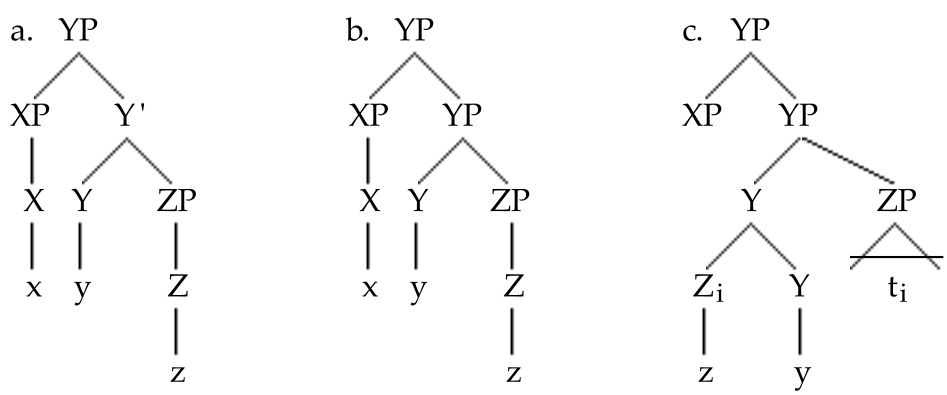
(2) <u>Asymmetric C-Command</u>

X asymmetrical c-commands Y iff X c-commands Y and Y does not c-command X.

(Kayne 1993:2)

Informally, we want it to be the case that of any two terminals, one is dominated by a non-terminal which asymmetrically c-commands a non-terminal dominating the other, but not vice versa. Consider now the trees in (3).

(3)



Both c-command definitions yield equivalent results for the specifier-head-complement structure in (3a). In particular, XP asymmetrically c-commands Y and Y' asymmetrically c-commands X, so that each of the terminals *x* and *y* are dominated by a non-terminal which asymmetrically c-commands a non-terminal dominating the other terminal. But things are different with respect to the adjunction structures in (3b,c). According to

Chomsky's c-command definition, the lower YP-segment asymmetrically c-commands X and XP asymmetrically c-commands Y in (3b), a situation that is similar to the one just described in connection with (3a). In (3c) on the other hand, neither of the terminals *z* and *y* is dominated by a non-terminal that asymmetrically c-commands a non-terminal dominating the other terminal, since Z_i and the lower Y-segment (symmetrically) c-command each other. According to Kayne's c-command definition, XP still asymmetrically c-commands Y in (3b). (Note that no category dominates XP and the last sub-clause of (1b) is therefore vacuously satisfied.) But there is now no non-terminal dominating *y* which asymmetrically c-commands a non-terminal dominating *x*: Since only categories may c-command and since no segment of the c-commander may dominate the c-commandee, neither the lower YP-segment nor YP as a whole c-commands X. For the same reasons, neither the lower Y-segment nor Y as a whole c-commands Z_i in (3c). Z_i ends up asymmetrically c-commanding Y (both are dominated by the same categories, YP and Y').

In the above discussion of the competing c-command definitions, dominance figured prominently. To be able to refer to the set of terminals that a non-terminal dominates, Kayne introduces the concept of an image under dominance of the non-terminal category X, d(X), as defined in (4a). The dominance images of non-terminals in (3a,b) are listed in (5a,6a), respectively. Kayne then extends the image concept from single non-terminals to ordered pairs of non-terminals (cf. (4b)). Let us assume that the ordering relation in <X,Y> is "X asymmetrically c-commands Y". The dominance images of pairs of non-terminals ordered by asymmetric c-command in (3a,b) are listed in (5b,6b), respectively. Finally, the image concept straight-forwardly extends from ordered pairs of non-terminals to sets of ordered pairs of non-terminals (cf. (4c)). Let us assume that A is the maximal set of pairs of non-terminals <X,Y> where X asymmetrically c-commands Y. The dominance images of the sets of pairs of non-terminals ordered by asymmetric c-command in (3a,b) are given in (5c,6c), respectively.

(4) a. d(X) = the set of terminals that the non-terminal category X dominates
b. d<X,Y> = the Cartesian product of d(X) and d(Y) = the set of all ordered
pairs{<a,b>} such that *a* is a member of d(X) and *b* is a member of d(Y).
c. d(A) = the union of all d<X,Y> for <X,Y>∈ A

(5) a. d(YP) = {x,y,z}
d(Y') = {y,z}
d(Y) = {y}

b. d<XP,Y> = {<x,y>}
d<XP,ZP> = d<XP,Z> = {<x,z>}
d<Y',X> = {<y,x>,<z,x>}

d(XP) = d(X) = {x}
d(ZP) = d(Z) = {z}

d<Y,Z> = {<y,z>}
c. d(A) = {<x,y>,<x,z>,<y,x>, <z,x>,<y,z>}

(6) a. d(YP) = {y,z}
d(Y) = {y}
d(XP) = d(X) = {x}
d(ZP) = d(Z) = {z}

b. d<XP,YP> = {<x,y>,<x,z>}
d<XP,Y> = {<x,y>}
d<XP,ZP> = d<XP,Z> = {<x,z>}
d<Y,Z> = {<y,z>}
c. d(A) = {<x,y>,<x,z>,<y,z>}

The last of the concepts underlying the Linear Correspondence Axiom is that of a linear ordering. A linear ordering has the three defining properties in (7), where L is the ordering relation in question and S is the set of elements under consideration (L linearly orders the elements in S iff (7a-c)).

(7) A linear ordering has three defining properties:
a. It is transitive, i.e. $xLy \wedge yLz \Rightarrow xLz$
b. It is total, i.e. $x \in S \wedge y \in S \Rightarrow xLy \vee yLx$
c. It is antisymmetric, i.e. $\neg (xLy \wedge yLx)$ (Kayne 1993:2)

A simple example of a linear ordering is the relative ranking of the figure skaters Oksana Baiul, Nancy Kerrigan and Chen Lu by the judges of the 1994 Winter Olympics represented in figure 1. A more formal representation of the judges' decision would be {<Baiul,Kerrigan>,<Kerrigan,Lu>,<Baiul,Lu>}. This ordering is linear since it is transitive (given that Baiul was better than Kerrigan and Kerrigan was better than Lu, Baiul was better than Lu) and total and antisymmetric (for any two of the three skaters, one was better than the other but not vice versa). It is important to note that a linear ordering based on a hierarchical relation such as "*x* was a better figure skater than *y*" does not impose any restrictions on its representation in time and space. At the medal ceremony (cf. figure 1), the best skater is located high and in the middle, with the runner-up lower and to her left and the third-placed contestant lowest and to her right, but this arrangement is of course purely conventional and any other agreed upon order would do just as well. To put it differently, the pair <Baiul, Kerrigan> translates into "Baiul was a better figure skater than Kerrigan", but not into e.g. "Baiul spatially [or temporally] precedes [or follows] Kerrigan". In section 4, this fact will be crucial to my argument that (contrary to what is claimed by Kayne) the Linear Correspondence Axiom does not impose any restrictions on underlying word order.

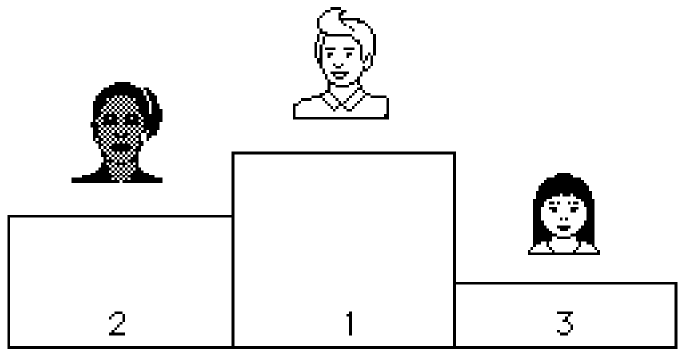


Figure 1: A Linear Ordering

Given our understanding of asymmetric c-command, dominance image and linear ordering, the Linear Correspondence Axiom can now be introduced without further ado.

(8) <u>Linear Correspondence Axiom</u>
Let P be a phrase marker, T the set of P's terminals and A the maximal set of ordered pairs {<X,Y>} such that X and Y are non-terminals in P and X asymmetrically c-commands Y. Then d(A) is a linear ordering of T.

The Linear Correspondence Axiom excludes (3a) and admits (3b) as possible phrase structures. To see this, consider again the images of their maximal sets of pairs of non-terminals ordered by asymmetric c-command in (5c) and (6c). The d(A) of (3a) in (5c) is not a linear ordering of the set of (3a)'s terminals, since it is not antisymmetric: It contains not only <x,y> and <x,z> but also <y,x> and <z,x>. The d(A) of (3b) in (6c) on the other hand is a linear ordering of the set of (3b)'s terminals, since it is transitive, total and (due to the absence of <y,x> and <z,x>) antisymmetric. An immediate consequence of the Linear Correspondence Axiom is therefore that all specifiers (including subjects) must be adjuncts to XP instead of sisters of X' under XP. This result is both welcome (in that all categories are now either heads or phrases) and problematic (in that the D- and S-structure positions of subjects cannot be straight-forwardly distinguished from those of adjuncts). In the next section, I will briefly discuss three other consequences of the Linear Correspondence Axiom for clause structure, all of which I take to be unambiguously positive. What I have in mind is the fact that the Linear Correspondence Axiom goes a long way towards deriving X'-theory, structure preservation and the head movement constraint. Section 3 can be skipped by readers who are already familiar with Kayne's paper and who are interested solely in my arguments against the alleged effects of the Linear Correspondence Axiom on word order.

3. Clause Structure

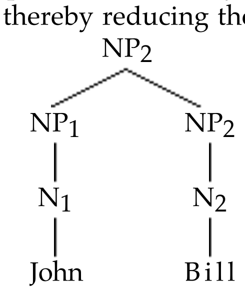
3.1. XP → X (ZP)

The Linear Correspondence Axiom automatically requires every maximal projection to directly dominate exactly one head, a requirement that has to be stipulated in traditional X'-theory. A single maximal projection cannot directly dominate two heads (cf. (9a)), since in this case neither terminal would be dominated by a non-terminal which asymmetrically c-commands a non-terminal dominating the other terminal: Both A and d(A) would be empty (i.e. non-total) and d(A) would not be a linear ordering of the terminals (cf. (10a)). This problem does not arise if the second head projects a phrase of its own as in (9b), where V (dominating *see*) asymmetrically c-commands N (dominating *John*) and no non-terminal which dominates *John* asymmetrically c-commands a non-terminal which dominates *see*. d(A) is hence a linear ordering of *see* and *John* (cf. (10b)). Conversely, a single maximal projection cannot directly dominate two maximal projections as in (9c), where NP_1 asymmetrically c-commands N_2 and NP_2 asymmetrically c-commands N_1, and d(A) is hence not a linear ordering of the terminals because it is not antisymmetric (cf. (10c)). This problem vanishes if we mediate the conjunction of NP_1 and NP_2 via the head J (for junctor) as in (9d), in which case JP dominates NP_2 (the complement of J) but not NP_1 (the adjunct to JP) and d(A) is a linear ordering of the terminals (cf. (10d) and the discussion of (3b,6) above).[2]

---

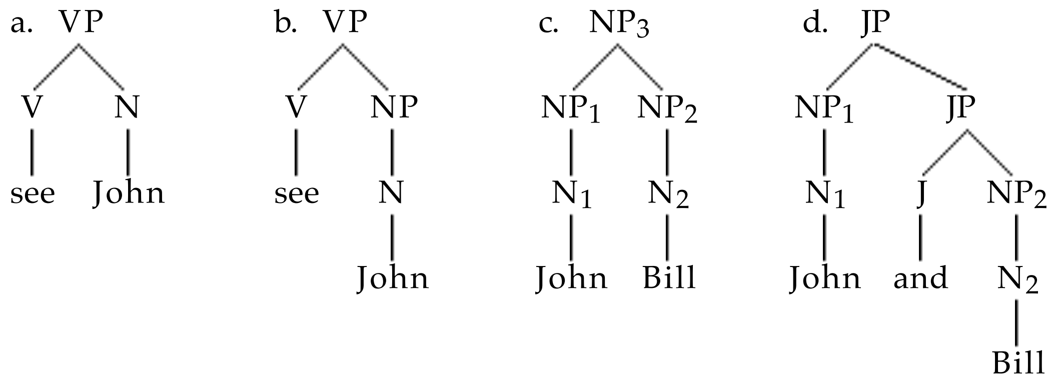
2 While the Linear Correspondence Axiom derives the X'-theoretic stipulation that every maximal projection directly dominates exactly one head, it does not derive the fact that phrasal coordination requires a junctor such as *and* (contrary to a claim in Kayne (1993:8)). Thus the Linear Correspondence Axiom allows the adjunction structure in (i) with the (image of) the set of pairs of non-terminals ordered by asymmetric c-command in (ii) resulting in a linear ordering of the terminals *John* and *Bill*. (Thanks to Tilman Becker for pointing this out to me.) Principles of Case-theory might exclude (i), but the same principles are likely to exclude (9c) as well, thereby reducing the role that the Linear Correspondence Axiom plays in X'-theory.

(i)



(ii) $A = \{<NP_1,NP_2>,<NP_1,N_2>\}$ d(A) = {<John,Bill>}

(9)



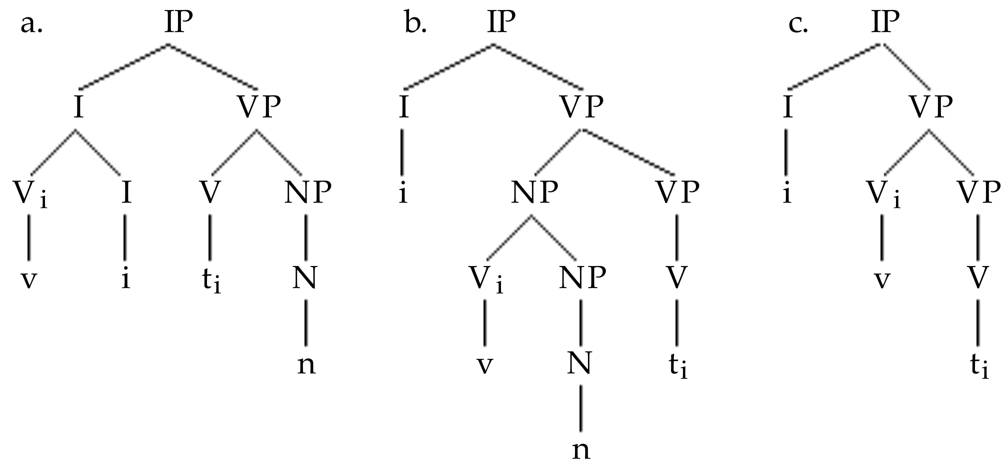
(10)
a. $A = \varnothing$ $\quad d(A) = \varnothing$
b. $A = \{\langle V,N\rangle\}$ $\quad d(A) = \{\langle see,John\rangle\}$
c. $A = \{\langle NP_1,N_2\rangle,\langle NP_2,N_1\rangle\}$ $\quad d(A) = \{\langle John,Bill\rangle,\langle Bill,John\rangle\}$
d. $A = \{\langle NP_1,JP\rangle,\langle NP_1,J\rangle,\langle NP_1,NP_2\rangle,\langle NP_1,N_2\rangle,\langle J,N_2\rangle\}$
$d(A) = \{\langle John,and\rangle,\langle and,Bill\rangle,\langle John,Bill\rangle\}$

3.2. Structure Preservation

The Linear Correspondence Axiom allows heads to adjoin to other heads but not to maximal projections (with two important exceptions discussed below) and maximal projections to adjoin to other maximal projections but not to heads. This desideratum is known as "structure preservation" in traditional theory, where it has to be stipulated.

Adjunction of a head to a head leads to a well-formed tree (cf. (11a)) whose d(A) is a linear ordering of its terminals (cf. (12a)). Adjunction of a head to a maximal projection that is c-commanded by another head is illicit, regardless of whether that maximal projection is the specifier of the complement of the higher head (cf. (11b)) or the complement of the higher head itself (cf. (11c)). In either case, the problem is that since I and V_i, i.e. the only non-terminal categories dominating either *i* or *v* but not both, symmetrically c-command each other, d(A) does not establish a ranking between the two heads (cf. the fact that neither <i,v> nor <v,i> is contained in the d(A)'s of (11b,c) in (12b,c)). Being non-total, d (A) is not a linear ordering of the terminals.

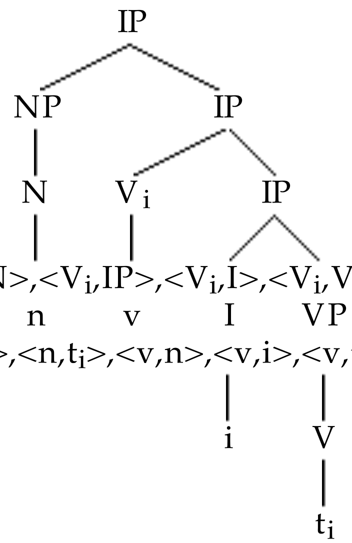
(11)



(12) a. $A = \{\langle V_i,I\rangle,\langle V_i,V\rangle,\langle V_i,NP\rangle,\langle V_i,N\rangle,\langle I,V\rangle,\langle I,NP\rangle,\langle I,N\rangle,\langle V,N\rangle\}$
$d(A) = \langle v,i\rangle,\langle v,t_i\rangle,\langle v,n\rangle,\langle i,t_i\rangle,\langle i,n\rangle,\langle t_i,n\rangle\}$
b. $A = \{\langle I,N\rangle,\langle I,V\rangle,\langle NP,VP\rangle,\langle NP,V\rangle,\langle V_i,NP\rangle,\langle V_i,N\rangle,\langle V_i,VP\rangle,\langle V_i,V\rangle\}$
$d(A) = \{\langle i,n\rangle,\langle i,t_i\rangle,\langle n,t_i\rangle,\langle v,n\rangle,\langle v,t_i\rangle\}$
c. $A = \{\langle I,V_i\rangle,\langle V_i,VP\rangle,\langle V_i,V\rangle\}$
$d(A) = \{\langle i,t_i\rangle,\langle v,t_i\rangle\}$

The Linear Correspondence Axiom however allows two cases of head-adjunction to a maximal projection, namely head-adjunction to the highest maximal projection (cf. (13)) or its specifier (which shares all the relevant properties of the first case and is therefore not discussed below). As indicated, d(A) is a linear ordering of the terminals. Kayne (1993:22) excludes (13) by assuming 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)" . In (13), V_i lacks a specifier: Since NP and V_i aren't dominated by any categories, they symmetrically c-command each other. But Kayne's requirement that heads must have specifiers is stipulative and his definition of specifier is ad hoc. What remains is that the Linear Correspondence Axiom excludes most but crucially not all cases of head-adjunction to a maximal projection.

(13)

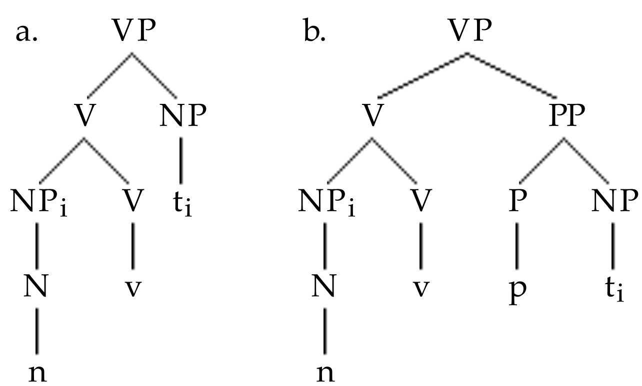


A = {<NP,IP>,<NP,I>,<NP,VP>,<NP,V>,<V_i,N>,<V_i,IP>,<V_i,I>,<V_i,VP>,<V_i,V>,<I,V>}

d(A) = {<n,i>,<n,t_i>,<v,n>,<v,i>,<v,t_i>,<i,t_i>}

Adjunction of a maximal projection to a maximal projection leads to a well-formed tree (cf. (3b,9d)) whose d(A) is a linear ordering of its terminals (cf. (6,10d)).[3] Adjunction of a maximal projection to a head is illicit, regardless of whether the adjunct is the complement of the head (cf. (14a)) or a phrasal part of that complement (cf. (14b)).[4] In the first case, d(A) is not total since it contains neither <v,t_i> nor <t_i,v> (cf. (15a)) due to the fact that V and NP symmetrically c-command each other. In the second case, d(A) is not antisymmetric since it contains both <n,p>, <t_i,n> and <p,n>, <n,t_i> (cf. (15b)) due to the fact that NP_i asymmetrically c-commands P and NP and PP asymmetrically c-commands N.

(14)



---

[3] Both (3b) and (9d) illustrate phrase-adjunction to the root node. It is easy to show that phrase-adjunction to a specifier or complement is also well-formed. To conserve space, I will not go into this matter.

[4] Kayne considers only the second case, although the two cases violate the Linear Correspondence Axiom in interestingly different ways.

(15) a. $A = \{\langle NP_i,V\rangle,\langle NP,N\rangle\}$ $\quad d(A) = \{\langle n,v\rangle,\langle t_i,n\rangle\}$
b. $A = \{\langle NP_i,V\rangle,\langle NP_i,P\rangle,\langle NP_i,NP\rangle,\langle V,P\rangle,\langle V,NP\rangle,\langle PP,N\rangle\}$
$d(A) = \{\langle n,v\rangle,\langle n,p\rangle,\langle n,t_i\rangle,\langle v,p\rangle,\langle v,t_i\rangle,\langle p,n\rangle,\langle t_i,n\rangle\}$

3.3. Head Movement Constraint

Travis's Head Movement Constraint in (16) requires head-to-head movement to be strictly cyclical without any leaps over intermediate heads. This allows us to read syntactic structure off morphological structure: If the morpheme order Stem^Affix$_1$^Affix$_2$ can be derived only via cyclic raising of the stem, then Affix$_2$ must be higher in the tree than Affix$_1$.[5]

(16) Head Movement Constraint
An X^0 may only move into the Y^0 which properly governs it.(Travis 1984:131)

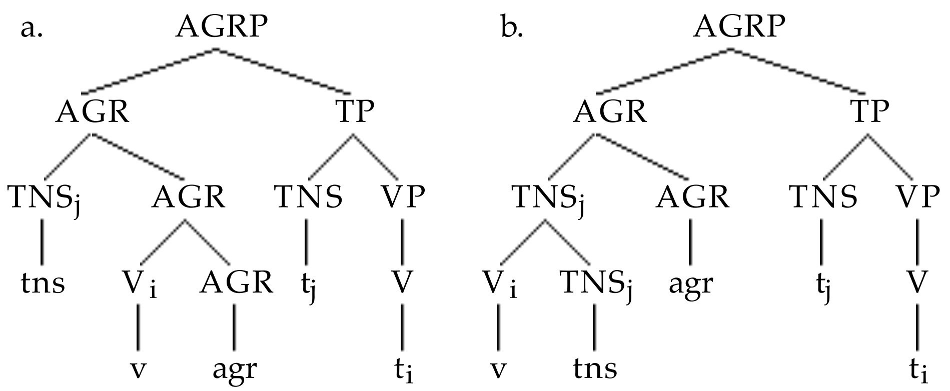
It is usually assumed that the Head Movement Constraint (and the Mirror Principle, cf. footnote 5) follows from the Empty Category Principle, but Johnson (1992) and Rohrbacher (1993) show that this is not the case: A head can skip over another head and still govern its trace, provided that the intervening head has adjoined to the target of long head movement (cf. (17a)). The Linear Correspondence Axiom on the other hand rules out structure like (17a): Since neither of the two terminals *tns* and *v* is dominated by a non-terminal that asymmetrically c-commands a non-terminal dominating the other terminal, (17a)'s d(A) in (18a) contains neither <tns,v> nor <v,tns>. It is hence non-total and not a linear ordering of (17a)'s terminals.[6] By comparison, cyclic head movement results in the well-formed tree in (17b) with the d(A) in (18b), a linear ordering of (17b)'s terminals.

---

5 See also the Mirror Principle of Baker (1985) and its discussion in Baker (1988) and Speas (1990).

6 This is but a special case of the Linear Correspondence Axiom's prohibition against multiple adjunction to the same category. Along the same lines, no two maximal projections can adjoin to the same phrase and as a consequence, nothing can adjoin to a clause containing e.g. a subject. It remains to be seen whether this is a welcome result.

(17)



(18) a. $A = \{\langle TNS_j,AGR\rangle,\langle TNS_j,TNS\rangle,\langle TNS_j,VP\rangle,\langle TNS_j,V\rangle,\langle V_i,AGR\rangle, \langle V_i,TNS\rangle,\langle V_i,VP\rangle,\langle V_i,V\rangle,\langle AGR,TNS\rangle,\langle AGR,VP\rangle,\langle AGR,V\rangle,\langle TNS,V\rangle\}$

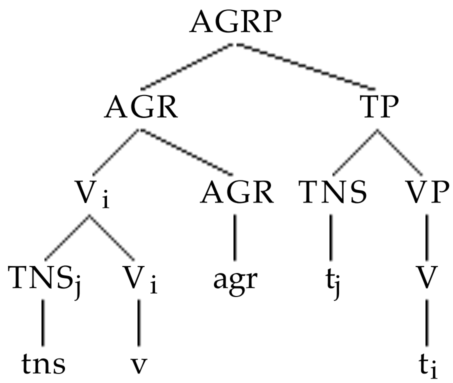
$d(A) = \{\langle tns,agr\rangle,\langle tns,t_j\rangle,\langle tns,t_i\rangle,\langle v,agr\rangle,\langle v,t_j\rangle,\langle v,t_i\rangle,\langle agr,t_j\rangle, \langle agr,t_i\rangle,\langle t_j,t_i\rangle\}$

b. $A = \{\langle V_i,TNS_j\rangle,\langle V_i,AGR\rangle,\langle V_i,TNS\rangle,\langle V_i,VP\rangle,\langle V_i,V\rangle, \langle TNS_j,AGR\rangle,\langle TNS_j,TNS\rangle,\langle TNS_j,VP\rangle,\langle TNS_j,V\rangle,\langle AGR,TNS\rangle, \langle AGR,VP\rangle,\langle AGR,V\rangle,\langle TNS,V\rangle\}$

$d(A) = \{\langle v,tns\rangle,\langle v,agr\rangle,\langle v,t_j\rangle,\langle v,t_i\rangle,\langle tns,agr\rangle,\langle tns,t_j\rangle,\langle tns,t_i\rangle, \langle agr,t_j\rangle,\langle agr,t_i\rangle,\langle t_j,t_i\rangle\}$

But the Linear Correspondence Axiom leaves one loophole for non-cyclic head movement: If the skipped-over head adjoins to the moved head itself instead of to the target of long head movement, the d(A) of the resulting tree is a linear ordering of its terminals (cf. (19)). It may be possible to exclude (19) by appealing to a version of the strict cycle condition according to which head-adjunction creates an opaque structure of which no part can be targeted for further head-adjunction. It is not entirely clear whether this condition should also exclude phrasal adjunction to specifiers (including subjects) and adjuncts, and care must be taken to ensure that it does not exclude (17a) along with (19) if there is to be an independent role for the Linear Correspondence Axiom in the derivation of the Head Movement Constraint.

(19)



$A = \{\langle TNS_j, V_i\rangle, \langle TNS_j, AGR\rangle,$ $TNS_j, TNS\rangle, \langle TNS_j VP\rangle, \langle TNS_j, V\rangle,$ $\langle V_i, AGR\rangle, \langle V_i, TNS\rangle, \langle V_i, VP\rangle,$ $\langle V_i, V\rangle, \langle AGR, TNS\rangle, \langle AGR, VP\rangle,$ $\langle AGR, V\rangle, \langle TNS, V\rangle\}$

$d(A) = \{\langle tns, v\rangle, \langle tns, agr\rangle, \langle tns, t_j\rangle,$ $\langle tns, t_i\rangle, \langle v, agr\rangle, \langle v, t_j\rangle, \langle v, t_i\rangle, \langle agr, t_j\rangle,$ $\langle agr, t_i\rangle, \langle t_j, t_i\rangle\}$

This section has shown that the Linear Correspondence Axiom at least partially derives important properties of clause structure, a promising result. In particular, it excludes both double-headed and headless phrases (but see footnote 2), it prohibits head-adjunction to maximal projections other than the root-node as well as phrase-adjunction to heads, and it blocks circumventing the Head Movement Constraint by adjoining a skipped-over head to the target of long head-movement. In the next section, I will turn to the effects of the Linear Correspondence Axiom on word order. I will argue that here the results are less promising.

4. Word Order

4.1. Only SVO and OVS are possible, NOT!

Consider a simple transitive VP containing a subject NP_1 (now understood to be an adjunct to VP), a verb V and an object NP_2 dominating the terminal nodes n_1, v and n_2, respectively. Kayne (1993:22-23) makes the following claim with respect to this scenario:[7]

> "Since [the subject] NP_1 asymmetrically c-commands V, i.e. A contains $\langle NP_1, V\rangle$, it follows that d(A) contains $\langle n_1, v\rangle$. Similarly, since V asymmetrically c-commands N_2 [the lower non-terminal category of the object], d(A) contains $\langle v, n_2\rangle$. It therefore follows that with respect to the ordering of terminals, n_1 and n_2 are on opposite sides v... Thus a theory based on the Linear Correspondence Axiom and the definition of c-command [in (1b)] is now seen to yield highly specific implications about word order... The conclusion so far is that of the six permutations of verb, subject and object, only two are permitted

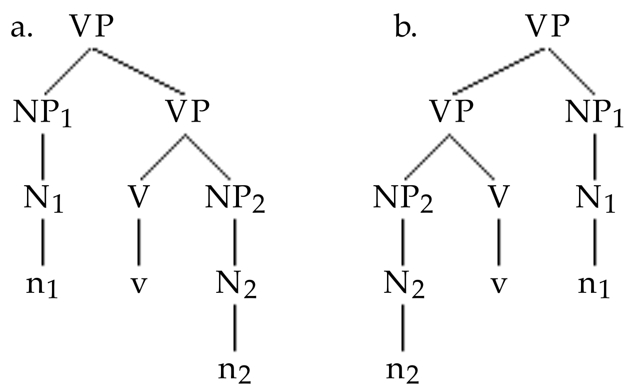
---

[7] Kayne's claim covers all specifier-head-complement configurations. I have changed some of the labels in the quotation in order it to tailor it to the concrete example, but this does not affect the argument. Moreover, whether specifier and complement are simplex (as assumed here) or complex is irrelevant as well.

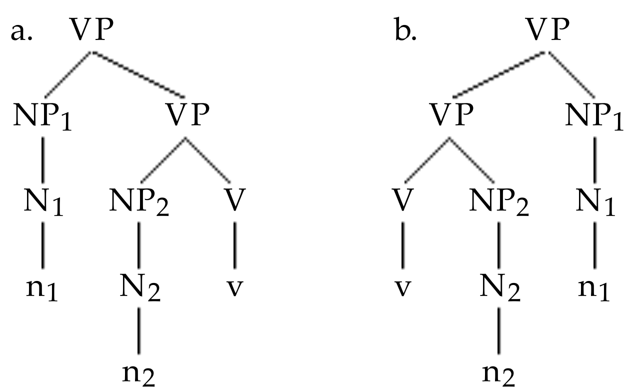
by the theory, namely SVO and OVS. The other four (SOV, OSV, VSO, VOS) are all excluded by the requirement that specifier and complement be on opposite sides of the head."

This conclusion is however clearly wrong: A theory based on the Linear Correspondence Axiom and a category/exclusion-based definition of c-command simply does not impose any restrictions on word order and in particular, it does not require that subject and object be on opposite sides of the verb. To see this, consider not only the 'good' SVO and OVS trees in (20), but also the allegedly 'bad' SOV and VOS trees in (21).

(20)



(21)



In all of these trees, asymmetric c-command establishes the same hierarchical relation between their non-terminals (cf. (22a)) and as a consequence, they share the same d(A) (cf. (22b)). Since this d(A) is a linear ordering of their terminals, Kayne's theory allows all of the four trees in (21,22), including those with the word order SOV and VOS.[8]

---

8 I assume that the underlying orders OSV and VSO are independently ruled out in any reasonable theory, presumably by a constraint against intersecting branches. It is important to note though that without such an independent constraint, the Linear Correspondence Axiom does not rule out these orders, and that it therefore has no effects on word order whatsoever.

(22) a. $A = \{<NP_1,VP>,<NP_1,V>,<NP_1,NP_2>,<NP_1,N_2>,<V,N_2>\}$
b. $d(A) = \{<n_1,v>,<v,n_2>,<n_1,n_2>\}$

Kayne's argument that $<n_1,v>$ and $<v,n_2>$ together require n_1 and n_2 to be on opposite sides of v in underlying word order seems to be based on the implicit assumption that for any linear ordering L, $<x,y> \in L$ necessarily translates either always into "*x* precedes *y* in the real world" or always into "*x* follows *y* in the real world". Yet this is not a valid assumption, as should be clear from the discussion in section 2. In fact, $<x,y> \in L$ fixes the relative sequential order of *x* and *y* only if the ordering relation itself is sequential in nature, but not if it is structural. In our case, where L = d(A) and the latter is based on the purely structural relations dominance and asymmetric c-command, $<x,y> \in d(A)$ restricts the structural relation between *x* and *y* but not their relative sequential order.[9] To take a concrete example, the pair $<n_1,v>$ from d(A) in (22b) translates into "n_1 is dominated by a non-terminal which asymmetrically c-commands a non-terminal dominating *v*", but not into "n_1 spatially [or temporally] precedes [or follows] *v*". Just as the linear ordering {<Baiul,Kerrigan>,<Kerrigan,Lu>,<Baiul,Lu>} established by the relation "*x* is a better figure skater than *y*" can correspond to the arrangement in Figure 1, where both Baiul and Lu follow Kerrigan, the linear ordering $\{<n_1,v>,<v,n_2>,<n_1,n_2>\}$ established by the relation " *x* is dominated by a non-terminal which asymmetrically c-commands a non-terminal dominating *y*" can correspond to the tree in (21b), where both n_1 and n_2 follow v instead of being on opposite sides of the latter, or any other of the trees in (20,21).

This sub-section has yielded an important result. Let me summarize it as follows:

(23) A theory based on the Linear Correspondence Axiom and (asymmetric) c-command does not make any predictions with respect to word order. In particular, such a theory does not rule out SOV or VOS languages.

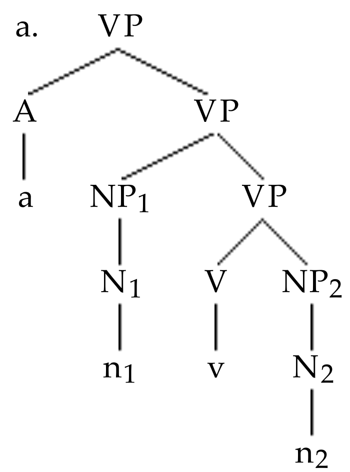
---

9 We could of course change this by adding another axiom to the theory:
(i) Asymmetric c-command implies precedence.
(i) is in fact taken from the handout of a tutorial on the antisymmetry of syntax Kayne gave on March 17, 1994 at the Seventh Annual CUNY Conference on Human Sentence Processing (which I did not attend). It is important to note that (i) is not appealed to in the written version of the paper, neither explicitly nor implicitly: If Kayne had taken (i) for granted in the paper, he would not have had to introduce the abstract beginning node *a* to exclude OVS as a possible underlying word order (cf. the next sub-section). Moreover, adding (i) to the theory amounts to stipulating SVO as the universal underlying word order instead of deriving this generalization directly from an independently motivated axiom such as the Linear Correspondence Axiom. At least to my mind, this is an important difference. Stipulating SVO as the universal underlying word order is a theoretically uninteresting move, deriving it directly from an independently motivated axiom would have been theoretically interesting..

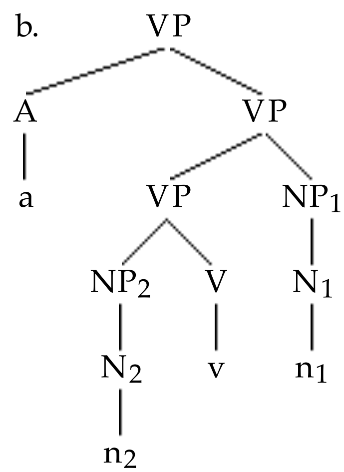
4.2 <u>Only SVO is possible, NOT!</u>

Kayne notes that while there is a root-node for dominance (i.e. a node that dominates all other nodes in the tree), there is no comparable root-node for asymmetric c-command (i.e. a node that asymmetrically c-commands all other nodes in the tree). In the trees in the preceding sub-section for example, NP_2 has the widest asymmetric c-command range, but even it does not c-command N_2 since it dominates it. Kayne introduces such a root-node for asymmetric c-command in the form of an abstract beginning node *a* that is adjoined to the top of the highest projection. Note that this move requires an additional stipulation, namely that adjunction of the abstract beginning node *a* to the highest projection may violate structure preservation, i.e. unlike all other heads, a does not need a specifier (cf. the discussion in section 3.2). We arrive at (24ab,25ab) instead of (20ab,21ab) as the trees for SVO, OVS, SOV and VOS, respectively. These trees share the set of pairs of non-terminals ordered by asymmetric c-command in (26a) and the dominance image of that set in (26b).

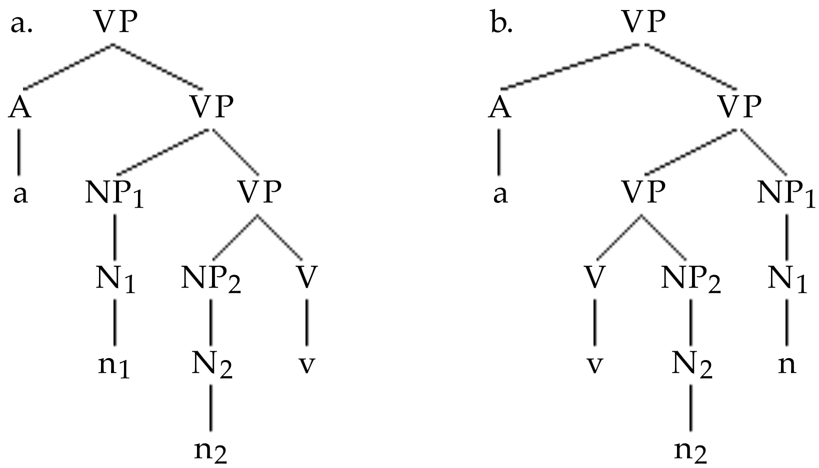
(24) a.



b.



(25)



(26) a. A = $\{<A,VP>,<A,N_1>,<A,V>,<A,NP_2>,<A,N_2>,<NP_1,VP>,<NP_1,V>,$
$<NP_1,NP_2>,<NP_1,N_2>,<V,N_2>\}$

b. $d(A) = \{<a,n_1>,<a,v>,<a,n_2>,<n_1,v>,<n_1,n_2>,<v,n_2>\}$

Kayne (1993:26) now gives the following argument for excluding not only the SOV and VOS trees in (25), but also the OVS tree in (24b):[10]

> "The question is whether <x,y> is '*x* precedes *y*' or '*x* follows *y*'. Assume the latter. [d(A) contains $<a,n_1>$, <a,v> and $<a,n_2>$]. So that if <x,y> is '*x* follows *y*', we conclude that '*a* follows n_1, *v* and n_2'. But *a* is the abstract beginning terminal. Thus we have a contradiction. Therefore <x,y> cannot be '*x* follows *y*', but must rather be '*x* precedes *y*' [and as a consequence, only the SVO tree in (24a) is admissible]."

This argument suffers from the same shortcoming as the one in section 4.1. Crucial to the argument is once more the implicit assumption that $<x,y> \in d(A)$ in (26b) translates either always into "*x* precedes *y*" or always into "*x* follows *y*", yet this assumption is unfounded. Without additional information, we can say about each $<x,y> \in d(A)$ only that "*x* is dominated by a non-terminal which asymmetrically c-commands a non-terminal dominating *y*", but nothing else. In particular, we can't say anything about the relative sequential order of *x* and *y*, which must be fixed by other factors (i.e. not by dominance and asymmetric c-command) and might quite plausibly be determined differently for each such pair. In the case of $<a,n_1>$, <a,v> and $<a,n_2>$, the fact that *a* is the abstract beginning node independently requires *a* to precede n_1, *v* and n_2. But this does not mean that in the case of $<n_1,v>$, $<n_1,n_2>$ and $<v,n_2>$, n_1 must also precede *v* and n_2 and *v* must also

10 Again I have taken the liberty to change some labels in the quotation in order to tailor it to the examples at hand.

precede n_2. So far, nothing in the theory determines the relative sequential order of these terminals. Hence all the trees in (24) and (25) are compatible with the Linear Correspondence Axiom, since they share the d(A) in (26b) which is a linear ordering of their terminals. Let me again summarize my findings:

(27) A theory based on the Linear Correspondence Axiom, (asymmetric) c-command and an abstract beginning node which asymmetrically c-commands all other nodes does not make any predictions with respect to word order. In particular, such a theory does not rule out VSO, SOV or VOS languages.

We are of course free to look outside Kayne's theory for independent reasons supporting SVO as the universal underlying word order. One area where such reasons could come from is parsing. Thus it would not be unreasonable to make the following two claims about parsing:

(28) **Hierarchical and Sequential Implicatures for Parsing**

a. <u>Parsing is Top-to-Bottom</u>
If a category of the projection x heads (i.e. X or XP) asymmetrically c-commands a category of the projection y heads (i.e. Y or YP), then x is parsed before y.

b. <u>Parsing is Left-to-Right</u>
If x occurs to the left of y, then x is parsed before y.

The Hierarchical and Sequential Implicatures for Parsing together exclude all word orders except SVO. In the OVS tree (20b) and the VOS tree (21b), NP_1 asymmetrically c-commands N_2 and V and n_1 must therefore be parsed before n_2 and v. Yet n_2 and v occur to the left of and must therefore be parsed before n_1. Likewise, in the OVS tree (20b) and the SOV tree (21a), V asymmetrically c-commands N_2 and v must be parsed before n_2. Yet n_2 occurs to the left of and must therefore be parsed before v. No such contradiction arises in the SVO tree (20a).[11]

Note that the Hierarchical and Sequential Implicatures for Parsing yield this result regardless of whether the Linear Correspondence Axiom holds and whether there is an abstract beginning node. My conclusion in (27)

---

11 Like the Linear Correspondence Axiom, the Hierarchical and Sequential Implicatures for Parsing rule out all structures with two terminals x and y where a category of the projection x heads asymmetrically c-commands a category of the projection y heads and vice versa. As a consequence, they require specifiers to be adjuncts (cf. the discussion of (3ab) in section 2) and maximal projections to directly dominate a head (cf. the discussion of (9c) in section 3.1). By the same token, they prohibit adjunction to a head by the complement of the complement of that head (cf. the discussion of (14b) in section 3.2).

that Kayne's syntactic theory doesn't restrict word order thus remains valid. In fact, if something like (28) is on the right track, it is not syntax at all but rather parsing that shoulders the burden of restricting word order.

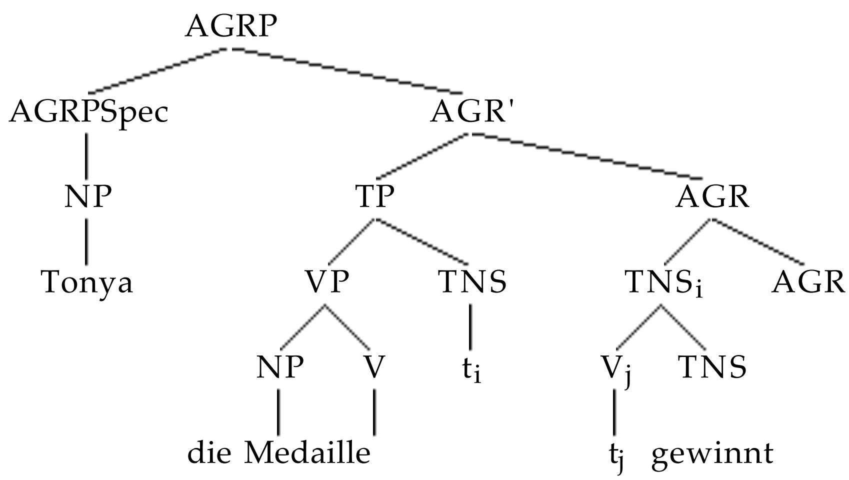
I however believe that at least one of the Hierarchical and Sequential Implicatures for Parsing is wrong and that there are other underlying word orders besides SVO. Empirical evidence supporting the latter belief comes from German. I will briefly discuss this evidence in the next and final section of this paper.

5. <u>German as an SVO Language</u>
5.1. <u>Simple Embedded Clauses</u>

I agree with Kayne that German should have verb movement to the highest inflectional head if its overt subject-verb agreement in person (and number) is any indication (cf. Rohrbacher (1993)). Then the problem arises how to account for the clause-final position of the finite verb in embedded clauses (cf. (29)).

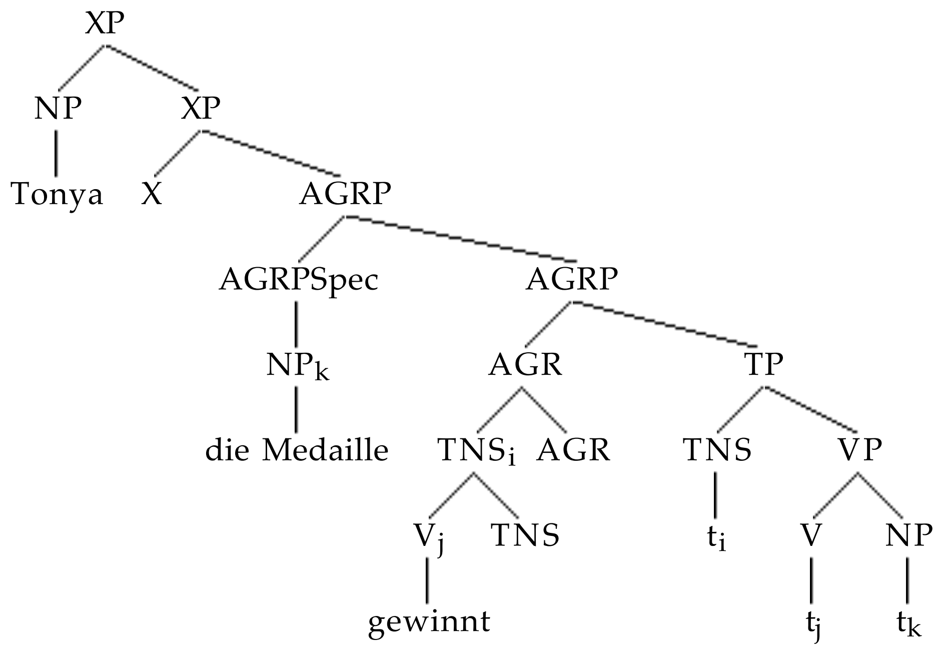
(29) a. ... daß Tonya die Medaille gewinnt.
that T. the medal wins.
"... that Tonya wins the medal."
b. *... daß Tonya gewinnt die Medaille.

The traditional view holds that German is underlyingly an SOV language (cf. Bach (1962)) and that the inflectional projections are also right-headed (cf. Grewendorf (1988:150)). According to this view, the clause-final position of the verb in (29) is the result of string-vacuous verb movement to the right (cf. (30)).

(30)


Kayne notes that in a where neither right-headed projections nor rightward movement are available, the verb in (29a) must have moved to a clause-medial AGR and the object must have scrambled over it. Moreover, since Kayne excludes multiple adjunction to the same category (cf. footnote 6), the subject and the scrambled object cannot be in the same minimal maximal projection. It follows that the subject is not located in AGRPSpec, but in some higher specifier. All this is indicated in (31).

(31)



The next sub-section presents additional facts which strongly suggest that (30) and not (31) is the correct analysis.

5.2. <u>Placement of Adverbs and Separable Verb Particles</u>

In contrast to the closely related German, Yiddish has been traditionally analyzed as an SVO language with left-headed inflectional projections (cf. Santorini (1989)). Both languages have separable verb prefixes which in matrix clauses appear away from the verb if the latter has moved to COMP (cf. (32)) but on the verb if an auxiliary has moved to COMP and the verb has stayed in situ (cf. (33)).

(32) a. Abraham schickt den Brief nicht weg.
A. sends the letter not away
b. Avrom shikt nit avek dem briv. (Yiddish)
A. sends not away the letter.
"Abraham doesn't mail the letter."

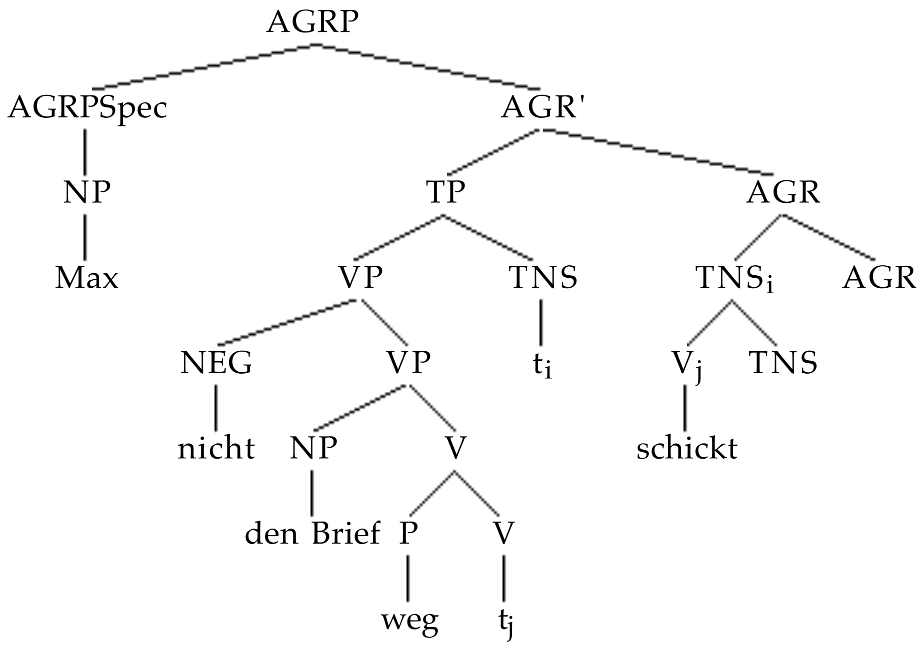
(33) a. Abraham hat den Brief nicht weggeschickt.
A. has the letter not awa
b. Avrom hot nit avekgeshikt dem briv.
A. has not away-sent the lette
"Abraham hasn't mailed the letter."

In complementizer-introduced embedded clauses without auxiliaries, the parallelism collapses. In German, the prefix appears on the verb (cf. (34a)) whereas in Yiddish, it appears away from the verb (cf. (34b)). These examples also illustrate a further difference between the two languages in this context: In German, the sentential negation marker and other sentential adverbs occur to the left of the verb, whereas these elements occur to the right of the verb in Yiddish.

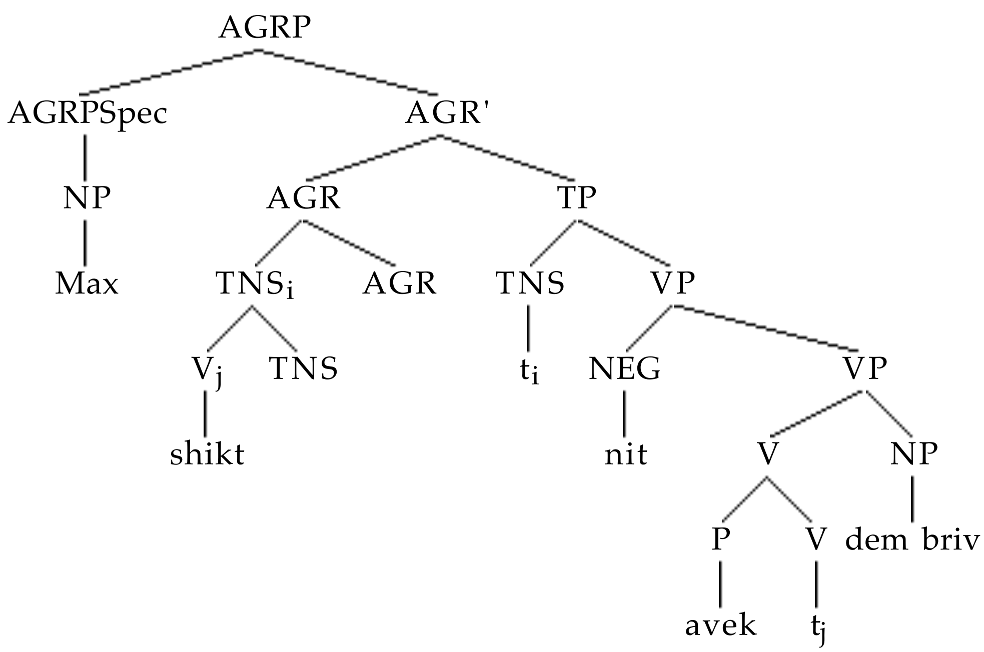
(34) a. Abraham bedauert daß Max den Brief nicht weg
A. regrets that M. the letter not away-sends. (Ger
b. Avrom bedoyert az Max shikt nit avek
A. regrets that M. send not away the lette
"Abraham regrets that Max doesn't mail the letter."

In the traditional SOV/SVO-analysis, these differences follow directly from the difference in underlying word order (German SOVTnsAgr versus Yiddish SAgrTnsVO). In German, where the verb has undergone string-vacuous V to AGR raising to the right, it remains to the right of negation and continues to immediately follow the particle (cf. (35), which does not take into account scrambling of the object). In Yiddish, where the verb has undergone V to AGR raising to the left, it has in the process crossed both negation and particle which now appear to its right (cf. (36)). It is immaterial for this SOV/SVO-analysis whether negation is adjoined to VP as shown in (35,36) or heads a phrase of its own between AGRP and TNS as proposed in Chomsky (1989).

(35)

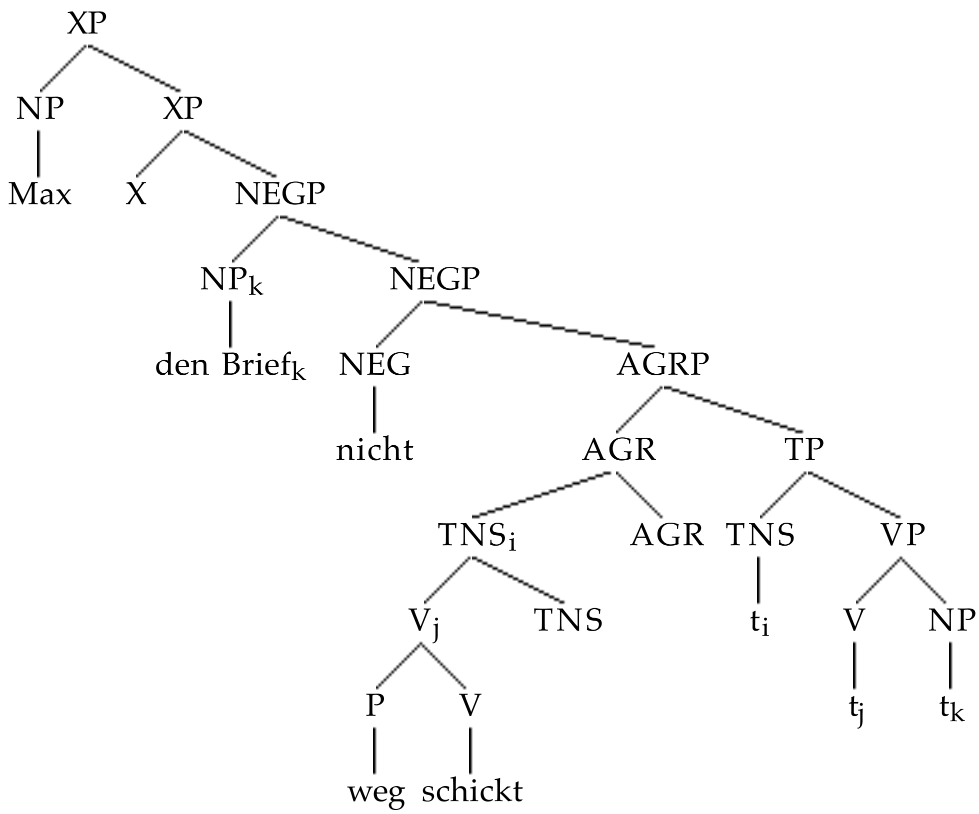


(36)



The uniform SVO analysis on the other hand must resort to two separate ad hoc assumptions to capture these two differences. Not only must it be stipulated that negation is generated higher in German than in Yiddish. In addition and more importantly, it must be stipulated that while any verb movement can strand the particle in Yiddish, only verb movement to COMP but not verb movement to AGR can do so in German. This second stipulation is particularly unattractive. The tree that the uniform SVO analysis assigns to the Yiddish example is essentially the one in (36). The tree that it assigns to the German example is given in (37).

(37)



It thus appears that unlike a theory which allows both SVO and SOV (and possibly other permutations) as underlying word orders, a theory which allows only SVO (or only SOV) as the sole admissible underlying word order cannot handle the German/Yiddish contrast in a satisfactory fashion. I conclude contra Kayne that the more permissive approach is the correct one and that there is no universal underlying word order.

In this paper, I have argued that no universal underlying word order can be derived from independently motivated aspects of syntactic theory (e.g. Kayne's Linear Correspondence Axiom) and that any universal underlying word order would be incompatible with word order differences between the otherwise closely related languages German and Yiddish. This conclusion might be conceptually undesirable, yet I believe that it is sound both from a theoretical and from a empirical point of view.

References

Bach, Emmon (1962). "The Order of Elements in a Transformational Grammar of German." *Language* 38:263-269.

Baker, Mark (1985). "The Mirror Principle and Morphosyntactic Explanation." *Linguistic Inquiry* 16:289-297.

Baker, Mark (1988). *Incorporation. A Theory of Grammatical Function Changing* (Chicago: University of Chicago Press).

Chomsky, Noam (1986). *Barriers* (Cambridge, Mass.: MIT Press).

Chomsky, Noam (1989). "Some Notes on Economy of Derivation and Representation." *MIT Working Papers in Linguistics* 10:43-74.

Grewendorf, Günther (1988). *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse* (Tübingen: Gunter Narr Verlag).

Johnson, Kyle (1992). "Head Movement, Word Order and Inflection." Manuscript, University of California, Irvine.

Kayne, Richard (1993). "The Antisymmetry of Syntax." Manuscript, City University New York.

Rohrbacher, Bernhard (1993). *The Germanic Languages and the Full Paradigm: A Theory of V to I Raising.* Ph.D. Dissertation, University of Massachusetts, Amherst.

Santorini, Beatrice (1989). *The Generalization of the Verb-Second Constraint in the History of Yiddish*. Ph.D. Dissertation, University of Pennsylvania.

Speas, Margaret (1990). "Functional Heads and the Mirror Principle." Manuscript, University of Massachusetts, Amherst.

Travis, Lisa (1984). *Parameters and Effects of Word Order Variation*. Ph.D. Dissertation, Massachusetts Institute of Technology.