

MIRROR THEORY*

Michael Brody

This text was completed in 1997. Most of chapter 3 (with minor changes) will eventually appear in *Linguistic Inquiry*, other parts elsewhere. The analyses in 1.4 and 4.2 need some revisions, the former because of literature that came out since 1977, the latter because it does not deal with the question of how the V+Infl morphological unit is created. (Note added in 1999.)

1997

Department of Phonetics and Linguistics

University College London

Gower Street, London WC1E 6BT

m.brody@ling.ucl.ac.uk

Contents

Introduction

Chapter 1. Remarks on Perfection in Syntax

1.1. Introduction

1.2. Perfection in living systems

1.3. Perfection in syntax: Move

1.4. Perfection in syntax: representations and derivations

1.5. Bare checking theory, successive step and partial "movement" chains

1.6. Summary

Chapter 2. Properties of projections and of C-command

2.1. Introduction

2.2. Insert and Project

2.3. C-command

2.4. Attempts to explain c-command

2.5. Precedence and time

Chapter 3. A radically restrictive theory of syntactic representations

3.1. Introduction

3.2. The standard explanation of the mirror principle and its problems

3.3. Mirror Theory

3.4. Specifiers and complements

3.5. C-command and antisymmetry

3.6. Summary

Chapter 4. "Roll-up" structures and morphological words

4.1. Introduction

4.2. Adverbials on the right

4.3. V-raising in Hungarian

Introduction

Perfect syntax assumes that the syntactic system of natural language is a subsystem of the mind-brain of humans of a kind that has not been distorted either by accidents of its own evolutionary history or by other evolutionary contingencies indirectly affecting it. Research in syntax in the last decades, particularly within the framework of the Principles and Parameters and the Minimalist approaches suggests strongly that syntax may be among those, perhaps less common, aspects of a living organism that are perfect in the same sense in which natural science takes certain aspects of our non-living environment to be perfect: describable with a non redundant and simple set of primitives and axioms. In other words we expect that the specific and idiosyncratic properties of the system should be encapsulated in these primitives and axioms of the theory rather than appear as stipulated exceptions and supplementary statements and generalizations to the system the basic principles generate. In this spirit I shall argue here against some aspects of the minimalist framework that involve the assumption of externally forced or other syntactic imperfections. The main proposal of this work however is a theory of syntactic and morphological representation that is highly restrictive both

conceptually and empirically. This theory is put forward as a central subcomponent of an eventual full theory of perfect syntax.

Imagine a "proto-language" composed of minimal sound-meaning units that are not further divisible. Each such unit U is an expression E , although not each expression is a unit. To allow the language to have an infinite set of expressions we introduce the recursive step: a unit may be added to an expression to form another expression:

$$(1) E+U=E'.$$

This creates representations like (2)

$$(2) [_{E''} [_{E'} [_{E'} [_{E'} U] +U'] +U''] +U'''].$$

Consider the order of E and U in (1). Take left to right order on the page to express time in its normal direction. Suppose that we do not abstract away from the fact that expressions are constructed in time, ie. the representations are generated and/or interpreted left to right. Then the order in (1), that is U following E seems more natural than the converse since it need not involve storage in the construction of expressions. (1) creates representations like (2), the inverse

order rule in (3) would create representations like (4)

(3) $U+E=E'$

(4) $U''' + [_E U'' + [_E U' + [_E U]]]$

In (2) U is E , E (already constructed)+ U' is E' , E' (already constructed)+ U'' is E'' , etc, whereas on the inverse order in (4) $U''' + E''$ cannot be generated/interpreted until E' is, E' cannot be constructed until E is, and so on.

Let us explore (1) further. Notice that another way of thinking of the same system would be to assume that the representation is in fact just the string in (5)

(5) $U+U'+U''+U''' \dots$

and take the expression-bracketing of (2) to be the way (5) is interpreted. The string in (5) can then be thought of as follows. Take a U that (immediately) precedes U' to be (immediately) contained in U' . Now the string is ordered also by the containment relation: each unit immediately contains the previous one. In other words we take (5) to be equivalent to (6) where units are constituents, their respective position expressed by the position of their label.

(6) $[[[U] \ U'] \ U''] \ U''']$

Let us also define here the specifier-head relation. Thinking in terms of precedence and containment, there are two ways of doing this. We can take the spec of U' to be the U that immediately precedes U' or the U that is immediately contained in U' . (Note that under both definition we take U to be an element that may contain other units.)

The system generated by (1) is antisymmetric; it provides a linear ordering of the units in the representations. It is also one in which (7) holds

(7) if x precedes y then x is not less deeply embedded than y .

This is in contrast to Kaynean syntactic structures, where the corresponding implication is that "if x follows y then x is not less deeply embedded than y ". This "proto-language" system seems to match quite closely however some central aspects of morphology, namely suffixation generally understood. In other words it fits those aspects of morphology that fall under the mirror principle. Accordingly I shall refer to the system characterized so far as "proto-morphology".

Proto-morphology is fairly powerful: it can represent

complex words like for example Hungarian (8) or (9). Pretend counterfactually, for the sake of simplifying the example, that the indivisible sound-meaning units are as indicated:

(8) i-tat-tam

drink-cause-past(1sg)

"(I) made (him/her/arb) drink"

(9) olvas-ni akar-tam

read-INF-want-past(1sg)

"(I) wanted to read"

Let us return now to the mirror image of proto-morphology, the system that has representations like (4) or in terms of containment representations where the first unit is the highest and the last one the most deeply embedded. Inverse proto-morphology loses the advantage of not requiring storage, but has the type of antisymmetric structure characteristic of syntax. Let us take inverse proto-morphology to be "proto-syntax".

Directionality considerations having to do with specifiers introduce an element of a(nti)-symmetry into this otherwise symmetric system. Suppose not only the symmetry but also its directionality, as expressed in the definition based on

precedence (rather than containment) of the notion of specifier above is real. In other words a specifier must precede the head, the head has a single "valence" both in morphology and in syntax, positioned "on its left" to bind a preceding spec. We could speculate that this is because of the proto-morphologically successful but ultimately, in the larger syntactic context, failed attempt to avoid storage. Perhaps, and perhaps more plausibly, the real reason of this directionality constraint is not language internal and like a deeper explanation of the symmetry of the system remains to be provided by the eventual integration of syntax with the brain sciences, and biology, -- where symmetries and directionality effects are both rather commonplace.

Although proto-syntax must involve storage, it can gain a different and apparently greater advantage. By mirroring a proto-morphological structure proto-syntax frees up the spec-positions associated with the units in the representation. Take these mirrored spec's to still remain specifiers (proto-) morphologically but not syntactically. Proto-morphological specifiers show up on the opposite side of the associated head due to the mirror relation: call them (proto-)syntactic complements. The positions vacated by complements can thus be occupied by other structures of proto-syntax, in other words by mirrored structures of proto-morphology. Thus the directionality

condition captured in the definition of specifier interacts with the mirror image symmetry between proto-morphology and proto-syntax in a way that introduces and to that extent explains the asymmetry between these two otherwise symmetrical components: in proto-syntax, but not in proto-morphology, the possibility arises of additional specifiers.

For example $V+v+Infl$ in English is a(simplified) proto-morphological structure. Its mirror image $Infl+v+V$ is the (similarly simplified) complement line of the sentence. Spec positions of $Infl$, v and V , units in this complement line can again be filled by proto-syntactic structures that correspond to parts of the subject, object etc. The operation is recursive, as the inserted proto-syntactic structures are also inverse proto-morphological strings, their units also have their spec-positions freed for further syntactic specifiers. The major claim of this monograph is that the resulting system, call it extended proto-syntax, is in fact the correct theory of the syntactic "phrase"-structure of human language.FN1

Footnote to the Introduction

FN1 I do not intend any historical claim perhaps implied by the terms of "proto-syntax", "proto-morphology" etc. to be taken seriously. The discussion in the text is meant as a logical reconstruction not as a historical one.

Chapter 1.

Remarks on Perfection in Syntax

1.1. Introduction

I argued in earlier work that the assumption of the minimalist framework that various aspects of syntax are due to imperfections forced on this modul of the grammar by other systems it interacts with is neither empirically justified nor conceptually desirable. I refer to the alternative of the minimalist framework that rejects this assumption as perfect syntax. In section 1.2 I comment on some general ontological issues arising from the apparent perfection of a system of biological organisms. I argue that the fact that a biological system is perfect, in the same sense in which we assume the existence of perfect systems in non-living nature, is not necessarily a surprising fact and is certainly not in need of an ontological apology. In section 1.3 I sketch some of the major consequences of the rejection of the hypothesis of externally forced syntactic imperfection that have to do with the concept of "move", consequences that contribute to a substantive characterization of the theory of perfect syntax. In the next two sections I discuss some apparently correct predictions that follow from additional necessary restrictions in

this approach. Perfect syntax integrates the theory of "lexico-logical form" of Brody 1995a in that it must reject the derivational-representational duplication of the minimalist approach. I argue in 1.4. that in addition to the fact that the theory predicts the reconstruction asymmetry between adjuncts and arguments, the additional reconstruction asymmetry between "referential" and "non-referential" phrases (Heycock 1995) also follows in this more restrictive approach. I show that despite appearances, the two asymmetries do not motivate maintaining the less restrictive and methodologically objectionable derivational-representational system. In 1.5 I spell out those predictions of bare checking theory (Brody 1997a), another near-necessary consequence of the perfect syntax view, that make it possible to understand better certain basic properties of successive step and partial "movement" structures.

1.2. Perfection in living systems

Throughout the history of generative grammar, the methodology of the natural sciences has been conspicuously successful in syntactic theorizing. Search for symmetrical and other elementary patterns, and the elimination of redundancy lead to better predictions, deeper understanding and new insights with

amazing regularity. Syntax apparently can successfully aim for perfection in the same sense in which physics does, by looking beyond the often chaotic surface appearance to search for underlying simplicity and elegance. This state of affairs was often taken to be in conflict with the basic assumption in the field, that the object of research is an element of the human mind-brain. As a biological system, syntax could be expected to exhibit redundancies and more generally the inelegant organization that the "tinkering" of evolution is typically thought to produce.

The conflict between what the linguist finds and what the biologist might expect the linguist to find may only be apparent, however. Syntax is not at all the sole instance where some aspect of a living organism exhibits a perfect structure. Flowers are biological systems, but their petals are apparently often organized into patterns elegantly generateable from a nonredundant base. Furthermore, some patterns of living organisms, show clear structural correspondences with non-living structures. For example the scales of reptiles or the skin of a pineapple have strongly similar features to the structure of crystals. Given that the laws of physics establish boundary conditions for both the organic and the inorganic world, it in fact need not be particularly surprising to find that biological organisms sometimes exhibit perfect structures of the type that

can arise where no life is present.

It is sometimes suggested that the appearance of conflict is due to the fact that unlike biology, research in syntax analyses the relevant aspects of the mind-brain as a black box, abstracting away from the instantiating mechanisms. It is thought that the analysis is at a higher "functional" level, where perfection is to be expected. Grammar is taken to have the status of a theory of the kind that in a different domain would take for example the kidney to be an unarticulated filter and characterizes this organ in terms of input (blood) and output (blood and urine) measurements, --call this a virtual kidney. Biology on the other hand, it is suggested, takes account of the microstructure of the relevant portion of the mind-brain and the kidney, in the latter case taking note of the component processes and mechanisms like glomerular filtration, tubular secretion and reabsorption etc.FN1

It is of course unquestionable that current syntactic theory is abstract in the sense that it has little to say about the underlying instantiating mechanisms. But as the earlier examples suggest, perfection does not appear to be necessarily incompatible with life on any level. Just as we should not feel compelled to say that flower petal or reptile scale arrangements are compatible with living organisms only because there may be complex and inelegant systems and mechanisms determining or

underlying these patterns, syntactic perfection does not seem to be in need of a similar apology.

At the same time the virtual kidney metaphor seems rather misleading for at least two reasons. First, it suggests that grammar, like the virtual kidney does not describe a real object. This is true only in the rather uninteresting sense, in which for example biology, to the extent that it is not equivalent to physics, does not study a real object either, --biology looks at say the glomerulus as a black box, in abstraction from the instantiating atoms, quarks, strong forces or whatever.

Secondly, consider the suggestion that the fact that current syntax provides a higher level "functional" analysis, not only makes perfection in grammar compatible with this system being a property of a living organism, but makes perfection expected. Here we need to separate two issues: abstractness and functionality.

Taking the issue of abstractness first, it should of course be possible for some object that has an imperfect structure on some level of analysis to show perfection on a higher level. But this is clearly not a generally necessary or expected state of affairs. The periodical table for example is abstract in roughly the same sense as current grammar: it says nothing about the instantiating mechanisms. But this makes it less rather than more perfect: various unexpected aspects of the table only become

regular and comprehensible when the underlying physical mechanisms are taken into account.

Could perfection in syntax be due to to the fact that it describes a common function of a (possibly a heterogenous overlapping etc.) set of systems of the mind-brain? This again is very dubious already on general grounds: some systems of living organisms are very well adapted to carry out a given function in a given set of circumstances while other such systems are not.

In addition since an argument from common function assumes the existence of such a common function, it invites the question of what this function is. The usual answer is that grammar serves the function of communication. This is presumably true only in the rather weak sense in which say, insects have the function of fertilizing plants. But whatever the strength and the correct interpretation of this answer, there are good standard reasons (like the unusability of multiple center-embedding or the various structural and other ambiguities, etc.) that appear to show that grammar is not a perfect system for this function.

To summarize, the natural language related syntax patterns of the mind-brain do not appear to be a mysteriously unique among systems of living organisms in their perfection. For this among other reasons, apparent perfection provides no basis for negative conclusions regarding the ontological status of grammar.

1.3. Perfection in syntax: Move

Since Chomsky's early work the phrase structure and the transformational systems have been taken to be the two main components of transformational generative syntax.^{FN2} Starting with Chomsky 1973, the early inventory of transformations has been reduced to a non construction-specific rule of Move or Affect (the latter to encompass both movement and deletion), and conditions on it. The elimination of PS rules essentially in favour of lexical projectional requirements, proposed within the later principles and parameters framework (Chomsky 1981) made it possible to view the phrase structure component as consisting of a general rule of Project, again together with conditions restricting its application.

In the more recent minimalist theory (Chomsky 1995), phrase structure, in abstraction from Move, is taken to be a near perfect system. Move on the other hand is conceptualized as an imperfection, one that the external interacting systems force on syntax. This approach assumes that optimally lexical items would be assembled into syntactic structures by a relatively simple operation of Merge. This operation recursively links two or more elements each of which may be a lexical item or a structure earlier created by Merge. Simplifying somewhat, Merge forms a set that includes these elements and which is labelled (projected) by

one of the elements to which Merge applies. Notice though, that at least labeling is an obvious deviation from perfection. The alternative approach to phrase structure in chapter 3 below will address this problem among many others.

In the present chapter I will not consider matters that have to do with phrase structure, but will continue to sketch issues that situate the theory of phrase structure to be developed in chapter 3 in a general framework. Consider then how Move is viewed in the minimalist setting. In contrast to Merge, Move is taken to be due in its entirety to "extraneous" conditions..., conditions imposed on [the computational system] by the ways it interacts with external systems. That is where we would hope the source of "imperfections" to lie, on minimalist assumptions" (Chomsky 1995, p. 317)

Similarly, once we have Move, the minimalist approach takes covert, post-Spellout point movement of features (in the terminology of Brody 1995a expletive-associate chains) to be the optimal case. Before the Spellout point, movement involves full categories together with their phonological material. One assumption concerning pre-Spellout point movement (recently supported by Lasnik 1997) is that this is due to PF requirements.FN3 "...only PF-convergence forces anything beyond features to raise." Again this has been viewed as an imperfection: "If that turns out to be true or to the extent that

it does, we have further reasons to suspect that language "imperfections" arise from the external requirement that the computational principles must adapt to the sensory-motor apparatus, which is in a certain sense "extraneous" to the core systems of language as revealed in the [lexical input to LF] computation" (Chomsky 1995, p.265).

The view that external systems force imperfections on syntax is rather surprising. I think the desirability of this idea fades when it is compared with an alternative that is based on more standard methodology. According to this alternative picture, the observed imperfections are only apparent, and the fact that they show up is due to the interaction of otherwise "perfect" subsystems. This view is of course more restrictive and therefore preferable if facts allow us to maintain it. I shall continue to refer to the framework that rejects the idea of (externally forced) syntactic imperfections as "perfect syntax"(Brody 1997a, to appear). I take syntax to be perfect in the sense in which for example the propositional calculus is perfect: a system with a simple set of primitives and axioms. Of course the metaphor is not meant to imply that the nature of syntax is not an empirical question: syntax is one modul of a larger system of our mind-brain.

In the setting of the minimalist program arguments based on general considerations similar to those in the physical sciences,

like simplicity, symmetry, nonredundancy etc. are at last wholeheartedly accepted. Provisos that these features may be surprising properties in biological systems are not taken to weaken the force of such arguments any more and the issue such provisos raise is placed where it appears to belong: it is "a problem for biology and the brain sciences, which, as currently understood, do not provide any basis for what appear to be fairly well established conclusions about language" (Chomsky 1995, p.2). But this background would seem to remove any remaining general motivation to deviate from the standard methodology, according to which "data imperfections" do not point to imperfections in the underlying systems, but rather result either from the interplay of perfect systems or simply from our incomplete understanding of these.

Furthermore, in order to deviate from the optimal assumption of syntactic perfection, it would be necessary to demonstrate that this is in conflict with external requirements. But not only has this not been demonstrated, it is very unlikely (without of course being logically impossible), given the current state of our knowledge, that anything of the sort could be convincingly shown. This is because the system of perfect syntax is not a priory given, whenever a perfect system is in conflict with external requirements, there is always a possibility that a different perfect system would not create the conflict.

Consider the assumption that Move is an imperfection forced by LF checking requirements. If Move was really an imperfection, we would like to know, why it cannot be avoided by say, freely deleting checking features or by checking features being always generated in a position that is accessible without Move, etc. Further possibilities are easily imaginable and numerous: in order to demonstrate conflict between external requirements and Move-less perfection all would need to be excluded on principled grounds. Nothing like this has been established, or looks demonstrable. Remember also, that although the result would in some sense be interesting, it would not be a desirable one. (Similar comments apply in the case of the idea that overt Move of categories and phonological material is a forced imperfection. How do we rule out overt movement restricted to the Spellout component, how do we rule out free deletion of strong features etc, etc.)

Perfect syntax can be thought of as a theory that attempts to take the minimalist program at least in certain respects to its logical conclusion. It is therefore interesting to observe, that already at this early stage it is clear that the outlines of this approach will look rather different from the minimalist picture. The step of disallowing the remaining (forced) syntactic imperfections of the minimalist program might seem to be minor,

but it naturally leads to the rejection of many of the central differentiating features of this theory. Let me try to indicate here very sketchily some major differences.

First, since perfect syntax does not allow imperfections, Move cannot be one. It is therefore necessary to find a different conceptualization for this relation. Given the representational nature of perfect syntax, one obvious alternative is the copy relation. Move creates LF copies in the minimalist framework; -- properties of Move can in principle be taken to be properties of the relevant copy relation. Importantly, unlike Move, the copy relation does not seem to be an imperfection: it appears to be a necessary feature in syntax, given its relation to the lexicon. Lexical items must be in a copy relation with their correspondents in a syntactic structure. An entry does not disappear from the lexicon when used syntactically.

The approach naturally leads to the (partial) elimination of the Move relation from syntax. Suppose that the conceptual-semantic interpretive system optionally takes two (or more) identical elements to be in the relevant "Move" relation (abstractly understood, not involving movement). In most, and perhaps all cases, there will be independently motivated principles violated if the wrong choice is made. In particular the generalized projection principle, GPP (Brody 1995a) will require that only the most deeply embedded one of the copies

occupy a selected or selecting thematic position.FN4 This ensures in most core cases that copies not in a chain relation cannot be misinterpreted as co-members of some chain.

It is sometimes assumed that the Spellout rule needs to know about the movement history of a derivation, since it needs to distinguish copies that are traces of Move from others. If the abstract "Move" relation is established only at a post-syntactic level L, then L and PF will need to be linked by a global recoverability condition. But such a recoverability condition between PF and some non-phonological level may be necessary anyway, given the natural inclusiveness condition of Chomsky 1995, that prevents lexicon-extraneous marking of elements as having participated in a Move relation.

Secondly, the minimalist strategy appears to be to attribute nonoptimal syntactic properties to the Move relation, presumably on the assumption that Move is an imperfection in any case, hence the fact that it shows non-optimal properties is not surprising. The more restrictive framework of perfect syntax cannot proceed this way. There are reasons to think that some of the putative "imperfect", ie. stipulated, properties of Move are not real. Brody to appear argues for example that "chain-uniformity", the requirement that chain co-members all are of the same X'-projection level dissolves (or, more precisely, becomes a special case of the identity/nondistinctness condition on chain co-

members) given an improved theory of phrase structure. The theory of syntactic representation to be developed in chapter 3 below, in which X'-projections are eliminated, will likewise have no place or need for such a condition. Other properties attributed to Move may be more general and not Move-specific properties.

"Last resort", to take another case, the principle that requires that each application of Move be triggered by some checking relation translates in a representational framework as a requirement that all non chain-root positions be licensed by a checking relation. This immediately suggests the generalization: all chain positions, ie. all positions in syntactic representations, must be licensed. By the GPP only chain-roots are licensed by selectional (or modificational see fn.4) relations, non-chain roots therefore can only be licensed by checking relations (assuming the nonexistence of other types of licensing). Thus "last resort", instead of being a stipulated property of Move, becomes just a particular case of the general principle of "full interpretation". (For full interpretation, see Chomsky 1986, 1995).

Or take c-command. In Brody to appear I argued that the fact that chains are restricted by this relation is due to information transmission (feature percolation) in syntactic trees being directional: this can only proceed upwards, from a position P to some P' that c-commands P. This structural directionality is the

major factor that creates/explains the GPP, since it constrains the inheritance of thematic information. Chains that violate the c-command requirement will automatically violate the GPP. The directionality condition may, but does not need to be chain specific. While one could attempt to reduce all relations --like anaphora, bound variables, quantifier scope etc.-- that c-command applies to to chain relations, it is quite possible to argue that c-command (or feature percolation) is not only relevant for chains, but applies more widely, to these various other related types of interpretive relations.

It remains to ask (i) why a directionality constraint exists at all and (ii) why it takes the particular form (c-command) it does. As to the first question I shall be agnostic: I shall take this to be axiomatic, along with the different (but perhaps related) linear directionality requirement to be discussed in chapters 2 and 3 below. Directionality is widespread in biological organisms, in particular both amino-acids, and DNA have a directional organization. (Of two mirror image forms only one exists in nature.) It seems neither too fanciful nor too implausible to tentatively expect that the explanation of the existence of the directionality constraints will be provided by the eventual harmonization of linguistics with the brain sciences and biology. The second question on the other hand, I shall address in some detail in the next two chapters.

In the theory to be developed below in chapter 3 I will argue that chains in the general case involve the relation of domination: a set of features dominates another identical one and ultimately the category, again with an identical feature set, that is the most deeply embedded member of the chain. Thus structural directionality can apply in its simplest form, distinguishing only "up" (towards the root node) and "down" in the tree. It is only in the case of the so-called "overt" chains that in addition a spec-head relation must be established between another (identical) category and a feature set, that dominates the category that is the deepest member of the chain. Note that neither domination nor the spec-head relation is likely to be a chain specific concept

In sum perfection in the domain of chain theory does not appear to be obviously out of reach either.FN5

1.4.Perfection in syntax: representations and derivations

The third major difference between the minimalist theory and perfect syntax I discussed in detail in earlier work (Brody 1995a, to appear) perfect syntax cannot adopt the redundant derivational-representational organization of the standard minimalist framework in which the central concept of chain/Move

is duplicated by being captured both representationally and derivationally. There are no theoretical considerations and practically no direct empirical arguments in favour of such a theory. Perhaps the strongest case that has been constructed is one based on the analysis of reconstruction effects given by Heycock 1995. Heycock adopts Lebaux's (1989) treatment of the adjunct-argument reconstruction asymmetry exemplified in (1) and (2):

(1) a. *[The claim that John_x was asleep], he_x was not willing to discuss t

b. [The claim that John_x made], he_x was not willing to discuss t

(2) a. *[Which fan of Madonna_x] did she_x like best t

b. [Which fan near Madonna_x] did she_x like best t

An R-expression in the complement ("that John was asleep", "of Madonna") in the a. examples appears to trigger sharper principle C effects in the reconstructed position than the R-expression embedded in the adjunct ("that John made", "near Madonna") in the b. examples. Working in a derivational framework, Lebaux assumes that as a consequence of the projection principle, arguments but not adjuncts must be present throughout

the derivation, in other words that adjuncts may be added after movement has taken place. He takes principle C to hold at all derivational points. Thus there will be a derivation of the b. examples that does not violate principle C, namely the one where the adjunct containing the relevant R-expression has been inserted after movement of the modified category. The a. examples, where the R-expression is in an argument will necessarily violate principle C at the pre-movement stage.

Heycock notes an interesting further reconstruction asymmetry, between what is traditionally and (as she notes, citing Kroch 1989) rather sloppily called "referential" and "non-referential" elements

(3) a. [Which stories about Diana_x] did she_x most object to t

b. [Which lies aimed at exonerating Clifford_x] did he_x
expect t to be effective

(4) a.*[How many stories about Diana_x] is she_x likely to invent
t

b.*[How many lies aimed at exonerating Clifford_x] is he_x
planning to come up with t

(3a) has an interpretation as indicated in (5), and (3b) is parallel. Here a referential phrase is fronted and these sentences exhibit no principle C violation. This is in contrast

to the amount questions in (4) where the non-referential interpretation of the fronted category is most natural, as made explicit for (4a) in (6):

(5) "There is a set of stories about Diana; what subset of that set did Diana most object to?"

(6) "There is a number such that Diana is likely to invent that number of stories about Diana; what is that number?"

To account for both the adjunct-argument and the referential-non-referential contrasts, Heycock proposes to conjoin Lebaux's theory of adjuncts being insertable at later stages of the derivation with the hypothesis that nonreferential phrases necessarily reconstruct at LF.

The interpretation in (5), slightly more formally stated in (7a) can be associated with (7b), which results under the copy theory of movement if the copy in the trace position deletes by LF, (curly brackets indicate deletion):

- (7) a. Which y : y story about Diana, Diana most objected to y
b. [Which stories about Diana] $_y$ she most objected to {which stories about Diana} $_y$

The interpretation in (6)/(8a) can be associated with (8b), which

results in the copy theory framework by deleting the restriction in the higher copy and the quantifier in the lower one.

(8) a. Which y (y a number), Diana is likely to invent y stories about Diana

b. [How many y { y stories about Diana}] she is likely to invent [{how many y } y stories about Diana]

Heycock argues that in order to capture both the argument adjunct and the referential non-referential asymmetry, principle C has to hold at all derivational stages including LF. In other words she needs to supplement her LF-representational account of the referential-nonreferential asymmetry with Lebaux's derivational theory of the adjunct argument asymmetry. If principle C held only at LF, then on Heycock's other assumptions an R-expression in a referential phrase would always escape principle C if the referential phrase is moved higher than the putative binder of the R-expression. Thus her LF account of the referential non-referential asymmetry cannot make the distinction between R-expressions in adjunct and in arguments, since both the adjunct and the argument may be internal to a referential phrase. With principle C holding only at standard LF representations and without any further elaboration, the expectation would be that an R-expression fails to trigger principle C effects independently

of whether it is inside an adjunct or an argument as long as it is in a referential phrase that has been moved over the putative binder of the R-expression. This appears to be incorrect, as the examples in (1) and (2) above show.

Under Chomsky's (1995) theory, principle C must, but according to Heycock's evidence --as she notes--, cannot hold only at LF. Chomsky captures the argument adjunct asymmetry by postulating a "preference principle" that results in arguments but not adjuncts normally occurring in their reconstructed position at LF.⁶ Thus in the examples in (1) and (2) the R-expression in the complement but not in the adjunct will trigger principle C. On Heycock's analysis of the referential non-referential asymmetry reconstruction results in a distinct interpretation. She notes that given this assumption the preference principle gives rise to the incorrect expectation that questions containing complements allow only the non-referential reading. (Actually this depends on whether the preference principle overrides interpretive requirements or conversely. In the former case we indeed get the prediction that phrases with complements must be referential. In the latter, the prediction is different, although equally incorrect: that there is no adjunct argument asymmetry under reconstruction, that an R-expression triggers principle C under reconstruction iff it is inside a non-referential phrase.)

It is interesting and perhaps revealing to note that this rare attempt to show that syntax has to make use of both derivations and representations is incompatible with basic assumptions in the minimalist framework, namely with the restriction that syntactic conditions can only hold at the interface level. I would like to show that given the assumptions about LF representations in Brody 1995a, the problems Heycock raises do not arise and her clearly undesirable conclusion need not follow.

Let us first note with Heycock a slight complication: amount questions can also have a referential reading. Under this reading an embedded R-expression will not trigger principle C under reconstruction:

(9) a. ?[How many stories about Diana_x] was she_x really upset by
t

b. [How many lies aimed at exonerating Clifford_x] did he
claim that he_x had no knowledge of t

Heycock does not explain how paraphrases of (9a) like say (10a) or (10b) would relate to the syntactic derivation without reconstruction indicated in (11).

(10) a. "There is a set of stories about Diana that Diana was

really upset by; what is the cardinality of that set?"

b. Which y : $y = |z|$: z stories about Diana, Diana was upset by $|z|$

(11) [How many stories about Diana] _{z} was she really upset by {how many stories about Diana} _{z}

For concreteness, I will assume that "how many" (corresponding to "which y " in (10b)) is extracted further to have large enough scope and that at LF it quantifies over the same variable, z , as the one associated with the chain indicated in (11). These assumptions give a structure like (12), which after the indicated deletions plausibly represents the intended meaning.

(12) [How many z] [{how many z }: z stories about Diana] _{z} was she really upset by z {how many z : z stories about Diana}

According to the partially determined full interpretation (PDFI) hypothesis outlined and defended in Brody 1995a, at least in A'-chains all chain internal copies are present at the interface level, LF. Binding theory operates at this level, while deletion of some of the copies and with it disambiguation of quantificational structure takes place at some later stage of the interpretation. In this theory both the adjunct argument and the referential non-referential asymmetry can straightforwardly be

treated at LF. The argument adjunct asymmetry was treated in this theory essentially in terms of the assumption that the projection principle requires selected complements but not non-selected adjuncts to be present in all chain co-members. This is essentially an adapted version Lebaux's hypothesis, see Brody 1995a for detail and advantages of the reformulation.

Let us summarize Heycock's data considered above: principle C reconstruction effects can be avoided only by R-expressions in adjuncts, but only if the adjunct is in a "referential" argument. Therefore all that remains is to show that an R-expression in a non-referential adjunct cannot make use of the option of not being present in the lower chain position. If in the nonreferential reading of (4a) the adjunct is added only in the higher copy, then instead of (8b) we get (13):

(8b) [How many y { y stories about Diana}] is she likely to invent
[{how many y } y stories about Diana]

(13) [How many y { y stories about Diana}] is she likely to invent
[{how many y } y stories]

But (13) cannot express the right meaning, namely "what is the number that she is likely to invent that many stories about Diana". This is particularly clear when the (post-LF) deletions in the semantic component indicated in (13) are taken into

account. The meaning associated with (13) can only be "what is the number that she is likely to invent that many stories".

In fact on the near inevitable assumption that chain internal deletion is restricted to copies, (presumably a recoverability effect,) we cannot in fact derive the wrong meaning either: "y stories about Diana" cannot delete in the semantic interpretation of (13), since there is no controller for this deletion. Notice that if "y stories" in the higher copy is deleted then "about Diana" remains semantically unlinked. Crucially, "y stories about Diana" here does not equal "y stories AND y about Diana": recall that y is a number.

No such problem arises on the referential readings. Take the referential reading of the amount question in (9a). Here if the adjunct is present only in the higher copy, then instead of (12) we get (14):

(12) [How many z] [{how many z}: z stories about Diana]_z was she really upset by z{how many z: z stories about Diana}

(14) [How many z] [{how many z}: z stories, z about Diana]_z was she really upset by z{how many z: z stories}

Deletion of "how many z" and "z stories" in the reconstructed position is possible under identity. This deletion will not leave the adjunct, "about Diana" semantically unconnected with the rest

of the structure. The referential reading of (9a) asks for the cardinality of the set composed of objects that are stories, that Diana was really upset by and which are about Diana. A similar analysis is appropriate for the questions involving fronted referential phrases in (3). Thus under the PDFI approach of Brody 1995a, --now viewed as an aspect of the theory of perfect syntax-- the existence of the two reconstruction asymmetries follows from basic semantic and syntactic principles, like the recoverability principle and the GPP. No empirical motivation appears to remain for the theory where chains and move redundantly duplicate each other and the binding theory applies to both.

1.5. Bare checking theory, successive step and partial "movement" chains

There is a consequence of the rejection of the derivational-representational grammar architecture which seems to be important enough to merit being considered to be a fourth central difference between the minimalist theory and perfect syntax. Chomsky (1995) introduces a difference between interpretable and non-interpretable checking features. Non-interpretable features are visible for the derivation, but cannot be tolerated at LF, given the principle of full interpretation. Interpretable features can be present both during the derivation and at the

interface. Interesting consequences follow from this distinction.

But a distinction in such terms is natural only in the framework, where both representations and derivations are postulated. Non-interpretable features are derivational in the sense of not being allowed to be present at the (sole) syntactic representation LF, while interpretable features can exist both derivationally and representationally. Since interpretable features presumably cannot be eliminated from the theory, perfect syntax must dispense with non-interpretable ones. The assumption that all syntactic features must be semantically interpretable is in fact a natural further restriction on syntactic primitives and as such it is desirable independently of the derivational representational duplication issue. The hypothesis that there are no solely PF motivated syntactic features takes seriously the spirit of the Chomsky's (1995) remark according to which the "core systems of language" involve the lexicon-LF interaction, the Spellout system being extraneous. An approach based on this more restrictive hypothesis has the nontrivial task of accounting for the predictions achieved through the +/- interpretable distinction, but without assuming it.

The essence of bare checking theory is that lexical items in a sentence might duplicate certain features, since information is stored in the lexicon item by item. Chain formation or checking occurs because such duplication must be eliminated: it

violates FI. For example in a question like (15), there is a WH feature at least on some element of the (split) C node and also on some element of the wh-phrase. We can assume that in principle a WH feature is legitimate anywhere in the scope of a predicate that selects a question (or under a root node if the main clause is interpreted as a question).

(15) I wonder who +WH Bill saw (who)

The multiple copies of the +WH feature must be eliminated by LF: interpretively the sentence contains a single question. Taking spec-head relations to fall under an extended notion of chains, call it the CHAIN relation, the only way to eliminate the multiple copies is to relate them via CHAINS. In (15) this can be achieved by the formation of a chain that links the phrase in the spec position of the +WH C head with the wh-phrase in situ. But all three +WH features now merge in a CHAIN, ie. either through a chain or a spec-head relation. Notice that under standard assumptions about covert chains and checking, the account is the same whether or not wh-"movement" is "overt", ie. whether or not the wh-phrase occupies the in situ or the A'-position at spellout. The difference is only that in the covert case, we have a category without phonological material or simply a (non-categorial)feature bundle in spec-head (or adjunction) relation

with the +WH C and in a chain relation with the wh-in-situ. (In the theory to be developed in chapter 3, the covert wh-chain will involve no such phonologically zero element in spec-head relation with the +WH C, rather the +WH feature will be taken to form a chain directly with the wh-phrase it dominates.)

I will not compare bare checking theory here with the checking theory of the minimalist framework, see Brody 1997a for discussion. In what follows I shall outline instead the treatment of successive step and partial "movement" that this theory makes possible.

Successive step chains are a major problem in the minimalist framework, especially in the "chapter 4" system of Chomsky 1995: each step of movement needs a trigger, --or in terms of the revision of last resort above in section 1.3, each chain member must be licenced. Thus the conjunction of the last resort type assumption with the standard checking theory effectively rules out the empirically well-motivated phenomenon of successive step chains. In (16) for example, under standard assumptions the intermediate copy of "who" cannot be triggered/licensed: "think" does not select a question, hence the intermediate C is not +WH.

(16) I wonder who +WH Mary thinks (who) John will meet (who)

Under bare checking theory, the fact that the intermediate

verb does not select an interrogative does not necessarily entail that the C node selected by it cannot be marked +WH. If this +WH element merges with other +WH nodes via some CHAIN, then it might relate to a question in some other part of the structure. In other words, we can analyze the successive step configuration in (16) as in (17).

(17) I wonder who +WH Mary thinks (who) +WH John will meet (who)

Here all +WH markings (five altogether) merge to form a single question. Given semantic selectional requirements, this question must be associated with the predicate "wonder".

Notice that a structure like (18) continues to be excluded:

(18) *John thinks who +WH left

(18) is ruled out since "think" selects a proposition but a +WH marking must be interpreted as a question. In (18) there are no other semantically licit +WH features with which the illicit +WH in the embedded clause could merge.

This approach helps also to understand the peculiar behavior of partial "movement" constructions, where a wh-phrase surfaces under a predicate that does not select a question. In the German

sentence in (19) for example, the contentful wh-phrase "wer" occurs in the non-interrogative complement of the verb "glaubst".

(19) Was glaubst du wer gekommen ist

What believe you who come is

"Who do you believe came"

Bare checking theory can immediately explain how this otherwise puzzling state of affairs can exist. If the C head of the complement clause is +WH then the obligatory presence of the complement internal wh-chain is expected.^{FN7} Just as in the case of successive step chains, the matrix verb tolerates the +WH complement even though it selects a non-interrogative clause. This is because at LF all wh-features merge through CHAIN construction, hence the structure will contain only a single question. Selectional requirements only allow this question to be a matrix interrogative.

There is a debate in the literature between the "direct" and the "indirect dependency" approach to partial movement structures (see especially the papers in Lutz and Muller (1996) and references cited there). The indirect dependency approach is characterized by the conjunction of two claims. First that the two wh-phrases in (19) are not directly connected, rather, both are associated with their respective CPs and the (trace of the)

expletive scope-marking wh-phrase in the matrix clause is associated with the whole of the embedded clause. Secondly and in accordance with this hypothesis a semantics is provided, in which (19) is analyzed as a combination of two questions (which turns out to be equivalent to a single one): a matrix question over propositions and one corresponding to the embedded sentence over individuals (Dayal 1994, 1996).

The solution based on bare checking theory is incompatible with the combined multiple questions semantic analysis. Bare checking theory necessarily claims that semantically there is only a single question in the sentence throughout, --this is the assumption that forces the construction of the wh-CHAIN. And in fact the semantics of the indirect dependency approach is highly dubious independently of any considerations having to do with checking theory. This is because it allows clauses selected by predicates that do not select interrogatives to have the semantic force of questions, -- an implausible assumption that is very nearly a direct contradiction.

On the other hand the indirect dependency analysis is strongly motivated syntactically. The evidence for the claim that in a structure like (19) the expletive scope-marking wh-phrase in the matrix clause is linked to the whole embedded clause is very clear at least in languages like Hindi or Hungarian (cf.eg. Mahajan 1996, Horvath 1995). Thus in Hungarian for example, the

case of the scope marking wh-phrase corresponds to the case that the (overt) expletive associated with the embedded clause receives in the corresponding non-interrogative construction:

- (20) a. **Mitol** felsz hogy kit latott Mari
what-from are-afraid-you that who saw M
"Who are you afraid that Mary saw"
b. Felsz **attol** hogy Mari latta Petert
Are-afraid-you that-from that M saw P
"You are afraid that Mary saw Peter"

- (21) a. **Mi** nyilvanvalo hogy kit latott Mari
What+NOM is-obvious that who saw M
"Who is it obvious that M saw"
b. Nyilvanvalo **az** hogy Mari latta Petert
Is-obvious that+NOM that M saw P
"It is obvious that Mary saw Peter"

- (22) a. **Miert** ment el Janos mert kivel beszelt Mari
why(=what-for) went away J because with whom talked M
"With whom did John leave because M talked"
b. Janos **azert** ment el mert Mari beszelt Peterrel
J that-for left away because M talked P-with
"John left because Mary talked to Peter"

There is thus an apparent contradiction: we are led to the indirect dependency analysis on syntactic grounds that we must also reject for semantic reasons. But the contradiction is not real, since the two aspects of the indirect dependency analysis are not necessarily connected. There is no inconsistency in rejecting the semantic assumptions of this theory at the same time as accepting the solid syntactic evidence for the claim that the wh-phrases in this construction are not directly linked in a chain.

In fact the expletive is probably associated not with the clause but with the head of clause. Chomsky and Lasnik 1978 pointed out the link between the pleonastic "it" and the overt/ness of complementizer of the associated clause in English, ie. (23a-e). Note also the grammaticality of the structure with the wh-phrase in spec-C and non-overt complementizer, (23f):

- (23) a. It is obvious that Mary left
b. It is impossible for Mary to go
c. It is unclear if/whether Mary should go
d. *It is obvious Mary left
e. *It is unclear Mary should go
f. It is unclear who Mary saw

This generalization seems related to the fact that the wh-chain in the embedded clause in partial movement constructions like (19) must be overt, --something that the analysis based on bare checking does not entail by itself.

If the analysis of (19) is along the lines of (24), where the A-position copy of "was" is linked to the embedded +WH head, then the full wh-CHAIN will include both the higher and the lower wh-phrase, their traces and both the higher and the lower +WH C head.

(24) Was +WH glaubst du (was) wer +WH gekommen ist

Thus in our present terms the two wh-phrases are co-members of a CHAIN but not of a chain.

There is a construction that appears to be difficult to treat under these assumptions. This is the partial movement construction where the embedded clause is a yes-no question as in the Hungarian example in (25):

(25) ?Mit gondolsz, hogy elment-e Janos

What think-you that away-went-whether J

"What do you think about the question of whether John left?"

This construction is sometimes taken to provide crucial evidence in favor of the interpretation in terms of two interacting questions, characteristic of the indirect dependency approach. Interpreted as a single question on the pattern of partial movement structures involving constituent questions, it should be a yes-no question on the matrix clause: "Do you think that John left". This prediction is clearly incorrect. Since the structure corresponding to (25) is apparently impossible in German, but available in Hindi^{FN8}, it is also suggested by several authors that a different analysis of the syntax and/or semantics of partial movement may be appropriate for different languages, --an undesirable conclusion.

On slightly closer examination it appears that structures like (25) are misanalyzed as partial movement, --at least in Hungarian. As the English gloss already hints, the Hungarian equivalent of the verb "think" also allows both a direct object and an "about"-phrase as complements. Thus we can take (25) as the version of the word by word translation of its English gloss, in which the string "about the question" is silent.

(26) Mit gondolsz, (arrol a kerdesrol) hogy elment-e Janos

What think-you (that-about the question-about) that away-
went-whether J

"What do you think about the question of whether John left?"

Evidence for this assumption is provided by the fact that both the non-elided version of (26) and also (27), where only the sentence introducing pronoun is present are also grammatical.

(27) Mit gondolsz, (arrol) hogy elment-e Janos

What think-you (that-about) away-went-whether J

"What do you think about whether John left?"

Deletion in (27) would raise no recoverability issues.

Structures like (26)-(27) are possible without the matrix wh-phrase, showing clearly their independence from partial movement:FN9

(28) Mari ezt gondolta arrol (a kerdesrol) hogy elment-e Janos

M this thought that-about (the question-about) that away-went-whether J

"Mary had this thought about (the question of) whether John left"

Furthermore and crucially, where the argument structure of the verb does not make an alternative analysis like (27) available, structures that look like partial movement with yes-no questions are clearly ungrammatical in Hungarian:

(29) a. *Mitol felsz hogy elment-e Janos

What-from are you afraid that away-went-whether J

b. *Mi nyilvanvalo hogy elment-e Janos

What is-obvious that away-went-whether J

In contrast genuine partial movement structures corresponding to the examples in (29) are all perfect:

(30) a. Mitol felsz hogy ki ment el

What-from are-you afraid that who went away

b. Mi nyilvanvalo hogy ki ment el

What is-obvious that who went away

The same misanalysis appears to create the impression that partial movement is possible where the scope marker participates also in a multiple-wh construction. Thus the examples in (31) which, with the silent sentence introductory pronoun, masquerade as partial movement can have multiple wh, but (32), the genuine partial movement construction cannot:

(31) a. ?Ki mit mondott, (arrol) hogy ki ment el

Who what said (about) that who went away

"Who said what about who left"

b. ?Ki mondott mit, (arrol)hogy ki ment el

Who said what...

(32) a. *Ki mitol felt hogy ki fog elmenni

Who what-from was-afraid that who will away-go

b. *Ki felt mitol, hogy ki fog elmenni

who was-afraid what-from...

In (32) the wh-pronoun "mit" is an expletive in (32) it is not. But in both cases, whether the wh-pronoun "mit" is in situ or in the preverbal A'-position, makes no grammaticality difference.

This means that no question arises as to why a scope marker in a partial movement construction cannot occur in situ in Hungarian. Since this language moves at least one wh-phrase to A'-position, and the scope marker of the partial movement construction cannot participate in multiple wh structures, it necessarily occurs in an A'-position. Perhaps in relevant respects the same conclusions carry over to languages like German and explain the ungrammaticality of (33):

(33) *Wer meint was wann sie gekommen ist

Who thinks what when she come is

1.6. Summary

I sketched in this chapter some major restrictive characteristics of the theory of perfect syntax. In this theory (a) move or rather the chain relation cannot be conceptualized as an imperfection, and (b) there cannot be syntax internal stipulative conditions on it. (c) The theory cannot assume the existence of both representations and derivations, (d) which naturally leads to the elimination of semantically non-interpretable features from syntax. I have also discussed some advantages of the assumptions in (c) and (d), where these restrictions of the theory seemed to be explanatory in that they appeared to force analyses with correct descriptive consequences.

A fifth difference between perfect syntax and the minimalist approach, that I will not discuss in detail here, concerns economy conditions. In the minimalist system interface conditions will be satisfied "as well as possible", as measured and ensured by economy conditions. But if there are no (externally forced) imperfections, no economy conditions may be necessary to ensure that the conditions are satisfied optimally. If there are no imperfections, there is no need to measure and compare degrees of imperfection. Such a conclusion is consistent with the fact that it is often easy to restate most of the global economy conditions

as computationally simpler "default" or licensing conditions, preventing optionality. It is also noteworthy that during the development of the minimalist approach the role of economy conditions seems to have shrunk considerably.FN10

In sum perfect syntax, in contrast to the minimalist program makes the optimal assumption and aims to eliminate forced and unforced imperfections from syntax as a matter of principle. As we shall continue to see also in later chapters, when the consequences of this more restrictive approach are examined, it turns out to lead to a picture of syntactic competence that is rather different from that of the minimalist framework.

Footnotes to Chapter 1

FN1 Li 1997.

FN2 This section includes a reworked version of some paragraphs of Brody 1998.

FN3 A different alternative assumption also in Chomsky 1995 is that syntactically special "categorial" features force overt movement.

FN4 The GPP should be extended to cover also semantic relations like adverbial modification: modifier and modifiee are also restricted to chain-roots.

FN5 A number of issues, mostly having to do with questions of locality, will nevertheless remain, --a topic I will not address here.

FN6 See Brody 1995a for problems with the preference principle.

FN7 Note that overttness of the chain of the contentive wh-phrase does not follow, --see the text below for a possible

direction to explore to explain this.

FN8 Horvath marks as ungrammatical Hungarian constructions like (25). They seem to me and others I asked acceptable. See next note for comments.

FN9 There is a difference however: the expletive "arrol" ("that-about") is deletable in (27) but not in (28). Perhaps this deletion is actually ungrammatical also in (27) (hence Horvath's judgement reported in the previous note) and becomes acceptable only through analogy/ confusion with genuine partial movement.

FN10 For example Chomsky 1995 argues forcefully in chapter 4 that one of the paradigm cases of economy in chapter 3 of the same volume, the shortest derivational steps restriction is in fact not an economy condition. For some additional comments on economy principles see Brody 1997a, for critical reviews Sternefeld 1996 and Johnson and Lappin 1997.

Chapter 2.

Properties of projections and of C-command

2.1. Introduction

In this chapter I intend to prepare the ground for the theory of syntactic structure to be presented in chapter 3. Section 2.2 recapitulates the theory of phrase structure in Brody (to appear) the starting point for the theory to be presented later. In section 2.3 and 2.4 I discuss critically Epstein's (1995) and others' derivational characterization and attempted explanation of the c-command relation. I also provide a different, nonstandard way to relate c-command and precedence but argue that none of the approaches reviewed here are genuinely explanatory. In section 2.5 I argue that Kayne's attempt to relate to the direction of time the fact that in syntax asymmetric c-command appears to map to the precedence relation (rather than to its inverse) between terminals is unsuccessful.

2.2. 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, ie. 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 categories (copies of an LI) each immediately dominating the next. The most deeply embedded element, x^{\min} , is 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.FN1

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 and 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.FN2

Projection is local in the sense that a category C and its (direct or indirect) projection cannot be separated by a category N of a different type (ie. 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.FN3 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.^{FN4} 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 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 2.3 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 chapter 3 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 some aspects of the relation of the theory outlined above to Kayne's antisymmetry hypothesis and to his linear correspondence axiom (LCA).

2.3. 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): FN5

(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):FN6

(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 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 situation 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 (ie. 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.^{FN7} There is nothing however in this state of affairs that would be "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 (ie. 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 than 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 eg. 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 postulating an m-command relation between the head and its spec is not beyond doubt.

The clearest representant 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). (This problem will disappear however in the theory to be developed in chapter 3 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 framework presently considered, syntactic computation reduces to the insert relation (which can in principle be construed 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

Ie. 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) the filter 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). FN8

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 by the definition in (4). Given (4), projecting heads do not c-command anything. These categories are parts of PLs, they 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 chapter.

2.4. Attempts to explain c-command

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 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) (p.25.), it disallows

relations between elements that at any point in the derivation were members of different unconnected subtrees.

(8)[Epstein's(27)] 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 (not= T_1) and
- ii) T_2 is a term of K_2 (not= 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 y c-commands x, since in such a configuration no unconnected subtrees have been formed.FN9 In other words (8) does not capture the asymmetry of the c-command relation. Epstein attributes the asymmetry to his 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 not= T", ie. that the top node of the an unconnected tree does not count as an unconnected element. But this means that "K

not=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 not=T" restriction by noting about the top nodes (to be related by merge/move) of the unconnected trees, ie. about K_1 K_2 , that "each equals a root node, neither has undergone Merge or Move, hence each is (like a lexical entry) not 'yet' a participant in syntactic relations" (p.26.) FN10 FN11 In other words, the two instances of the "K not=T" stipulation in (8i) and (8ii) can be exchanged for an additional fourth subclause as in (8'):FN12

(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, ie. 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 again 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 representation there is a subtree which properly contains (ie. 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 1997a), since the derivational account needs to assume these. Furthermore, the easy translateability 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, --eg. 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.^{FN13} 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 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 constructed for c-command.

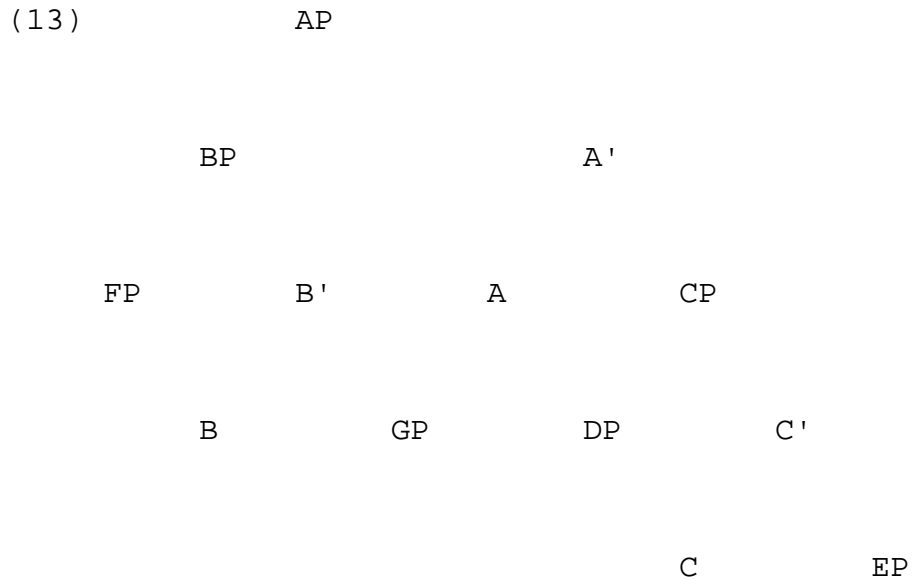
The remaining problem is of course that linear order appears to be too weak to express hierarchical relations. Linear order creates a string, a one-dimensional structure, while standard syntactic hierarchical relations create a two-dimensional tree. This can be remedied however if the theory can make 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 (ie. 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):



- (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 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 (eg. 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 chapter 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 requires, why have both notions as primitives (or as independently defined concepts)? (15) defines c-command in terms of precedence, hence the relationship of the two notions need not and cannot be stated here 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, ie. 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 chapter I will turn to a theory in which they are. In that theory however the LCA will not be necessary to ensure antisymmetry, as we shall see.

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 to 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 (ie. 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 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 chapter will make possible a different and more promising approach.

2.5. Precedence and time

The discussion in the previous section assumed the universal spec-head-complement order in syntactic representation. In the theory to be developed in chapter 3, the fact that the complement and the specifier are on opposite sides of the head will be a consequence of the mirror principle, but the relative order of the specifier and the associated head will essentially remain a stipulated primitive. In Kayne's (1995) theory the LCA ensures that the spec and the complement are on opposite sides of the associated head. Thus both in Kayne's LCA approach and under the theory to be developed below, head-complement order will follow once specifier-head order is set. In this section I should like to discuss critically Kayne's attempt to derive the specifier-head order from independent considerations.

Before turning to that issue, let us note that the fact, that spec and complement are on opposite sides of the head follows from the LCA, only if the assumption in (14) is given:

- (14) the specifier asymmetrically c-commands the head and the
head asymmetrically c-commands elements of the complement

In the theory summarized in section 2.2 above, spec, head and

complement are sisters. In the theory to be developed in chapter 3 the status of the head will be somewhat special, but with respect to the specifier and the complement the sisterhood property will carry over. To ensure that spec and complement are on opposite sides of the head in the theory sketched in section 2.2, we can simply substitute "is adjacent to in the linear order" for "asymmetrically c-commands" in (14). Thus an assumption like (14) seems necessary in some form both in Kayne's theory and under the approach of section 2.2. The fact that specifier and complement are on opposite sides of the head will follow however directly from the independently motivated mirror principle in the framework of chapter 3.

Let us turn to the question of setting the order of the specifier and the head. The specifier will precede or follow (and the complement accordingly follow or precede) the head, depending on whether asymmetric c-command between two nodes A and B is mapped by the LCA into the relation of "a precedes b" or "a follows b" between the terminals dominated by A and B respectively. Kayne has put forward a "possible explanation" of the fact that the relevant relation is precedence and thus the specifier must precede the head (and hence the complement follow it), which is also based in part on the LCA. This explanation will not be reproducible in terms of the theory to be developed

in chapter 3. In this section I would like to show that this entails no comparative loss of explanatory power, as Kayne's solution turns out not to be a real one.

His account is essentially as follows. (i) Assume the existence of an abstract node A that asymmetrically c-commands every node in the tree and the existence of an abstract terminal t that A dominates. (ii) Associate a time slot with each terminal and then associate with each time slot the string of terminals "produced" up to that time. Thus the string of terminals abcdz is mapped to the set of strings in (15):

(15) a, ab, abc, abcd, abcdz

(iii) Assume that a fixed relation (precede or follow) must hold between t and every terminal in every substring and that this relation cannot be satisfied vacuously, (the latter assumption Kayne does not explicitly mention but obviously intends to make).

Given the assumption in (iii), t must be a and not z, and the relation must be precede and not follow. Hence "A asymmetrically c-commands X" maps to "t=a precedes x" (where x is the terminal(s) dominated by X) and not to "t=a follows x" and more generally "Y asymmetrically c-commands U" maps to "y precedes u". (Strictly speaking the more general statement does

not follow here. The relation could be for example "y precedes u if y is the first element of the string otherwise y follows u". The more general statement follows if the choice is restricted to "precede" and "follow", a restriction Kayne's text makes in the following terms "more specifically the question is whether <x,y> is "x precedes y" or "x follows y".)

On the further assumption (iv) that spec asymmetrically c-commands the head (and the head asymmetrically c-commands categories in the complement), ie. (14) above, spec-head (-complement) order follows. Kayne takes his explanation to relate this order to the asymmetry of time, since the order "derives from the hypothesis that [(15)] (rather than a sequence of substrings working backward from the final terminal) is the correct way of representing the relation between terminals and time slots" (p.38).

The explanation appears to be circular however on at least two counts. Firstly, if the order is to be derived from the hypothesis in (ii) that each terminal-time is associated with the string that precedes it (ending with the terminal in question) then this assumption needs to be motivated independently: (ii) is not an obviously better stipulation than spec-head(-complement) order. Stipulating (ii) in fact begs the question. This is

because the hypothesis that (15) "is the correct way of representing the relation between terminals and time slots" does not follow from (nor is it made natural by) the asymmetry of time. The asymmetry of time is perfectly compatible with associating with each terminal-time the sequence of terminals of the sentence not yet "produced". This can be seen for example from the fact that there is no contradiction in associating with each terminal-time both the string of terminals in the sentence produced up to that time and also the string of terminals of the sentence not yet produced. In other words the question of whether asymmetric c-command maps to "precedes" or "follows" is not answered by noting that terminals are "produced" in an order that corresponds to the asymmetric direction of time.

Secondly, the corresponding problem arises also for the premiss in (i). The stipulation that t , the terminal dominated by the highest node of the tree (ie. A) must (nonvacuously) be in a fixed relation with every terminal in every substring in (15) is again not an improvement over stipulating that LCA maps to precedence, as it has no independent motivation. And again, the assumption begs the question: why is it the terminal t dominated by the highest node A , rather than say the terminal t' , dominated by the most deeply embedded node, Z , that must be in a fixed relation with every terminal in every substring in (15)? On that assumption, not any more or less natural than the earlier one,

asymmetric c-command must again map to the "follows" and not to the "precedes" relation.

Notice thirdly, that any conclusion that follows from the LCA with respect to the order of specifier, head and complement, follows only on the assumption in (iv)=(14) above, that the specifier asymmetrically c-commands the head, and the head asymmetrically c-commands categories in the complement. But again, as noted in Brody to appear, substituting "precedes" for "asymmetrically c-commands" in (14) ensures spec-head(-complement) order. As there is no a priory reason to prefer the version of the axiom in (14) with asymmetric c-command over the one that refers to precedence, we have one further reason to question the derivation of the specifier-head(-complement) order from the LCA and the asymmetry of time. The approach needs to assume a near equivalent of the principle it attempts to explain.

In sum, the fact that the theories discussed in this work take the order of the specifier and its head to be an unexplained primitive does not make them inferior to existing alternative approaches, since there exist no successful attempts to explain why this order rather than the converse obtains. As noted in the introduction and in section 1.3, it does not seem implausible that the real answer to this problem might well have to wait for

the eventual integration of linguistics with other branches of biology and psychology.

FN1 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 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 and chapter 3 below for a theory where the distinction is eliminated.

FN2 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.

FN3 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.

FN4 Cf. Brody (to appear) for some discussion.

FN5 For present purposes "term" can be taken as a synonym of "constituent".

FN6 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').

FN7 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. Brody (to appear).)

FN8 As pointed out by Higginbotham and May, structures where the determination of the reference is grounded and thus has a starting point are fine. Cf:

(i) [_y His_x mother] likes John_x, [_xher_y secretary]

FN9 More precisely no unconnected subtrees have been formed that properly include x or y.

FN10 More precisely, K_1 , K_2 have not yet undergone merge or the merge part of move.

FN11 Notice that "syntactic relation" here means: not yet part of the tree, and not as before, c-command.

FN12 Again, read "merge part of move" for "move" in (8').

FN13 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.

A radically restrictive theory of syntactic representations

I shall argue in this chapter for a theory of syntactic structure that eliminates phrasal projections altogether. I shall first enumerate some problems with the standard explanation of the mirror principle based on the head movement constraint (HMC) in section 3.2. In section 3.3 a system is presented, in which there is no syntactic distinction between words and phrases and the mirror principle is axiomatic, --the mirror theory. The core structure of the sentence is argued here to reduce to (1) in which the structure of the word "V-v-Infl" is syntactically expressed directly by the (inverse order of the) (boxed) complementation line:

$$\begin{array}{ccc} \text{Su} & & \mathbf{v} \\ & \searrow & \\ & \text{t}_{\text{SU}} & \mathbf{V} \end{array}$$

The only primitive relation between elements in syntax and morphology will be the spec → head relation, where spec (and its constituents) are ordered to precede the head. I take the mirror principle to license some morphological spec → head orders to be (geometrically) mirrored in syntax. The head-complement relation is just a reverse order (morphological) spec-head relation. This gives also spec-head-complement order. Morphological words (MWs) can be formed via morphological spec-head relations. Members of morphological words (heads) can form spec-head relations with other MWs. Eg. in (1) Su (which abbreviates a set of MWs) is the spec of Infl, and Infl is a head, a member of the MW: V-v-Infl. An MW is then spelled out (in the morphological spec-head order) in the position of the highest strong head (or, lacking a strong head, in the lowest position), ie. in (1) "V-v-Infl" in Infl in French but in v in English. Given an adverb in the spec of some functional projection F between v and Infl (Cinque 1997), we derive the correct word order, Su Adv V Obj in English and Su V Adv Obj in French. The section concludes with a list of advantages of the mirror theory view. Due primarily to its restrictiveness (there is only a single primitive configurational relation), the theory explains many generalization ranging from locality of head chains to the main effects of Kayne's LCA and

various additional properties of "phrase"-structures having to do (in standard terms) with phrasal and even "extended" projections.

In the version of mirror theory to be defended here, complementation is restricted to mirrored MWs. If X and its argument Y do not form a morphological word then Y cannot be the complement of X, hence Y must be a spec of X or the spec of some element of (a decomposed) X in a lower shell. Section 3.4 discusses some (apparently inferior) alternatives to this version of the theory and some consequences and advantages of this view.

Section 3.5 presents an additional advantage of mirror theory, namely that it might make it possible to systematically factor c-command into two more simple relations: the spec-head relation and domination (in fact also an extended spec-head relation in the proposed system). As in mirror theory a head and the phrase it projected are represented as a single node, the analysis of MWs does not need to refer to c-command, elements of MWs establish a domination relation with each other. Syntactic spec-to-spec chains (the equivalents of XP-chains of standard frameworks) can then also be analyzed as composite units. In the case of a two-member chain this unit is composed of a chain between a "phrase" (ie. a head and its constituents) and some dominating head H, FN1 and another "phrase" in a spec-head relation with H. Some independent evidence is presented for this approach.

3.2. The standard explanation of the mirror principle and its problems

Cinque (1997) develops a strong empirical case for the claim that there is a correspondence in universal grammar between the hierarchy of spec positions that he argues adverbs occupy and the hierarchy of clausal functional projections. In the process, he provides much additional support for a bakerian mirror principle. According to one version of 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.

(2) olvas-hat-om
read-permissive -1sg, present

(3) TP

T

ModP

	Mod	VP
-om	-hat	olvas

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 (eg. Starke 1994, Rizzi 1995, Cinque 1997).

The mirror principle is often attributed to the strict locality of head movement/chains (Baker 1988). This involves two assumptions. 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.FN2 For example there cannot be a three membered head chain [x^1 , x^2 , x^3] where x^1 has a host y , and x^2 has another host z . x can only "move" further together with its host. The partial structure in (3) cannot be completed with a chain as in (4), only with the chain structure in (5). (Phonologically not expressed traces, ie.

non top chain members, are in square brackets, top members of nontrivial chains in curly brackets.)

(4) TP

T	ModP	
	Mod	VP
{olvas}-om	[olvas]-hat	[olvas]

(5) TP

T	ModP	
	Mod	VP
{olvas-hat}-om	[{olvas}-hat]	[olvas]

It is easy to see how these assumptions can entail the mirror principle. If successive heads roll up, the last element and each unit so created subsequently moving to the left of the immediately higher one (ie. respecting the HMC), the resulting order of the heads will be the exact inverse of original

(syntactic/complement)order.

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 that an analysis in terms of phrasal movement can ultimately be given for such cases.FN3 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 highest member of the chain of the apparently incorporating element is in fact in a lower position.)

Another issue has to do with the fact that considerations pertaining to head-chains, the HMC and the mirror principle clearly do not exhaust the set of ordering requirements of morphology. Prefixes in general, like eg. Romance clitics, and certain types of compounding (eg. French "ouvre-boite", etc.) quite clearly do not fall under these principles, at least not in

the same way as suffixes like the ones discussed so far do. It would be incorrect to take the mirror principle to require that morphological order systematically mirrors the syntactic order in all cases. The appropriate domain of application for the mirror principle needs to be defined. The characterization above restricted the principle to apply to just those affixes that form head-chains linking their word-internal and their syntactic, complement internal positions. I will offer a somewhat different characterization in the context of the theory to be developed below.

But even after putting aside the directly empirical issues, questions remain about the HMC based derivation of the mirror principle. First of all in syntax the information that explicates the structure of words is expressed both word-internally (ie. X^0 -internally) and by the phrasal order given 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, IP, and a lower projection of V, VP. It is not obvious that the account of this duplication, based on the conspiracy of the HMC and the no excorporation condition qualifies as a genuine explanation of this pervasive parallelism. Relating the phrasal and the word-internal orders in this way makes the correlation

somewhat accidental, and invites the question: why should it be the case that these two in principle unrelated conditions force grammar to express redundantly the same ordering twice, both in terms of the phrasal complementation structure and morphologically, in terms of word structure. If both the HMC and the no excorporation condition could convincingly be reduced to a simple theory of locality, then this point would become weaker, but still not all of its force will be taken away. (Inverse) phrasal order and morphological order seem to be just two sides of the same coin. The question still remains, if locality is not used here to ensure the correspondence of some order with itself. In other words we might expect that a better account might somehow capture the identity of the two orders, and in this way explain their correspondence by in fact making an explanation unnecessary.

Further problems arise 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 (3) "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. eg. 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.FN4 The V+Infl unit forms a chain though with the guest-of-inflection position(s)FN5 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 three times in syntax: in the structure of complements, on the lexical item that is to check functional heads and on the heads that are to be checked. 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 checking theory raises at least two apparently serious questions. First of all this theory does not in fact resolve the problem that configurations where the phonological position of a head is lower than the syntactic position of its suffix 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 (3) will be along the lines of (3'). (Traces and deleted functional heads are in square brackets.)

(3') TP

T

ModP

Mod

VP

{ $\hat{\text{olvas-om}}$ }-[$\hat{\text{om}}$] [$\hat{\text{olvas-om}}$]-[$\hat{\text{om}}$] [$\hat{\text{olvas-om}}$]

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 then the internal suffix could be checked 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 transparent 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 that refers to 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. FN6 FN7

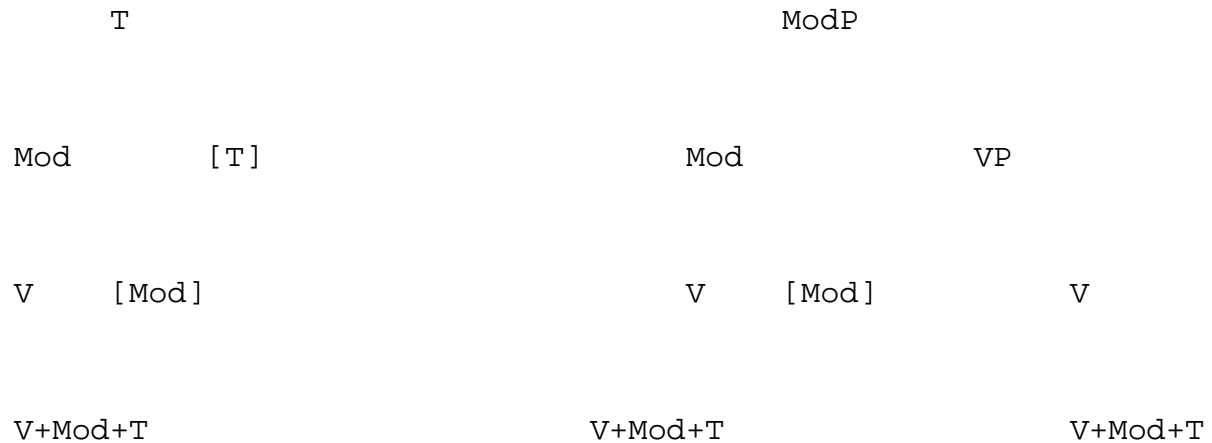
The second problem with checking theory has to do with the fact that it ensures (in conspiracy with the X'-theory of phrasal projection) that word structure and (inverse) phrasal complementation structure match. The order of host suffixes must match (the inverse of) the word-internal order of suffixes. If each suffix projects a phrase and each phrase is projected by a suffix, then this order of host suffixes then corresponds to the order of phrasal complements.

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 phrase structure rules. Similarly, we might want to search for a theory that makes it possible for the complementation structure and the morphological structure encapsulated in the structure of words not to be generated independently. Notice that this problem is related to but distinct from the one raised above in connection with the HMC account of the mirror principle. There the problem was, that it might be better to avoid syntactically representing the ordering information in duplicate. In this way we might hope to explain in a strong sense the observed correspondence. The present point is a different one, namely that under checking theory the principles generating the duplicate information are themselves undesirably duplicated.

A third, additional 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 apparently unavoidable assumptions introduce a further systematic set of otherwise unmotivated duplications. A fuller structure of (3) for example, will be (6):

(6) TP



{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 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. There is no evidence however for this additional duplication, which, given the triplication of this information that checking theory creates,

makes the analysis quadruplicate the features in question. Furthermore, again, the problem is not just that (unlike in the case of trace-copies) presence of the duplicate (quadruplicate) information in the syntactic representation is not independently motivated. The additional and probably more serious issue is that three distinct sets of principles generate the same structure. The duplication created by the series of word-internal dominating nodes appears unavoidable in standard approaches, 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 (a) words and chain members are constituent nodes, (b) nodes are labelled and by one of their constituents (c) non chain-tails cannot label (attributed to the GPP in Brody 1994, 1995a, to appear, see also the "target projects" requirement of Chomsky 1995)

There are some 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, FN8 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 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 (cf. section 3.5 below). It is perhaps suggestive also, that none of the theories, reviewed in chapter 2, that attempt to reduce c-command to more basic notions appear to be able to allow for c-command out of adjunction.

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, ie. 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.

There are other differences in the behavior of head-chains and XP-chains that can easily be stipulated to exist, but are again somewhat difficult to understand if the relation standardly expressed in terms of head-chains is essentially of the same type as the one captured in terms of XP-chains. Head chains, in contrast to XP-chains typically show forced reconstruction

effects. For example, in (7) "not" has scope over the auxiliary "can/need", and the structure is often analyzed accordingly in terms of the auxiliary forming a head chain whose tail is lower than negation (cf. eg. Roberts 1996, Cormack and Smith 1996).

(7) John can't/needn't do that

In the case of XP-chains, such configurations normally result in ambiguity: scopally the chain can be interpreted both in its head or in its tail position. (Cf. eg. "who does everyone like t", where as is well known "everyone" may or may not take scope over "who". The situation is essentially similar with A-chains, cf. eg. Hornstein 1994.) (7) however shows no similar ambiguity, it cannot be interpreted with possibility/obligation taking scope over negation.

Another apparently relevant difference concerns VP-ellipsis. In this construction the antecedent of the ellipsis site cannot contain the trace of head chain, as shown by (8), but an XP-trace creates no violation, (9), --cf. Lasnik 1994 (for the observation and a different analysis) and Roberts 1996. In (8a-c) "have/be" is in T, in (8e) the "be" of the first conjunct is lower but still outside the VP, as a result of Akmajian and Wasow's (1975) "be-shift", as Roberts notes in this context:FN9

- (8) a. *John was {t here} and Mary will (be here) too
 b. *John has {t left} but Mary shouldn't (have left)
 c. *John hasn't {t a driver's license} but Mary should (have a driver's license)
 d. John does not {have a driver's license} but Mary should (have a driver's license)
 e. ?*The children have been {t very good} here. I wish they would (be very good) at home
- (9) a. I wonder which book John bought t and you asked which one Mary did (buy t)
 b. John was seen t and Mary was (seen t) too

Again, the distinction can be easily stated, but nevertheless remains unexpected under the theory where head-chains and XP-chains are the same kind of object.

3.3.Mirror Theory

The problems just listed might be taken to motivate the search for an alternative view of the mirror principle and of head chains. I will now present such a theory below, which expresses the "head chain" relation differently. This theory takes the mirror principle to be a more basic generalization than the HMC

and derives from it certain properties having to do with excorporation, c-command and locality, currently attributed to head chains and constraints on them. Let us approach this framework by looking first 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 2.2. 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^0 s) 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} really necessary? In other words, do PLs exist, is the postulation of the PL-structure justified?

Focussing first on the distinction between words and phrases, consider the basic structure in (10).

(10) XP

ZP X YP

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 XP and YP on the other. As noted in Brody (to appear) and in chapter 2 above, 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 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. More precisely, the assumption is that spec and its constituents precede, while the complement and its constituents follow the head.

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 (10). 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, (10) reduces to (11):



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 (i.e. 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.

(11) can be viewed as eliminating the apparent conflict between the long tradition of dependency theories, see eg. Hudson (1990) and references cited there, and phrase structure theories of syntactic representations.^{FN10} 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.^{FN11}

Consider next the distinction between X^{\min} and X^0 . If X^{\min} s are lexical elements, then this distinction captures the

difference between words that are assembled in syntax and those that are put together in the lexicon. Two questions arise: First, does the distinction really exist? It does not if either all words are assembled syntactically or if all of them are put together in the lexicon. Secondly, even if the distinction between X^{\min} s and X^0 s exists, it does not follow that it must be made explicit in syntactic representations. Take V+Infl as a simplified example. Suppose that instead of analysing this as in (12) we attribute to it the simpler structure (13), where X^0 and X^{\min} levels are not distinguished.

(12)

	Infl ⁰
V^0	Infl ⁰
V^{\min}	Infl ^{min}

(13)

	Infl
V	Infl

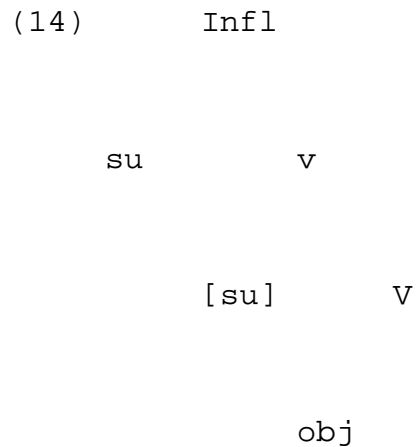
However, even without looking at empirical phenomena that might be taken to motivate the X^0 - X^{\min} distinction (like the distinction between inflectional morphology and the incorporation

of elements that are members of open classes, a matter to which I will return below) a problem already arises. Although neither the word vs phrase nor the lexical item vs word distinction seems necessary at least in the 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 (11), 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 morphological 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 (11), 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.

At the same time, each of these nodes can, as usual, have their own specifiers. In (14) for example, these spec's are the

subject, its trace and the object:



Call the potentially decomposable morphological unit formed in (14) by V, v and Infl a morphological word (MW). Given representations like (14) the fact that two elements are part of a single morphological unit will not be necessarily explicit in syntactic structure: the two elements need not collect under a special type of syntactic node: the X^0 . This need not cause any morphological problems in itself, in fact it might eliminate yet another quasi-redundancy. The morphological distinction between free and bound forms is necessary independently of whether we have a syntactic distinction between words and phrases. Syntactically separate heads, like eg. the English tense or genitive can be obligatorily affixal morphologically and elements like Hungarian verbal prefixes which appear to form a lexical, non-compositional X^0 unit are able to behave as independent free

forms. Thus the distinction between morphologically free and bound forms seems necessary whether or not there is an additional word-phrase distinction in syntax.

Let us ask next what licenses the syntactic complement structure like the one in (14), where V is the complement of v and v is the complement of Infl for cases where the morphological 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. Suppose that the single primitive relation of the morphological 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 topologically mirrored morphological spec-head relation, ie. it is an ordinary spec-head relation in inverse order. Certain morphological spec-head relations are expressed as head-complement relations in syntax.

As exemplified in (14), additional spec-head relations can then be licensed in syntax between elements of MWs that have free spec valences (are not morphologically specified as being the spec of anything) and elements of (syntactically mirrored) morphological words.

I state the mirror principle that inverts the morphological spec-head order in (15):

(15) The Mirror Principle

the syntactic relation of "x complement of y" is identical
to an inverse order morphological relation "x spec of y"

(Universal) spec-head-complement order does not need to be specifically stipulated: it follows from the primitive spec-head order (spec precedes head) and from the (equally axiomatic) mirror principle, which reverses this order in syntax in some of those cases where it exists also morphologically.

So, given the mirror principle, the morphological and therefore the spellout order of two elements, x,y in the syntactic complement relation (hence also co-members of a morphological word) is the inverse of their syntactic order. Notice that the mirror principle is not a biconditional, it does not require that all MWs be expressed in the mirrored syntactic form. And note also that given the mirror principle in the form in (15), there is no need to postulate two symmetric representations, one syntactic the other morphological. The morphological representation is simply the inverse order mirrored construal of the syntactic complement line.

Consider next the spellout question of mirrored MWs: Which element of the morphological word represents the spellout position? Here let us adopt the standard account. This takes

place in the position of the deepest unit of the mirrored morphological word if none of its other elements has a "strong feature". If some do, then spellout takes place in the highest strong position. Thus both "overt" and "covert" head chains correspond to MWs. In (14) for example take a VP-adverbial like "often" to be in the spec of some head F, between Infl and v (Cinque 1997). V-v-F-Infl is then spelled out in the position of Infl in French and in the position of v in English, preceding and following the adverbial respectively.

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 spelled out in the position of S but where the complement structure is S-X-H, ie. where S is separated from the complement that mirrors (syntactically represents) H by another head, X. This is impossible by hypothesis (ie. by the mirror principle): no such complement structure could have been created, since the complement structure must mirror the morphological structure, -- here H+S.

Having eliminated head chains in favour of morphological

words, 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 can exist, these link syntactic spec positions.FN12 Head chains correspond to morphological words, which are not chains, hence they do not need to share properties like c-command and locality with ("phrasal") chains. Similarly, no scopal chain-ambiguities are expected. In "John needn't come," syntactically the auxiliary is in a single position below negation. Its presence above negation is purely a morphological, or rather spellout matter.

The fact that VP-ellipsis tolerates traces of XP-chains but not traces of head-chains in the antecedent also becomes more transparent. Under the present theory X^0 -chain relations and XP-chain relations are treated quite differently, hence their different behavior is expected. Major differences between the concept of the morphological word and that of the (XP-)chain provide various ways of accounting for the contrast in behavior. One possibility might be to consider each chain member to have its own spellout position. Normally the spellout position of traces (non-head chain members) will not be used as a consequence of independent interacting factors, like for example Case theory in the case of an A-chain trace. But the assumption that each chain member has its own spellout position is necessary to treat cases where for example overtly realized resumptive pronouns show

up in the position of traces.

We can then assume that the antecedent of the ellipsis site must contain a spellout position of the elements that are subparts of the antecedent. This seems to be a natural hypothesis especially if ellipsis involves a morphological/spellout operation, like destressing. It follows that the antecedent may contain traces/copies but not proper subparts of MWs whose spellout position is elsewhere, outside the antecedent.

The theory has other advantages. Certain questions that arose in the minimalist and earlier phrase structure systems, and which the theory summarized in section 2.2 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^0 s, hence no possibility of one type dominating the other in an illegitimate configuration. The syntactic representation can be interpreted as consisting either of only X^0 -type or of only XP-type elements.

Recall also that when a category is interpreted as a word, all spec-head relations, mirrored (ie. 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 morphological words, all chains (targeting syntactic spec positions) 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. In accordance with this understanding of what "phrasal" means, in what follows I shall use terms like "VP" to refer to the V node together with its constituents (in other words the nodes it dominates). This use of the term "phrase" should of course not be taken to imply the existence of a V-projection, ie. of a phrasal category distinct from the V head.

3.4. Specifiers and complements

Without any auxiliary assumption, the hypothesis that the syntactic head-complement relation corresponds to a morphological spec-head relation entails that all complements that do not form an MW with the element on which they depend must be reanalyzed as

syntactic specifiers. This conclusion does not appear to create major problems for arguments of lexical heads, like for example direct objects of verbs. These must now be specifiers not only in their chain top but also in their chain tail position. Some theories of aspect like eg. Borer 1993 allow the assignment of semantic roles in Case-checking spec-positions. In the context of the generalized projection principle (Brody 1995a) this leads directly to the conclusion that objects must be "base-generated" in their Case-checking spec positions, ie. that they do not have a lower (complement) chain-tail. (Cf. Arad 1996 for a theory of aspect related to Borer's that embraces this conclusion.) If on the other hand objects have a lower "VP-internal" chain tail, that can still be a spec in a structure that would correspond to a multiple layer VP (with a decomposed verb interpretation of the Larsonian shell structure, on this matter cf. Brody 1995a, to appear):

(16) v

 V

 Obj

In (16) Obj is the spec of V and a member (constituent/dependent)

of the complement of v, hence it follows v and precedes V. This may be an appropriate structure for a sentential complement for example, if this does not form a Case checking chain. Further decomposition of the verb may be necessary to accomodate additional complements like datives, obliques adverbials etc, but this is no different from the situation in other binary branching theories.

How about those complements of functional categories, that are neither arguments of lexical heads nor appear to form a morphological word with the selecting head? Take the English complementizer node "comp" and "infl," ignoring their internal composition for present purposes. It may be possible to say that when the comp node is empty, as in "**comp** Mary **has** left," the element in infl forms a morphological word (corresponding to a covert head chain) with the comp node that is spelled out in infl. Mirroring this MW makes infl the syntactic complement of comp. However infl and comp can also be spelled out independently as in "**that** Mary **should** leave." In order to treat "should" as the syntactic complement of "that" here, it would be necessary in present terms to postulate a morphological word that includes both. This would correspond to the standard covert head chain linking these two elements. But while nothing in principle prevents us in the standard framework from referring to the less

restricted notion of covert head chain, it would seem to empty the notion of morphological word of most of its content if we allowed a such a unit to be spellt out as two nonadjacent, independent and morphologically non-interacting segments.

We are thus led to assume that contrary to the generally held view, infl must be a specifier of comp, rather than its complement. Similarly, assuming that in the above example the main verb "leave" does not form an MW with the auxiliary "should", the verb must be analyzed as a specifier rather than the complement of "should". (Again, for the sake of the example I ignore decomposition of these elements and other potentially intervening heads.) Note that from the point of view of the validity of this conclusion it is immaterial if the auxiliary is a higher verb or if it fills a position in the extended projection of "leave" (cf. eg. Cinque 1997 for discussion of these possibilities.)

The conclusion that those complements that constitute separate morphological words are in fact always specifiers is not strictly speaking forced. But I shall accept it, essentially because it appears to be preferable to the alternatives that would avoid it. Before exploring the consequences of this conclusion further, consider first the following alternative approach, which I shall call the "extended word theory." The

complement relation between two functional or a functional and a lexical head 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 referred to as extended projections and which, as noted above, recent research suggests have a universal structure. One could assume^{FN13} that it is in fact "extended words" corresponding to extended projections, rather than morphological words that complement relations mirror.^{FN14}

This theory of extended word mirroring could then have two different versions as concerns the status of morphological words. One possibility would be to create morphological words along the usual lines by creating chains. Given the antisymmetric nature of the structures, this would regularly involve remnant "movement" chains. Alternatively an extended word could be thought of as an abstract morphological word, larger than the unit that morphology can submit to phonology/spellout. This larger unit would be related to the pronounceable MW via the same spec-head relation that relates elements of the smaller pronounceable units, --spec-head is the only configurational relation that exists. In this version the (rightmost) element E_n of an MW L_1 that is not the spec of anything internal to L_1 can become the spec of the (leftmost) element E_1 of another MW L_2 , ie. of the element that

has no spec internal to its own MW L_2 . If the links that create extended words are taken to be morphological then the whole of the extended word will be mirrored in syntax as a set of complement relations.

Both versions of the extended word theory seem dubious. Under the first version, the grammar must contain a separate component where extended words are assembled, --general parsimony considerations strongly militate against postulating such an, otherwise unmotivated, component. Notice that the status of extended words and morphological words is very different: MWs are assembled (or, equivalently here, their structure constrained) in the morphological component, an independently necessary modul of the grammar. Additionally, if MWs are assembled via syntactic chain construction, then in addition to the mirror principle, which ensures that extended words are mirrored as complement series, the HMC and the no excorporation condition need to be reintroduced to constrain the structure of morphological words. But as we have seen above (cf. the discussion following (25)), these conditions are redundant in the context of the mirror principle in a theory based on the mirroring of morphological words.

The second version of the extended word theory, according to which extended words are abstract morphological units does not improve matters much. This is because whatever initial

plausibility the idea that morphology allows larger units than it can submit to phonology (ie. to spell out) might have, incorporation phenomena tell us that extended words are not larger than what can phonologically present itself as a morphological word. N and V incorporation involve MWs that can span several extended words (cf. eg. Baker 1988 or section 4.2 below for examples). Hence on the second version of the extended word theory we would be left without any principled reasons for why extended words do not necessarily form phonologically observable morphological words.

This argument against the second version of the extended word theory might be taken to be weakened by the fact that the present framework in principle provides the usual two major options for the treatment of open class incorporation structures, which can be assimilated either to syntactic chains (that correspond to XP-chains in standard theories) or to MWs (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 incorporation involves morphological words then a unit consisting of an incorporated element and its host will normally involve two "extended words".

Consider version 2b of the extended word theory in which extended words are created nonsyntactically and mirrored as before, and thus inflectional morphology is treated in terms of

morphological words but open class incorporation is analyzed differently, as involving syntactic chains and thus syntactic spec positions. This theory is also unlikely to be correct, however. In the next chapter 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. If inflection but not open class incorporation is treated in terms of morphological words, their similar strict locality behavior would not be captured. On the other hand if both phenomena are expressed in terms of morphological words, then their strict locality follows from the mirror principle in the present theory.FN15

Given the foregoing considerations, I shall at least tentatively conclude that the mirror theory should hold in its strong form: a category C can be the syntactic complement of another, C', only if C is the morphological specifier of C', ie. C and C' form a morphological word.

The relation between the syntactic spec and the head is biunique. Hence if the category C that would standardly be

treated as the complement is analyzed as the spec of C' (where C' is a functional category or a (decomposed) lexical category (segment)), C' cannot have another syntactic spec. An element that standardly would have been taken as its spec will now have to occupy the spec position of some higher (functional or lexical category segment) head. This, perhaps radical, conclusion is corroborated by independent evidence. Cinque (1997) argues that there exists an Agr-type "DP-related" head that dominates each functional element. Certain distributional generalizations (like eg. the fact that the subject in Italian must precede all lower adverbials but may either precede or follow all higher ones) suggest that the subject can occupy positions intermingled with the adverbial positions of the universal hierarchy. Cinque notes further, that an Italian finite verb can be found in each position between a subject and a following adverb. For example in (17) the finite verb "ha" is between the subject "Gianni" and the adverbial "rapidamente"/"raramente".

(17) a. *Gianni **ha** rapidamente alzato di nuovo il braccio*

"G. Has quickly raised again his arm"

b. *Gianni **ha** raramente rifatto tutto bene*

"G. has rarely done everything well again"

On the assumption that the subject in these sentences is in

between the adverbial positions of the universal hierarchy, this indicates that a head position (call it "Agr") also exists between each such subject position and the following adverb.

There is also evidence (Koizumi 1993, Bobaljik 1995) for an Agr_O position lower than the small v head associated with the base position of the subject and similarly for Agr_{IO} lower than the theta position of the object. Thus there is independent reasons to make the assumption, essentially forced by the restrictive nature of mirror theory, that lexical and functional heads are alike in being at least potentially dominated by an Agr-type (whatever that exactly means) head. This head can host specifiers standardly associated with the functional or lexical head the Agr-type element immediately dominates.

Note that this conclusion does not entail that all categories that are standardly treated as specifiers, ie. sisters of some intermediate projection level will now necessarily be specifiers of an Agr-type node. All that follows from mirror theory is that if a category X had standard spec and complement Y and Z respectively, and under mirror theory Z is now the syntactic spec of X (in the case where Z and X do not form an MW, so Z cannot be the complement of X) Y cannot also be the spec of X. But it may not only be the spec of the Agr-type head associated with X, but also the spec of some other higher head. One such case may be instantiated by subjects in VPs if these are

indeed in the spec position of a small v head dominating V.

Thus the standard phrasal structure in (18) will generally correspond to the structure in (19) under mirror theory.

(18) XP

 YP X'

 X ZP

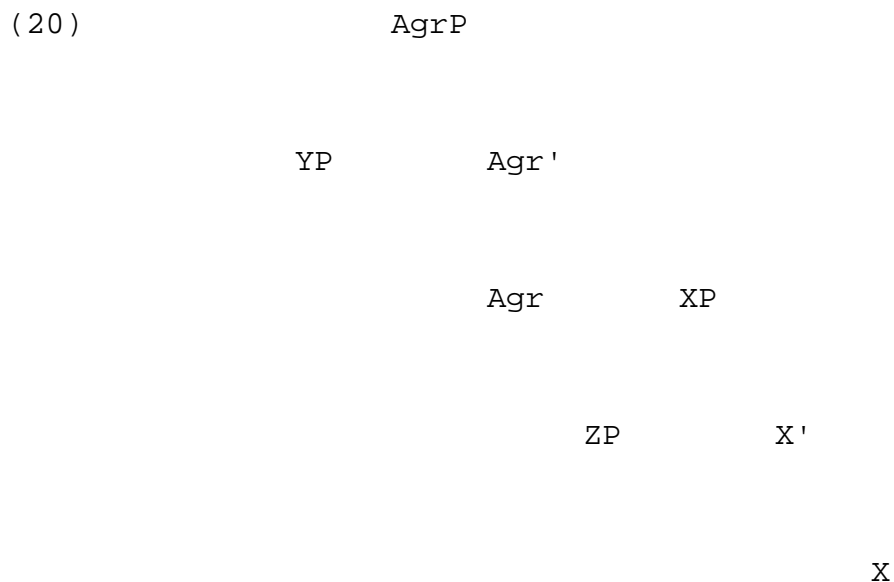
(19) Agr

 Y X

 Z

Agr and X must form a morphological word here, which on the assumption that Agr is strong will be spelled out in Agr, giving spec-head-comp order.FN16 Even though not strictly speaking incorrect, it is somewhat misleading to relate (18) to (19). (19) does indeed express the spec-head-complement structure that has been expressed standardly as (18), but (19) expresses the claim that the spec and the complement are spec's of two related but

distinct heads. Thus it in fact corresponds more closely to a structure that in standard terms would look like (20)



Let us return finally to the concept of extended projections. Recall that these correspond to extended words in the present theory which eliminates phrasal projection altogether. If a category can be a complement only if it forms a morphological word with the element it is the complement of, then extended words, cannot generally be thought of as a series of heads, each the complement of the next. It will remain true, however, that extended words must correspond to a series of heads where each dominates the next (each depends on the previous one), but the dominance/dependency relation can now involve not only a

morphological spec-head relation that corresponds syntactically to a head-complement link but also a syntactic spec-head relation.

This has an immediate advantage in the analysis of those structures that under Kayne's antisymmetry hypothesis have to be treated in terms of a phrasal "roll-up" chain structure. Kayne (in derivational terminology) suggested (21) as one possible structure for sequences of inflectional morphemes in head-final type languages. (The other alternative is leftward head-raising with complements shifted across the spellout position of the head+suffixes unit, a configuration that in terms of the mirror theory is analyzable as a morphological word, --cf. section 4.3 below for more detail and a specific case.)

(21) ...[/XP [/YP ZP Y (ZP)] X (YP)] U (XP)

In (21) the complement of Y is shifted to Y's spec position and YP, a complement of X, is rolled up as a whole into the spec position of X. Similarly XP, U's complement, which contains all the elements so far described turns up in the spec of U. If Y is a verb and X and U inflectional suffixes, then the verb will follow its complement, ZP, but it will precede the suffixes, with which it will not form a constituent. Cinque (1997) notes, quoting Mahajan (1989), that in Hindi the sentence final sequence

of verbs carrying the functional suffixes indeed do not form a constituent, just as Kayne's suggested analysis of this language predicts. This appears to rule out also the alternative of simple leftward head movement analysis.

Cinque observes further that in Hindi, the order of the V+functional suffix series combinations are the mirror image of what he argues is the universal order of these elements, exhibited directly by English and Spanish for example. The order (22), which he argues is universal is exemplified in (23a) and (23b):

(22) Tense > Perfect Aspect > Progressive Aspect >Voice >V

(23) a. These books **have been being read** all year

b. Esos libros **han estado siendo** leídos todo el año

Compare the Hindi examples in (24)

(24)a. Kis-ko raam-ne socaa ki siitaa-ne dekhaa thaa
 Who Raam thought that Sita see-ANT be-PAST
 'Who did R. think that S. had seen?'

b. Raam rotii khgaataa rahtaa thaa
 Raam bread eat PROG be-PAST
 'R. used to keep on eating bread'

Cinque's point, that the inverse order of the series of main V+functional suffixes follows directly from the Kaynean roll up structure, may be taken as further evidence for the phrasal roll-up analysis. Similar phrasal roll-up structures have been proposed also for various configurations, like sentence-final adverbial clauses, final stacked adverbials and PPs, etc. in head-initial type languages, --cf. eg. Kayne 1994, Cinque 1997, Barbiers 1995 and section 4.2 below for more detailed discussion.

As is widely acknowledged, the approach is quite problematic however, since the chain structures it postulates have no independent motivation. In particular, no systematic set of "triggers" for these movements have been found, --in more representational terminology, we have no principled account of what licences the non-tail positions of these chains. Furthermore, there is in fact no genuine independent evidence for the presence of the tail positions of these chains in the relevant structures. In other words it would be preferable to generate the roll-up structures directly, with the complements starting out in spec positions and eliminating the chains linking spec's and complements. We can then take whatever selectional relation was taken to license the complement to in fact license the same element in the spec position. This eliminates both the "movement-trigger" problem and the problem of lack of direct evidence for the chain-tail in complement position. In order to

do this, it is necessary to reject the assumption that each element of an extended projection must be the complement of the previous one. But this is exactly the proposal I arrived at above on independent grounds in the more restrictive framework of mirror theory: that in the series of elements corresponding to an extended word, where each dominates the next, both morphological spec-head (ie. syntactic head-complement) and syntactic spec-head relations are legitimate. The restrictions of the theory thus again forces an empirically justified analysis.

3.5 C-command and antisymmetry

The relation of c-command ceases to be necessary as it applied to head chains: elements of MWs are in a dependency/domination relation with respect to each other. The remaining conditions that involve c-command can also be restated to refer to the dependency/domination relation. The dependency or domination relation is the simplest non-local structural relation between a pair of categories definable on syntactic trees (apart from the relation between any two nodes, if that is taken to be a structural relation); --the structural "equivalent" of precedence. For example principle C can be taken to prohibit an R-expression from having an antecedent that dominates it (or

equivalently, on which it depends), --presumably an Agr-type node. If the syntactic spec of this Agr is taken to pick up the reference of the head, then this spec 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 chapter 2, a central issue about c-command, which was not resolved there in a satisfactory way, is the strange asymmetric stipulation in the definition of this relation: x c-commands y iff the category **immediately** dominating x dominates, \pm **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. As indicated, the conclusion is made possible by the mirror theoretical analysis of head chains as morphological words.

Some evidence for factoring the notion of c-command into the domination and the spec-head relation is provided by the properties of the non-distinctness requirement that chain-members are subject to. Lower members of a chain may optionally omit adjuncts, but not arguments, present in the highest member of an "overt" chain. Thus recall (section 1.4) for example, that the "reconstructed" trace/copy position triggers no principle C violation in cases like (25) in contrast with (26):

(25) Which claim that John made do you think he later denied?

(26) ?*Which claim that John was asleep do you think he later denied?

On the other hand it has been proposed that "covert" chains involve only a subset of the features of the contentive in the tail of the chain, cf. Brody 1995a, Chomsky 1995. Thus in "covert" chains the lowest element must be the most fully specified one, higher members are feature sets, while in "overt" chains the highest member must carry the full specification and lower members are (potentially) less fully specified categories. We can make sense of this situation in terms of the assumption that the concept of chain refers to the relation between a constituent and one or more sets of features that dominate this constituent. So in a chain it is always the lowest element that

is the most fully specified one. The approach instantiates the idea that covert chains only have certain features of the contentive element in their non-tail positions differently both from the theory of Chomsky 1995 and from that in Brody 1995a, and to appear. Cases standardly treated as "overt chains" will additionally involve spec-head relations with the feature sets in the chain. Presumably for reasons of recoverability, the highest, normally phonologically overt, specifier constituent must be more fully specified than either the other lower specifiers of the same chain or the contentive element of the chain itself.

Consider (27), a potentially problematic construction for the proposed approach to c-command, discussed in Reinhart (1983) and more recently in Kayne (1994):

(27) Every eight year old girl's father thinks she is a genius

Here the quantifier phrase "every girl" can bind the pronoun "she", but it is not the specifier of any head that dominates this pronoun. Under Kayne's theory of specifiers as adjuncts, "every eight year old girl" can be taken to be the highest specifier in the DP and c-command "out of" the DP, hence bind the pronoun. Kayne relates (27) to (28) and (29). In (28) he takes the negative, in (29) the wh-phrase phrase to similarly c-command

out of the DP. He assumes that this makes it possible for the negative to licence the polarity item and the wh-phrase to serve as the spec of the +WH C.

(28) Nobody's articles ever get published fast enough

(29) We know whose articles those are

Given that the present theory dispenses with adjunction, to achieve the same pronoun binding effect in (27), it is necessary to assume that the QP can be linked to some higher position, perhaps to the spec of Beghelli and Stowell's (1996) DistP, in any case to a position external to the DP. The possibility (according to Kayne only in some registers) of (30), where the QP cannot be analysed as c-commanding "out" of the DP from its overt position suggests that some version of QR is in fact necessary to treat these structures.

(30) The father of every eight year old girl thinks that she is a genius

Kayne proposes that the possibility of (30) is due to the option of covertly moving the QP to the highest spec position inside the DP, from which it would c-command just as in (27). The same option can save (31), by assimilating it to (28). But the

approach leaves it open why (32) cannot be similarly saved, by being reduced to (29):FN17

(31)The author of no linguistics article ever wants it to go unread

(32) a.*We know articles by who(m) those are
b.*I wonder people from what city the game is likely to attract

If however we assume that the universal and the negative quantifiers are linked to a higher position in order to satisfy some requirement in spec DistP and spec NegP respectively, then (32) remains without any obvious motivation to be so linked. This provides the necessary distinction. Notice that the wh-element cannot link to spec-(+WH)C, since that position is occupied by the larger wh-phrase that properly includes it. In order to allow (29) it is necessary to postulate also feature percolation or some equivalent way of wh-phrase construction, but this must remain restricted to spec positions.

A similar solution is necessary for relatives, given (33):

(33) the man pictures of whom proved that he was innocent

Given Kayne's antisymmetric analysis of relatives, "the man" does not form a constituent in (33), hence it cannot serve as the antecedent of the pronoun. Kayne assumes that relatives have a [_{DP}D CP] structure, with the highest determiner "the" in D. According to this analysis in (33) "pictures of whom man" raises from subject position to spec-CP and then "man" raises further to the spec of the originally raised larger constituent. Again, given the lack of adjunction in the theory defended here, I am forced to assume that "man" raises higher than the spec of this constituent, so that it can c-command the bound pronoun. Presumably it raises to the spec of some element of the split complementizer space. Notice that both under this view and under Kayne's it is necessary to stipulate that an additional option of linking the wh-element to a higher position is available in relatives but not in interrogatives (cf. the contrast between (33) and (32)).

The assumption that the constituent "man" in (33) is higher than the (highest) spec position of "pictures of whom" entails that "pictures of whom proved that he was innocent" is a constituent, which Kayne's structure denies. The present approach is supported then by (34) and (35) which suggest that the traditional assumption that the relative wh-phrase together with the following relative clause form a constituent that does not include the "head" noun is correct:

- (34) a. the man who Mary saw and who Betsy liked
 b. *the man who Mary saw and man who Betsy liked
- (35) a. This is the man, I believe, who Mary liked
 b. *This is the, I believe, man, who Mary liked

Notice that it would be implausible to defend the constituent structure Kayne proposes by assuming that the niching of the phrase "I believe" in (35) is governed by conditions stated on phonological or intonational phrases, which need not correspond to syntactic constituent structure. First this would seem to be irrelevant for the constituency evidence from co-ordination in (34). Secondly under this kind of defense of Kayne's assumptions, given (35) above or (36) below, phonological phrasing would have to be able to separate the specifier of the specifier of a constituent and attach it to the preceding phonological phrase. But then we would expect, counterfactually, the parallel phrasing in (37) to be also possible.

- (36) The dog [which bit the horse] [which kicked the mouse]
 (37) *John thought my [brother said your][friend left]

Finally the antisymmetric property of representations is also ensured by mirror theory. While this is intuitively clear,

Kayne's LCA, which relates (asymmetric) c-command and precedence could not be adopted here. Given mirror theoretical structures, neither the standard definition of c-command, nor the domination relation that I proposed as an improved alternative, can be straightforwardly mapped to precedence relations between terminals. But under mirror theory there are simply no means provided from which non-antisymmetric structures could be built. Hence no overriding condition like the LCA is necessary to ensure the antisymmetry effects. As in the case of the structure of crystals for example, the properties of the basic building blocks determine the limits of variability of the composed larger structures.

More specifically, the LCA ensures that specifier and complement are on different sides of the head: this follows here from the mirror principle directly. A stipulation/axiom to the effect that the specifier precedes the head is necessary in both frameworks (cf. section 2.5). The LCA entails binary branching: the mirror theory does not provide means to violate this restriction: for each head only a single spec-head and a single mirrored lexical spec-head (ie. complement) are allowed as syntactic relations. The LCA ensures that projection lines always branch rightwards, in other words that only the complement can be on the right of the head, specifiers and adjuncts must be on the left. It rules out furthermore multiple adjunction to the same

element. In the present theory adjunction is eliminated (cf. Sportiche 1994, Brody 1994, to appear and Cinque 1997 for arguments), hence the issue of multiple adjunction does not arise. Similarly, projection lines are also dispensed with, hence rightward branching reduces to specifier-head-complement order, which, as just noted, mirror theory ensures.

3.6. Summary

I recapitulate the major general characteristics (I) and advantages (II) of mirror theory:

I.(a) The only primitive relation between elements in syntax and morphology is spec \rightarrow head, where the spec (and its constituents) precede the head.

(b) By the mirror principle, some morphological spec \rightarrow head orders can be (geometrically) mirrored in syntax. The head-complement relation is just a reverse order (morphological) spec-head relation.

(c) Members (heads) of morphological words can form spec-head relations with other morphological words.

(d) An MW is spelled out (in the morphological spec-head order) in the position of the highest strong head (or, lacking a strong head, in the lowest position).

II.(a) Single primitive configurational relation: spec-head.

(b) Locality and no excorporation properties of head-chain type relations follow. All head -complement links must match (are identical to) a(n inverse) spec-head link in the word structure: structures corresponding to excorporation or non-local head-chains cannot be created.

(c) C-command problems of head-chains do not arise: morphological words involve domination.

(d) Antisymmetry effects guaranteed (no means to violate LCA requirements, hence no need for the LCA).

(e) No categorial projection, hence uniqueness and locality issues of categorial projection (Chomsky 1995, Brody to appear) do not arise.

(f) No word-phrase difference in syntax, hence no extended structure preservation (Chomsky 1995, Brody to appear) question.

(g) Single expression of word-structure in syntax (in the case of suffixes the complementation structure). All three duplications (listed in section 3.2 above) eliminated.

(h) Apparent conflict between dependency and constituent structure frameworks resolved. (Structures are interpretable as a dependency diagram or as a constituent structure.)

(i) The theory forces the independently motivated (Cinque 1997) presence of "Agr-type" heads dominating each functional element F with a spec and a complement with which F does not form

an MW. (The complement must be the syntactic spec of F, hence the "standard" spec of F must in fact be the spec of a higher head.)

(j) The theory forces a weaker characterization of "extended projections", where these must correspond to a set of nodes each dominating the next, but not necessarily in the "complement of" relation. Given the evidence from "phrasal roll-up" structures this is again apparently a correct conclusion, -- reached on principled grounds.

(k) The notion of c-command becomes unnecessary not only for head chain type relations, but given the sharing of features between spec and head, c-command may be unnecessary in general: it can be decomposed into the domination and the spec-head relations. The structural requirement on syntactic spec-to-spec chains can also reduce to domination. Such an analysis, under which each chain member must simply dominate the next is possible only under mirror theory where the relevant heads will dominate the tail of the chain.

Footnotes to chapter 3

FN1 A chain with more than two members will accordingly involve a phrase and several heads each dominating the next.

FN2 I argue in Brody 1995a, and 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.

FN3 For some discussion see Phillips 1996, especially his note 17 on p.191. and the text to which his note relates.

FN4 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.

FN5 A set of chains in a roll-up structure when the analysis is detailed enough to take account of more than one Infl position.

FN6 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. Further auxiliary assumptions would be necessary to make this account compatible with bare checking theory.

FN7 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 also needs to assume additionally the HMC and the no excorporation requirement (or the equivalent in note 6), still would not help with most of the other problems raised. I will

therefore propose a more radical solution below.

FN8 As noted (note 6), under a (non-standard) version of checking theory the excorporation prohibition may be dispensable, but the suggestion there does not explain why there is a difference between X^0 and XP chains in this regard any more than other solutions in the literature. For example Baker (1988) suggested prohibiting word-internal traces. But if separate heads can come together to form a word in syntax, why can they not separate again? Note also that under checking theory excorporation would not result in a word-internal trace anyway.

FN9 In Lasnik's example reproduced in (i) as Roberts notes "be-shift" occurs only at the ellipsis site and not at the antecedent. (This is shown by the grammaticality of "Being obnoxious, John never was" on the assumption that VP-preposing of constituents including a "be-shifted" "be" is impossible.)

(i)*John was being obnoxious and Mary will (be obnoxious) too

Roberts does not make it explicit which principle he takes to exclude (i) on his assumption that V-traces are prohibited in the antecedent of VP-ellipsis. Since "be-shift" must take place in the second conjunct, I shall assume, as seems natural, that the

ungrammaticality of (i) with no "be-shift" in the antecedent is due to lack of parallelism.

FN10 For other more recent attempts to simplify the theory of phrase structure in terms of dependences 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.

FN11 Hudson 1990 defines a constituent as a category together with its dependents (and assumes that the theory makes no use of this notion as he treats movement/chains in alternative ways.). His is thus a conceptually quite different notion of constituent from the one outlined in the text.

FN12 But see section 3.5 below for a different approach to "XP"-chains.

FN13 I made a version of this assumption in an earlier version of this work (Brody 1997b).

FN14 In the present theory "extended projections" become

"extended words" since not even non-extended phrasal projection is taken to exist.

FN15 This explanation is also lost if both inflectional morphology and open class incorporation are assimilated, as is standardly done, to (XP-)chains.

FN16 Note the possible analysis of "VP-internal" SOV: same structure , ie. (19) but with weak Agr.

FN17 Kayne makes essentially the same objection to a May-style, freely adjoining QR: this could incorrectly save (32b).

Chapter 4

"Roll-up" structures and morphological words

4.1. Introduction

This chapter contains two case studies. The first one in section 4.2 elaborates the analysis of phrasal roll-up structures suggested towards the end of section 3.4. The solution is applied to the case of sentence final adverbials which are shown to provide further evidence for the analysis and hence indirectly for the strict version of mirror theory that forces it. Section 4.3 looks at Hungarian prefix chains, and argues that some of these involve syntactic spec-targeting chains and others are better treated in terms of morphological words. Mirror theory, in which morphological words are necessarily local in a much stricter sense than the one in which syntactic chains are, provides a framework where apparently central aspects of the rather complex behavior of these prefixal elements can be understood.

4.2 Adverbials on the right

Recall the distinction between the LCA and the antisymmetry hypothesis from chapter 2. That phrase structures conform to the requirements of the LCA, the antisymmetry hypothesis, is distinct from the claim that this is due to the LCA, an additional issue. As noted in Brody to appear, the LCA creates some redundancies given the theory of phrase structure outlined in section 2.2. For example uniqueness of projection does not follow in general from the LCA but this condition predicts the effects of uniqueness for the special case when two heads X and Y are both immediately dominated by a phrase. (There will be no pair (C, C') of constituents then, related by asymmetric c-command, such that C dominates X and C' dominates Y). Similarly, extended structure preservation, another consequence of the system outlined in 2.2, follows from the LCA only in part. A phrase P dominated by (adjoined to) a word that has a complement C violates the condition, since categories embedded in the complement would be asymmetrically c-commanded by P but categories in P would also be asymmetrically c-commanded by C.FN1 Auxiliary stipulations are necessary however to derive the other half of extended structure preservation, ie. the fact that words cannot be immediately dominated by a phrase that they did not project.FN2

As we have seen problems relating to projections, like uniqueness and structure preservation disappear in the more restricted framework of mirror theory, which dispenses with

categorial projection completely. As discussed above in section 3.5, this theory furthermore enforces the antisymmetric effects of the LCA by not providing elements and relations with which a non-antisymmetric structure could be built. The theory thus directly inherits many of the problems that Kayne's antisymmetry hypothesis faces. In what follows I would like to look at one particular area that appears to have caused difficulties for the antisymmetry approach, --namely the position of English sentence final adverbial clauses (cf. Williams 1994, Brody 1994, Manzini 1995, Hornstein 1994). I shall go through a number of arguments that all seem to suggest that contrary to what might be expected from mechanically applying the antisymmetry hypothesis, these adjuncts must be higher than preceeding VP-internal elements. I will argue that earlier suggestions and partial solutions to these problems are not fully adequate. I will provide further evidence that the particular member of this family of analyses that mirror theoretical considerations led to in section 3.4 above is in fact the correct solution.

As is well known principle C tests show that adverbials like in (1) are not c-commanded by V-complements:

(1) We sent him_x there in order to please John_x's mother

Faced with such examples one logically possible approach is to deny that principle C tests for c-command. Haider (1993) took such a position arguing that the hypothesis that principle C operates under c-command leads to contradiction. In his example in (6) the pronoun her must be disjoint from the name Mary in the complement clause but not from Mary in the extraposed relative.

(2) Someone has told her_x [who Mary_x met] [that Mary*_x will inherit the castle]

Haider argues that no standard phrase structure could ensure lack of c-command of the extraposed clause at the same time as c-command of the more peripheral complement. But in fact there are a number of analyses compatible with the observation in (2) and a c-command dependent principle C. For example the relevant structure could be (3), with an RNR derivation for the complement clause:

(3) Someone has told [her_x (that *Mary_x will...)][who Mary_x met] [that Mary will inherit the castle]

(4) *[Which claim that John_x was asleep] do you think he_x denied t(which claim that John_x was asleep)

Since principle C is sensitive to elements in A'-trace positions,

as exemplified by the standard wh-movement case in (4), disjointness from the name in the complement clause can be determined in the trace position and the extraposed and complement clause could be stacked higher than and on the right of V and its complements.

A different Kaynean intraposition analysis that respects antisymmetry is indicated in (5). Kayne (1994 p.122-123) considers a different example without a complement that includes a potentially coreferential category but his analysis carries over essentially unchanged. He argues that in a structure like (5) the pronoun her is in an A'-position and has a trace position to the right of the extraposed (in Kayne's theory: stranded) clause. He points out that principle C holds in the reconstructed positions of A'-chains and we may add that it does not see elements in A'-position. Hence it is this trace rather than the overt position of the pronoun that counts for principle C, and therefore disjointness from the name in the complement but not from that in the stranded ("extraposed") clause is predicted.

(5) Someone has told [her [who Mary_x met] (her_x)[that *Mary_x will inherit the castle]

Haider's example thus does not show that the assumption that

principle C is sensitive to c-command leads to contradiction, even within an antisymmetry framework. Let us continue with the standard assumption concerning the relation of principle C and c-command which entails that the adverbial in (1) is not c-commanded by the object of the VP. The standard solution of adjoining the adverbial higher, while not available in the antisymmetry framework, appears to be corroborated by much other evidence. I shall next consider briefly some of this evidence.

The anti c-command requirement of parasitic gaps also diagnoses the complement-adjunct distinction.

- (6) a. Who did you hire t after you talked to pg
b. *Who t went home after you talked to pg

In (6a) the object trace appears not to c-command the parasitic gap in the adjunct, while in (6b) it c-commands the parasitic gap in the complement, resulting in ungrammaticality.

Kayne's (1994) suggestions concerning the apparent discrepancy between the parasitic gap data and the antisymmetry hypothesis are problematic. One of his suggestions is to return to the 0-subjacency account of parasitic gaps, but as noted by Manzini (1994, 1995), 0-subjacency accounts seem flawed in general since they would allow extraction from adjuncts. Another

suggestion of Kayne's is that the parasitic gap related operator should be taken as an object-oriented pronominal that moves to matrix AgrO and is disjoint from the subject in (6a) due to the effect of principle B of the binding theory. In addition to being rather stipulative, this suffers from the same problem as the 0-subjacency account: it involves movement out of the adjunct. Kayne's third suggestion is intraposition of the adjunct around the primary wh-trace, parallel to intraposition of a second complement around a "heavy" XP in "inverse heavy shift" constructions. As noted again by Manzini (1995), the problem here is that the trace is not heavy in any sense, hence a focusing/defocusing rule like (inverse) heavy shift is inappropriate and should be inapplicable.

Quantifier scope data is also sensitive to the difference between complements and adjuncts. Hornstein 1994 observes that (7) and (8) where the existential quantifier is taken to bind the pronoun in the adjunct and in the complement respectively, different scope relations can obtain. In (7) the existential unambiguously takes wide scope over the universal, but (8) is scopally ambiguous, here the universal can also take wide scope/distribute over the existential. Thus under the construal where the existential binds the pronoun, for example in (8a) a different person might have asked each attendant, but in (7a)

there must be a single person who danced with every woman, not a different one for each.

(7) a. Someone_x danced with every woman before he_x left the party

b. I got someone_x to review every brief without PRO_x leaving the office

(8) a. Someone_x asked every attendant if he_x could park near the gate

b. John got someone_x/at least one patron_x to tell every critic that he_x hated the play

Hornstein explains this in terms of his A-movement theory of quantifier scope. We can abstract away here from the details of this theory, since the observations suggest that the adjunct must have a different position from the complement under any theory that achieves inverse scope in such examples via reconstruction of the existential to a lower position under the universal. As long as the adjunct is higher but the complement is lower than this reconstructed position, the contrast between (7) and (8) will follow: the existential in the reconstructed position will be able to bind the pronoun in the complement but not in the adjunct.

Additional related problems for the antisymmetry hypothesis arise from structures with multiple adverbials(cf. Andrews 1983, Pesetsky 1989, Cinque 1997). On the assumption that when these are sentence final each c-commands the one that precedes it, the scope relations among these adverbials appear to be straightforwardly determined by c-command in both pre-V and sentence-final position. In both (9) and (10) the interpretation is unambiguous, with twice having higher scope in (9) and intentionally in (10). Thus the sentences in (9) describe two cases of intentional knocking while those in (10) refer to an intention involving two knocks.

- (9) a. John knocked on the door intentionally twice
b. John twice intentionally knocked on the door

- (10) a. John knocked on the door twice intentionally
b. (??)John intentionally twice knocked on the door

The last set of data in the present inventory that shows that adjuncts and complements cannot be in the same position has to do with movement and deletion tests for constituent structure. On the assumption that complements but not final adjunct clauses are lower than V, there will be a constituent that excludes the

adjunct but includes the complement. The existence of a constituent that includes V and its complements but excludes the adjunct clause appears to be confirmed by VP-deletion (11), (12) and fronting (13). (Recall that in mirror theory there are no phrasal projections, we now take the term VP to refer to V and all categories it dominates.)

(11) Mary sent him there in order to please John's mother and Klara did in order to upset her.

(12) Although Mary did in order to upset John's mother, Klara sent him there in order to please her.

(13) ...and send him there Klara did, in order to please John's mother

Examples parallel to (11)-(13) can be constructed also with complement remnants, but the acceptability status of such examples appears to be different. (In the case of deletion (14), (15) we get "pseudo-gapping" structures, on these cf. Williams 1995):

(14) ? Mary sent John and Klara did Bill

(15) ?Although Mary gave the book to John, Klara did to Bill

(16) ? ...and give the book Mary did to John

All these examples might be analyzed as involving shifting of remnants to the front of the V followed by remnant VP-fronting/deletion. But whatever the analysis is that makes deletion and fronting of what appear to be VP subparts in (14), (15) and (16) possible, these structures are somewhat marked. A configurational difference between adjuncts and arguments can explain why VP fronting/deletion is not similarly marked with non-complement remnants.FN3

The combined evidence from principle C, parasitic gaps, quantifier and adverb scope, VP fronting/deletion clearly shows that sentence final adverbials and adverbial clauses are not in the same position as complements. Let us consider the type of solution to the antisymmetry problem this fact creates that many have proposed. This involves creating the difference between complements and adjuncts via placing some part of the structure that includes the VP-internal material into a spec position higher than the adverbial.

For the result clause structures in (17) Kayne suggests such a structure:

(17) a. She_x has so much money now that Mary_x is the envy of all her classmates

 b. [[_y She_x has so much money now] [_c that] Mary_x is the envy of all her classmates t_y

Kayne proposes that there is no disjointness effect in (17) because the string "she has so much money now" is in fact an embedded clause that surfaces in spec-C (as indicated). The pronoun she then does not c-command Mary; --and this is true both for its surface and trace-internal position.

Putting aside for the moment^{FN4} the question of what triggers movement to spec-C, a problem that would be relevant for both the adverbial clause in (1) and (17), it is easy to see that an analysis along these lines, which would take the main clause to be in the spec-C (or spec-P) of the in order to clause as in (18) could not work for the case in (1).

- (18) a. in order to please John's mother [we sent him there]
b. [we sent him there] in order to please John's mother

The analysis looks dubious to start with since it inverts the intuitively clear main and subordinate clause status in the structure and provides no plausible source for the main clause. The incorrectness of the inversion of the main/subordinate status can be shown in cases where the structure is embedded further. In (19a) selectional properties of the main predicate show that the main clause of (1) is in fact the main clause. If this was in the highest spec (spec-C or spec-P) of the adverbial clause, then we would expect the adverbial clause and not the clause in its spec-

C to be interrogative. The clause in spec-C would clearly be too far from the matrix predicate for wh-selection.

- (19) a. John wondered who we sent there in order to please him
b. John believes him to have been sent in order to...

Similar comments apply in the case of a matrix ECM predicate as in (19b). If the main clause was in spec position, its own spec would be too far from the matrix Case assigner to end up accusative, whether ECM applies via government or movement to spec-AgrO.

These problems will not arise under a different version of the hypothesis (essentially following the the proposals of Barbiers 1995, and Cinque 1997 among others), according to which in (1) it is the VP and not the CP (ie. in mirror theory the V and not C, together with the nodes it dominates) of the main clause that occurs in a spec position. This spec position would be lower than the subject of the main clause, and thus lower than C and I and their spec's which take their usual place. The problems just raised in connection with the CP shift to spec-C of the in order to clause then would not arise, as the spec-C and the subject of the main clause would occupy their standard position, making wh-selection and ECM structures feasible. To

solve the problem of lack of source for the VP assume that the in order to clause is in a spec position of some functional head. This is natural, given Cinque's (1997) theory in which all adverbials that precede the thematic position of V are in the spec position of their own dedicated functional head.^{FN5} In the spirit of Barbiers and Cinque take then the VP to raise in order to serve as a subject of the adjunct clause, understood as a predicate.

Moving the VP to a spec position above the adverbial answers all the problems listed above in connection with clausal adverbials. The object of the V will not c-command then the adverbial clause, hence there will be no c-command violation in (1) and the anti c-command requirement in the parasitic gap structure (6a) is respected. (6b) will continue to violate the requirement since the subject trace in spec-I will c-command the adverbial. Assuming that the universal can distribute over the existential quantifier only if this latter is reconstructed into its theta position under V entails that in examples like (7) under inverse scope the existential cannot bind a pronoun in the adjunct: the reconstructed existential is now embedded in a spec node. In other words, if the existential binds the adverbial clause internal pronoun then it is too high to be reconstructed under V for inverse scope. There is no problem with binding the pronoun in the complement in (8) under reconstruction since the

complement is dominated by V (in standard terms: internal to the VP) in spec-Pred. (The same account of the complement adjunct contrast will work also if reconstruction targets a position higher than V, say the "share" head of Beghelli and Stowell (1996), as long as instead of V this higher head (and therefore everything this head dominates) is taken to occupy the relevant spec position.)

Furthermore since the VP in spec-Pred is a constituent that includes the complements but not the adverbial clause, the analysis can straightforwardly distinguish the fully grammatical VP deletion/fronting cases in (13) and (14) that leave the adverbial clause behind and which involve deletion/fronting of a constituent from the semi-grammatical cases in (15) and (16) which either do not operate with constituents, or more likely, act on remnant constituents and which involve markedness.

The analysis raises other questions however. Manzini (1995) objects to this kind of approach on the grounds that it would create a unique case where a nontrivial chain is created for the sake of satisfying a checking requirement involving predication. It is not obvious how strong this objection is. One might imagine that the presence of clausal subjects in spec-I is triggered by predication, especially in a framework like that of bare checking theory (section 1.5 above), in which all syntactic features must

be semantically interpretable, and therefore heads of A-chains cannot be licensed by uninterpretable Case. Cf. Brody 1997 for some discussion of this matter. At least one other likely candidate for a treatment in terms of predication would be the head of the relative under the raising analysis of relatives forced in the antisymmetry framework (Kayne 1994).

But there is another, related problem. To see this, it is necessary to establish that the adverbials are in spec position contrary to the suggestion that the V(P) originates as (or in) the complement of the adverb (eg. Sportiche 1994). Cinque provides evidence against this latter assumption based on the existence of head positions in Romance in between adverbials. But, in any case, the proposal of considering adverbs as heads would not generalize to clausal adverbials. The analysis could not take the adverbial clause to be the complement of the adverbial, as it assigns this position to the VP-trace (ie. to the copy of the V which together with the nodes it dominates occupies the theta-related chain-tail position). Given that adverbials must be in spec positions, it is natural to take the V(P) to be the spec of a higher node, call it Pred. The complement of Pred is the adverbial suffix in the spec of which the free morpheme adverb sits. In turn the complement of this adverbial suffix includes the V(P) trace position. The heads Pred and Adv are empty here.

(20)

I

Pred

"VP"

F

Adv/Adv clause

t(VP)

Thus under this analysis it is not enough to add predication to the set of triggers, but additionally an otherwise not obviously necessary functional head, Pred, must be postulated..

A further problem arises, in the case of adverbials apparently stacked on the right periphery of the sentence as in (9a) and (10a). In the antisymmetry framework examples like (9a) and (10a) are usually treated in terms of roll-up structures, that is in terms of successive incremental intraposition. Thus it is assumed that the adverbials are stacked on the left of the V and the VP raises first to the left of the lower adverbial A and then the constituent it forms with A raises to the left of the second adverbial. This ensures that the underlying scope order is the same in the (a) and (b) examples in (9) and (10) in spite of the surface word order difference.

The successive incremental intraposition analysis exhibits also the problem just discussed in connection with adverbial clauses. The only likely candidate for triggering the intrapositions in the roll-up structure is predication: the raised category must serve as a subject for the adverbial predicate. Thus the lower adverbial is taken to be predicated of the VP and the higher one of the constituent that includes the VP and the lower adverbial. This means that an otherwise unmotivated Pred head must be present on top of each adverb.

Notice that this consequence cannot be avoided by assuming that stacked adverbs as opposed to adverbial clauses are in head position (and thus establish a spec-head predication relation with the VP directly). As exemplified by constructions like (21), the adverbial can have its own complement distinct from the VP also when there is more than one adverbial stacked at the sentence periphery.

(21) John kissed Mary recently more quickly than/before Bill

And there may be more than one adverbial on the right periphery, with inverse scope order even when one or more of these are clausal:

(22) John kicked the door twice (intentionally) in order to

irritate Mary

(23) John left when Mary arrived in order to please her

The interpretation of (22) is that "It was in order to irritate Mary, that John intentionally kicked the door twice", ie. "in order to> intentionally>twice". In (23) again, the rightmost "in order to" clause takes scope over the temporal adjunct clause.

Hence the analysis that takes the VP-trace to be the complement of the adverbial is unlikely to be helpful and we can safely assume that the analysis in (20), with the adverbial in spec position, is essentially correct also for structures that contain adverbials without complements.

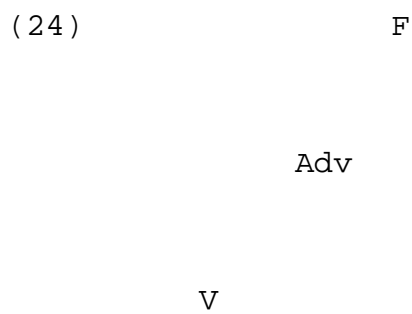
The problem (additional to the adverbial clause case) that this analysis of multiple adverbials creates is that we seem to be without means to rule out the case where both adverbials are predicated of the same VP, ie. to rule out successive step non-incremental VP-intraposition. (Note that a single subject can in general have more than one predicate as for example in structures with NP-chains.) If successive non-incremental VP-intraposition was allowed, then it should be possible for example in (9a) for the VP to "move" to a spec-Pred above "twice", and then "move" further to the spec-Pred above "intentionally", resulting in the interpretation on which "intentionally" has higher scope than "twice", which is unavailable in (9a).

In other words, the postulated mechanism of successive incremental VP-intraposition is quite unlike other known cases of XP-"movement". Unlike other known cases of XP-chains, it is not allowed to be multiple membered (no "successive non-incremental movement" is possible, but it can have the incremental movement structure that other well-established cases of XP-chains cannot have. In fact incremental VP-intraposition is more similar to head-movement in the standard minimalist framework than to XP-movement. It is far from clear, how such a similarity could be expressed and motivated. In chapter 3 I argued against the existence of head-chains in syntax. If this is correct, then there remains no likely candidate in syntax to assimilate incremental intraposition to.

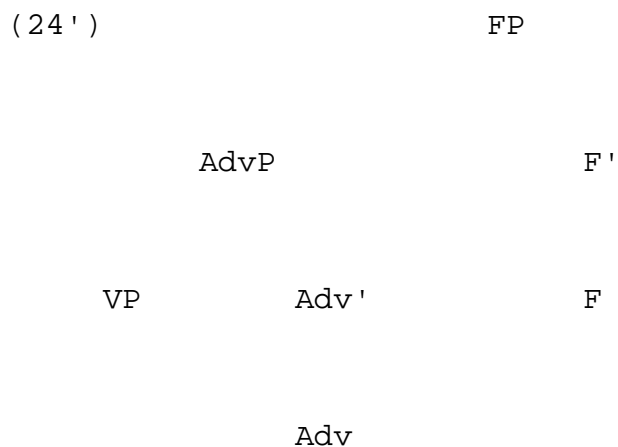
These considerations lead to the hypothesis, that structures involving adverbials on the right periphery should involve no chains at all.

An interesting consequence of the assumption that the VP can be "base generated" in a spec position, more precisely that the tail of a (normally trivial, one-member) VP-chain can be in a spec-position is that it can resolve a (near-)conflict between the two hypotheses concerning the status of the adverbs that were discussed above. The assumption that the adverbs are or can be in head position and the VP is or can be in their spec (cf. eg.

Sportiche 1994) is not in fact strictly incompatible with the status of adverbs as a spec's (Cinque 1997). Suppose that the VP is in the spec of an adverb that itself occupies the spec position of the associated and dedicated functional head, as in (24):



The structure in (24) corresponds to the more standard (24'):



If the VP forms a chain with a trace under F (F' in (24')) then

the configuration in (24)/(24') is possible only under a Kayne type definition of c-command, allowing spec's to c-command "out" of spec's they are dominated by. A simpler notion of c-command, such as the ones discussed in chapters 2 and 3 would not allow this. But if (24)/(24') represents the "base" position of the VP, ie. if there is no trace under F/F', then the c-command issue will not arise at all.

Notice that in order to allow (24)/(24'), it is necessary to weaken the notion of "extended projection" (or "extended word" in mirror theory). The extended projection/word of V in (24)/(24') includes F. An extended projection now must be allowed to continue not only via complements, but also via spec's. (Note furthermore that on the assumption, which I keep, that the extended word cannot skip categories (each must immediately dominate the next), the extended word of the V in (24) includes also the Adv head.)

The problems of the successive incremental intraposition account do not carry over to this analysis. No Pred needs to be postulated, since as we have just seen, the VP can be in the spec of the associated adverbial. There is no issue of excluding non-incremental intraposition, VPs do not "move" to spec-Adv at all, either because of the c-command problem or because this is the position where the adverb and the VP are in a modification relation and this relation should be governed, along with

selectional relations, by the Generalized Projection Principle (GPP) or both.

The structure of a clause with more than one sentence final adverbial will then be (25), which corresponds in the standard frameworks to (25'):

(25)

F*

Adv*

F'

Adv

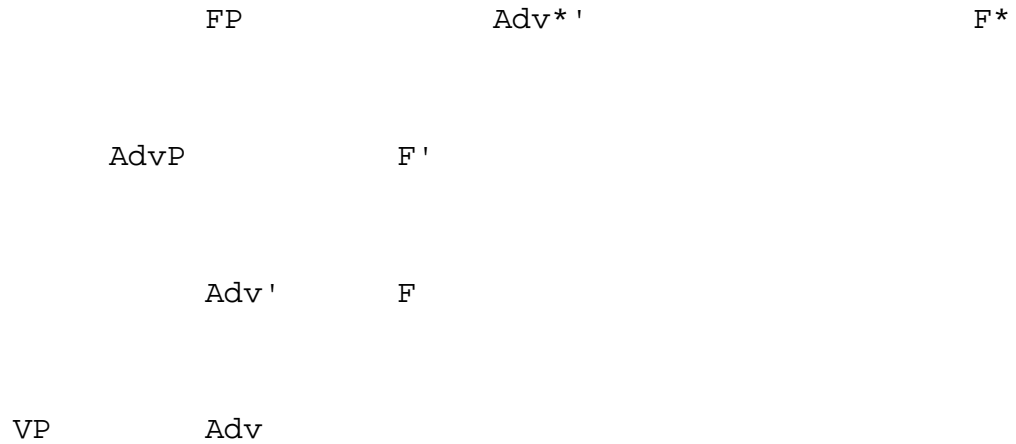
V

(25')

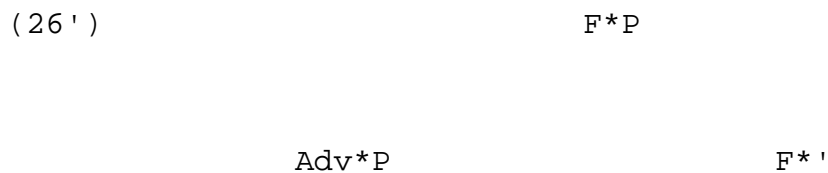
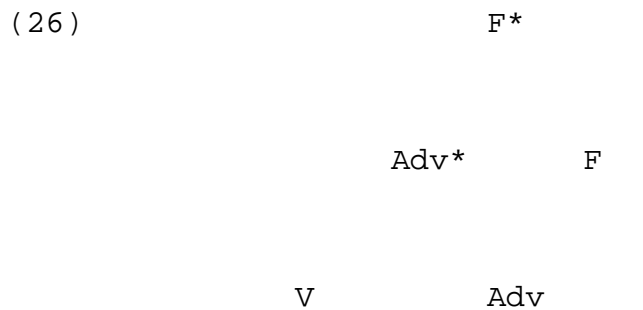
F*P

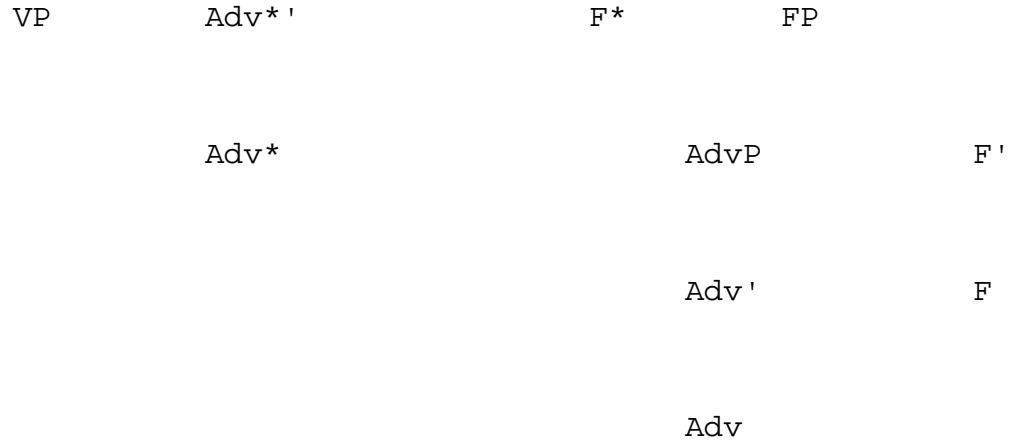
Adv*P

F*'



A structure like (26)/(26'), corresponding to the successive non-incremental intraposition case cannot arise: this would violate universal constraints on possible extended projections, --the universal (domination) order is (27) by hypothesis.





(27) F*... >...F... > ...V

To summarize: Independently of mirror theoretical considerations, all theories of right-peripheral adverbials that assume that the VP forms a chain with a position on the right of these adverb(s) are problematic. In particular the solution that assumes VP shift to a spec position raises the issue of the unmotivated nature of the postulated host head and the questions of why, no successive non-incremental step VP-chains can be constructed and what makes successive incremental chains possible. The possibility of successive roll up chains and the impossibility of non-incremental successive step structure are properties that are not shared by better established "phrasal" chains. If the VP originates in a spec position on the left of the right peripheral adverb(s) then the universal order of functional heads must be defined on a series of nodes where each

immediately dominates the next, but without the additional requirement that each must be a complement of the previous one. This is exactly the conclusion that the strict version of the mirror theory in chapter 2 lead to. The behavior of right peripheral adverbials therefore provides further evidence for mirror theory, and in particular for its most principled variant.

4.3. V-raising in Hungarian

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 involve syntactic spec positions and head-chain type relations. Recall that under mirror theory chains that involve (directly, or if the suggestions in section 3.5 above are correct, indirectly) syntactic spec positions are the equivalent of standard XP-chains. These are the only type of chains under this theory, head-chain type relations correspond to 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.FN6

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.

- (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 (ie. 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 focussed element (capitalized) in the preverbal focus position characteristic of Hungarian and the VM "szét" ("apart") appears to incorporate into the 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. As noted in the previous section in analyses with successive cascading XP intraposition, 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 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 (ie. 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. FN7

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:

- (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:FN8

- (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 not just don't 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 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 lowest infinitive the MW or (syntactic-spec targeting) chain was constructed to satisfy the requirement of the VM, ie. (36). Given some version of last resort FN9, 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 (2) 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, focussing 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 focussing 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 focussing involves a focus head, F. Phrasal focussing in Hungarian always involves contrast or at least an "exhaustive listing" sense, hence it is natural to assume that 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, ie. to form the MW: "akar-->ni-->fog-->ok-->F" linking them. FN10

(38) and (39) differ in that in the former the highest member of the string of infinitives has been focussed, 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 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 MWs follows from the basic assumption of 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.

(40)

@

T(-om)

szet

fog-

@

-ni

akar-

@

-ni

kezd(e) -

@

-ni

szed-

(szet)

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, in standard terms, 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 +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 preceeding 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", ie. 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") , ie. 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, ie. 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. FN11

(45) @

F

MOST T(-om)

fog-

@

-ni

akar-

-ni

kezd(e)-

-ni

szed-

szet

Footnotes to Chapter 4.

FN1 As the argument crucially uses the presence of the complement C, the derivation of the lack of word-dominated phrases needs the assumption, as Kayne notes, that the source for the phrase P must be internal to the complement C. Ie, if there is no complement then there is no source for the phrase. This might still be problematic if empty categories, hence empty complements do not need to be ordered by the LCA.

FN2 On these see Kayne 1994 section 3.7.

FN3 The strength of this argument is weakened by the apparent existence of another reason for the difference. As Phillips 1996 points out there seems to be a preference for fronted VPs to be potentially complete. He gives examples like the following (the goal argument is optional with "give" but obligatory with "hand"):

- (i)?... and [give candy] he did to the children on weekends
- (ii)*... and [hand candy] he did to the children on weekends

FN4 Another issue (raised eg. by Manzini 1995 against the family of solutions I am considering here) that we can also put

aside here is that of extraction.

(i) Who did you say that we sent there in order to please
John's mother

(ii)?*"Who did you say that we sent him there in order to
please t

Extraction on this analysis must be possible from a left branch
as in the case of (1) but not from the main branch (ii). But at
least the former is a general problem in the antisymmetry
framework, not directly linked to the analysis of (1). Indeed the
problem arises in the minimalist framework even without
antisymmetry given object shift to AgrO. In these approaches
extraction from left branches will often be necessary creating
problems for connectedness (Kayne 1983, Manzini 1992) type
approaches to locality.

FN5 Cinque in fact assumes also another, VP internal,
source for "circumstantial" adverbials.

FN6 Most Hungarian examples that follow are from Szabolcsi
1996 and from references cited there.

FN7 It is not clear if the tensed deficient verb will have

the requirement (34) for VMs when there is a focussed 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 structures 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 spelled 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).

FN8 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) (ie. "will begin to want to...") is not available. The restriction does not appear to be simply that only infinitives without

complements can serve as VMs: infinitives with noninfinitive
complements can be spelled 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

FN9 Recall from chapter 1 that last resort does not need to
be a derivational condition. Translated to 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/ in syntactic structures) are licensed by some (checking
or thematic) relation.

FN10 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 hívott fel
John's friend called up

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

FN11 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 focussing structures in (ibcd) 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, eg. 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.

References

- Akmajian, Akmajian and Thomas Wasow. 1975. The constituent structure of VP and AUX and the position of the verb BE. Linguistic Analysis 1:205-245.
- Andrews, Avery. 1983. A note on the constituent structure of modifiers. Linguistic Inquiry 13:313-317.
- Arad, Maya. 1996. A bi-directional view of the syntax-lexicon interface. Ms. UCL.
- Baker, Mark. 1988. Incorporation, a theory of grammatical function changing. Chicago: University of Chicago Press.
- Barbiers, Sjeff. 1995. Extraposition and the interpretation of X-bar structure. GLOW newsletter 34.
- Barrs, Andrew and Howard Lasnik. 1986. A note on anaphora and double objects. Linguistic Inquiry 17:347-54
- Beghelli, Filippo and Tim Stowell. 1996. Distributivity and negation: the syntax of "each" and "every". in Anna Szabolcsi ed. Ways of scope taking. Kluwer, Dordrecht.
- Bobaljik, Jonathan. 1995. Morphosyntax: the syntax of verbal inflection. Doctoral dissertation. MIT.
- Borer, Hagit. 1993. The projection of arguments. in University of Massachusetts, Occasional Papers in Linguistics 17, GLSA Amherst, MA.

- Brody, Michael. 1980. On circular readings. The Linguistic Review Vol.2.
- Brody, Michael. 1990. Some Remarks on the Focus Field in Hungarian UCL Working Papers in Linguistics Vol.2, University College London.
- Brody, Michael. 1994. Phrase Structure and Dependence. Ms., University College London.
- Brody, Michael. 1995a Lexico-Logical Form: A radically minimalist theory. Cambridge, Mass.:MIT Press.
- Brody, Michael. 1995b. Hungarian focus and bare checking theory. In Arbeitspapiere des Sonderforschungsbereichs 340, University of Tübingen
- Brody, Michael. 1997a. Towards Perfect Chains. in Liliane Haegeman ed. Handbook of Syntax , Kluwer.
- Brody, Michael. 1997b. Mirror Theory in UCL Working Papers in Linguistics Vol.9, University College London.
- Brody, Michael. to appear. Phrase Structure and projection. Ms., University College London, to appear in Linguistic Inquiry 29.
- Chomsky, Noam. 1981. Lectures on Government and Binding. Dordrecht: Foris.
- Chomsky, Noam. 1986. Barriers. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1995. The Minimalist program. Cambridge, Mass.: MIT Press.

Chomsky Noam and Howard Lasnik. 1978 Filters and Control.

Linguistic Inquiry

Cinque, Guglielmo. 1997. Adverbs and functional heads, a cross linguistic perspective. Ms. University of Venice.

Cormack Annabel and Neil Smith. 1996. Checking features and split signs. Ms. University College London.

Dayal Veneeta. 1994. Scope marking as indirect wh-dependency. Natural Language Semantics. 2.2:137-170.

Dayal Veneeta. 1996. Scope marking: in defense of indirect dependency. in Lutz Uli and Gereon Muller eds.

Epstein, Samuel. 1995. Un-Principled Syntax and the derivation of syntactic relations. Ms. Harvard University.

Groat, Erich. 1995. On the redundancy of syntactic representations. GLOW abstract. Tromso, Norway.

Haider, Hubert. 1993. Detached clauses - the later the deeper. Ms. University of Stuttgart.

Heycock, Caroline. 1995. Asymmetries in reconstruction. Linguistic Inquiry 26:547-570.

Higginbotham, James and Robert May. 1981. Questions, quantifiers and crossing. The Linguistic Review 1:41-80.

Hornstein, Norbert. 1994. LF: The grammar of logical form. From GB to minimalism. Ms., University of Maryland.

Horvath, Julia. 1995. Partial wh-movement and wh "scope-markers". in Kenesei, Istvan ed. Approaches to Hungarian. Vol.5, JATE,

- Szeged.
- Hudson, Richard. 1990. English Word Grammar. Blackwell.
- Johnson, David and Shalom Lappin. 1997. A critique of the minimalist program. *Linguistics and Philosophy* 20: 272-333.
- Kayne, Richard. 1994. *The asymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Koizumi, Masatoshi. 1993. Object Agreement Phrases and the Split VP Hypothesis In MITWPL vol.18. MIT, Cambridge, Mass.
- Koopman, Hilda. 1994. Licensing heads. in. Lightfoot David and Norbert Hornstein eds. Verb movement. CUP.
- Koopman, Hilda. 1996. The spec head configuration. Ms., UCLA.
- Koopman, Hilda and Anna Szabolcsi. 1997. The Hungarian verbal complex. Ms. UCLA.
- Kroch, Anthony. 1989. Amount quantification, referentiality and long wh-movement. Ms., University of Pennsylvania, Philadelphia.
- Larson, Richard. 1988. On the double object construction. Linguistic Inquiry 19:335-391.
- Lasnik, Howard. 1994. Verbal morphology: Syntactic structures meets the minimalist program. Ms. University of Connecticut.
- Lebeaux, David. 1989. Relative clauses, licensing and the nature of the derivation. Ms., University of Maryland.
- Li, Yafei. 1997. An optimized universal grammar and biological redundancies. Linguistic Inquiry 28:170-178.

- Lutz, Uli and Gereon Muller eds. 1996. Papers on wh-scope marking. Arbeitspapiere des Sonderforschungsbereichs 340, University of Tübingen.
- Mahajan, Anoop. 1996. Wh-expletives and the syntax of partial wh-movement. in Lutz Uli and Gereon Muller eds.
- Manzini, Maria Rita. 1992. Locality. Cambridge, Mass.:MIT Press.
- Manzini, Maria Rita. 1995. Adjuncts and the theory of phrase structure. Ms. University College London.
- Pesetsky, David. 1989. Language particular processes and the earliness principle. Ms., MIT.
- Phillips, Colin.1996. Order and Structure. Doctoral dissertation. MIT.
- 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.
- Roberts, Ian. 1996. "Have/be"-raising, move and procrastinate. Ms. University of Stuttgart.
- Sportiche, Dominique. 1994. Adjuncts and Adjunction. ms. UCLA.
- Starke, Michal. 1994. On the format for small clauses. Ms.

- University of Geneva.
- Sternefeld, Wolfgang. 1996. Comparing reference sets. Seminar for Sprachwissenschaft report, University of Tübingen.
- Szabolcsi, Anna. 1996. Verb and particle movement in Hungarian. Ms. UCLA.
- Williams, Edwin. 1994. Thematic Structure in Syntax. Cambridge, Mass.:MIT Press.