UCL Working Papers in Linguistics 9 (1997)

Mirror theory

MICHAEL BRODY

Abstract

The paper argues for a theory of syntactic structure in which there is no distinction between word and phrasal level categories: the same category will be interpreted as a head in some respects and as a phrase in others. The theory takes the mirror principle to be axiomatic, analogous in status to the aspect of the projection principle that ensures that arguments of a head are projected from the lexicon. Similarly to this principle, the syntactic complement structure series is projected via the mirror principle from lexical-morphological word-internal structure. Given the lack of a syntactic word/phrase distinction, subparts of extended projections can serve as the (only) syntactic representation of words.

1 Introduction

I argue in this paper for a theory of syntactic structure in which there is no distinction between word and phrasal level categories: the same category will be interpreted as a head in some respects and as a phrase in others. The theory takes the mirror principle to be axiomatic, analogous in status to the aspect of the projection principle that ensures that arguments of a head are projected from the lexicon. Similarly to this principle, the syntactic complement structure series (i.e. essentially the 'extended projection') is projected via the mirror principle from lexical-morphological word-internal structure. Given the lack of a syntactic word/phrase distinction, subparts of extended projections can serve as the (only) syntactic representation of words.

In section 2 I summarize the phrase structure theory of Brody (to appear). Section 3 discusses attempts in the literature to derive the c-command relation from more basic considerations and suggests a way of reducing this notion to two linear orderings, one of which is precedence of terminals. Section 4 presents the outlines of mirror theory in which again the basic primitive, the spec-head relation, is characterized non-structurally, in terms of linear precedence (spec precedes the head). One problem that arises particularly sharply in mirror theory is the treatment of Bakerian open class incorporation

phenomena. Should this be analysed as involving phrasal chains (e.g. Koopman and Szabolcsi 1997) or in terms of the notion that will correspond to head chains in mirror theoretical terms? Section 5 provides some evidence for the latter option via an analysis of Hungarian verb raising constructions.

2 Insert and project

I will take as starting point the theory of phrase structure in Brody (to appear), a version of which I will summarize in this section. In this theory phrase structure involves exactly two relations: project and insert. The former of these two notions, project, corresponds to the relation of a category to its head, the element that projected it. The latter notion, insert, is the relation between an element E immediately dominated by a category of which E is not the (zero or intermediate level) head, i.e. which was projected by some element external to E.

Project licenses immediate domination relations between multiple copies of features of a single lexical item (LI). This results in objects that I refer to as projection lines (PLs). A projection line consists of a set of categories (copies of an LI) each immediately dominating the next. The most deeply embedded element, X^{min}, is normally (except in the case of empty traces) a full copy of the LI. Each element except the most deeply embedded in the projection line is the direct projection of the category it immediately dominates and each (except, presumably the element immediately dominating X^{min}, the direct projection of X^{min}) is also an indirect projection of X^{min}.

A PL is a (quasi-)lexical object: it involves only a single lexical item. We can think of the PL structure as the form in which LIs are presented to syntax. PLs are formed presyntactically, either within the lexicon or as part of the operation that selects lexical items for inclusion in a syntactic structure. On the assumption that not all complex words are lexical items and thus X^{min} level elements, as suggested for example by cases of productive noun and verb incorporation, in addition to the distinction between the lexical item and its projections it is necessary to distinguish also word (X^0) and phrasal (XP) level projections. One way to do this is to assume that phrasal projections are copies of the syntactic features of the LI while word level projections consist of both syntactic and morphological features.[1]

[1]Some distinction additional to the relational definitions of projection levels between words and phrases is necessary also in the minimalist framework of Chomsky 1995. In this theory if X^{min}s are zero

Assume further (cf. Brody 1994, to appear) that intermediate phrasal and (following essentially Kayne 1994) word internal projections do not exist. This means that PLs can consist at most of an X^{min} dominated by an X^0 dominated by an XP. Such a move has certain explanatory advantages, in particular it explains the widely held assumption that no chains can be constructed with intermediate level projections as their members. The elimination of intermediate projections essentially forces an approach where transitive heads are split into at least two elements that each project a phrase: the specifier is that sister of the higher element that does not contain the lower element, while the complement is the sister of the lower element. In many cases the analysis is independently motivated. For example, in the case of a VP the subject will be the sister of the higher small v head, the object the sister of the lower V. Notice also that the resulting structures are ternary branching.

Insert licenses immediate domination relations between a category of a projection line PL_1 and another projection line PL_2. Thus for example the relation between V and its dominating VP and the D of the subject of the verb and its DP is licensed presyntactically by project, the immediate domination relation between VP and DP (assuming the VP-internal subject hypothesis) by insert. The general condition on phrase structures is that all PL external immediate dominance relations must be licensed by insert.

The separation of immediate dominance relations into the two domains of project and insert provides a way to understand three observations concerning phrase structure that no other theory of phrase structure can so far satisfactorily explain. These have to do with the locality and uniqueness of the projection relation and what is sometimes referred to as extended structure preservation: the fact that heads cannot be dominated by phrases projected by some different head and phrases cannot be dominated by words.[2]

Projection is local in the sense that a category C and its (direct or indirect) projection

---

level projections, then both non X^{min} X^0 and X' categories are defined similarly as intermediate level projections. See Brody (to appear) for discussion.

[2]In the theory of Chomsky 1995, where a category can sometimes be both a minimal and a maximal projection in certain configurations, a head can sometimes (when it is also a phrase) be dominated by a phrase that it has not projected. But the fact that heads cannot generally be dominated by phrases surfaces in this theory also: first in the restriction that a head can only be dominated by such a phrase when the head is itself also a phrase, and secondly in the prohibition of movement of a word level category to a position immediately dominated by a phrase.

cannot be separated by a category N of a different type (i.e. which is not a projection of C). In other words if N is not a projection of C and N dominates C then N must dominate all projections of C. There cannot be for example a DP that dominates a V and which the VP that this V projected dominates. Projection is also unique: a category cannot inherit properties from more than one element. There are no NVPs for example dominating and inheriting properties from both a V(P) and an N(P). Notice that uniqueness and locality are true for both XP and X^0-level projections.

Locality and uniqueness of projection both follow immediately from the separation of phrase structural relations into the (quasi-)lexical project and (syntactic) insert, the latter a relation between these (quasi-)lexical objects. Since only PLs involve projection, and PLs contain the features of only a single LI, projection must be unique. For the same reason it must be local. The PL of some lexical item LI cannot be interrupted by a projection of some different LI, since the quasi lexical PL is nothing but a special structure given to a single lexical item — only one LI is accessible for the projection relation.[3] The intuition behind these accounts is that locality and uniqueness of projection is due to projection being an essentially non-syntactic, lexical-entry internal process. Unlike the syntactic insert relation, project does not relate lexical items to each other — hence it is inevitably local and unique.

Extended structure preservation also becomes transparent in this framework. That a phrase cannot be dominated by a word has been attributed to morphology not being able to tolerate non X^0 elements. This leaves the other, symmetric half of the condition, that a word cannot be immediately dominated by a phrase projected by some distinct element, open. Various theories have been proposed to fill this gap.[4] The solution of excluding words from syntax in parallel to the exclusion of phrases from morphology seems impossible within standard frameworks. Words appear to play an important role in syntax, and most relevantly they can be (immediately) dominated by a phrase that they project.

Once project and insert are separated, the morphological condition can straightforwardly be stated as a biconditional: apart from the case of projection it is just as true that words cannot be immediately dominated by phrases as that phrases cannot

---

[3]Technically, locality follows from the definition of project (or merge/move in the standard minimalist framework) without making project presyntactic. This fact provides no explanation of locality, however, since these relations/operations could equally be defined in a way that does not entail locality. In other words the desired consequence follows from the definitions by stipulation only.

[4]Cf. Brody (to appear) for some discussion.

be immediately dominated by words. We can assume that insert applies modularly: it can relate phrases only to phrases and words only to words. In other words insert does not license immediate dominance relations in either direction between words and phrases.

While this theory based on project and insert is fairly successful in making some central properties of phrase structure understandable, as might be expected it also raises several issues. One area concerns the appropriate definition of c-command relations for the structures postulated. In section 3 I will address this problem together with the more general question of whether the notion of c-command itself may be a consequence of more general considerations.

Another set of questions relate to PLs and the distinction between the two kinds of elements of PLs: words and phrases. The status of extended projection and its relation to phrasal projection is also left open so far, an issue that appears to be related to the word-phrase distinction.

The particular instantiation of the word-phrase distinction proposed above also has a problem that is potentially revealing. If words have both morphological and syntactic features and phrases only the latter, then even phonologically empty X^{min}s must always have some morphological feature, otherwise the distinction between the X^0 and the phrasal projection of an empty element could not be made. Thus the proposal that elements that have a morphological feature are taken to be word level might reduce to the less appealing claim that word level projections must be marked as not being phrasal (or phrases need to be marked as not being words). Recall that despite appearances to the contrary, this problem is shared with the minimalist approach. In section 4 I will present a more restrictive theory that resolves this problem by completely eliminating the word-phrase distinction from syntax. In the next two sections I shall also discuss the relation of the theory to Kayne's antisymmetry hypothesis and to his linear correspondence axiom (LCA).

3 C-command

Epstein in a very influential paper (1995) pointed out that in the cyclic derivational framework of the minimalist approach, c-command can be defined as in (1):[5]

[5]For present purposes 'term' can be taken as a synonym of 'constituent'.

(1) x c-commands all and only the terms of a category y with which x was paired by merge or by move in the course of the derivation

He compared (1) with Reinhart's representational definition, which I restate in (2):

(2) x c-commands y iff
  a. the first branching node dominating x dominates y, and
  b. x does not dominate y, and
  c. x does not equal y

Epstein claimed that the derivational definition in (1) answers certain questions concerning properties of the relation that are 'unanswerable given the representational definition of c-command' (p.19). Before looking at this claim, notice that (2) can be made more easily comparable to (1) if restated as (3) in a form parallel to (1):[6]

(3) x c-commands all and only the terms of its sister

He suggests that (1) explains that (a) x appears to c-command whatever the **first** (and not fifth nth etc.) branching node dominating x dominates since 'this is the projected node created by pairing of x and y...' Furthermore x does not c-command (b) the first branching node dominating x, (c) nodes dominated by x and (d) x itself in each case the reason being that x was not paired with the category in question by merge or move during the derivation.

But the derivational definition in (1) appears to give us neither more nor less insight into why these properties characterize c-command than the representational definition in (3). We can say without any loss (or gain) in understanding that x appears to c-command whatever the **first** (and not fifth nth etc.) branching node dominating x dominates since 'this is the node that dominates (all and only) x and its sister y'. Similarly instead of saying that x does not c-command itself, the nodes dominating it and the nodes it dominates because x was not paired with these we can say without any apparent loss of insight that x does not c-command these because these are not its sisters.

Epstein suggests also that the fact that c-command makes reference to branching can

---

[6]Or, if binary branching is not assumed then:

(3') x c-commands all and only the terms of its sisters

Note that sisterhood is taken not to be reflexive in (3)/(3').

be explained in a framework where 'Structure Building (Merge and Move) consists of Pairing, hence it invariably generates binary *branching*.' Again, in reality, this point is in fact neutral with respect to the issue of whether syntax should be constructed as a representational or derivational system. The assumption that pairing by merge and move is always binary is an additional assumption, there is nothing in the notion of concatenation that would force this operation to always be binary. The syntactic concatenation could in principle operate on any number of elements. This would allow also the unary operation alongside the binary ternary etc. options. But just like the concatenation operation can be restricted to be binary, correspondingly, the branching of trees can be restricted to the binary option, ensuring the same result in representational terms: the elimination of non-branching nodes (along with the elimination of other n-ary branching for $n \neq 2$).

Additionally, Epstein argues that the representational definition of c-command is inconsistent with the independently motivated hypothesis of the invisibility of intermediate projections. He considers the example of the category that is the sister to a VP-internal VP-spec subject — I will refer to this as V'. If V' is invisible for the computation of c-command relations then the elements contained in it (the verb and its complement) will c-command the subject and also the categories the subject contains. This is undesirable. On the other hand, Epstein points out that the situations is different if c-command relations are determined derivationally by (1). Then assuming that the intermediate projection V' can ultimately neither c-command nor be c-commanded (i.e. if its c-command relations established by (1) are eliminated) then the subject will asymmetrically c-command the verb and its complement as required by Kayne's LCA. Notice that if V' is fully visible to c-command relations then the subject and V' will symmetrically c-command each other, creating problems for the antisymmetry hypothesis.

Given the assumption of antisymmetry, it seems necessary to assume that V' or more generally intermediate projections (or lower adjunction segments) are visible for the computation of c-command relation, but cannot themselves c-command or be commanded.[7] There is nothing however in this state of affairs that would be

[7]Brody 1995a, to appear, argued and I assumed above that the best hypothesis to explain the invisibility of intermediate projections (for chain theory) is that they do not exist. This is not strictly relevant here since the same kind of problems would arise given the split structure without intermediate projections or adjunction I proposed for the VP. (The phrase dominating the lower VP and the head projecting the higher VP must not count for c-command in that system if antisymmetry is correct — cf.

'incompatible' with a representational view.

Consider instead the weaker claim that this behaviour of intermediate projections can be naturally attributed to the assumption that at the point in the derivation where a category becomes an intermediate projection (i.e. once it projects further) its c-command relations become invisible (it neither c-commands nor can it be c-commanded) but nevertheless during the earlier derivation it has already participated in determining c-command by other nodes (it counts for the calculation of c-command by these).

The problem with this line of argument is that the interpretation of 'becoming invisible' is not antecedently given, it is not any more natural to understand invisibility as entailing only the loss of ability to c-command and be c-commanded then to understand it as the loss of any c-command related role (including the role in the calculation of c-command relations between other nodes). Thus again the advantage of the derivational approach is only apparent. The statement that intermediate nodes participate in the calculation of c-command relations by other nodes but they do not participate in c-command relations themselves is not improved upon by saying that this latter property arises at a point in a derivation where the nodes become intermediate nodes (project further).

Notice also two problems with the strict pairing approach to c-command in the definitions in (1) (and also (3)). One of these is that this approach does not seem to provide any obvious way to reconstruct the notion of m-command. M-command is standardly used e.g. for the relation between the head and its spec. Epstein suggests that the m-command relation between heads and their spec should be eliminated and the (c-command) relation between the projection of the head and the spec should be used instead. But as we have just noted Epstein makes the standard assumption that intermediate projections do not c-command, so this alternative is actually not available. This problem may not be serious however, since the necessity of an m-command relation between the head and its spec is not beyond doubt.

The clearest representative of the second problem is head-chains. Under the strict pairing approach a 'moved' head which forms a constituent with another host head will not be able to c-command its trace(s). (Again this problem will disappear however in the theory to be developed in section 4 below.)

Let us return now to the theory of phrase structure set out in the previous section. The intuition expressed by Epstein's c-command definition in (1) is that this relation is determined by the operations of syntactic computation. In the present framework syntactic computation reduces to the insert relation (which can in principle be construed

---

Brody (to appear)).

representationally or derivationally). The obvious way to adapt Epstein's tempting intuition to this framework is to assume that c-command is determined by the insert relation:

(4) for x, y categories,
x c-commands all and only the categories dominated by y, to which x is related by the insert relation

That is, x c-commands all nodes dominated by the category y to which x is related by a non PL-internal (non-lexical) immediate dominance relation.

This formulation has certain advantages. Given the definition in (4) every category will c-command itself and the nodes it dominates. This I take to be an advantage rather than a drawback, since it eliminates a redundancy. Prohibition of c-command in these configurations is to prevent a category serving as an antecedent for itself or for categories it contains. The exclusion of such a configuration has also been attributed to the so called 'i-within-i' filter (cf. Chomsky 1981) that prevents an antecedent from (reflexively) dominating its anaphor. We know that the i-within-i filter cannot be reduced to the properties of c-command listed in (2b) and (2c) above since alongside structures like (5) it appears to generalize to rule out also structures like (6).

(5) $[_x$ it$_x$'s picture]

(6) $[_y$ His$_x$ mother] likes $[_x$ her$_y$ secretary]

The i-within-i filter appears to be ultimately a property of the interpretive component that prevents circular, nongrounded series of antecedents (Brody 1980, Higginbotham and May 1981).[8]

Notice, however, that this approach inherits the two problems with the pairing approach to c-command noted earlier. The first problem was c-command of the spec by the head. This is in fact made slightly worse now, because by the definition in (4) projecting heads do not c-command anything since these categories are parts of PLs, they

[8]As pointed out by Higginbotham and May, a structure where the determination of the reference is grounded, has a starting point are fine. C.f.:

(i) $[_y$ His$_x$ mother] likes John$_x$, $[_x$ her$_y$ secretary]

do not participate in an insert relation. Hence they do not establish a c-command relation with their complement either. One approach might be to define m-command as a relation between PLs as in (7).

(7) A category x m-commands all and only the categories dominated by its PL

Complement and specifier are now both m-commanded by the head but can be distinguished in terms of their different positions in the tree. By the definition of c-command in (4) — just like in terms of the earlier definitions — the specifier but not the complement c-commands a higher projection of the head in the standard approaches, or the lower phrasal shell in the theory of Brody 1995a, to appear.

This would still leave the second problem open: how to ensure c-command between members of head-chains. Various technical solutions come to mind which I will not explore since the problem will not arise in the more restrictive framework of the next section.

The various definitions of c-command — as Epstein notes in connection with his cyclic derivational version — do not explain why c-command exists, they just state its properties. The question remains why certain — or perhaps all — syntactic relations are restricted by c-command? Why cannot categories establish the relation with any other category in the tree? And if the set of categories with which a given element can establish a (relevant) relation is to be restricted, why is it restricted precisely in the way the definition of c-command states, rather than in one of the infinitely many other imaginable ways?

Epstein offered an explanation within the cyclic derivational framework he adopted. This is based on two assumptions that he refers to as (a) the first law/the unconnected tree law, and (b) the law of pre-existence. The unconnected tree law states that a syntactic relation can only hold between elements that are members of the same tree and excludes relations between elements of unconnected trees. 'Derivationally construed', as in (8) (Epstein's (27), p.25), it disallows relations between elements that at any point in the derivation were members of different unconnected subtrees.

(8) T_1 can enter into a c-command (perhaps more generally, *syntactic*) relation with T_2 only if there exists NO DERIVATIONAL POINT at which:
- i) T_1 is a term of K_1 ($\neq T_1$) and
- ii) T_2 is a term of K_2 ($\neq T_2$) and
- iii) there is no K_3 such that K_1 and K_2 are terms of K_3

Given the cycle, the condition in (8) prevents sideways c-command between two elements x and y. In all such configurations cyclicity allows only derivations in which two unconnected subtrees have been formed at some stage that properly contain x and y respectively.

Notice that 'derivationally construed' actually adds another assumption to the unconnected tree law, namely that lack of (c-command) relation at any derivational level freezes and cannot be overridden later:

(9) If there was no (c-command) relation at any given point in the derivation between terms x, y, there cannot be a relation later.

(8) still allows x to have a relation to (c-command) y where in fact y c-commands x, since in such a configuration no unconnected subtrees have been formed.[9] In other words it does not capture the asymmetry of the c-command relation. Epstein attributes this to the principle of derivational pre-existence (10), that disallows x c-commanding y on the grounds that y was not present when x was introduced.

(10) x cannot bear a relation to y when y is nonexistent

Given the assumption that the lack of a relation at a derivational point cannot be remedied at a later stage, i.e.(9), (10) entails the exclusion of what we might call upward or reverse c-command.

On closer examination, the condition in (8) does not actually explain however the impossibility of sideways relation. The intuitive content of the condition is that two categories unconnected at any point in the derivation cannot enter into a (c-command) relation. But in fact all merged/moved categories were unconnected before merger, still all can c-command the appropriate nodes. In order to allow categories to c-command at all, it is necessary to add the stipulation in (8i,ii) that '$K \neq T$', i.e. that the top node of the an unconnected tree does not count as an unconnected element. But this means that '$K \neq T$' in fact just encodes the difference between c-command and lack of it and instead of an explanation we have only another way of stating (part of the) c-command configuration.

Epstein comments on the '$K \neq T$' restriction by noting about the top nodes (to be related by merge/move) of the unconnected trees, i.e. about K_1 K_2, that 'each equals a root node,

[9]More precisely no unconnected subtrees have been formed that properly include x or y.

neither has undergone Merge or Move, hence each is (like a lexical entry) not 'yet' a participant in syntactic relations' (p.26).[10]/[11] In other words, the two instances of the 'K≠T' stipulation in (8i) and (8ii) can be exchanged for an additional fourth subclause as in (8'):[12]

(8') T_1 can enter into c-command (perhaps more generally, *syntactic*) relation with T_2 only if there exists NO DERIVATIONAL POINT at which:
- i) T_1 is a term of K_1 and
- ii) T_2 is a term of K_2 and
- iii) there is no K_3 such that K_1 and K_2 are terms of K_3 and
- iv) merge/move has already applied to $T_1$and T_2

The intuition (8') expresses is that two terms that are integrated into some subtree by merge/move cannot form a relation if at any point in the derivation **after** they have been so integrated they are unconnected, i.e. they are members of distinct subtrees. With the addition of (8'iv), (8') states that if applying merge/move to two elements x, y does not result in a subtree of which both are terms, then x does not c-command y. So, inverting the conditional, if x c-commands y then merge/move applying to x and y must have resulted in a subtree that includes both — in other words either x or y must have been merged with some tree that included the other.

The explanation of the definition in (1) involves then breaking it up into two parts: x c-commands y if neither of the following two situations obtains: (a) there is no derivational point at which x, y have been integrated into unconnected structures and (b) there is no derivational point at which x is present/integrated but y is not. We can now bring the two parts of the account (8') and (10) together again, since in both cases what is crucial, is that there is a derivational point at which a (sub)tree exists into which x is integrated but y is not. But whether or not we make this improvement, the account provides no evidence for derivations, since it can be easily restated in representational terms.

Instead of referring to a derivational point at which there is a (sub)tree into which x is integrated but y is not, we can say that x cannot c-command y if in the single syntactic

---

[10]More precisely, K_1, K_2 have not yet undergone merge or the merge part of move.

[11]Notice that 'syntactic relation' here means: not yet part of the tree, and not as before, c-command.

[12]Again, read 'merge part of move' for 'move' in (8').

representation there is a subtree which properly contains (i.e. contains but is not equal to) x but not y. Instead of rationalizing that all derivational stages must be checked for x-y connection and where no c-command holds there was one at which x was in a (sub)tree that did not contain y, we can presume that all subtrees in the representation must be checked for x-y connection and we have no c-command where we find one in which they are unconnected. (Note also that the representational version is in fact preferable, if the bottom to top derivation and the cycle have no independent motivation (Brody 1995c), since the derivational account needs to assume these. Furthermore, the easy translatability of the account into non-cyclic representational terms provides some additional evidence against these constructs.) But until we have an explanation of why a relation cannot be established at a later derivational stage that connects the relevant subtrees that were unconnected earlier (or, in representational terms, why the connection must hold in all subtrees), it will remain debatable for both the representational and the derivational versions to what extent the account explains and not just rephrases Reinhart's definition.

In contrast to the clear exposition of the nonexplanatory nature of the definition in (1) in Epstein's paper, this definition is itself sometimes taken to provide a sufficient explanation of c-command. Thus for example Groat 1995 states that 'c-command arbitrary as a representational definition, it is explainable as a property of the derivation' Take a configuration like (11), where Z c-commands A B C, A,B does not c-command Z.

(11) $Z+[_C\ A\ B]$

According to Groat this 'follows straightforwardly if the relations formed by [merge] are in fact properties of the operation. Z is merged, hence Z is in relation with $[_C$ A B]. A B were not merged with Z, hence they are not in relation with Z'.

But notice that we need to decide if merge/move applies to trees or to categories. If the former then in (11) Z merges with C, hence Z does not c-command A and B. If the latter then say $[_Z$ D E] merges with $[_C$ A B], and D and E are incorrectly predicted to c-command A and B. In neither case do we get the desired result. We can of course stipulate c-command again, slightly differently, by saying that a category merges with a tree, or add Epstein's derivational pre-existence principle for the same effect.

Properties of c-command cannot be explained simply by reference to the merge operation since c-command is an asymmetric relation, while there is nothing in the notion

of merger that would force it to create the required asymmetry. If we are to derive c-command, we need to reduce it to some non-symmetric notion. One such is linear precedence. In the remainder of this section I will attempt to relate precedence and c-command in a nonstandard way. I do not intend the suggestions to follow as an explanation of c-command, but simply an alternative way to relate this notion to linear precedence.

A strong tendency in generative grammar has been to eliminate reference to linear order and substitute this with hierarchical notions. The motivation has generally been the assumption that hierarchical relations in contrast to linear precedence have universal applicability, the latter therefore was assumed not to play any role at semantically significant levels of syntactic analysis. Furthermore it seems to have been taken for granted that c-command cannot be reduced to linear order, while the opposite unification seemed feasible.

In more recent work by Kayne (1994) and the literature it inspired, the approach that eliminates linear order from syntax has gained further momentum. According to Kayne's hypothesis linear order is universally fully determined via a basic general principle (his Linear Correspondence Axiom, LCA) by asymmetric c-command relations between the nodes of the tree. Thus apart from the overriding principle of the LCA, syntax never needs to make reference to linear relations, these can always be expressed in hierarchical terms. It is interesting that this further radical development of the idea that for syntactic purposes linear order reduces to hierarchical relations actually takes away most of the original motivation for taking hierarchy rather than order as the basic syntactic notion.

Under Kayne's hypothesis, spec/adjunct, head and complement are universally so ordered. Word order differences between languages and constructions will be due to terminals occupying different positions in structures — e.g. a head might precede a spec in one language and follow it in another because it may be in a higher head position than the spec in one and in a lower one in the other.[13] But if the order of the grammatical relations (spec-head-complement) is constant and universal, then reducing the notion of c-command to linear order cannot be ruled out on the grounds of the assumed lack of universality in how grammatical relations are linearly ordered. It is sometimes suggested that order is somehow intrinsically irrelevant for the level of logical form, but if order

[13]A further related general issue is if the different positions of terminal elements invariably arise as a result of movement/chain construction, and if so whether movement/chain structure is also universal, languages differing only in whether they choose the overt or the covert movement/chain option — whatever the eventual correct instantiation of this distinction proves to be.

can be universal, then it is difficult to see why this should be necessarily so. We do not know a priori what relations structure the LF level — in principle, linear order seems as good as c-command.

But in fact there is a good reason why linear order is a better candidate for being the right primitive: it is independently necessary. We know that at least at the PF level linear order is indispensable, whereas no similar argument can apparently be constructed for c-command.

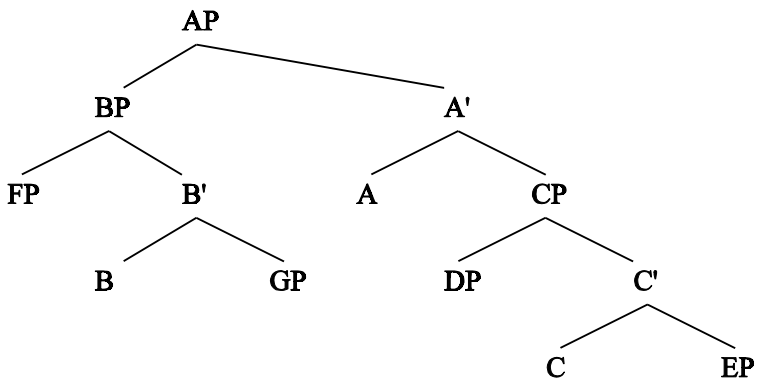
The remaining problem is of course that linear order appears to be too weak to express hierarchical relations. Linear order creates a string, a one-dimensional structure, while standard syntactic hierarchical relations create a two-dimensional tree. This can be remedied however if the theory can make reference to two independent orders. In a derivational framework, one order is given by the step by step derivation, the other can be linear order.

Consider a top-down derivational system that starts at the root node of the tree and expands each node step by step. Expanding a node is equivalent to inserting its immediate constituents. To have an ordering, suppose that the system invariably expands the rightmost node among those it has already created, in accordance with (12).

(12) Expand rightmost

To repeat: (12) means that at each step the expansion procedure takes the rightmost among those nodes that need expanding (i.e. have constituents distinct from themselves, are not terminal) that the procedure created by its earlier applications. Thus for a while the procedure moves to the right, then backtracks to expand a further level of embedding and so on. To take a concrete example, consider the derivation of the tree in (13):

(13)



(14) i. AP → BP, A'
ii. A' → A, CP,
iii. CP → DP, C'
iv. C' → C, EP
v. BP → FP,B'
vi. B' → B,GP

The order of the derivation is indicated in (14). (14i) expands the root node, (14ii) the rightmost of the resulting nodes, A'. Of the nodes now present, again the rightmost, CP is expanded (14iii), and then again the rightmost: C' (14iv). Taking A, DP, C and EP to be terminals, backtrack to the next rightmost, namely BP. (14v) expands BP and then (14vi) expands B', always the rightmost expandable node. Since FP, B and GP are again terminals, the derivation terminates.

Let us refer to the point in the derivation at which a given node enters the tree as its insertion point structure. Notice now that in the derivation obeying (12) of the tree (13), all nodes that precede x at its insertion point c-command (symmetrically or asymmetrically) x. The symmetrical cases are c-command of the X' level by the subject and c-command of the complement by the head. Suppose we are willing to give up the assumption that c-command (symmetric or asymmetric) by an element that follows counts as c-command. This step seems to have no serious untoward consequences, and as the literature on the syntactic relevance of the 'precede and command' conjunction demonstrated (e.g. Barrs and Lasnik 1986 and later developments in Larson 1988 and Kayne 1994) seems to be in the right direction. We can now define c-command in terms

of precedence, as in (15), the configuration of 'precede and c-command'.

(15) x c-commands y iff x precedes y in the insertion point structure of y

As noted earlier, such a definition is made possible by the two orderings generated by the expand rightmost principle and by precedence in the insertion point structure.

The definition in (15) does not allow Kaynean c-command by the adjoined element 'out of' the adjunction structure. Thus as before an X^0 element forming a nontrivial chain is still predicted, apparently incorrectly, not to c-command its traces — see next section for a theory where the problem will not arise. (Note also that a head does not c-command its spec either, since its spec precedes.)

A theory that defines c-command as in (15) cannot meaningfully adopt the LCA since the resulting system would be circular. In Kayne's system c-command is defined independently of precedence and the LCA relates the two notions. This raises the question: if there is as strong a relation between the two notions as the LCA require, why have both notions as primitives? (15) defines c-command in terms of precedence, hence the relationship of the two notions need not and cannot be stated in addition by another condition (15) however does not entail many of the effects of LCA, in particular it does not ensure that syntactic trees are antisymmetric. Notice that there are two related but different issues here. One is if trees are indeed antisymmetric in Kayne's sense, i.e. if it is true as a descriptive generalization that they obey the LCA. Assuming that this is the case then there is a second issue, namely whether the LCA is part of the theory, whether antisymmetry is due to the LCA or to some other consideration. In Brody (to appear) and above I discussed a system of phrase structure in which trees are not necessarily antisymmetric, but in the next section I will turn to a theory in which they are. In this theory however the LCA will not be necessary to ensure antisymmetry.

Note finally, that while the derivational expand rightmost principle makes it possible to reduce the c-command relation to linear precedence, the derivation this theory presupposes is quite dubious. First, although the derivation need not involve the rule of move, hence it need not be subject to the criticism in Brody 1995a relating to the chain/move duplication, the result of eliminating c-command in this way can only be achieved if it is a step by step process. The definition in (15) must refer to intermediate structures of the derivation. This is a weakening of the theory of syntax for which no independent motivation exists. Secondly, it seems difficult to reconcile the idea that syntactic structures are in some sense projected from the lexicon with the top down

derivation that is reminiscent of the operation of the phrase structure rules of early transformational grammar. One possibility might be to sidestep the problem, and assume that top-down stepwise application is a property of the semantic interpretive component. It would then be the interpretive rules that 'read' the structure in accordance with the 'expand rightmost' principle. C-command would be defined in terms of precedence on partially constructed interpretations.

But it is in fact unnecessary to express the needed ordering in terms of an otherwise unmotivated derivation, syntactic or interpretive, since again the proposal can be put in representational terms. We might for example define an abstract order of nodes in the representation to achieve the effects of the expand rightmost principle. Assume that nodes of the tree are ordered in the following fashion: every node N precedes the left branch node of N which in turn precedes the right branch node of N, but expansions of the node on the left branch of a node N must follow all the recursive expansions (i.e. expansions and expansions of expansions and ...) of the right branch of N. Call this abstract order an s-string, and precedence on the s-string: s-precedence. These restrictions appear to unambiguously define s-strings. For (13) they give (13'):

(13') AP, BP, A', A, CP, DP, C', C, EP, FP, B', B, GP

AP, BP, A' corresponds to the node-spec-complement order of the root node. A' expands as A, CP, CP as DP, C', C' as C, EP. So far we only looked at top nodes on left branches but no expansion of left branch nodes. In (13) this only occurs in the case of BP, the only top node on a left branch that is expanded. All expansions of BP must follow all recursive expansions of A', its sister right branch node. FP and B' then follow EP, the last element of the recursive list of expansions of A'. FP precedes B' since the left branch of a node N (BP) precedes the right branch of N. Finally left and right branch nodes, B, GP of B' follow in that order.

Another order is given by the linear ordering of terminals, call it l(inear)-string. This will, as usual, define an ordering, call it l-precedence, also for nonterminals. Take a non-terminal, T_1, to l-precede another, T_2, if the terminals dominated by T_1 precede those dominated by T_2. We can now define c-command as a conjunction of s-string and linear precedence: x c-commands y iff x both l- and s-precedes y. However, given the lack of independent evidence for the existence of s-strings, this translation of c-command does not qualify as an explanation any more than the other alternative reconstructions of the notion reviewed earlier. As we shall see, the framework to be outlined in the next section will make possible a different and more promising approach.

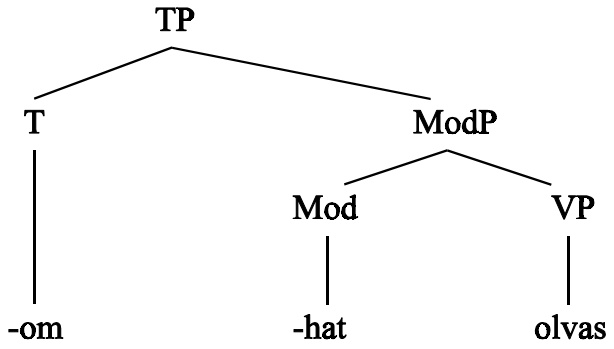
4 A (strongly) minimalist theory of syntactic representations

Cinque (1997) provides strong support for Baker's mirror principle. According to this principle, where words syntactically move to a host with which they form a unit, the order of morphological affixes mirrors the syntactic order of the relevant heads. Thus for example the syntactic order of T, permissive suffix and V mirrors their overt morphological order in Hungarian.

(16)

olvas	hat	om
read	permissive	1sg. present

(17)

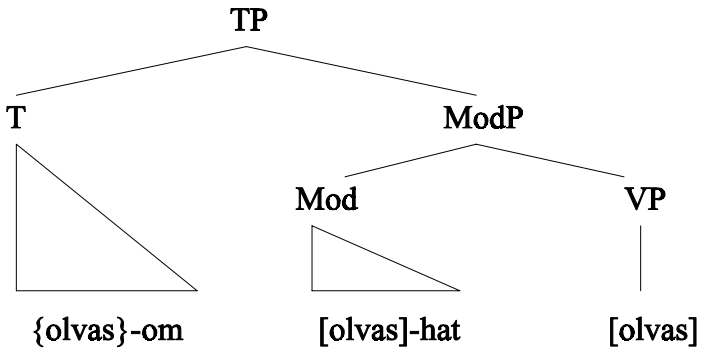


The importance of the mirror principle is enhanced by recent work that argues explicitly or suggests strongly that languages and constructions choose elements and segments from a universal and universally ordered series of functional projections (e.g. Starke 1995, Rizzi 1995, Cinque 1997).

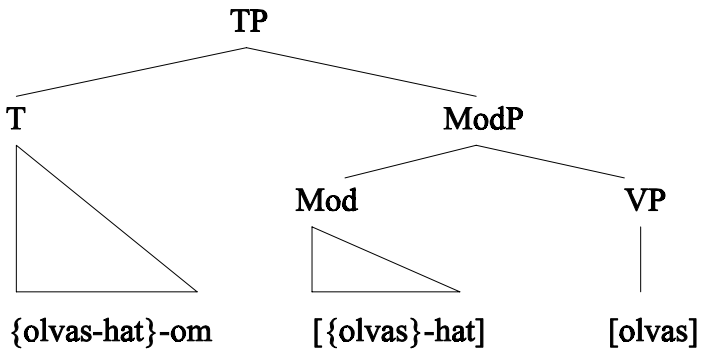
Since Baker 1988, the mirror principle is often attributed to the strict locality of head-movement/chains. This involves two assumption. One of these is that in a head chain the top element of the chain (left-)adjoins to a host that is the nearest c-commanding head — essentially the head movement constraint, HMC. The other assumption is that excorporation is prohibited, so head chains must be of the 'roll-up' type, where a head x rolls up into y and then the resulting xy unit rolls up into z, and so on. There is no non

roll-up successive step head chain.[14] For example there cannot be a three membered head chain [x1, x2, x3] where x1 has a host y, and x2 has another host z. x can only 'move' further together with its host. The partial structure in (17) cannot be completed with a chain as in (18), only with the chain structure in (19). (Phonologically not expressed traces, i.e. non top chain members, are in square brackets, top members of nontrivial chains in curly brackets.)

(18)



(19)



It is easy to see how these assumptions can entail the mirror principle.

Various empirical questions have been raised concerning the strict locality requirement of the HMC. For example it has been argued in connection with Romance and Slavic languages that head chains can sometimes cross more than one head. I will assume here

[14]I argue in Brody (1995a, to appear) that the cycle is an unnecessary construct. Accordingly, I make the terminological adjustment and refer to excorporating 'successive cyclic' XP and X^0 movement/chains as successive or stepwise chains.

that an analysis in terms of phrasal movement can ultimately be given for such cases.[15] Koopman 1994 proposes that host heads can excorporate. Again, I will tentatively assume that they cannot. The relevant structures might involve phrasal chains (cf. Koopman and Szabolcsi 1997), with a phrase in spec of the host rather than a word adjoined to the host, in which case no complex head is created from which the head would excorporate. (Another logical possibility is that there is no excorporation because incorporation into the apparently excorporating host in the relevant cases has never taken place: the chain of the apparently incorporating element in fact begins in a lower position).

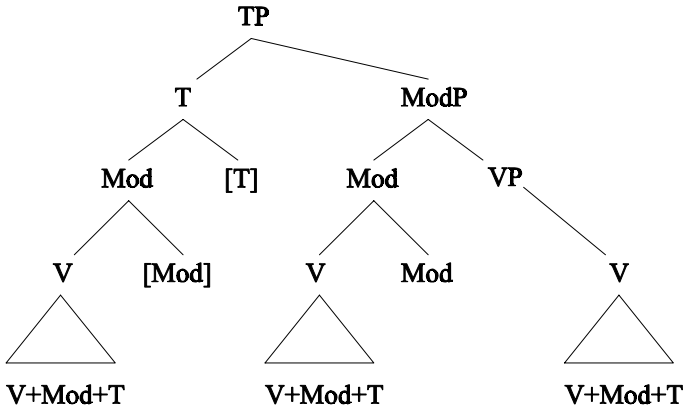
But even after putting aside the directly empirical issues, serious questions remain about the HMC based derivation of the mirror principle. A major and general problem is that in syntax the information that explicates the structure of words is expressed both word-internally (i.e. X^0-internally) and by the (inverse) structure of complementation. For example given a word consisting of say a V and an Infl morpheme in that order, the associated complementation structure will be constructed from a projection of Infl and a lower projection of V. Under most versions of the standard approach to head chains and the mirror principle, word structure and complementation structure are then required to match. There is a problem here reminiscent of the duplication that phrase structure rules created in theories antedating the principles and parameters approach. In these theories the number and type of arguments of a lexical head were specified in the lexicon and also by the appropriate phrase structure rules, which were then required to match. The assumption that phrase structure is projected from lexical properties resolved this duplication by eliminating PS rules. Similarly, we might want to search for a theory that makes it possible to project complementation structure directly from the lexical information encapsulated in the structure of words, making it possible to eliminate the independent construction of complementation structures. (Notice that there are in fact two related but distinct problems here. One is that there is no evidence for syntactically representing the information in duplicate, the other that principles generating the information are themselves duplicated.)

A second difficulty for the head movement explanation of the mirror principle arises from the fact that the guest and the host head must form a unit. In phrase structure grammars this must be a labelled constituent, necessarily labelled by the host. These

[15]For some discussion see Phillips 1996, especially note 17 on p.191. and the text to which the note relates.

apparently unavoidable assumptions introduce a further systematic set of otherwise unmotivated duplications. A fuller structure of (17) will be (20):

(20)



{olvas-hat-om-[hat]}-[om] [olvas-hat-om]-[hat] [olvas-hat-om]

The checking operation will result in the deletion of the square bracketed functional heads and neither the V head of VP nor the Mod head of ModP surface, since these are traces. We need not worry at this point about duplications introduced by traces and the checking heads, as we may consider these to be motivated on independent grounds. But there is another duplication in the structure that appears more difficult to defend, namely the repetition of the 'V+Mod+T' series in the set of heads dominating this unit in its chain top position in T. This duplication may appear to be a technical issue of little consequence. But there is no evidence for this additional duplication, which, given the additional duplication by the complementation structure makes the analysis triplicate the features in question. Furthermore, again, the problem is not just that presence of the duplicate (triplicate) information in the syntactic representation is not independently motivated, but that two (three) distinct sets of principles generate the same structure. The duplication created by the series of word-internal dominating nodes appears unavoidable in the standard approach as it is the consequence of certain basic assumptions which are distinct both from those that determine word-internal morphological order and those that define the complement series/extended projection. These assumptions are that (i) words and chain members are constituent nodes, (ii) nodes are labelled and by one of their constituents (iii) non chain-roots cannot label (attributed to the GPP in Brody 1994, 1995a, to appear, see also the 'target projects' requirement of Chomsky 1995).

As just noted in passing, yet another duplication arises from checking theory. The explanation of the mirror principle at least in the crude form given above predicts that a complex word composed from a host suffix and a chain-forming guest (with a lower trace) will appear in the syntactic position of the host suffix. Thus in the example (17) 'olvas' will surface in TP and not in VP. But as is well known especially since the Emonds-Pollock analysis of the verbal complex in French and English, the phonological position of a word often does not correspond to the syntactic pre-suffix position. For example the verb in English precedes its inflection(s), hence on the account of the mirror principle just outlined, it should form a chain whose top member is the guest of the higher host inflection. Pre-VP adverbs and negation show that the verb in English remains in the VP (cf. e.g. Pollock 1989, Chomsky 1995).

The most popular resolution of this problem is checking theory (Chomsky 1995). According to checking theory, the verb will be introduced into the syntactic tree together with inflection, and remains in place in syntax.[16] The V+Infl unit forms a chain though with the guest-of-inflection position(s)[17] and through this chain the V+Infl unit can check the specifications of the Infl node(s) ensuring that they are identical to its own. A necessary additional assumption is that a checked duplicate (functional) head deletes at some point in syntax or phonology.

Thus, checking theory introduces yet another duplication of the word structure. The information is now reproduced altogether four times in syntax: in the structure of complements, on the lexical item that is to check functional heads, on the heads that are to be checked and in the word-internal set of nodes that eventually come to dominate the checking lexical item. The duplication involved in checking theory can perhaps be defended by reference to checking in other chain types. For example in the case of wh-movement the postulation of such duplications (wh-features on both the host C node and on the wh-phrase) has been characteristic of the standard analysis long before checking theory and can indeed be argued for on independent grounds. (See also Brody 1995b for some independent evidence for this duplication specifically for head chains.)

But thirdly and importantly, the checking theory does not in fact resolve the problem

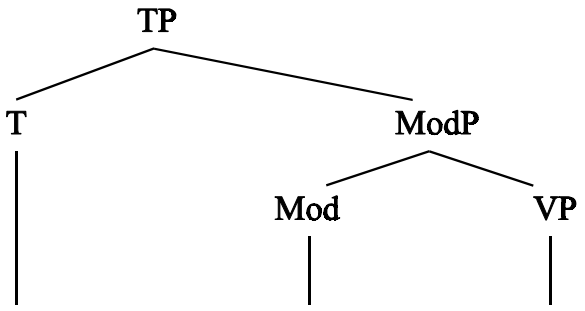
---

[16]In Chomsky 1995 chapter 3, where checking theory is introduced, the verb would have remained in place only in overt syntax. In Brody 1995a and also in chapter 4 of Chomsky 1995 there is no covert displacement of phonological material and the verb remains in situ throughout.

[17]A set of chains in a roll-up structure when the analysis is detailed enough to take account of more than one Infl position.

that configurations where the phonological position of a head does not correspond to the syntactic pre-suffix position raise for the locality based explanation of the mirror principle. This is because given the duplication that this theory introduces, we now need an auxiliary assumption to ensure the mirror principle effect. Given the checking approach the structure of (17) will be along the lines of (17'). (Traces and deleted functional heads are in square brackets).

(17')



{olvas-hat-om-[hat]}-[om] [olvas-hat-om]-[hat] [olvas-hat-om]

Let us think of this structure in derivational terms for a moment. Given checking theory to get all and only the correct suffix orders it is necessary to stipulate additionally, that checking must proceed in strict order starting from the innermost suffix on the complex lexical element. The impossible '*olvas-om-hat' could also arise from the same syntactic structure in the above example if the external suffix could be checked when this unit moves to the lower functional head and the internal suffix in the second movement step. Let us put aside the problem that starting from the innermost element is rather unexpected for a quasi-morphological operation (Pollock 1993) and concentrate on the requirement that checking order must respect the order of suffixes. That this requirement is distinct from the question of whether checking starts with the innermost or the outermost suffix is obvious with three or more suffixal elements. But the ordering requirement amounts to a stipulation that is not obviously better than stipulating the mirror principle itself: the mirror principle is also just an ordering statement based on suffix order.

The point becomes perhaps even more obvious if we return to the representational framework. The ordering statement in the representational approach cannot refer to

earlier and later applications. The statement that the innermost suffix must be checked first will have to be translated as saying that the innermost head must be checked by the lowest head among those that host a member of the chain of the lexical head+suffixes unit. The requirement that checking order respects the order of suffixes becomes the condition that the inverse order of the syntactic heads that host a member of the head+suffixes unit corresponds to the order of suffixes — in other words the residue of the mirror principle itself with locality for the chain stated separately.[18]/[19]

There are three additional problems that have to do directly with properties of head chains, and therefore are problems also for the explanation of the mirror principle which crucially involves such chains.

Firstly, the mirror principle will follow from locality only if excorporation is impossible, but it is not clear why in general it should be impossible. Wh- and NP-'movement' XP chains can be successive in a non roll-up fashion. Why should head movement be different? While various technical and partial answers exist,[20] we appear to have no clear understanding of the reasons for this prohibition that needs to hold for all head chains if the mirror principle is to be attributed to locality, but appears to hold only for head chains.

The second problem is the one we came across in a different connection in the previous section: assuming that c-command must hold between chain members, head chains

---

[18]Given checking theory, the prohibition on excorporation can be translated as the requirement that in each non-root position of a head chain some suffix must be checked. Thus a version of checking theory might allow a successive step (non roll-up) head chain of for example the V+v+Infl unit, linking the V, the word-internal v-spec and the word-internal Infl-spec positions. The chain of V+v+Infl still cannot have a member in an additional (word-internal spec) position between V and v or between v and Infl given the requirement that in each (non-root) position some suffix must be checked.

[19]The comments about checking theory in the text refer to the standard version. One can imagine an improved version that avoids some of the problems raised. For example, given checking theory, the matching requirement on word structure and complementation structure (the first problem in the text) can be eliminated. If complements are generated in a random order, the correct complement order will be forced by the requirement, which as we have seen restates the mirror principle, that checking order must respect the order of suffixes. Such an approach, which I think would be a step in the right direction, does not help however with most of the other problems raised. I will therefore propose a more radical solution below.

[20]As noted in footnote 18, under a (non-standard) version of checking theory the excorporation prohibition may be dispensable.

necessitate the introduction of a more complicated and more stipulative definition of c-command. In particular it is necessary to allow for c-command 'out of' certain types of constituents, namely the constituent created by the host and the top of the chain of the guest head. Kayne (1994) defines c-command in such a way that c-command out of adjunction is allowed, but the evidence for this modification that does not involve head chains remains inconclusive. It is perhaps suggestive also that none of the theories, reviewed in the previous section, that attempt to reduce c-command to more basic notions appear to be able to allow for c-command out of adjunction.

Finally a further difficulty has to do with the somewhat idiosyncratic nature of locality involved in head movement. A and A'-chains cannot cross A and A' positions that may be occupied by a potential antecedent — the relativized minimality generalization of Rizzi (1990). But head chains typically cross a head: namely the host of the chain-top. There are various ways to make XP and X^0 chains more similar here. Rizzi for example appears to assume that the host does not count as an intervener for the chain of the guest because it c-commands the guest: a category is a potential intervener for a chain-link only if it c-commands the lower chain member but not the higher. It is interesting to observe that the solution is incompatible with the adjunction structure of words and the Kaynean definition of c-command: it is crucial in Kayne's theory that neither the lower nor the higher segment of an adjunction host c-command the adjoined element.

Other approaches are available, for example to define interveners as (XP or X^0-internal) spec's, i.e. to ignore head positions for relativized minimality. But I shall instead take the facts at face value, as another difference between head chains and XP-chains.

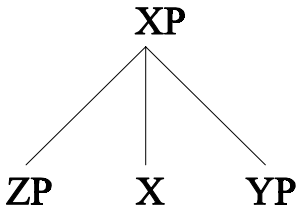
The problems just listed might be taken to motivate an alternative view of the mirror principle. Below I will develop a theory that eliminates head chains, takes the mirror principle to be the basic generalization and derives from it certain properties having to do with excorporation, c-command and locality, currently attributed to head chains and constraints on them. I shall approach this task by looking at the distinction between words and phrases in the theory of phrase structure of Brody (to appear), a version of which I summarized above in section 2. Recall that phrases in this theory were created from lexical items by the rule of project, which forms projection lines PLs. Some elements on the PL are phrases, other (lower) elements are words (X^0s) the lowest one the lexical item (X^{min}).

I argued that a PL should contain only a single phrasal node, an assumption that lead to the postulation of a tripartite shell structure for spec-head-complement structures. Similarly I adopted the view that there is only a single X^0 node on a PL, eliminating the nonmaximal phrasal and the intermediate (non-highest among the X^0-projections) X^0

levels. Let us now ask the further more radical question: are the remaining distinctions between XP, X^0 and X^{min}, in other words the postulation of the PL-structure, really necessary?

Focusing first on the distinction between words and phrases, consider the basic structure in (21).

(21)

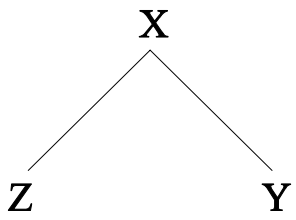


Here X projected a phrase XP creating a (partial) PL consisting of an XP immediately dominating X. Insert licensed the phrase to phrase immediate domination relations between XP and ZP on the one hand and YP on the other. As noted in Brody (to appear), spec-head-complement order follows from Kayne's LCA only if the relevant structural asymmetries are stated: spec asymmetrically c-commands the head which asymmetrically c-commands material in the complement. One alternative discussed there was to state spec-head-complement order directly. As we saw in the previous section c-command can be expressed in terms of linear orders. Taking this lead, let us temporarily assume spec-head-complement order as a primitive (subject to some simplification later), with a view to eventually deriving the major effects of the LCA from the theory.[21]

Given this approach to spec-head-complement order, there seems to be no compelling reason to distinguish syntactically XP and X, in other words to retain the PL of X in (21). A single node can just as well serve as the syntactic representation of both a phrasal category and of its head. Applying the argument also to ZP and YP, (21) reduces to (22):

[21]More precisely, the assumption is that spec and its constituents precede, while the complement and its constituents follow the head.

(22)



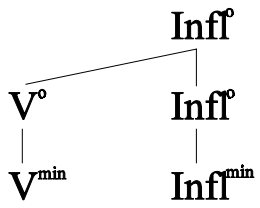
Thus as far as the word-phrase difference is concerned, there is no need for the ultimately somewhat strange operation of project or its counterparts — the set forming and labelling effects of merge and move — in the minimalist framework. As far as the word-phrase distinction is concerned, there is no need to create copies of a lexical item and establish immediate domination relation between these copies. A single copy of a lexical item can serve both as a head and as a phrase.

(22) can be viewed as eliminating the apparent conflict between the long tradition of dependency theories, see e.g. Hudson (1990) and references cited there, and phrase structure theories of syntactic representations.[22] Taking X to stand for a phrase, the lines connecting nodes can stand for immediate dominance relations. Taking X as a head, the lines express dependencies.

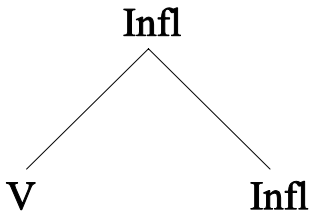
Consider next the distinction between X^{min} and X^0. If X^{min}s are lexical elements, then this distinction is necessary only if there are words that are assembled in syntax and not in the lexicon, in other words if not all X^0s are X^{min}s. For cases of simple inflectional affixation the lexicalist hypothesis is widely accepted, and is indeed presupposed by the various versions of the minimalist checking theory. Take V+Infl as a simplified example. Instead of analysing this as in (23) we can attribute to it a simpler structure (24), where X^0 and X^{min} levels are not distinguished.

[22]For other more recent attempts to simplify the theory of phrase structure in terms of dependencies cf. Brody 1994 and Manzini 1995. This latter work, like Hudson's and others' in the dependency grammar tradition also dispenses with phrasal nodes, but adopts the assumption of Brody 1994 that all dependencies in the syntactic representation exhibit left-to-right order which I crucially reject here in favour of a principled alternative.

(23)



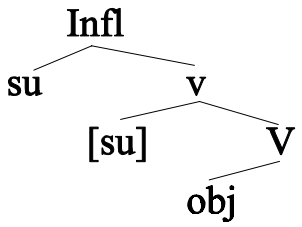
(24)



However, even without looking at the cases that might be taken to motivate the X^0 - X^{min} distinction (to which I will return below) a problem already arises. Although neither the word-phrase nor the lexical item-word distinction seems necessary at least in these elementary cases, abolishing both appears to make it difficult to provide a structure for complex words. Given a genuinely minimal analysis of the spec-head-complement structure like (22), there appears to be no place in syntax to express word-internal structure.

But this is not quite true. In fact, the impossibility of expressing word-internal structure in syntax in the traditional X^0-internal format is an advantage, since it eliminates a redundancy. As noted above, this syntactic configuration simply reduplicates lexical information that is reduplicated in syntax also in another way: in the (inverse) order of functional and non-functional projections. Given standard phrase structure trees, phrasal nodes intervene between segments of the word in the representation the inverse order of projections provide, making these relations perhaps less suited to express word-internal structure. But the impoverished theory of (22), presents no such problems. Thus for example the lexical V+v+Infl structure will be expressed in syntax as Infl taking a v and v in turn taking a V complement. Each of these nodes can, as usual have their own specifiers, say the subject, its trace and the object:

(25)



Call a potentially decomposable lexical item a lexical word (LW). Let us ask next what forces the syntactic structure where V is the complement of v and v is the complement of Infl for cases where the lexical word consists of a V which is the spec of a v which is the spec of an Inf. The answer of course is the mirror principle.

Thus suppose that the single primitive relation of the lexical and the syntactic representations is the spec-head relation. In this relation spec precedes the head. The mirror principle, an axiom of the system, ensures that the complement relation is nothing but a topological mirror image of the lexical spec-head relation, i.e. it is an inverse order spec-head relation. Spec-head relations established in the lexicon will turn up as head complement relations in syntax. Additional spec-head relations can be licensed in syntax between elements of lexical words that have free spec valences (are not specified in the lexicon as being the spec of anything) and elements of (syntactically mirrored) lexical words.

(Universal) spec-head-complement order thus does not need to be specifically stipulated: it follows from the primitive spec-head order and from the (equally axiomatic) mirror principle.

Consider next the spellout of mirrored lexical words. One question is, which element of the lexical word represents the spellout position? Here we can adopt the standard account: this takes place in the position of the deepest element of the mirrored lexical word if none of its other elements has a 'strong feature'. If some do, then spellout takes place in the highest strong position. Thus in (25) for example take a VP-adverbial like 'often' to be in the spec of some head α, between Infl and v projection (Cinque 1997). V-v-α-Infl is then spelt out in the position of Infl in French and in the position of v in English, preceding and following the adverbial respectively.

The other spellout question is why the spellout order of elements of a lexical word corresponds to the lexical rather than the mirrored syntactic order. This state of affairs is natural in a distributed morphology type approach, where the mirrored relations between features of the lexical entry would contribute to building the syntactic structure

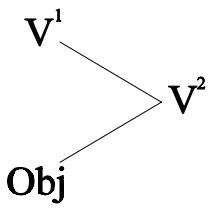
and the lexical entry is then inserted as is (i.e. in the lexical order) after the syntactic structure has been built, but before the morphological component.

In the standard framework the mirror principle follows from the HMC and the prohibition on excorporation only with major and numerous difficulties, as we have seen. In the proposed system the mirror principle trivially entails the effects of the HMC and the no excorporation requirement. Crossing over an intermediate head by a nonlocal step or via excorporation would correspond to a structure where say a head H with a suffix S is spelt out in the position of S but separated from the complement that mirrors (syntactically represents) H by another head. This is impossible by hypothesis (i.e. by the mirror principle): no such complement structure could have been created, since the complement structure must mirror the structure of the lexical item — here H+S.

Note also that having eliminated head chains in favour of lexical words (and the similar but more inclusive constructs of extended and morphological words to be discussed below), the c-command and locality problems of head chains cease to be problems. In the present theory only chains corresponding to XP-chains in standard frameworks exist, these link syntactic spec positions. Head chains correspond to lexical words, which are not chains, hence they do not need to share properties like c-command and locality with (phrasal) chains. (This is not to say of course that there may not be a more inclusive concept that both chains and lexical words fall under.)

Certain points having to do with complements need immediate clarification. First, if the complement relation is a mirrored lexical spec-head relation then standard complements, like for example direct objects of verbs, must be specifiers not only in their chain top but also in their chain root position. Some theories indeed propose to base generate objects in positions that correspond to their minimalist Case-checking spec-positions (e.g. Arad 1997). If they have a lower 'VP-internal' chain root, that can still be a spec in a structure that would correspond to a multiple layer VP (with a decomposed V, cf. Brody 1995a, to appear):

(26)



In (26) Obj is the spec of V^2 and a member (constituent/dependent) of its complement hence it follows V^1 and precedes V^2. This may be an appropriate structure for a sentential complement for example that does not form a Case checking chain.

How about complements of functional categories, like (the final element of) C or auxiliaries? It would result in large scale distortion of structures if their complements were systematically treated as spec's — a move for which there appears to be no evidence. Such relations have long been assumed to be somehow different in kind from ordinary complement relations, cf. Grimshaw 1991 for a theory that is explicit on this point. It is widely accepted that such projections and their complements form larger units, which Grimshaw called extended projections and which, as noted above, recent research suggests have a universal structure. I will accordingly assume that 'extended words' (EWs) can be formed from lexical words. The relation between the elements of the lexical words forming an extended word can be only spec-head, the only configurational relation that exists. The (rightmost) element E_1 of a lexical word L_1 that is not the spec of anything internal to L_1 can become the spec of the (leftmost) element E_2 of another lexical word L_2, i.e. of the element E_2 that has no spec internal to its own lexical word L_2. Assume further that the links that create extended words are lexical, hence the extended word will be mirrored in syntax as a set of complement relations.

Consider next incorporation of open class elements, like for example the noun+verb unit in a case of productive noun-incorporation. The present framework provides the usual two major options for the treatment of such relations. These could be assimilated either to syntactic chains (that correspond to XP-chains in standard theories) or to lexical/extended words (which occupy the place of head-chains). If open class incorporation involves chains, then the incorporated element will be in a syntactic spec position. If it is of the lexical/extended word type then a unit consisting of an incorporated element and its host will normally involve two extended words. Call this unit a 'morphological word'. A morphological word is thus formed via a nonsyntactic spec-head relation from two extended words in the same way as an extended word is formed from (two or more) lexical words.

The spec-head relation involved in creating a morphological word is neither syntactic, nor strictly speaking lexical. It seems better therefore to speak of the spec-head relation as being either syntactic or morphological, where the lexical spec-head relation will fall under the morphological type, the type that is mirrored in syntax by complements. In fact it is unclear if the notion of lexical word is necessary at all given the notion of morphological word, a lexical word may simply be a morphological word that happens to be internal to an extended word. I will tentatively assume so.

Consider the following two alternative directions. One might pursue a theory in which extended words are created nonsyntactically and mirrored, but eliminate both lexical and morphological words and analyze the relevant structures as syntactic chain constructs. Alternatively, one might retain the lexical (mirrored) nature of lexical words but analyze incorporation as involving syntactic chains and syntactic spec positions. In the standard framework these options correspond to analyzing incorporation or both incorporation and inflection in terms of XP–chains instead of head-chains. In the next section I will provide some evidence from Hungarian for what is in fact a traditional assumption: head chain type relations like inflection and incorporation obey a stricter locality requirement than XP-chains. Lacking the successive step option, an antecedent in the former relations must surface in a position that is strictly local to its trace. This property is shared by inflectional morphology and open class incorporation, but not by other (XP–) chains. Given lexical and morphological words, their strict locality follows from the mirror principle in the present theory, an explanation that would be lost if these were assimilated, as is standardly done, to (XP–)chains — an assimilation that the alternatives sketched would only strengthen.[23]

Syntactic structures make it possible to recursively differentiate extended words from nonwords. The root node with all its complements is an extended word, any spec of any element of this extended word together with all its complements is an extended word and so on. But the syntactic structure now provides no means of distinguishing extended words from lexical words, i.e. without the addition of some extraneous marking, the boundaries of lexical words making up an extended word are not visible in syntax.

This provides another argument for linking the present theory and distributed morphology. The late lexical insertion of distributed morphology would provide the needed information of lexical/morphological wordhood for morphology. Thus in such a framework a syntactic specification of the distinction between lexical/morphological words and extended words is unnecessary and therefore otiose: it would only reduplicate independently available lexical/morphological information.

The theory has various additional advantages. Certain questions that arose in the minimalist and earlier phrase structure systems, and which the theory summarized in

[23]This is not to say that a syntactic spec can never form a word with the head it is the spec of. But we now expect this to be only a phonological unit — a 'phonological word'. If morphological processes can also be shown to apply to such units, then the notion of 'morphological word' in the text and the concept of 'word for morphology' will not coincide.

section 2 above made some headway in providing a solution for, simply do not arise — the optimal situation. Since there is no phrasal projection and no projection lines — or in minimalist terms merge does not create sets distinct from the elements merged, and therefore also does not label any such units — the issues of uniqueness and locality of projection do not arise.

The extended structure preservation problem also essentially disappears: there is no syntactic distinction between XPs and X^0s, hence no possibility of one type dominating the other in an illegitimate configuration. A category is always taken to be an X^0-type element by morphology. It was necessary to ensure that a set of categories is interpreted as an X^0-type element by morphology in all and only the appropriate configurations — when these form a lexical/morphological word. This issue is remotely related to but different from and much wider than the extended structure preservation question of the earlier framework.

Recall that when a category is interpreted as a word, all spec-head relations (mirrored, i.e. complement, or not) express dependencies. When the categories are interpreted as phrasal constituents, the relations correspond to constituenthood. Since head-chains reduce to a spellout issue of the mirrored lexical/morphological words, all remaining chains (with traces/copies) correspond to the phrasal chains of the minimalist framework and its predecessors. (Phrasal) chain construction is possible, because categories and relations between them can also be interpreted as phrases and constituents respectively.

As for c-command, the relation ceases to be necessary as it applied to head chains: elements of morphological words are in a dependency/domination relation with respect to each other. Perhaps the remaining cases of c-command can be taken to be the appropriate combination of the syntactic version of the basic spec-head relation and the dependency/domination relation. For example principle C will prohibit an R-expression from having an antecedent that dominates it (or equivalently, on which it depends) — like for example an Agr/Infl node. If the syntactic spec of this Agr/Infl is taken to pick up the reference of the head, then it in turn cannot corefer with the R-expression. The structural requirement on chain construction might similarly reduce to the simple notion of dependence/domination. In a wh-chain for example the Q head associated with the wh-phrase can be taken to form a chain with the (wh-feature of the) trace/copy wh-phrase which it dominates. The antecedent wh-phrase will then not be a member of the chain itself, but a constituent (whose highest category is) in a syntactic spec-head relation with the chain.

As we saw in section 3, the central issue about c-command, which is so far unresolved in a satisfactory way, is the asymmetry of the definition of this relation: x c-commands

y iff the category immediately dominating x dominates, +/- immediately, y. If the approach suggested in the previous paragraph proves feasible, then the conclusion will be that the strange asymmetry was an artefact of coalescing two distinct relations to which in fact different constraints refer: the domination/dependency relation and the syntactic spec-head relation.

Finally antisymmetry of representations follows from the fact that for each head only a single spec-head and a single mirrored lexical spec-head (i.e. complement) are allowed as syntactic relations. No overriding condition like the LCA is necessary to ensure antisymmetry. There are simply no means provided from which non-antisymmetric structure could be built. Like for example in the case of the structure of crystals, the properties of the basic building blocks determine the limits of variability of the composed larger structures.

To summarise:

(i) The only primitive relation between elements in syntax and the lexicon is spec → head.

(ii) Lexical spec-head orders are (geometrically) 'mirrored' in syntax. (Reverse order spec-head relation is the complement relation involved in extended projection. Other complements are specs in lower shells.)

(iii) Lexical atoms can be composed into lexical words, LWs, (words with inflection), extended words, EWs, (corresponding to extended projections) and morphological words, MWs (corresponding to incorporation of open class elements and perhaps subsuming also LWs) via lexical spec-head relations.

(iv) Members of EWs (heads) can form syntactic spec-head relations with other EWs.

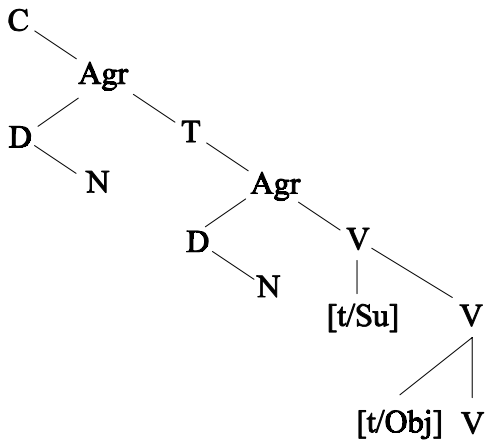
(v) An LW/MW is spelt out (in the lexical spec-head order) in the position of the highest strong head (or, lacking a strong head, in the lowest position).

For example,

(27) a. LWs: V → V → v → Agr → T → Agr, C, N, D, N, D
EWs: V → V → v → Agr → T → Agr → C, N → D, N → D

b. Mirrored in syntax:
C ← Agr ← T ← Agr ← v ← V ← V
D ← N, D ← N

c. Members of EWs can have and can be syntactic spec's:



I shall refer to the theory just summarized as the mirror theory. The main general advantages of the mirror theory are the following:

- All syntactic and morphological structures are built from the single grammatical relation: spec-head.
- Locality and no excorporation properties of head-chain type relations follow. All spec-head links in the word-structure must match a(n inverse) spec-head link in the complementation structure: structures corresponding to excorporation or non-local head-chains cannot be created.
- C-command problems of head-chains do not arise.
- Antisymmetry guaranteed (no means to violate it, hence no need for the LCA).
- No categorial projection, hence uniqueness and locality issues of categorial projection do not arise.
- No word-phrase difference in syntax, hence no extended structure preservation question.
- Single expression of word-structure in syntax (the complementation structure). All three duplications ((abc) under 'problems' above) eliminated.
- Apparent conflict between dependency and constituent structure analysis resolved.

5 V-raising in Hungarian: a case study

As noted, the version of the mirror theory outlined in the previous section entails the traditional generalization in its strict form: head 'movement' type relations are highly local. This generalization is stipulated by the HMC, and follows from relativized minimality only with some difficulty, as we have seen. Strict locality follows from the general idea of relativized minimality only weakly in any case — as witnessed by the proposals to modify relativized minimality in various ways to allow various types of apparent HMC violations. Evidence for the local nature of head-chain relations (morphological words) therefore constitutes evidence for the theory in which complementation structures mirror extended/morphological words and morphology spells out continuous subparts (lexical/morphological words) of such complementation structures. This theory dispenses with head chains but entails strict locality of head-chain type relations.

Hungarian verbal modifier and more generally prefix incorporation presents a challenge to the assumption of HMC locality, which at least superficially looks sharper than the relatively minor violations of the HMC in certain Slavic and Romance constructions. Unlike these structures, Hungarian prefix incorporation presents itself as an apparently long distance head-chain phenomena, although with various idiosyncrasies. Recently Szabolcsi (1996), and Koopman and Szabolcsi (1997) have argued for treating the relevant structures in terms of XP-chains. In what follows, I shall propose an alternative analysis that makes use of both chains that target syntactic spec positions (the equivalent of XP-chains, the only type of chains under the mirror theory) and head-chain type relations — mirror theoretically: morphological words. I shall argue that the data becomes understandable, once these two relations are separated and the traditional assumption concerning strict locality of head-chains/morphological words is made.[24]

Hungarian has a class of verb-associated elements, usually referred to as verbal modifiers, (VMs) that includes verbal particles, small clause predicates, bare nouns etc., that appear to be able to form long distance chains.

[24]Most Hungarian examples that follow are from Szabolcsi 1996 and from references cited there.

(28) Szét fogom akarni kezdeni szedni [szét] a rádiót
Apart will-I want-INF begin-INF take-INF [apart] the radio
'I will want to begin to take apart the radio.'

The verbal particle 'szét' ('apart') in (28) belongs to the verb 'szedni' ('take-INF'), but surfaces separated from it by a string of verbs. As noted by Szabolcsi 1996 these verbs do not form a reanalyzed complex, additional nonverbal material can intervene between them. The question therefore arises: Are these VM-chains long distance head-chains? Or are they pied piped XP-chains? (Given antisymmetric structures such phrasal chains will necessarily involve remnant movement.) Szabolcsi assumes the latter on the basis of the non-local nature of the relation. One might argue, however, that examples like (28) indicate precisely the untenability of the assumption that head-chains must be strictly local, and the necessity of allowing non-local chains (either via long steps or via excorporation).

I shall provide some evidence below that (28) indeed involves XP-chains (i.e. chains and not morphological words under mirror theory). But potential impressionistic support for the contrary position might be examples like (29), which show that the successive roll-up chain structure characteristic of head-chains is also sometimes an acceptable alternative option. In (29) there is a focused element (capitalized) in the preverbal focus position characteristic of Hungarian and the VM 'szét' ('apart') appears to incorporate into the verb 'szedni' ('take-INF') creating the unit 'szétszedni' ('apart-take-INF'), (29a), which incorporates further into 'kezdeni' ('begin'), (29b), and the unit 'szétszedni kezdeni' ('apart-take-INF-begin-INF') appears to incorporate into 'akarni' ('want-INF'), (29c):

(29) a. MOST fogom akarni kezdeni {szét}szedni [szét] a rádiót
NOW will-I want-INF begin-INF {apart} take [apart] the radio-ACC
b. MOST fogom akarni {szétszedni} kezdeni [szétszedni] [szét] a rádiót
c. MOST fogom {szétszedni kezdeni} akarni [szétszedni kezdeni] [szétszedni] [szét] a rádiót
'I will want to start to take apart the radio NOW'

The option of roll-up chain structure is not one that standard wh/NP-movement chains have. Note also that in analyses with successive cascading XP intraposition, e.g. Barbiers 1995, Cinque 1997, noncascading successive intraposition must be sharply prohibited. This is in contrast to the situation here, as (28) shows.

Szabolcsi (1996) and Koopman and Szabolcsi (1997) take both (28) and (29) to

involve phrasal chains. The prediction of the mirror theory is that the non-local chain in (28) cannot be head-chain type: MWs cannot skip heads. The roll-up structures of (29) however can be treated in terms of MWs since the chains here (which MWs will replace under mirror theory) exhibit strictly local links. I shall provide some evidence below that the prediction is correct: nonlocal relations like (28) are (syntactic-spec targeting) chains and not MWs. Furthermore I shall argue that the Hungarian verb raising paradigm can be understood only if the roll-up structures in (29) are treated in terms of a head-chain type relations — as MWs.

Before proceeding, we need to take account of some data that apparently complicate the situation further. The VM, like the phrase in successive step wh/NP chains, cannot stop in most intermediate positions. But unlike these, it can stop in the lowest of these:

(30) MOST fogom (?*szét) akarni (?*szét) kezdeni szét szedni [szét] a rádiót
NOW will-I apart want-INF begin-INF take-INF [apart] the radio
'I will want to begin to take apart the radio NOW'

Let us start by asking why the VM appears in front of the finite verb in (28)? Verbs in Hungarian fall into two types, some like 'fog' ('will') require a VM in a neutral (i.e. without a focus type operator) sentence, others like 'utál' ('hate') do not allow one. Thus taking an example with 'fog', (31a) is grammatical but (31b), where no VM precedes this finite verb is not. This is in contrast to the next example with 'utál', here (32a) where the VM precedes the verb is ungrammatical, but (32b) is fine. This consideration excludes also (32c), but both (31c) and (32c) are ruled out additionally, because a nontrivial roll-up structure is too large to qualify as a VM. (33) shows that infinitivals can also serve as VMs.

(31) a. Haza fogok menni
Home will-I go-INF
'I will go home'
b. *Fogok haza menni
c. *Hazamenni fogok

(32) a. *Haza utálok menni
Home hate-I go-INF
'I hate to go home'
b. Utálok hazamenni
c. *Hazamenni utálok

(33) Úszni fogok menni
swim-INF will-I go-INF
'I will go swimming'

Let us refer to verbs of the class to which 'fog' belongs as deficient verbs. It appears then that when a deficient V is tensed, it needs a VM like 'haza' ('home') or the infinitive 'úszni' ('swim-INF'):

(34) A tensed deficient V needs to be immediately preceded by a VM

It is necessary to refer to tense in (34), since all the verbs in (30) belong to the deficient class. (34) clearly cannot be allowed to refer also to infinitivals in general, since then a VM would not only be allowed in intermediate positions, but would actually have to occur in all intermediate positions — an incorrect prediction.[25]

Consider next the question of why the VM can show up in the position nearest to the lowest of its chain? Notice that the VM can surface in the lowest position of its chain only if it is an infinitive:

---

[25]It is not clear if the tensed deficient verb will have the requirement (34) for VMs when there is a focused element in the clause.

(i) MOST fogom szét [fogom] próbálni szedni a rádiót

Szabolcsi 1996 and others take structures of this type to be impossible, but (i) seems acceptable to me. If there are grammatical structure like (i) then we may assume that (34) needs to be satisfied by some member of the chain of V. In (i) 'szet' satisfied (34) for the trace of 'fogok'. The verb is then spelt out in the position of F, the (empty) focus morpheme.

If structures like (i) are invariably bad then (34) must be restricted to verbs that are not associated with focus. If in such structure the tensed verb has a focus feature (Brody 1990, 1995b), then we may take this feature to transform a deficient V into a nondeficient one, or essentially equivalently, serve as a VM for (34).

(35) a. MOST fogok akarni kezdeni hazamenni \\úszni menni
NOW will-I want-INF start-INF home-go-INF\\swim-INF go-INF
'I will want to start to go home\\to go swimming NOW'
b. MOST fogok kezdeni akarni *menni haza \\menni úszni

We can make sense of this data if we assume first that infinitives are optionally taken to be VMs and secondly that there is a requirement also on VMs that requires them to be supported:[26]

(36) A VM must be supported by a verb on its right.

It is clear that the requirement to occur in the antepenultimate position cannot be due to a requirement of the host verb. As the bad cases of (30) show, deficient infinitivals do not require a VM to precede them. But the VM cannot remain in situ. Note additionally, that when the VM shows up to the left of a verb, even if it is to satisfy its own requirement, the host verb must be of a type that licences the VM:

(37) a. MOST fogok akarni utálni\\habozni\\elkezdeni úszni
NOW will-I want-INF hate-INF\\hesitate-INF\\away-start-INF swim-INF
'I will want to hate\\hesitate\\start to swim NOW'
b. *MOST fogok akarni úszni utálni\\habozni\\elkezdeni

Verbs like 'utálni\\habozni\\elkezdeni' ('hate-INF\\hesitate-INF\\away-start-INF') are not deficient; they don't just not require but do not even allow a VM to immediately precede them. (So the deficient V that obtained a VM like 'el-kezdeni', ceases to be deficient, it

[26]In a series of infinitives only the last one can be taken as a VM. For example (35a) cannot be analyzed as 'akarni' (want-INF) being a VM that is linked to a position under 'kezdeni' (begin-INF), the relevant interpretation of (35a) (i.e. 'will begin to want to...') is not available. The restriction does not appear to be that only infinitives without complements can serve as VMs: infinitives with noninfinitive complements can be spelt out higher, without any special (emphatic/contrastive) interpretation:

(i) MOST fogok(/fogom) olvasni akarni [olvasni] egy (/a) könyvet
NOW will-I read-INF want-INF a (/the) book
'I will want to read a/the book NOW'
(ii) MOST fogom közölni akarni [közölni] hogy Mari elment.
NOW will state-INF want-inf that Mary left
'I will want to state that Mary left NOW'

licenses no additional VM.) In mirror theoretical terms, this means that they can neither form an MW with a VM, nor do they license them as their spec (in V or T or any other associated functional head).

The assumptions so far cover (28) and (30). In the good version of (30) where the VM is to the left of the lowest infinitive the MW or (syntactic-spec targeting) chain was constructed to satisfy the requirement of the VM, i.e. (36). Given some version of last resort,[27] the VM can surface higher only if some requirement forces it to do so. The infinitivals in (30) carry no such requirement. In the sentence without focus in (28) the matrix tensed deficient verb has its own requirement (34), that justifies the presence of the VM preceding it.

Let us next turn to the evidence, that strictly local, head chain-type relations, MWs, are also involved in Hungarian verb raising. Consider (38) and (39). As Szabolcsi observes, focusing the verb allows different interpretive possibilities in the two structures:

(38) AKARNI fogok [akarni] kezdeni hazamenni
WANT-INF will-I [want-INF] begin-INF home-go-INF
'I will indeed want to begin to go home' or
'I will WANT to begin to go home' (and not, say, TRY to begin to go home)

(39) KEZDENI fogok akarni [kezdeni] hazamenni
BEGIN-INF will-I want-INF [begin-INF] home-go-INF
'I will want to BEGIN to go home' (and not, say, TRY to go home)
not: 'I will indeed begin to want to start to go home'

Both (38) and (39) involve focusing an element of a series of infinitives. Both structures are grammatical, but with differing interpretive options. As indicated by the glosses, (38) can be understood either in an emphatic or in a contrastive sense. (39) can only be understood contrastively, more precisely, in the sense of 'exhaustive listing.'

Assume that focusing involves a focus head, F. Phrasal focusing in Hungarian always involves contrast or at least an 'exhaustive listing' sense, hence it is natural to assume that

[27]Last resort does not need to be a derivational condition. Translated into representational terms, it requires all non-root positions of chains to be licensed by some (checking) requirement. Under the representational version, the relationship of last resort and full interpretation is immediately brought out, and it is natural to generalize the condition to require that all positions (in chains or syntactic structures) are licensed by some (checking or thematic) relation.

the syntactic spec-F position is always taken to imply contrast/exhaustive listing. It is then also natural to suppose that a category marked for focus will be interpreted as expressing emphatic focus unless it is chain-associated with the syntactic spec-F position. Let us assume this.

The contrast in interpretation between (38) and (39) will now follow directly if (syntactic spec targeting) chains and MWs are distinguished. The focus feature of the focus marked category needs to be checked, by the F head. One possibility is to form a chain linking the category to the syntactic spec of F position, the position associated with the contrastive/exhaustive listing interpretation. The other possibility is to make the focus marked category a lexical/morphological specifier of the F node, i.e. to form the MW: 'akar→ni→fog→ok→F' linking them.[28]

(38) and (39) differ in that in the former the highest member of the string of infinitives has been focused, while in the latter a more deeply embedded one. Hence if head chains, are local and there is no excorporation, then the interpretive properties of (38), (39) will immediately follow. In standard terms, (38) can involve either an XP chain or a (roll-up) head-chain (with 'akarni' incorporated into 'fogok' and 'akarni fogok' into F). Non-local (39) can only contain the former. Under the mirror theory (39) can only involve a chain structure while (38) has both this and the MW option. Recall that the locality of the internal structure of LWs/MWs is built into a basic assumption of the mirror theory.

It follows now not only from mirror theoretical considerations, but also from the argument based on (38), (39), that the non-local chain of the VM in (28) must also be a (syntactic-spec targeting) chain. Ignoring the question of the position and the chain of the object, the mirror theory analysis of (28) is thus along the lines of (40), where I indicate with an @ sign the spellout position of the MWs.

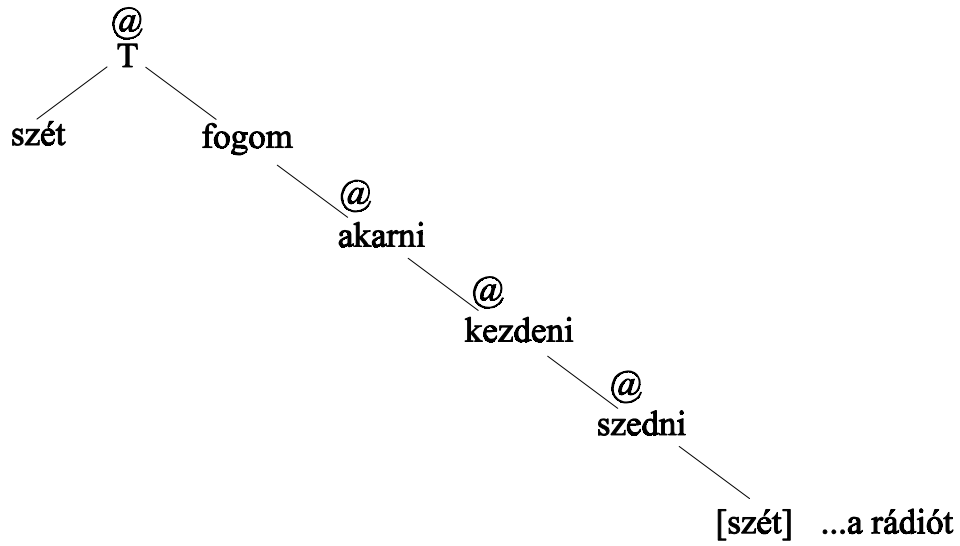
---

[28]Notice that in this MW the focus marked 'akar' is not directly the spec of F, but rather the spec of the spec of the spec of F. Putting aside the technical question (percolation or satisfaction of the checking relation at a distance) note that the problem here is a special case of a much more general issue. Thus in (i) the syntactic spec of F position is occupied by a nonfocus category whose spec is the actual focus.

(i) JÁNOS barátja hivott fel
John's friend called up

The same problem arises also in the definition of wh-phrases etc.

(40)



Szét fogom akarni kezdeni szedni [szét] a rádiót
Apart will-I want-INF begin-INF take-INF [apart] the radio
'I will want to begin to take apart the radio'

But we still have no reason to consider the chains involved in successive roll-up structures to be (syntactic-spec targeting) chains. As we shall see we in fact have good reasons not to. But let us ask first what forces the construction of roll-up chain structures, or MWs in the present theory? We know that the lowest element, a VM, must either form an XP or a head-chain to satisfy its own requirement (36). In present terms it must either form a chain or participate in an MW. How about the verb above the VM? Why does it have the same requirement in these structures? The obvious assumption is that the +VM feature percolates to it. While this must in some sense be right, the assumption cannot be quite correct: as we have also seen above in connection with (31c), (32c), the roll-up structure cannot be in the usual neutral VM position preceding the finite verb:

(41) a. *{Szétszedni} fogom akarni kezdeni [szétszedni] [szét] a rádiót
b. *{Szétszedni kezdeni} fogom akarni [szétszedni kezdeni][szétszedni] [szét] ... a rádiót

Suppose then that VMs have some feature, call it +prefix, that forces them to 'incorporate', i.e. form part of a larger

MW in which they precede some other element. We can then assume that it is this +prefix feature that percolates (optionally) to higher elements making them subject to the same requirement. The larger units created in this way do not qualify however as VMs, and only VMs are licensed in the tensed preverbal position of deficient verbs.

(42) a. +VM → +prefix
b. +prefix can percolate up MW-internally (optional)

We need accordingly to modify the licensing conditions: nondeficient infinitivals not only do not license a VM but more generally do not license a +prefix category on their left:

(43) *MOST fogok hazamenni kezdeni utálni
NOW will-I home-go-INF begin-INF hate-INF
'I will begin to hate to go home NOW'

Similarly, as (29) shows, licensing of a VM by verbs on their left is in fact more generally a question of licensing a +prefix marked element.

Given this much background, (44) provides the evidence that roll-up structures indeed involve MWs and not chains (head-chains and not XP-chains in standard terms).

(44) a. MOST fogok hazamenni kezdeni akarni
NOW will-I home-go begin-INF want-INF
'I will want to begin to go home NOW'
NOT: I will begin to want to go home NOW'
b. MOST fogok hazamenni kezdeni akarni [hazamenni kezdeni] [hazamenni] [haza]
c. *MOST fogok hazamenni kezdeni akarni [hazamenni] [haza]

As we have seen in (28) the VM can move long distance (presumably in successive steps) in a non roll-up fashion. But once a roll-up structure is formed as in (29), the top element of the roll-up structure can only form nontrivial chains by further roll-up, it is generally not allowed to form chains in successive non roll-up steps. For example (44a) can only be interpreted with 'akarni' (want-INF) having scope over 'kezdeni' ('begin-INF') , i.e. the structure must be the fully roll-up (44b) and not the partially roll-up (44c)

where 'haza' ('home') appears to incorporate into 'menni' ('go-INF') and then 'hazamenni' forms a chain that crosses the two heads 'kezdeni' and 'akarni'.

If these roll-up structures can only involve head-chain type relations, i.e. MWs, then the facts of (44) will automatically fall out, since MWs cannot exhibit non-local relations. The mirror theory analysis of (44) is then like (45), where again @ indicates spellout positions of complex morphological words.[29]

(45) a. MW: szét → szedni → kezdeni → akarni, fogom → F
EW: szét → szedni → kezdeni → akarni → fogom → F
b. Mirrored in syntax:
F ← fogom ← akarni ← kezdeni ← szedni ← szét
c. Syntactic spec's added:

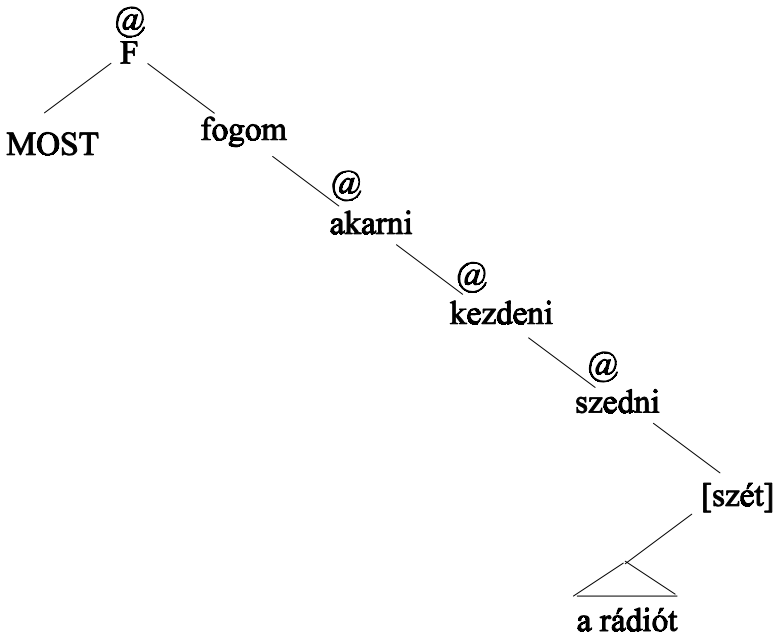
[29]In fact the generalization that roll-up structures are not able to form long distance chains does not hold in all types of structures. As noted earlier, nontrivial cascades like the ones in (29) do not qualify as VMs and therefore cannot appear before the finite verb in the VM position. The focusing structures in (ib), (ic) and (id) however contain a nonlocal chain of the roll-up structures 'szétszedni', 'szétszedni kezdeni' and 'szétszedni kezdeni akarni' respectively.

(i) a. SZÉT fogom akarni kezdeni szedni a rádiót
'I will want to begin to take APART the radio'
b. SZÉTSZEDNI fogom akarni kezdeni a rádiót
c. SZÉTSZEDNI KEZDENI fogom akarni a rádiót
d. SZÉTSZEDNI KEZDENI AKARNI fogom a rádiót

Consider the nonlocal relations between the focus position and the position where the roll-up structure is assembled. The prediction of the mirror theory is of course that these relations must be chains and not MWs.

Thus a roll-up structure can form a chain when such chain formation is triggered by some checking requirement, e.g. focus in (i). The chain structure in (44c) is not legitimate however, since the higher chain member satisfies no requirement (of its own or of its host) there.

theory



References

Arad, Maya (1997). A bi-directional view of the syntax-lexicon interface. Ms. University College London.

Baker, Mark (1988). *Incorporation: a theory of grammatical function changing*. University of Chicago Press. Chicago, Illinois.

Barrs, Andrew and Howard Lasnik (1986). A note on anaphora and double objects. *Linguistic Inquiry* 17. 347-54.

Brody, Michael (1980). On Circular Readings. *Journal of Linguistics Research* Vol.2. Bloomington, Indiana.

Brody, Michael (1990). Remarks on the order of elements in the Hungarian focus field. In I. Kenesei (ed.). *Approaches to Hungarian, vol. 3*: *structures and arguments*. 95-121. JATE, Szeged.

Brody, Michael (1994). Dependence and phrase structure. *UCLWPL 6*. University College London.

Brody, Michael (1995a). *Lexico-logical form. A radically minimalist theory*. Cambridge, Mass.: MIT Press.

Brody, Michael (1995b). Focus in Hungarian and Checking Theory. In I. Kenesei (ed.) *Approaches to Hungarian* 5. JATE, Szeged.

Brody, Michael (to appear). Projection and phrase structure. *Linguistic Inquiry* .

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo (1997). Adverbs and functional heads. A cross-linguistic perspective. Ms. University of Venice.
Epstein, Samuel (1995). Un-principled syntax and the derivation of syntactic relations. Ms. Harvard University. Cambridge, Mass.
Grimshaw, Jane (1991). Extended projections. Ms. Brandeis University.
Groat, Erich M. (1995). On the redundancy of syntactic representations. GLOW abstract. Tromsø, Norway.
Higginbotham, James and Robert May (1981). Questions, quantifiers and crossing. *The Linguistic Review* 1. 41-80.
Hudson, Richard (1990). *A word grammar of English*. Oxford: Blackwell.
Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
Koopman, Hilda (1994). Licensing heads. In Norbert Hornstein and David Lightfoot (eds.). *Verb movement*. Cambridge University Press.
Koopman, Hilda and Anna Szabolcsi (1997). The Hungarian verbal complex: Incorporation as XP-movement. Ms. UCLA.
Larson, Richard (1988). On the double object construction. *Linguistic Inquiry* 19. 335-391.
Manzini, M.Rita (1995). From merge and move to form dependency. In *UCLWPL* 7. University College London.
Phillips Colin (1996). Order and structure. Doctoral dissertation. MIT. Cambridge, Mass.
Pollock, Jean-Yves (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20. 365-424.
Pollock Jean-Yves (1993). Notes on clause structure. Ms. University of Picardie, Amiens.
Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, Mass.: MIT Press.
Rizzi, Luigi (1995). The fine structure of the left periphery. Ms. University of Geneva.
Starke, Michal (1995). On the format for small clauses. Ms. University of Geneva.
Szabolcsi, Anna (1996). Verb and particle movement in Hungarian. Ms. UCLA.