

Projection and Phrase Structure

Michael Brody

Perfect Syntax dispenses with the idea of externally forced imperfections in syntax. This article presents a system of principles relating (L)LF representations and lexical items that aims to be compatible with this assumption. The core of this theory is that phrase structures are viewed as projection lines (lexical items and their projections) linked by an Insert relation. This explains uniqueness and locality of projection, the fact that phrases and nonphrasal elements can immediately dominate each other only when they are part of the same projection line, and most effects of the ‘target projects’ requirement. I attribute a residue to the Generalized Projection Principle, for which I also provide an explanation. In addition, I explore various consequences of the present approach for the Move/Chain relation.

Keywords: adjunction, expletive-associate chains, feature movement, Generalized Projection Principle, perfection, phrase structure, pied-piping, projection, representational-derivational issue, ‘target projects’ condition

1 Perfect Syntax

Consider a rather standard system of grammar in which the relationship between meaning and sound is mediated by two interpretive systems applying to some interface representation(s) generated by syntax. Suppose that these apply to the same representation (say, the level Lexico-Logical Form (LLF) of Brody 1995a), clearly a desirable additional assumption. Such a view takes the Spell-Out component of the minimalist framework to correspond in the relevant respect to the post-LF semantic interpretation processes. The syntactic computation can then be viewed as having the task of determining certain basic properties of this interface representation, in particular those having to do with the relation between such representations and the elements provided by the lexicon, from which these are composed. Syntax generates (L)LF representations, either in a representational mode, putting interface constraints on possible relations between lexical items and (L)LF structures, or in a derivational idiom, assembling these structures from lexical items.¹

The main ideas of this article were presented in Brody 1994, read at the GLOW Colloquium in Vienna. The article itself was then written in its present form in the spring of 1995. Various versions were presented between 1994 and 1996 at University College London, MIT, UCLA, University of Massachusetts at Amherst, at the Universities of Vienna, Tübingen, Stuttgart, Potsdam, Budapest, Florence, and Venice, at the LOT Summerschool in Amsterdam, and at the SICOGG conference in Seoul. I am grateful to the audiences at these universities and conferences, to the *LI* reviewers, and particularly to Michal Starke for detailed comments and helpful conversations. For a theory that develops further the major themes discussed here, see Brody 1997.

¹ In Brody 1995a I argued against theories where central concepts are duplicated by being captured both representationally and derivationally. The choice between a purely representational and a purely derivational theory is of course a more difficult matter. Note that we cannot in principle exclude mixed theories either, when these succeed in avoiding

Given this background, consider a strong version of the minimalist hypothesis according to which (L)LF interface conditions reduce to “bare output conditions,” that is, conditions forced on (L)LF representations by the interpretive systems applying to them. The conjunction of this hypothesis with the assumption that the relations between lexical items and (L)LF structures are stated as interface conditions entails that there is no syntax at all that is part of human grammatical competence.

An alternative hypothesis might hold that there exists a separate component of syntax, a set of conditions that do not follow from the nature of the interacting interpretive systems. These conditions, relating lexical items and (L)LF representations, could largely, and perhaps completely, follow from the necessity of relating lexically stored elements and the representations that contain them.

Suppose that syntax, in the sense just characterized, exists. There is then an empirical issue regarding the nature of this system, which relates lexical items creating the (L)LF representation. Optimally, this system should be nearly trivial: we would hope that apparent complexities are indeed only apparent and are due either to independently motivated properties of the interpretive components or to the interaction of these with the syntactic module. The general methodological point is further reinforced by the expectation that the properties of this system are consequences of the need to relate lexical and (L)LF representations. We might thus expect to find a system that is significantly more perfect than the assembly system of the standard minimalist framework.

Even if the Chain/Move relation is taken to be part of this system, there should be no *syntax-internal* conditions on it (e.g., uniformity, Minimal Link Condition, C-Command, Last Resort); see Brody 1995b and the discussion of uniformity below. Furthermore, there should be no representational-derivational duplications of (nearly) identical concepts (e.g., Chain and Move or the representational definitions of well-formed syntactic objects in addition to actual derivations); see Brody 1995a and section 6 below, respectively. A more restrictive framework such as this also eliminates the possibility of using representational-derivational distinctions like deletion (interface invisibility) versus erasure (invisibility for the syntactic computation) that build on such duplications. Additionally, we will expect to be able to dispense with economy conditions and the serious computational complexity that some of these create. Restrictions like these are consequences of the assumption that apparent imperfections in the system relating (L)LF to its alphabet, the lexicon, result from syntax-external considerations. Let us call the theory meeting them *Perfect Syntax (PS)*. In the light of recent advances in the minimalist framework, the apparently ambitious program of PS seems quite reasonable. (See Brody 1995b for more discussion of this approach.)

In this article I discuss a system that could be part of such a theory of PS, and I empirically justify some of its restrictive aspects. I provide a set of necessary conditions on categorial projec-

the unwanted redundancies and duplications. A rather strong consideration that argues against the existence of syntactic derivations is the lack of genuine feeding-bleeding relations in syntax. We would expect these to be commonplace if derivations involving successive steps existed (see Brody 1995a,b for related discussion). The versions of the theory I present in this article—a single-step degenerate derivational alternative and a representational alternative—both exclude feeding-bleeding relations in principle.

tions (i.e., a theory of phrase structure) in section 2. In section 3 I derive basic conditions of this system from a theory of the assembly of syntactic structures or, more neutrally, from a theory of the relation between lexical items and (L)LF representations. I compare certain salient aspects of this theory with the corresponding properties of the standard minimalist framework (especially in section 5), arguing that the theory defended here is not only simpler but also more adequate in other ways.

In section 4 I turn to explaining the Generalized Projection Principle, the condition that entails that categorial projection, like selectional requirements, must involve the root position of chains. I discuss an explanation that I argue (in section 6) is superior to other recent accounts that give only partial solutions and assume the accidental conspiracy of unrelated principles.

2 A Minimal Theory of Phrase Structure

2.1 *The Principle of Categorial Projection*

Phrases (XPs), words (X^0 s), and their heads (X^{\min} s) in an X-bar projection appear to share properties—for example, being a nominal or a verbal element. It is often assumed that the shared properties of the phrase and the word are inherited from the head: category labels arise through projection, ultimately from a lexical element dominated by the labeled categories. It seems that every phrase and more generally every non- X^{\min} category must share properties with some lexical X^{\min} head; there are no “pure” phrases.² In other words, phrases and non- X^{\min} words can only arise through projection.

(1) *Principle of Categorial Projection (PCP)*

Every (non- X^{\min}) word and phrase is projected by a category that it immediately dominates.

The step from “Some phrases and non- X^{\min} words are projected” to “All phrases and non- X^{\min} words are projected” seems highly natural, although obviously it is not logically necessary. But given Chomsky’s (1995) general condition of inclusiveness (which entails that the interface levels consist only of arrangements of lexical features), the additional assumption in (1) seems unavoidable: all phrases, like everything else at the interface, must consist of lexical features.

Let us understand the term *projected* in (1) as “is a partial copy of.” Take phrases and non- X^{\min} words to be copies of features of lexical items. Presumably non- X^{\min} words are copies of syntactic and morphological features whereas phrases are copies of syntactic features only. X^{\min} itself is a full copy of a lexical item. (It is not the lexical item itself since it was copied from the lexicon into the syntactic structure rather than “moved” there.) A non- X^{\min} word or a phrase is directly or indirectly projected by an X^{\min} ; it is a partial copy that immediately or nonimmediately dominates X^{\min} . The PCP requires a category (word or phrase) to be directly projected by (be a

² It may be possible to consider a phrase to be “pure” if the properties determined by its head are marked on and thus attributed to the head only. Thus, a category that subcategorizes for a DP, for example, would in effect subcategorize for a “pure” phrase headed by D. I will not explore this possibility here mainly because it appears to be incompatible with Chomsky’s (1995) principle of inclusiveness.

partial or full copy of) the element that it immediately dominates. Since sentences are finite structures, only a finite set of copies can be created by reapplying the copy relation in this way. It follows that ultimately all non- X^{\min} categories are copies of the X^{\min} head ('head' in the X-bar-theoretical sense) of the construction, which is itself a copy of a lexical item.³

It seems to be an advantage of the PCP that it relates the fact that phrases have nonphrasal heads and the fact that phrases and their (phrasal or nonphrasal) heads share features. Compare the PCP with Kayne's (1994) approach, where the Linear Correspondence Axiom (LCA) entails that in certain configurations phrases have nonphrasal heads,⁴ but does not entail that phrases and their heads share properties.

That the PCP is stated in terms of immediate domination ensures the locality of the projection relation. We can state this requirement separately as in (2).

(2) *Locality*

If X^{\min} directly or indirectly projects an X^0 or an XP, then there is no category C such that X^0/XP dominates C, C dominates X^{\min} , and C is not a projection of X^{\min} .

(2) excludes configurations like (3a). (3b), where the lower YP may be interpreted as an intermediate-level projection or as a segment of adjunction, exemplifies (3a).

- (3) a. [_{XP}[_C X^{\min}]]
 b. [_{XP}[_{YP} X^{\min} [_{YP} Y^{\min}]]]

Locality is again not ensured by the LCA: there is an asymmetric c-command relation between X^{\min} and Y^{\min} in (3b). This orders the terminals as required by this condition.

As in Chomsky 1995, we can consider a phrase maximal if it is not immediately dominated by its copy, nonmaximal otherwise. Could the whole theory of phrase structure reduce to the PCP? It is clear that this well-formedness condition does not suffice; additional assumptions are necessary.

First of all we need to distinguish words (X^0 s) and phrases (XPs). It seems incorrect to consider all X^0 s to be lexical items given open-class incorporation phenomena showing syntactic movement properties. Let us proceed differently. Call the object headed by an X^{\min} and licensed by the PCP (i.e., the X^{\min} , the copy of X^{\min} , the copy of the copy, etc.—each related by the immediate domination relation to the next) a *projection line* (PL). Assume that there is a word-phrase boundary on the PL at some random point. Some corresponding additional assumption distinguishing words and phrases is also needed in the framework of Chomsky 1995, at least if we do not make the implausible move of considering all X^0 s to be lexical items. The relational

³ The projection or copy relation involved in the PCP is a two-place relation, and I assume that an element can be copied only once. In other words, there cannot be two distinct projections of some category C such that both dominate C but neither dominates the other; that is, there can be no 'upward-branching' projections. (Note here that we cannot attribute the exclusion of upward branching in general to Kayne's (1994) Linear Correspondence Axiom (LCA). For example, a terminal element immediately dominated by two distinct categories will not violate the LCA.)

⁴ A phrase that immediately dominates no nonphrasal element and at least two phrases that contain terminal elements violates the LCA. A headless phrase that either contains nothing, or contains a single phrase of a different type, or contains two or more phrases that in turn dominate nothing does not seem to violate the LCA.

definitions of this theory classify both a non- X^{\min} X^0 and a nonmaximal XP as intermediate projections. (They are both neither maximal phrases nor lexical items.) But they apparently must be distinguished somehow. To avoid violating inclusivity, let us assume that the word-phrase boundary is marked by the fact that words but not phrases carry morphological features. (Note that I do not adopt the view that a category may be classified in more than one way with respect to its projectional status; see section 5.)

Second, it must be ensured also that all and only non-word-internal heads project a phrase. Let us call this the *extended structure preservation restriction*.

- (4) *Extended structure preservation restriction*
- a. Every non-word-internal head projects some phrase.
 - b. No word-internal head projects a phrase.

Chomsky (1995) assumes that (4b) is due to a morphological condition: morphology does not tolerate phrases. Adopting a relational definition of projection levels, he rejects the assumption in (4a) and assumes instead that a non-word-internal head that has not projected is both minimal and maximal. Such elements can thus occupy specifier, complement, and X^{\max} -adjoined positions, which are reserved for maximal projections. He then rules out the presence of a ‘moved’ nonroot X^{\min} or X^0 in such positions using the principle of uniformity, which disallows chains whose members do not all have all the same projection-level characteristics. However, such an approach to (4) fails to capture the suggestive symmetry of this condition. (The same is true of Kayne’s (1994) LCA account of (4).) I shall provide additional arguments against the uniformity condition in section 5.

Third, the uniqueness of the relation between a phrase and a head needs to be ensured—say, as in (5).

- (5) *Uniqueness*
- Every phrase is projected by a unique category.

The uniqueness requirement ensures that a phrase cannot be projected by two heads. Thus, (5) excludes cases like (6a–b).

- (6) a. $*[_{X/YP} X Y]$
 b. $*[_{X/ZP} Z [_{XP} X]]$

In (6a–b) X/YP indicates a phrase that both X and Y have projected, that is, a phrase that shares properties with both. Notice that the LCA predicts this result only for the special case when the two heads are both immediately dominated by the phrase, as in (6a). The LCA rules out this structure since contrary to its requirement there is no pair (C, C') of constituents related by asymmetric c-command such that C dominates X and C' dominates Y . (According to Kayne’s theory, the terminals dominated by X and Y will therefore violate the requirement that all terminals need to be ordered by an asymmetric c-command relation between categories dominating them.)

The LCA will remain silent, however, about cases where multiple categorial projection does not occur in a configuration where more than one head is immediately dominated by the offending

phrase. For example, it allows (6b), a head-complement structure, where XP is the complement of Z. (Z asymmetrically c-commands X in (6b), ordering the terminals appropriately as required by Kayne's condition.) The uniqueness requirement on projection thus does not follow from the LCA in general.

The PCP includes the locality requirement and also entails that every non- X^{\min} X^0 and XP have an X^{\min} head, namely, the one that ultimately projected it. Together with the uniqueness and the extended structure preservation conditions, the PCP also entails that every phrase must have a unique head.

- (7) *X, when X is not phrasal and is immediately dominated by a word-external projection of Y, Y^n , unless $X = Y^m$.

In other words, a nonphrasal X, X^{\min} or X^0 , cannot be the complement or specifier/adjunct of some other projecting head Y. We note first that this $X^{\min/0}$ cannot be properly word-internal since by hypothesis it is dominated by a word-external projection of Y. A non-word-internal $X^{\min/0}$ distinct from Y cannot be immediately dominated by Y^n , since if $X^{\min/0}$ is not properly word-internal, then it must project some phrase XP by extended structure preservation, and this XP can be shown to intervene between Y^n and $X^{\min/0}$. We know that Y^n cannot intervene between XP and $X^{\min/0}$ by the locality requirement in (2)/the PCP. Furthermore, XP is distinct from Y^n , the phrase projected by Y, by uniqueness. Hence, Y^n cannot immediately dominate $X^{\min/0}$ —it can only immediately dominate an XP, which may in turn dominate $X^{\min/0}$.

To sum up so far: The PCP expresses the idea that syntactic categorial structure is projected from the lexicon. The PCP states that all syntactic categories are related to the lexicon: they are partial copies of (full copies of) lexical items (i.e., of X^{\min} s). That phrases and words must have an X^{\min} head follows from the PCP, given the independently necessary assumption that PLs must be of finite length. The extended structure preservation requirement ensures that a not properly word-internal lexical element must and that a word-internal element cannot project a phrase. That a phrase must not have more than one head will follow from extended structure preservation and the PCP together with the (subsumed) locality and (additional) uniqueness assumptions.

Of course, like uniqueness and locality, extended structure preservation is so far only stipulated, and all three are in need of an explanation. Before going further in trying to understand why these conditions on phrase structure should hold (section 3), I would like to comment on several concepts that current theories generally assume, the discussion of which I have so far avoided.

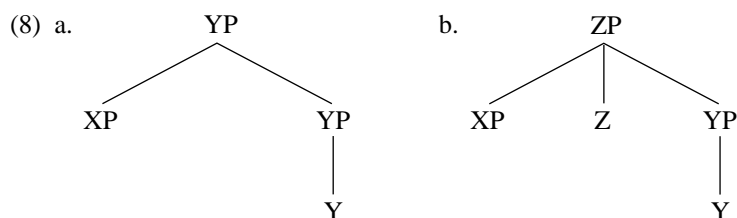
2.2 *Speculations on Adjunction and Intermediate Projections*

Notice first of all that the theory of phrase structure outlined in section 2.1 is neutral with respect to the question of binary branching and universal SVO order. A condition ensuring this (like the LCA) may or may not apply in addition to the PCP and related principles.

Current theories of phrase structure diverge in two major but related respects from a simple configuration where a phrase dominates a head and a number of other phrases. First, an intermediate X' level is assumed between the head and the phrasal node; and second, the configuration of

adjunction is allowed in addition. What is the status of adjunction and of intermediate projection levels given the theory of section 2.1?

These two configurations can be reduced to one if, as proposed by Kayne (1994), the intermediate X' level is treated as the lower segment of adjunction. It would be straightforward to graft a segment-category distinction, and with it a theory of adjunction, onto the theory of phrase structure as developed so far. But a simpler alternative might be to assume that there is no special adjunction configuration. Various arguments have been put forward that adjectives and adverbials, which have typically been treated as adjoined elements, must in fact occupy either the head or the specifier position of some higher projection (Sportiche 1994, Cinque 1993, 1995). Under this option, instead of left-adjunction of XP to YP as in (8a), we will have the configuration in (8b) with the higher head Z . Z may or may not be invisible and/or transparent for selection. (Selectional requirements may be satisfied here by the lower head Y .)



Consider next the question of intermediate projection levels. As noted, one possibility is to follow Kayne in treating the intermediate projection as a segment of adjunction. However, if adjunction does not exist, a different account is necessary. The PCP, as it stands, allows a PL to be composed of more than one phrasal node. A phrase does not have to immediately dominate the X^{\min} (or X^0) element of the PL that (indirectly) projected it; they can be separated by phrasal (and also nonphrasal) nodes of the same type. The system outlined above allows nonmaximal projections.

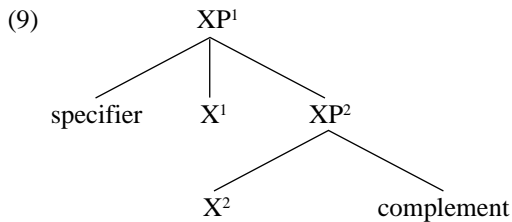
But let us consider an alternative theory. It is widely assumed that intermediate phrasal projections are not visible for the grammar; in particular, they cannot participate in chain formation. If correct, this fact could best be explained if intermediate phrasal projections did not exist at all. Let us suppose that they do not. A word can then project only a single phrase; at the phrasal level the PCP cannot reapply.

The corresponding restriction may be justified in the word-internal domain as well, given Kayne's (1994) arguments that only a single element can adjoin to a word. This leads to the assumption that a lexical X^{\min} can project only a single X^0 and a single XP . Suppose that this is correct. The PL of a lexical item X^{\min} may then consist of (a) only X^{\min} , (b) X^{\min} immediately dominated by X^0 or XP , or (c) X^{\min} immediately dominated by X^0 immediately dominated by XP . The relational definition distinguishing maximal and nonmaximal projections can be dispensed with; only the independently necessary distinction between words and phrases is needed.

The question then arises how specifiers and complements can be distinguished. For many cases the checking configuration will provide the answer: the specifier is the element that undergoes checking. This will need to be extended to specifiers of those projections that instantiate

adjunction in the impoverished system tentatively suggested above. But the specifiers of lexical categories may not participate in a checking relation with the lexical head. If so, then here a different solution is necessary.

We can differentiate specifiers and complements of lexical heads without postulating either adjunction structures or the existence of categories that are neither word-level nor maximal projections by an analysis partly in the spirit of Larson's (1988) work. Suppose that we take a phrase to consist of an internal XP that includes the head and its complements and an external XP shell that contains an empty head and the specifier or specifiers of X as in (9). The empty head X^1 and the lexical head X^2 are then taken to be related—suppose first as a head chain, (X^1, X^2) .



We could then take the specifier to be the sister of the higher head that does not contain the lower head; the complement(s) would be simply the sister(s) of the lower head. Notice that the tree in (9) is only partly Larsonian, since although it involves a higher shell, it is not binary branching.

The solution, as it stands, inherits a general problem of Larson's empty shell approach—namely, that it is incompatible with the Generalized Projection Principle (GPP; see Brody 1995a and section 4 below for discussion), which requires that the selectional properties of a head must be satisfied in the root position of its chain. In the case of (9), the problem is that the subject is not in the same phrase (XP^2) that contains the root position of the head chain. The specifier in (9) would therefore have to be selected from the position of X^1 , not the root position X^2 of the (X^1, X^2) chain. Furthermore, the higher head X^1 projects an XP, again in spite of not being in the root position of its chain, in violation of the “target projects” requirement.

One possible solution is to assume that the higher head creating the “empty shell” is in fact not empty but is itself an abstract lexical element, one that carries the appropriate categorial features and selectional requirements of the lexical item whose features are shared between a number of head positions. (This modification of Larson's approach is suggested in Brody 1993a, 1995a; see also Koizumi 1993, Collins and Thráinsson 1993, and Chomsky 1994, 1995 for similar proposals and additional argument.) Multiple-argument verbs under a Larsonian analysis would all require such a decomposition treatment.

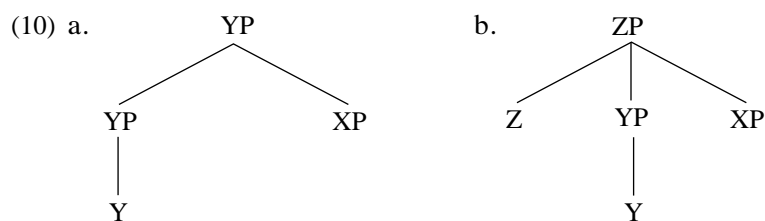
Let us apply this analysis to the present problem of eliminating the intermediate X' level in terms of a structure like (9). If X is decomposed into X^1 and X^2 , then the head chain that relates X^1 and X^2 would not be (X^1, X^2) , but presumably a (covert) (X^2, X^2) chain, where the higher member is in the word-internal checking domain of X^1 . If categories standardly taken as sisters of X' and sisters of X are distinguished as sisters of X^1 and sisters of X^2 , then simple transitive and intransitive heads must also decompose into two heads. For example, the verb *kick* would

have to be composed of an agent-selecting segment and a nonagentive KICK, something like the passive *was kicked*.⁵

It is interesting to observe that ternary structures like (8b) and (9) are compatible with the LCA if (a) instead of adjunction segments (intermediate projections) the complements themselves are taken not to be able to c-command and (b) the LCA is taken to apply only to nonempty terminals as proposed by Chomsky (1995) and Koopman (1996). (Perhaps (a) can be eliminated if there are no nonempty complements as suggested by Koopman.) Such an LCA would entail that in (8b), where the adjunct XP in specifier position is overt, the head Z must be empty. Similarly in (9), if the specifier is overt, then a phonologically realized head must appear in the lower X² position.

On the other hand, as we have seen, the LCA would create some redundancy with the phrase structure principles motivated above. Additional problems arise from structures usually analyzed as involving right-adjunction. This configuration cannot exist in a strictly binary-branching theory, where complements of embedded heads correspond to right-adjoined elements. As is well known, various tests suggest strongly that right-adjoined constituents are in fact higher than a general condition like the LCA allows them to be (see, e.g., Williams 1994, Brody 1994). Without the LCA these problems will not arise.

A treatment of right-adjunction compatible with the present framework might view the element A adjoined to constituent B as an additional complement of a *higher* head (rather than of a lower one as in the binary-branching account) whose preceding complement is B. Instead of structures like (10a), this treatment will produce ones like (10b).



As in the case of ‘‘left-adjunction,’’ the higher head Z may be invisible and transparent for selection.

If structures like (10b) exist, then the LCA cannot be a general condition on syntactic structures. The major residual effect of the LCA that we may need to find an alternative account for will be the universal word order restriction. Suppose that the hypothesis of underlying universal specifier-head-complement order is correct. This follows from the LCA, given some additional

⁵ In Brody 1994 I raised an apparent problem for this approach to the question of the intermediate X¹ level: with heads that assign no θ -role to their subjects, specifier and complements could be distinguished only at the price of postulating a fully empty head. For example, *seem* would have to decompose into a higher head that does not select its subject and that does not appear to contribute in any other way and a lower one that is exactly like *seem*. However, the problem arises only if the expletive subject is generated VP-internally. If a verb like *seem* simply has no VP-internal subject, then the question of how such a subject can be distinguished from the complements will never arise, and such considerations will never make decomposition necessary.

assumption that ensures that universally the specifier asymmetrically *c*-commands the head, which in turn asymmetrically *c*-commands elements embedded in the complement. To ensure this, we could stipulate (11), for example.⁶

- (11) A word-external element in the minimal domain of a head *H* can asymmetrically *c*-command *H* iff it is the specifier of *H*.

In the framework put forward here, specifier, head, and complement are sisters; hence, (11) must be rejected or modified. To ensure the specifier-head-complement word order, we can retain (11) but substitute *precede* for *asymmetrically c-command*. The stipulation appears to be necessary in some form whether or not the LCA is adopted. Formulated in linear terms, it is also sufficient to ensure—to state, in effect—the universal word order restriction.

3 Assembly of Syntactic Structures

3.1 Chain, Project, and Insert

The discussion of phrasal projection in section 2.1 has raised several questions. I would like to show that a version of the theory regarding assembly of syntactic ((L)LF) structures proposed in earlier work (Brody 1995a) provides straightforward answers.

This theory postulated three operations: Project, Chain, and Insert. Some additional operation selects a lexical item (LI) from the lexicon and creates a copy (LIC) for syntactic use. *Project* creates another copy of a subset of the features of this copy, F(LIC), and establishes the relation “immediately dominates(F(LIC), LIC).” Project can apply only once to any category, and it can create only one copy in any application. As we have seen, this ensures that there are no “upward-branching” PLs. However, Project can reapply to the F(LIC) it creates, thus creating a further copy. Single or multiple applications of Project that involve copies of features of a given LI result in a PL: a set of copies, originally of an X^{\min} (the LIC created by the selection operation), each related to the next by immediate domination. If the speculations in section 2.2 are correct, then the longest possible PL consists of three categories: an X^{\min} LIC, an F(LIC) with morphological and syntactic features (X^0), and an F(LIC) with syntactic features only (XP).

Chain creates a copy of a PL (multiple copies in the case of chains with more than two elements). This is either a copy of the whole PL (XP chains are *always* of this type, and if there is no excorporation of incorporated material, then *only* XP chains are) or a copy of the lower “morphological” part of the PL, up to the word-phrase boundary (X^{\min} and X^0 chains).

Chain and Project create the (syntactic) *input list*—a concept different from but related to the notion of “numeration.” The input list consists of a set of PLs. (Notice that a chain is not a member of the input list, but members of the chain—that is, the PLs—are members of the input list.) Although the objects in the input list (PLs) can be complex, they all involve copies of a single X^{\min} ; they are all copies of a single lexical item. The input list thus can be taken as the normal form in which LIs are presented to syntax.

⁶ In (11) I take the specifier to be the element that undergoes checking for the case of the minimal domain of functional heads, and the “external argument” for the case of (decomposed parts of) lexical heads.

Insert then applies to PLs, the elements of the input list. *Insert* establishes immediate dominance relations between members of distinct PLs. Relations created by *Project* remain fixed, however; they cannot be modified by *Insert*. *Insert* establishes a relation between two PLs and, like *Project*, it can apply only once to any given PL: a PL can be inserted into one PL only. ‘‘Upward-branching’’ structures are therefore impossible both PL-internally and PL-externally.

For a concrete example consider (12a), with the simplified structure in (12b).

- (12) a. Marie embrasse Pierre.
 Marie kisses Pierre
 b. $[_{IP} NP^{1'} V' + I [_{VP} (NP^1) (V) NP^2]$
- (13) a. Project: $NP^1 > N^1$, $NP^2 > N^2$, $VP > V$, $IP > I^* > I$
 b. Chain: $NP^{1'} > N^{1'}$, V'
 c. Insert all
- (14) a. V' , $VP > V$, $NP^2 > N^2$, $NP^{1'} > N^{1'}$, $NP^1 > N^1$, $IP > I^* > I$
 b. $IP > NP^{1'}$, $I^* > V'$, $IP > VP$, $VP > NP^1$, $VP > NP^2$
 (where $X > Y$ means ‘X immediately dominates Y’)

(Superscripts and stars in (12) through (14) are meant only as presentational aids.) *Project* creates the PLs in (13a), and *Chain* adds two more as in (13b): the higher members of the $(V', V) X^{\min/0}$ chain and of the phrasal XP chain ($NP^{1'} > N^{1'}$, $NP^1 > N^1$). *Chain* and *Project* in (13a–b) create the input list, an unordered set of PLs, shown in (14a). Finally *Insert* (13c) applies, relating elements in the input list by simultaneously establishing the further immediate dominance relations in (14b).

The theory is built on two core concepts: the concept of copy and the structural concept of immediate domination. Both concepts are involved in projection: a projection of an element is a (partial) copy that immediately dominates this element. Only the notion of copy is involved in lexical item selection and *Chain*, and only the notion of immediate domination is involved in *Insert*. As discussed in Brody 1995a, a major advantage of such a system is that the structure is built in one step; there are no intermediate syntactic structures (i.e., no structures distinct from LF where lexical items are related to each other). Notice that although the input list consists of structured objects (PLs), it is not a syntactic structure: all members of the input list and all (immediate dominance and copy) relations involve only a single lexical item. The theory is thus able to explain the basic minimalist generalization that no conditions can hold on noninterface structures: the generalization holds because noninterface structures do not exist.⁷

⁷ It is interesting to note that ordering *Project* so that it always precedes *Chain* would capture the ‘‘target projects’’ generalization. If *Project* cannot apply to copies created by *Chain*, then a ‘‘moved’’ category cannot provide a ‘‘label’’ for a higher node. We could order *Chain* and *Project* by assuming that the former but not the latter is a strictly lexical operation. *Project* would then create lexicon-internally the form in which lexical entries are presented to syntax (i.e., PLs), whereas *Chain* would be an operation between the lexicon and the input list (in effect, an aspect of the selection operation). An open question would remain in this approach, however: how to ensure, nonstipulatively, that copies of a PL are always inserted higher in the tree than the original PL, which would have to correspond to the most deeply embedded ‘‘trace.’’ In section 4 I develop a different approach.

A question that is only partly specific to this framework has to do with the notion of copy. Since XP chains are formed by copying a PL with a phrasal element, there must be a nondistinctness requirement on copies in chains to ensure that the same argument and selectional structure is inserted in all copies/members of this PL. (The problem arises also with non- X^{\min} X^0 chains.) For example, we need to ensure that the Principle C violation indicated in (15) can be ruled out at or beyond LF. This cannot be done if the chain member in the lower (bracketed) position is simply the PL projected by the highest head of the antecedent: that is, the PL $DP > D^{\min}$ projected by *which* in (15a) and the PL $PP > P^{\min}$ projected by *to* in (15b).

- (15) a. *which claim that John_x was asleep did he_x deny (which claim that John_x was asleep)
 b. *to John_x he_x gave a snake (to John_x)

The information that two (or more) elements are copies of each other must be available to the post-LF interpretive systems: copies that are members of the same chain must be distinguished at least at LF from accidentally identical categories that are not so related. Suppose, then, that structures in which two copies/chain members dominate distinct elements cannot be interpreted—in other words, that these are not proper copies. This nondistinctness condition will thus constrain the selectional requirements, and therefore both chain members in both (15a) and (15b) will dominate the same elements.⁸ Take for example (15b). Here the chain is formed on the PL that was projected by the preposition *to*—that is, on $PP > P^{\min}$. The preposition in both copies selects a complement that must be the same in both copies by the nondistinctness condition. The same account holds for (15a), with the selectional requirements of heads applying recursively. The head of the highest DP in the copy selects an NP, which must have been projected by the noun *claim*, given nondistinctness. This noun then selects a CP; and so on.⁹

3.2 Predictions for Phrase Structure

Let us return to the questions raised by the theory of phrase structure set out in section 2.

- (16) a. Why do all and only not properly word-internal categories project a phrase (extended structure preservation)?
 b. Why is projection local (no “foreign” elements may intervene between X^{\min} and any category it directly or indirectly projects)?
 c. Why is projection unique (each category is projected by a unique head)?

(16b) and (16c) receive an immediate answer, given the above theory of (L)LF assembly. Project applies before the syntactic structure is created (by Insert), and it applies separately to each LIC.

⁸ The condition probably does not require full identity. In principle, it allows elements like adjuncts to be missing from one chain member that are present in another (see Brody 1995a and section 3.3). Notice that an element may only be missing from “traces” and not from the head of the chain. Presumably this is due to a recoverability requirement.

⁹ In Brody 1994 I took *Near John he saw a snake* to be grammatical on the coreferential reading. If this reading is in fact not better than (15b), then it needs no additional comment. If, however, it is significantly better, then we can attribute this improvement to the option of not chain-relating the adjunct PP to the IP-internal position—an option not available in the case of the selected PP.

Hence, two LIs cannot be involved in projecting a given phrase, and no ‘foreign’ projection can ever intervene between an X^{\min} head and its projections in the input list. Since Insert cannot modify the relations established by Project, the conclusion carries over to fully formed syntactic representations.

As for extended structure preservation, recall that the impossibility of word-internal phrases, (4b), has been attributed to the fact that morphology does not tolerate such constituents. The symmetry of the extended structure preservation condition in (4) suggests an extension of this restriction to (4a). Consider the claim that parallel to (4b), (4a) holds because syntax does not tolerate nonphrasal elements. This would be an elegant modular solution, but the condition would clearly be incorrect: both phrases and nonphrasal elements (X^{\min} s and X^0 s/words) appear to play a role in syntax. But let us reconsider this idea in the context of the system of (L)LF assembly outlined above. The modular solution is made available here by the separation of Project, where words play a syntactic role, and Insert, where they do not. So let us assume that Insert is modular in the relevant sense.

- (17) Insert relates words to words (morphological application) and phrases to phrases (syntactic application).

(17) entails that all non-word-internal heads must project a phrase. If an X^{\min} projects only an X^0 and does not project a phrase (i.e., if the whole of the PL is subject to morphology), then only morphological Insert can apply to it; hence, it will be word-internal. It also follows from (17) that there can be no word-internal phrases; again, these could arise only if Insert were to combine words and phrases in a nonmodular fashion.

The theory of (L)LF assembly involving Chain, Project, and Insert was originally constructed as a system that can build syntactic structures in one step from input lists. Since it does not create intermediate syntactic structures, it explains their nonavailability. As I have just shown, the theory also explains three basic properties of the theory of phrase structure: extended structure preservation, uniqueness, and locality.

3.3 *F-Movement and Pied-Piping*

The account of chain formation proposed in Brody 1993a, 1995a, summarized and in some respects modified above, incorporates what is in effect a ‘pied-piping’ hypothesis. Both X^0 and XP chains involve linking elements (PLs) that were created (copied) from a single lexical item. ‘Pied-piping’ of the rest of the chain—that is, filling out all the copies with material additional to this highest head—occurs, as we have seen, because the selectional requirements apply recursively, subject to the nondistinctness condition on chain members.

Chomsky (1995) presents a different theory of movement and chain formation that shares the general idea of pied-piping with this account. He proposes that movement can only take place to establish a checking relation, and for this only a feature F needs to move. Movement of categories occurs only in the overt component of the grammar, and it occurs because F -movement pied-pipes the whole category. Such pied-piping in overt movement is forced by PF considerations.

(It is assumed in addition that certain features (formal features, FFs) are mechanically pied-piped in both overt and covert movement.)

For example, in *Whose book did you read?* the *wh*-feature must move to establish a checking relation with the corresponding feature on the C node. It must pied-pipe the word *who*; otherwise, the PF features of this word would be scattered at PF, a state of affairs naturally taken as resulting in an ill-formed representation. The genitive *'s* must also be pied-piped because of its affixal nature; thus, *whose* must move as a unit. But *whose* is not a syntactic object; it is not a constituent. Hence, the whole phrase *whose book* must move together.

Abstracting away from the difference between movement and chain formation (see Brody 1995a for discussion), we see that the two theories have much in common. Both assume that chains are formed on a single element of the head or phrase that ultimately is the member of the chain. In the account I proposed, this element is the head of the chain member; in Chomsky's account, it is the checking feature. The crucial difference appears to be that in the account defended here, pied-piping is a consequence of LF requirements, whereas in Chomsky's account, it is a consequence of PF conditions.

There are reasons to prefer the LF pied-piping approach. In the PF pied-piping theory, the question arises why pied-piping does not take place only in the Spell-Out component. Given a minimalist perspective, it is particularly difficult to understand why a PF requirement should force complications in the syntactic computation. But a theory according to which pied-piping takes place only in the Spell-Out component does not seem to be correct. The position of the "moved" phrase has syntactic and semantic effects. As an example, consider the contrast in (18).

- (18) a. Mary wondered which picture of herself John saw.
 b. *Mary wondered when John saw a/which picture of herself.

If anything beyond the *wh*-feature (or the formal features of the *wh*-word) remained in situ in syntax, then we would expect (18a) to behave syntactically and semantically in a way parallel to (18b). However, it does not. Thus, the contrast between (18a) and (18b) would be near-impossible to account for on the assumption that pied-piping takes place in the Spell-Out component. But this is the assumption that the PF-driven pied-piping hypothesis would naturally lead us to make.

On the other hand, there is evidence that pied-piping is LF-driven. The adjunct-argument asymmetry in reconstruction (Lebeaux 1989) follows from the nondistinctness requirement and projectional requirements discussed earlier.

- (19) which claim that John_x made did he_x later deny (which claim)
 (20) *which claim that John_x was asleep did he_x deny (which claim that John_x was asleep)

The Principle C violation in (19), where the relevant name, *John*, is inside an adjunct (the relative clause), is weaker than the one in (20) (= (15a)), where the name is inside a complement clause. As we have seen, selectional properties together with the nondistinctness requirement force the name in the complement to be present in the parenthesized copy in (20), and a Principle C violation results. In contrast, in (19) no selectional requirement forces the presence of the relative clause,

and the nondistinctness condition also allows its absence in the (parenthesized) copy. Hence, this sentence has a structure on which no Principle C configuration obtains. Clearly, the PF-triggered pied-piping account cannot capture such a distinction, but an appropriately constructed LF-triggered pied-piping account can.

Chomsky (1995) brings up another consideration: “The computation ‘looks at’ only F..., though it ‘sees’ more. The elementary procedure for determining the relevant features of the raised element α is another reflection of the strictly derivational approach to computation” (p. 269). For example, in (21) there is no question of determining where the *wh*-feature is located inside the complex *wh*-phrase *pictures of whose mother* since the computation looks at such features directly: pied-piping of the rest of the phrase is only an additional matter.

(21) Pictures of whose mother did you think were on the mantelpiece?

In reality, however, the elementary procedure for determining the relevant checking features is a property of the pied-piping theory. As we have seen, a representational pied-piping account is feasible (and also quite well motivated); hence, the question of derivationality does not seem relevant. Note, furthermore, that in any case the pied-piping account does not seem to achieve a genuine result here. This is because on this account also the relation between the XP (in (21) the *wh*-phrase) and the checking feature F (in (21) the *wh*) remains mysterious. This is of course true of both the LF- and the PF-triggered versions.

On the other hand, the PF-triggered pied-piping theory appears to create an additional problem within the standard minimalist framework in that it gives rise to a duplicate mechanism that appears conceptually and empirically unjustified. Consider a grammatical structure where movement without pied-piping has taken place. This could in principle be due not only to the covert nature of the movement but also, as Chomsky (1995:264) notes, to the failure of overt movement to pied-pipe, for whatever reason, as for example in Watanabe’s (1991) theory. There is no genuine evidence for making the theory more permissive in this way. See Brody 1995a for a critical discussion of Watanabe’s theory. (The problem is in fact more general, which strongly suggests that the covert component of syntax is superfluous. In the versions of the minimalist theory that allow the Spell-Out point to be distinct from (L)LF, empty categories can also be inserted both overtly and covertly.)

3.4 ‘Covert Movement’ Structures

Let us also consider briefly how overt and covert ‘movement’ structures can be distinguished in the present framework. The simplest assumption is that the distinction does not pertain to syntax at all, that it is only a matter of Spell-Out positions: in overt movement a higher copy, in covert movement a lower copy is subject to Spell-Out. It seems to me that in a framework that assumes that there are no covert \bar{A} -movement relations, there is little reason to depart from this simple hypothesis. If, however, there exist chains at LF corresponding to what used to be treated as covert \bar{A} -movement (see Brody 1995a,b), then the simple Spell-Out hypothesis will run into problems. For example, if the relation between the *wh*-in-situ and the [Spec, CP] where it is

interpreted is a chain relation, then [Spec, CP] must not contain a full copy of the *wh*-in-situ at LF.

- (22) a. John wondered which pictures of himself Mary bought (which pictures of himself)
 b. *John wondered which girl (which girl) bought which pictures of himself

If the [Spec, CP] of the embedded clause contained a full copy of the *wh*-in-situ *which pictures of himself*, then we would expect (22b) to be on a par with (22a); the anaphoric element should be appropriately bound by the matrix subject. But this is incorrect, strongly suggesting that the higher position in the chain of the *wh*-in-situ must not contain a copy. In earlier work (Brody 1993a, 1995a), I treated these structures in terms of what I called “expletive-associate chains.” Such chains expressed relations standardly treated in terms of LF movement. In expletive-associate chains the chain-forming associate always remains in situ and the higher positions in the chain are occupied, not by copies, but by an expletive element (or copies of this expletive). The expletive can carry features of the associate; this accounted for various “agreement” effects (e.g., checking of the *wh*-feature in covert *wh*-structures, or subject-verb agreement in *there*-associate structures).

Chomsky’s (1995) theory of covert movement as movement of only formal features (FF-movement) is very similar to this proposal. If we again abstract away from the representational/derivational difference, the major difference we find is that FF-movement is essentially head movement, whereas expletive-associate chains may be either head chains or XP chains. Without attempting to resolve the issue, I note that what evidence currently exists appears to favor the hypothesis that chains corresponding to covert movement relations can be phrasal.

- (23) There arrived three men.

Raising of FF(three men) in (23) to T violates the Head Movement Constraint, as Chomsky notes. The assumption that FF can be phrasal would avoid this problem. There are then two options: either FF is an additional specifier of T, or FF is identical to features of *there*. (The second option is equivalent to the expletive-associate chain solution modulo the representational/derivational issue.)

3.5 A Fully Representational Version

Given that the nondistinctness of chain members must be ensured by an interpretive condition, the Chain operation of the assembly system in section 3.1 appears to be partly redundant. The identity of the top PLs of chain members appears to be forced both by Chain and by the interpretive nondistinctness condition. But in fact the nondistinctness condition is not strong enough to ensure the identity of the top PLs of chain members. The nondistinctness condition ensures that the corresponding PLs in all chain members must be identical (i.e., based on the same LI). But as observed in section 3.3, it allows nonselected PLs (adjuncts) to be omitted. Now the highest PL in a chain member must be present in all chain members, but it is often not a selected element. Hence, the nondistinctness condition will not require this PL to be present in all chain members, and some additional principle that has the effect of Chain remains necessary.

A different problem with the assembly system proposed above is that the concept of “input list”—like that of “numeration”—is essentially a residue of D-Structure. If Chain could be eliminated from syntax, then the theory of (L)LF assembly could be stated as a fully representational condition along the lines of (24).

(24) *Structural licensing*

(L)LFs partition into PLs (quasi-lexical units). PLs are linked to each other by the modular immediate domination relation between their member categories.

The Chain operation (a) creates additional PLs, (b) links the additional PLs to other PLs created by Project, (c) ensures that the linked PLs involve the same LI, and (d) ensures that they contain an identical set of copies of the features of this LI. (The effect of the operation in (d) rules out nonuniform chains in which for example one member is an XP and another an X^0 .)

As for (a), Chain is of course not necessary to create additional PLs; this can be achieved by reselecting the relevant LI and, where necessary, reapplying Project. Linking elements that belong to the same chain (i.e. (b)) could be achieved by simply randomly marking elements at LF; in fact, Chain can be taken to reduce to this random marking. Notice though that both Chain and the LF random marking proposal still violate the inclusiveness principle according to which LF representations consist of nothing other than lexical features. A preferable option might be to assume that chains are not marked at LF and that they are constructed randomly, subject to nondistinctness and other conditions, by the interpretive principles applying to LF representations. This view will necessitate a global recoverability link between post-LF interpretation and PF, since the latter component needs the chain structure information to operate properly (e.g., for “trace-copy” deletion).

This leaves (c) and (d). As for (c), in addition to the nondistinctness condition that refers to chain-member-internal elements, we need (as noted above) a full identity requirement on the highest PL of the chain members. This full identity condition will ensure (d) as well. (As we will see in section 5, in the context of the proposed theory of phrase structure, it entails the major effects of the chain uniformity condition.)¹⁰

4 A Nonsyntactic Explanation of the Generalized Projection Principle

The discussion of categorial projection would remain incomplete without considering the Generalized Projection Principle (GPP), a major and pervasive condition, one effect of which is that categorial projection is restricted to root positions of chains. Although the existence of D-Structure as a distinct level of representation is quite dubious, there are few reasons to doubt the existence of the major generalization it expressed (Brody 1993b, 1995a; see also Chomsky 1993, 1994, 1995 for relevant discussion). This generalization, captured by the GPP, constrains the relation of chain members to their chain-external environment. It refers not only to categorial projection

¹⁰ The remarks in this section were prompted by several discussions of the derivational-representational issue with Michal Starke and by the comments of an *LI* reviewer.

but also to thematic selectional requirements, and in fact to syntactic and semantic selection in general. I assume therefore that the GPP is a principle of the interpretive component. All these requirements hold in the root positions of chains.

Thus, for example, a verb V^0 raised to some higher functional projection—say, C^0 —never projects a VP here: categorial projection holds only in the root positions of chains. Furthermore, a V in C never forces the specifier and the complements of C to satisfy the selectional requirements of V: selectional requirements hold only in chain roots. I argued in earlier work (Brody 1995a) that an appropriately formulated projection principle is both compatible with and necessary in a minimalist framework. I also attributed to the GPP the restriction against movement into a θ -position on the assumption that the GPP requires that selectional, including thematic, features not only must hold in but also must be satisfied by root positions. The statement of the GPP in (25) makes explicit that the principle refers to relations that link a chain member to elements external to its chain.¹¹

(25) *Generalized Projection Principle*

Projectional (categorial, thematic, selectional) features that link a member of chain C to its C-external environment must hold in and be satisfied by the root position of C.

In the rest of this section I would like to provide an explanation of the GPP, based on the account given in Brody 1995a. I will concentrate on selectional features first. Consider two chains that are to be related by a selectional feature F. It is natural to assume that F must identify all positions of the chain to which it is assigned. Assume, in addition, that all positions of the chain whose member assigns F must be marked as having assigned F. This second requirement is also natural in a framework like the present one, where chains consist of copies: all members of the chain of some X^0 element H are copies of H. Thus, if H has some selectional feature F, then all copies of H are naturally taken to have F, and in all these copies F must somehow be satisfied. In other words, I assume that the two chains will be properly related iff all members of both chains are appropriately identified as being related in this way.¹²

- (26) If a selectional (more generally, projectional) feature F of a member of chain C_1 selects (a member of) chain C_2 , then
- a. all members of C_2 must be identified as being selected by F and
 - b. all copies of F on members of C_1 must be identified as having been assigned.

Assume further that a projectional feature F, even if present on all chain members of C, can be assigned only from one position of C_1 and can be received only in one position of C_2 . Notice that projectional features behave differently in this respect from checking features. Unlike a

¹¹ Ignoring this aspect of the principle created complications in earlier versions of this article with respect to categorial projection. In an XP chain the XP is also projected in the nonroot position, apparently counterexemplifying a simple formulation of the GPP. But this XP is projected by a chain-member-internal element, the head of XP; hence, it is no more relevant to the GPP than the fact that the head of XP can also select its complement in the nonroot positions of this XP chain.

¹² Of course, either or both chains involved may be trivial, one-member chains.

projectional feature, a checking feature in a given chain can be multiply linked. Suppose finally that information can be transmitted in chains, or perhaps more generally in syntactic ((L)LF) representations, only bottom to top, from a position P to positions c-commanding P. It now follows from this general principle of information transmission that projectional features must link root positions of chains, as required by the GPP; if they linked a higher position, then the information that linking has been successful could not reach chain members that are located lower.

A particular implementation of this approach could take the interpretive process of information transmission to correspond to feature percolation. Let us say that an assignee position is selectionally identified if it has the appropriate selectional feature F, and the assigner position is selectionally identified if it has some feature S indicating that proper assignment¹³ has taken place. Like all features, F and S can percolate only to c-commanding positions. It follows, then, that the selectional feature F must be assigned to the most deeply embedded position in the assignee chain; otherwise, lower positions in this chain will not be selectionally identified. Similarly, F must be assigned from the most deeply embedded position of the assigner chain; otherwise, the feature S indicating that the selectional requirement F has been satisfied cannot percolate to all members of the assigner chain. All members (copies) of the assigner chain carry the selectional feature, which can only be satisfied through percolation of S under the assumption that a selectional feature can only be assigned¹⁴ once in any given chain.

Chains in which a nonroot position is selected (including ‘‘movement’’ to θ -positions) are now impossible: the selectional feature cannot percolate to the lower position of the chain, which thus fails to be identified. Conversely, no selection can take place from a nonroot position either. For example, a V raised to I or C now cannot select from the higher position of its chain since the information that this feature is satisfied could not reach the lower position of the chain.

The requirement that information transmission/feature percolation in chains is constrained by c-command corresponds to and extends the derivational principle excluding lowering applications of Move. In a framework that assumes the operation Move, a representation where the assignment of some projectional feature involves a nonroot chain position in violation of the GPP could have arisen in two ways: either through raising, in violation of the derivational equivalent of the GPP prohibiting movement into a position that involves selectional features, or through lowering from this position. Downward percolation of the selectional features corresponds to lowering in a system incorporating Move. This needs to be excluded in both frameworks. If it is, then in the present theory the GPP reduces to the principle in (26), namely, that all positions in a chain need to be selectionally (projectionally) identified. Thus, although the GPP follows from fairly simple chain-theoretical assumptions once the equivalent of lowering is excluded, the same explanation could not be translated into derivational terms in a system that assumes the operation Move. Excluding lowering rules would not help to explain why raising into a selected position is impossible.

¹³ Or checking. The account is neutral between checking and assignment technology. This is not to say that projectional features behave similarly to features usually called checking features. The point is only that the differences between the two types of features do not force the distinction to be made in terms of checking versus assignment.

¹⁴ Or checked; see footnote 13.

Since feature percolation is restricted by *c*-command, the assumptions that entail the GPP also entail, as a side effect, the *c*-command requirement on chains. Consider first an XP chain in which not all members of the chain are ordered by *c*-command. Any element that does not *c*-command the root of the chain *C* will not receive the selectional feature(s) that are assigned to the root element; hence, the non-*c*-commanding nonroot member of *C* will violate the identification condition in (26a). Similarly, in the case of an X^0 chain, the selectional feature(s) of the element that does not *c*-command the root element will not be satisfied, violating (26b).

Let us return to the effects of the GPP for categorial projection. Recall that the “target projects” generalization states that categorial projection is invariably initiated in the root positions of assigner chains. Categorial projection is thus similar to selection in this respect. Given the assumptions made in section 2.2, only a residue of the “target projects” problem remains. Given an XP or X^0 chain, a nonroot XP or X^0 cannot be dominated by a further XP or X^0 projection since a PL cannot consist of more than one XP/ X^0 . Since only X^0 and XP chains exist, the only residue of the problem is the case where the nonroot element of an X^0 chain is illegitimately dominated by an XP that this X^0 projected further.

The explanation of the GPP will immediately exclude this configuration if it is generalized from selectional features to cover categorial features as well—that is, features involved in categorial projection. Selectional features appear to be inherited from X^{\min} categories by X^0 elements, which in turn assign these to other categories. Let us assume that categorial projection works similarly: the categorial feature inherited from X^{\min} by X^0 is assigned/projected by X^0 to the category immediately dominating it. Like selectional features, categorial features licensing phrases can then be assigned only in a unique position of a given chain; this must be the most deeply embedded position, given (26) and the assumption concerning the direction of (chain-internal) information transmission.

The similar behavior of selection and categorial projection suggests that we should distinguish both of these features from the features usually referred to as checking features. As mentioned in footnote 13, projectional features can also be treated in terms of the checking technology—the real distinction is a different one. But I shall continue to call the two types projectional and checking features. Both checking and projectional features relate chains, rather than categories, but they appear to do so in different ways. Checking features are properties of chains. If such a feature is checked, then it is automatically taken to be checked in all copies (chain members) in the chain, independently of whether the copies are in a higher or lower position. A natural way of capturing this is to say that the checking feature is a property of the chain in syntax, rather than a property of the category from which it originates. By contrast, although a projectional feature also relates chains, it does not become a chain property but remains the property of the category that is lexically specified to carry it.¹⁵

Consider finally an aspect of the interaction of the GPP with the theory of (L)LF relations presented earlier. Take a “moved” non-chain-root word-external head *H*. Given the modularity

¹⁵ The similarity between selectional features and categorial projection casts some doubt on the idea that categorial features can serve as formal checking features. As we have seen, there is some reason to think that the features participating in categorial projection behave like other nonformal semantic selectional properties and unlike formal checking features.

of the Insert relation, H must project a phrasal node HP; Insert can only combine a phrase with a phrase. There are two options to consider: either the projected phrase HP may be internal to the phrase into which H was inserted, or it may force the category label of this phrase to be HP. Both options must be excluded, and indeed the GPP excludes them both: projection is restricted to root positions. Thus, the GPP and the modularity of the Insert relation together ensure that “moved” non-chain-root heads must invariably be head-internal. There is no contradiction here between modular Insert and the GPP. These principles only create a contradiction for word-external heads in non-chain-root positions. Insert requires all word-external heads to project a phrase, and the GPP restricts all projection to originate in root positions of chains. Hence, word-external heads that are not in root positions can neither project nor not project: they cannot exist.

5 Uniformity, Relational Definition(s) of Projection Levels, and the X^0 /XP Distinction

The system of phrase structure defended here assumes a nonrelationally defined distinction between X^0 s and XPs—that is, between words and phrases. Recall that I take words and phrases to be inherently different: only the former carry morphological features. It might appear that this assumption is in addition to those that the phrase structure theory of the standard minimalist framework has to make, where projection levels are defined relationally. But as noted in section 2.1, this is not the case. Consider the relevant definitions from Chomsky 1994, 1995.¹⁶

- (27) a. A maximal projection (X^{\max}) is one that does not project further.
 b. Minimal projections (X^{\min}) are the lexical items themselves.
 c. Intermediate projections (X') are elements that are neither maximal nor minimal.

Given a substructure S that is an intermediate projection (i.e., neither an X^{\min} nor an X^{\max}), (27) does not specify whether S is an X^0 or an X' (i.e., a nonmaximal XP). In all standard frameworks, however, a distinction is necessary: syntax and morphology appear to treat X^0 s and phrases differently. To take a relevant case, X^0 s but not intermediate phrasal projections are generally assumed to be available for chain formation (movement).

Let us next turn to the uniformity condition on chains, (28).

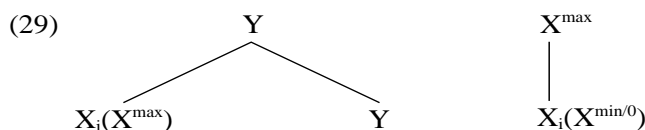
- (28) A chain is uniform with respect to phrase structure status.

Here the “phrase structure status” of an element is its (relational) property of being maximal, minimal, or neither. Since only X^0 and maximal XP projections are assumed to be accessible to the syntactic computational system, and hence for chain formation, the uniformity condition in (28) predicts that only (X^{\max} , X^{\max}) and (X^0 , X^0) chains exist.¹⁷

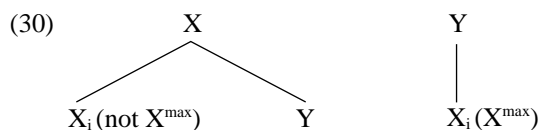
¹⁶ In fact, only the property of being a +/- maximal projection is defined relationally in (27)—being a lexical item is an inherent property.

¹⁷ In the richer system that assumes multiple adjunctions to a given X^0 , and thus multiple X^0 projections, parallel to the multiple XP projections created by multiple adjunctions/specifiers, probably X^0 chains should be restricted to maximal X^0 s, just as phrasal chains are restricted to maximal phrases. The problem would not arise in the present, more impoverished framework with a single X^0 and XP level for a given X^{\min} head. (An additional issue in both frameworks is whether X^{\min} chains exist.)

Given definitions like those in (27), it is easy to construct nonuniform chains. In (29) an $X^{\min/0}$ forms a chain with a copy that adjoined to or substituted into a word(X^0)-external Y^n , $Y \neq X$. Given (27a), this $X^{\min/0}$ will form a chain with an X^{\max} : here the first X_i of the chain is an X^{\max} and the second an $X^{\min/0}$.



In (30) a non-chain-root (“moved”) element X merges with some category Y , and then X (rather than the target of the operation) projects further, violating the “target projects” generalization and resulting in a nonuniform chain. In (30) the second X_i of the chain is an X^{\max} ; the first is not.



On usual assumptions, the structures in (29) and (30) are ill formed: minimal projections cannot move to (form chains with) positions that are not word(X^0)-internal, and it is always the target of movement that projects (only chain-root positions can project). The impossibility of configurations like (29) and (30) has been considered to provide evidence for the uniformity condition.

The behavior of clitics has been cited as additional empirical evidence for this system. According to the definitions in (27), a category can be both a minimal and a maximal projection: a lexical item that does not project further. Clitics appear to transparently instantiate this option, since they exhibit properties of both minimal and maximal projections. As X^{\min} s, they show up word-internally; but they seem to be linked to argument positions that are maximal. Furthermore, they can often form chains that ignore intervening heads, again suggesting (in the context of the Head Movement Constraint) that they are maximal. Thus, given the relational definitions in (27), clitics might be treated as both X^{\max} and $X^{\min/0}$ at the same time, accounting for their apparent dual nature.

The uniformity solution to the problems raised by the structures in (29) and (30) and by clitics faces numerous problems itself. Let us look at some of these before considering the alternative account that the present theory provides.

1. Ordinary X^0 head movement chains that target word-internal positions are nonuniform by the definitions in (27). Consider a chain where the root is a lexical item ($X^{\min/0}$) that projects further, whereas the nonroot word-internal member does not (hence X^{\max}). Such chains must be allowed, but they violate uniformity. Chomsky (1994, 1995) suggests that there is a special component “WI” at LF, where “independent word interpretation processes” apply. This then ensures that word(X^0)-internally the principles in (27) and (28) do not apply. WI is “a covert analogue of Morphology” (1995:322). But the reason for the existence of such a covert analogue of Morphology, and thus the status of WI, is unclear. Given the lack of evidence for such an

additional module, the WI hypothesis appears to amount to little more than a statement that head movement targeting a word-internal position is exempt from the uniformity requirement. But if so, then we cannot say that uniformity explains the impossibility of head chains like (29) where a nonroot member is word-external. The crucial distinction between good and bad cases here is the word-external versus word-internal contrast. Uniformity says nothing about this distinction, which is simply stipulated.¹⁸

2. If relational definitions do not apply word-internally, then the evidence clitics appeared to provide for them (namely, that clitics appear to have a different projectional-level status word-internally and word-externally) disappears. Since grammatical cases of head movement make it necessary to exempt word-internal structure from these definitions, the word-internal status of clitics becomes irrelevant and thus cannot be used to support the system.

3. The assumption that the relational definitions do not apply word-internally creates further problems. Consider the assumption that word-internal XP-adjunction is excluded in Morphology: “The morphological component gives no output (so the derivation crashes) if presented with an element that is not an X^0 or a feature” (Chomsky 1995:319). The question arises how Morphology will be able to tell what is an XP inside a word if contextual definitions do not apply inside a word. Clearly, some other characterization of minimal and maximal projections will be necessary. But the resulting system seems quite undesirable: why should we need two systems (one relational, one presumably not) to define projection levels? Differently put, why do we need contextual definitions of projection levels in addition to the apparently independently necessary inherent characterization?

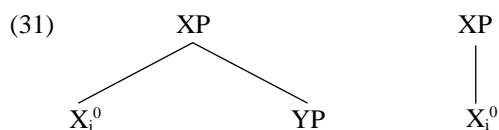
4. An additional curious feature of the relational-definitions-plus-uniformity theory is that according to this system, chains but not categories have to be uniform (recall that a nonprojecting lexical item is both an X^{\min} and an X^{\max}). This is of course logically possible: a chain member may be multiply characterized, but all chain members must have the same characterizations. But once we recall that characterizing an element as both minimal and maximal does not necessarily lead to contradiction and ungrammaticality, the uniformity assumption seems to lose much of its intuitive appeal.

5. As we saw in connection with the structure in (30), uniformity can be used in certain cases to ensure that the target rather than the “moved” (non-chain-root) category projects. (30) could also be excluded by noting that here the contextual definition of $+/-$ maximal projection makes the higher chain member in (30) a nonmaximal intermediate projection, which cannot be a chain member (Nunes 1995). But here uniformity captures only a small part of a much larger generalization. First, it cannot ensure generally that categorial projection always takes place in the root positions of chains. The generalization holds also for X^0 chains that involve word-internal positions. But uniformity is relevant only for phrasal movement. Given the system in (27) and

¹⁸ Note incidentally that in the standard minimalist framework even the status of “ordinary” morphology is rather unclear. Morphology must presumably be somewhere on the Spell-Out branch. Since conditions exist only at the interface in the minimalist grammar, Morphology would have to be located at the PF level, which does not have the structure necessary for this component to operate. In contrast, in the framework proposed in Brody 1993a, 1995a, where Spell-Out applies to the LF level, Morphology can be identical to WI and its principles will hold at this level.

(28), word-internally different principles must apply. Furthermore, if the extension of the GPP to categorial projection in section 4 is correct, then the condition that categorial projection always holds in the root positions of chains is only one aspect of a much larger generalization. Since the relevant general principle also constrains semantic properties like nongrammaticalized selection, it does not appear to be fully reducible to syntax.¹⁹

6. The violations in (29) and (30) can be combined, in which case the system incorrectly accepts the double infringement. Thus, if a “moved” X^0 , merged with some word-external element Y (as in (29)), happens to project further (as in (30)), then the chain can be uniform (X^0 , X^0), even though the structure involves both a word-external nonroot element of an X^0 chain and a “target projects” violation.



7. In my view, a further serious objection to the uniformity condition is that one would expect an optimally designed theory of syntactic computation simply not to make it possible to violate this condition: the theory should not provide devices that can create nonuniform chains. Relational definitions of projection levels seem dubious, since without them uniformity is also unnecessary in general. Chains are copies.²⁰ Given an intrinsic characterization of projectional status, the copy of an X^{\max} is an X^{\max} , the copy of an X^0 is an X^0 .²¹ This suggests that the grammar should contain no contextual definitions of projection levels. Since chains consist of copies, uniformity is unnecessary in general since there are no means to violate it—the optimal situation.

Consider the desirable effects attributed to uniformity. In the theory put forward here, these are entailed by independent principles. A structure like (29) is a special instance of illegitimate structures where a nonphrasal X^0 element is immediately dominated by some YP , not projected by it. This is ruled out by the modularity of the Insert relation. As for (30), this structure cannot arise under the present theory since a given PL can include only a single XP level, whether or not this belongs to a root or nonroot chain member.²² Finally, (31) is ruled out by the GPP: a nonroot $X^{\min/0}$ can neither select nor project an XP.²³

¹⁹ The fact that uniformity gives a partial account of one aspect of the GPP does not seem to be an argument in its favor. If anything, it is an argument against uniformity since it is seen to be redundant here; see section 6.

²⁰ That chains consist of copies is an independently necessary assumption in minimalist frameworks, where representational conditions, like the binding theory, can hold only at or beyond the interface level of LF. For example, as discussed in section 3.3, in order to rule out the Principle C violation indicated by the indices in (i), the trace must be (at least a partial) copy at LF (and perhaps beyond).

(i) whose_x mother did he_x like (whose_x mother)

²¹ Given the independently necessary full identity requirement on the top PLs of chain copies, ensured either by the Chain operation or by an interface condition; see section 3.5.

²² Alternatively, the impossibility of the configuration could be attributed to the GPP, parallel to the case of (31).

²³ There are various ways of treating clitics in the literature that take account of the fact that they are X^{\min} or X^0 elements linked to an XP position without assuming that the clitics themselves have dual projectional status. I know of no advantage that the dual-status assumption might have over such treatments.

6 Conspiracy Theories of the “Target Projects” Condition and the GPP

6.1 “Target Projects” in Chomsky 1994, 1995

In section 4 I considered (a) the requirement that categorial projection always takes place in chain-root position and (b) the restriction that selectional—in particular, thematic—properties must hold in and be satisfied by chain-root positions. As discussed, both generalizations may be effects of the GPP, itself a consequence of the principle restricting information transmission in syntactic structures.

Chomsky 1994 and Chomsky 1995 provide different explanations. In this section I discuss the approaches taken in these works.

I shall focus on the issue in (a) first and start with the solution proposed in Chomsky 1994. Here the generalization that categorial projection always takes place in root position results from conditions answering four distinct questions:

- (32) a. Why can a head in a substituted non-chain-root position not project?
 b. Why can a head in an adjoined non-chain-root position not project?
 c. Why can an XP in a substituted non-chain-root position not project further?
 d. Why can an XP in an adjoined non-chain-root position not project further?

Heads in nonroot position cannot project (32a–b) because the Head Movement Constraint would force such a raised head α to substitute into or adjoin to the α P, the phrase α itself projected. This is prohibited by the fact that such “self-attachment” would create an ambiguity: in such structures, the category/segment (α P) that dominates α in the nonroot position can be taken to have inherited its label from α in the nonroot position or from the α P that α projected in its root position. However, it is not clear why the ambiguity of “self-attachment” cases of adjoined and substituted projecting heads should create a violation.

In another case of adjoined nonroot heads, the prohibition concerning self-attachment is not relevant—namely, the case in which the head α in the nonroot position is adjoined to another head. This is the usual configuration of head chains and thus cannot be excluded in general. α in this adjoined nonroot position of course cannot project either. But it is not clear what excludes the configuration where the moved element projects instead of the head to which it adjoined.

Turning to the question of why nonheads (i.e., phrases) cannot project further in non-chain-root positions, the substitution case (32c) is ruled out by the principle known as Greed: “Move raises α only if morphological properties of α itself would not otherwise be satisfied in the derivation” (Chomsky 1994:43). In a configuration like (33), if XP^* raises to K and projects $XP+$, then XP^* ceases to be a maximal projection, given the relational definition of projectional status.

- (33) $*[_{XP+} XP_x^* [K t_x]]$

XP^* then will be “invisible for the computational system,” which sees only nonprojected elements and maximal projections, and therefore cannot “enter into a checking relation.”

But the raised XP in (33) could satisfy Greed before it projects (compare: “Adjunction to

X' by merger does not conflict with the conclusion that X' is invisible to [the computational system of the grammar]; at the point of adjunction, the target is an XP, not X'' ; Chomsky 1994: 32). At the “point of substitution,” then, the raised element is a maximal projection, not an X' .

Another reason is given in Chomsky 1994 for the ungrammaticality of (33), namely, the uniformity condition. This also rules out (33) since here the trace of XP^* is maximal (by hypothesis) but XP^* is not. But as we saw in section 5, there are strong reasons to reject the uniformity condition.

Finally, consider the facts that adjoined heads cannot project and nonheads (phrases) cannot project further (32b,d). An adjoined element that projects would create the following configuration:

$$(34) *[\alpha \alpha_x [K t_x]]$$

Chomsky assumes that the two-segment category in adjunction involves two elements that each have the status of a category: the lower segment and the two segments together. On this assumption, Full Interpretation (FI) is violated in (34). Whichever of the two categories, α or the two-segment element $[\alpha, \alpha]$, is taken to be the head of the chain whose root is t_x , the other element receives no interpretation at LF and thus violates FI. Chomsky concludes that the target must have projected. Taking $[\alpha, \alpha]$ to be the head of the chain is ruled out additionally by the uniformity condition. (This seems to work only where α is nonmaximal; if α is maximal, then so is $[\alpha, \alpha]$, and uniformity is not violated.)

However, whichever category projects, the fact that in adjunction structures there is only one LF role for the two-segment category, $[K, K]$, and the category corresponding to the lower segment, K , is a general problem in adjunction. (This appears to be recognized in Chomsky 1995.) To allow adjunction to heads, Chomsky invokes WI: the relevant restrictions again do not hold word-internally. For nonheads he suggests that this fact essentially restricts adjunction to non-thematic categories (plus some other restricted cases; see the appendix below). But if WI can neutralize the problem when a minimal projection α adjoins to X and X projects, it will also neutralize the problem if α projects. Hence, the conclusion that the nonroot element cannot project does not follow.

Similar comments hold for nonminimal projections. If a configuration in which the target of adjunction is in a nonthematic position is permitted because no problem arises with FI, then in the same kind of position the adjoined element should be able to project without violating this principle. This again is probably an incorrect result.

Additionally, the assumption that there are exactly three elements in adjunction structures with categorial status seems somewhat stipulative. Even granting that assumption, further questions arise. For example, it is not clear why α and $[\alpha, \alpha]$ could not jointly serve as the antecedent of the trace.

Later Chomsky approaches the generalization that phrasal projection always takes place from chain-root positions somewhat differently (see Chomsky 1995). He rejects earlier formulations of Greed and tentatively assumes that word-external adjunction does not exist (also see Brody 1994 and above). He bases much of the argument on the hypothesis that movement can take

place only to ‘immediately’ establish a checking configuration. This hypothesis is somewhat dubious since it conflicts with apparently well-established cases of successive-cyclic movement where intermediate landing sites appear to involve no checking (e.g., adjunct *wh*-movement in English). Given this assumption, however, word-external head movement cases will obey generalization (a), since if a moved head α projects, it necessarily establishes a head-complement relation with its target K : [$_{\alpha}$ α K]. Checking relations can only be established in specifier-head configurations. (Notice, though, that it is only because nonbranching projections do not exist in this system—in contrast to the theory of assembly adopted earlier, where they do exist—that the conclusion that K is the complement of α follows. In [$_{\alpha}$ [$_{\alpha}$ α] K], K would be a specifier.)

As before, the account does not extend to the word-internal domain. But the generalization (a) that a moved head does not project of course does. In the word-internal domain it is necessary to distinguish two cases: structures where the moved head H' adjoins to some other head H that projects a phrase HP , and structures where H' adjoins to some head H'' , where H'' itself is adjoined to H or to some other head adjoined to H . In the former case, if H' projects, the structure in (35) results since HP must have been projected before H' raised and projected.

(35) [$_{HP}$ [$_{H'}$ H' H]]

Chomsky proposes that the resulting structure is not well formed: HP has no appropriate head. In effect, he proposes a recursive filter (his (5)) that defines well-formed syntactic objects and rejects (35). This seems to be a dubious move. In a derivational theory the derivation necessarily provides a recursive definition of well-formed syntactic objects; no filter should duplicate this function.

Further conditions are required, as Chomsky notes, if H' illegitimately projects after raising and adjoining to some head H'' that is itself adjoined to a projecting head H . Here we get a structure like (36), where the considerations just reviewed are not relevant.

(36) [$_H$ [$_{H'}$ H' H''] H]

Turning to the question of why a moved XP cannot project further, Chomsky rejects the formulation of Greed that he earlier took to be relevant, and he attributes the impossibility of (33) to the uniformity condition, which I found numerous reasons to question in section 5.

Chomsky tentatively considers another approach as well, according to which the checking relation is asymmetrical and requires the checked element to be in the specifier position of the checking head. This would not be satisfied in (33). Here the head of K is taken to be the checking head and XP the checked element, but XP is not a specifier since its head projects the phrase; instead, K is the specifier of XP . Chomsky notes several problems with this approach, but in any case it would seem to contribute little to a solution: if checking is asymmetric in the way suggested, then the question still remains: why can a checking relation not be established in (33)? K cannot be checking XP now, but X could in principle check K , unless we stipulate that the target of movement must be the checker. But that is barely different from the original problem to be explained: the target must project.

6.2 GPP Effects in Chomsky 1994, 1995

Next let us consider the solutions for the other GPP effect: that thematic properties must hold in and be satisfied by chain roots. In his 1994 paper Chomsky does not discuss the problem of why the selectional (thematic, etc.) requirements of heads must hold in the root position of their chain, so I shall put this aside for the moment. He does, however, provide an account of the fact that thematic requirements must be satisfied by chain roots.

Consider the hypothetical verbs HIT and BELIEVE that assign a θ -role to their subject but no Case to their object. The GPP explains why such verbs cannot exist, given the independently motivated assumption that the Caseless object position in structures like (37) must form a chain with the subject position: such a chain involves a nonroot θ -position.

- (37) a. *John [_{VP} t' [HIT t]]
 b. *John [_{VP} t' [BELIEVE [t to VP]]]

In derivational terms the generalization translates as a ban on movement to thematic positions. Chomsky attributes this also to the principle of Greed. The DP *John* in (37) cannot raise to [Spec, VP] to pick up the unassigned θ -role, since it does not need to do so to satisfy its own requirements.

Even if the DP originates in a non- θ -position, Greed would prevent raising to a θ -position on the assumption that ‘the need for a theta role is not a formal property, like Case, that permits ‘last resort’ movement’ (Chomsky 1994:38). This explanation may not be general enough. First, the prohibition against movement to θ -positions holds also for θ -positions that are at the same time Case positions. As an illustration, consider the hypothetical preposition IN, which is like *in* except that it does not assign Case. This should allow a structure like (38a) (compare (38b)).

- (38) a. I gave John the study IN t.
 b. I gave John a book in the study.

The GPP predicts that structures like (38a) are ungrammatical. The explanation based on Greed does not have this consequence, unless not only structural Case positions but indeed all Case positions are taken to be systematically distinct from θ -positions.

Second, the Greed-based account allows movement to a θ -position when this is made necessary by some other principle. This again seems to be an incorrect prediction. For example, Relativized Minimality/the Minimal Link Condition (MLC) can force movement through a θ -position in a derivation in which a later step satisfies Greed. To see this, consider first Chomsky’s analysis of the ungrammaticality of (39).

- (39) *John reads often books.
 (40) [_{VP} John [_V v [_{VP₂} often [_V reads books]]]]

He suggests that (39) has the structure in (40), and this is ruled out since the adverbial in [Spec, VP₂] prevents raising of the object *books* to [Spec, Agr_O]. He writes, ‘Note the crucial assumption that the subject *John* is in [SPEC, VP]...otherwise that position would be an ‘escape hatch’ for the raising of *books*’ (p. 33).

Consider in this light (41), containing the verb HIT that assigns no accusative but is otherwise like *hit*. In (41) movement lands in a θ -position.

(41) John [_{VP} t [_{VP₂} often HITs t]]

Here the DP *John* must raise outside the VP in order to get Case. But then, as in the case of (39), Relativized Minimality/the MLC forces it through [Spec, VP], where it can pick up the subject θ -role. Thus, the nonexistence of a verb like HIT is not predicted.

As noted, in chapter 4 of his 1995 book Chomsky assumes that every application of Move must establish a checking relation, and he rejects the earlier formulation of Greed. Here he provides a different account of why movement cannot land in thematic positions, one that is meant to generalize to also answer the question why all thematic selection holds in chain-root positions:

With regard to assignment of θ -roles, the conclusion is natural in Hale and Keyser's [1993] theory. A θ -role is assigned in a certain structural configuration; β assigns that θ -role only in the sense that it is the head of that configuration.... Suppose β raises, forming the chain CH = (β, \dots, t)....[T]he chain CH is not in a configuration at all, so cannot assign a θ -role. In its raised position, β can function insofar as it has internal formal features: as a Case assigner or a binder. But in a configurational theory of θ -relations, it makes little sense to think of the head of a chain as assigning a θ -role. (p. 313)

First of all, it is not clear that it is any more natural (and it is certainly not independently motivated) to assume that a chain is "not in a configuration at all," rather than taking it to be in multiple configurations. More important perhaps is the question why "it makes little sense" to think of the head of a nontrivial chain as a θ -role assigner. In particular, why does it make little sense to think of the head of the chain as a θ -role assigner when apparently the root of the chain can be a θ -role assigner? "The trace *t* remains in the structural configuration that determines a θ -role and can therefore function as a θ -role assigner..." (p. 313). It is true that in its raised position β cannot assign a θ -role—or, more generally, activate any of its selectional features. But surely β is in *some* configuration in its raised position.

What we in fact need, then, is a characterization of configurations where selection is possible and those where it is impossible. An obvious suggestion is to assume that only categorially projecting heads can select. This will work once categorial projection itself is restricted to chain roots. Although perhaps not unnatural, it is clearly an additional stipulation that should be unnecessary. It is, if categorial projection and selection are both cases of semantic (nonformal) feature assignment constrained directly by the GPP, as argued in section 4.

Chomsky writes, "With regard to receipt of θ -roles, similar reasoning applies. If α raises to a θ -position Th, forming the chain CH = (α, t), the argument that must bear a θ -role is CH, not α . But CH is not in any configuration, and α is not an argument that can receive a θ -role" (p. 313). Again it is not clear why the chain is not in any configuration. But independently of this, the paragraph is difficult to interpret. If the chain and not α must bear the θ -role and the chain cannot, then the θ -Criterion is violated also in the grammatical cases where the chain would receive the θ -role in its root position. Given minimalist assumptions, α cannot receive the θ -role before it moves; to say that it did so would be to resurrect D-Structure (on the reasons for rejecting

D-Structure, see Brody 1993b, 1995a, Chomsky 1993). Thus, we seem to be left here without any account of the selectional effects of the GPP.

Appendix

Chomsky (1994, 1995) suggests a theory that radically restricts word-external adjunction but (in 1995 tentatively and partially only) retains this configuration in cases where the target has no θ -role (expletive-associate chains) or where in his derivational system the adjunct is not present at LF (intermediate traces deleted by LF and ‘‘semantically vacuous’’ scrambling where LF reconstruction eliminates the scrambled element). These cases do not seem to provide strong motivation for retaining the configuration. LF adjunction of the associate to its expletive chain-mate is a problematic and probably unnecessary operation. The agreement facts, which constitute the main evidence for this operation, can be accounted for without actual displacement of the associate (see Brody 1993a, 1995a). This is also recognized in Chomsky 1995; see section 4.2 above.

The necessity of adjoined intermediate traces in nonuniform chains is equally moot (see, e.g., Manzini 1992). Notice that the best and perhaps only strong evidence for their existence involves reconstruction effects (Lebeaux 1989, Barss 1986). For example, in (42) binding of the anaphor appears to be licensed from the position that is internal to the intermediate trace/copy.

(42) Which picture of herself did John think [t [Mary told Bill to buy t]]?

But in the minimalist framework such evidence supports the copy theory of movement. Given the basic assumption of this framework that conditions (like the binding theory) hold only at and/or beyond the interface levels, evidence like (42) shows that contrary to Chomsky’s suggestion in this context, intermediate traces/copies must be present at LF (see Chomsky 1993, Brody 1995a).

As for scrambling, an alternative treatment of radical reconstruction may be to consider it to be stylistic displacement—that is, to consider that it takes place in the Spell-Out component (see Aoun 1995). Chomsky suggests that LF reconstruction will provide an account of the contrast he finds between English topicalization cases like (43a) and (43b). Assuming that these work like examples that involve scrambling, the expectation is that (43a), the adjunct case, is worse since forced reconstruction in this example will create a configuration that violates Principle C. Since the fronted phrase *which pictures of John’s brother* is not an adjunct in (43b), this example will not be similarly excluded.

- (43) a. Pictures of John_x’s brother, he_x never expected that I would buy.
 b. Which pictures of John_x’s brother did he_x expect that I would buy?

But this approach would make it difficult to account for the *wh*-chain reconstruction effects like the contrasts between (44a) and (44b) that depend precisely on forced reconstruction of the *wh*-phrase. (The selected clausal argument but not the unselected adjunct internal to the fronted phrase must be present in the reconstructed copy, as shown in section 3.3; also see Lebeaux 1989, Chomsky 1993, Brody 1995a, for different ways of instantiating this idea.)

- (44) a. Which claim that John_x made did he_x deny?
 b. ?*Whose claim that John_x was asleep did he_x deny?

Thus, the evidence for an account involving forced reconstruction of elements adjoined to semantically nonvacuous categories seems unconvincing. Its support for the more general claim that word-external adjunction exists in syntax (under restricted circumstances) is therefore weak.

References

- Aoun, Joseph. 1995. Interception and reconstruction. Talk presented at the Vienna Round Table Colloquium, University of Vienna.
- Barss, Andrew. 1986. Chains and anaphoric dependencies. Doctoral dissertation, MIT, Cambridge, Mass.
- Brody, Michael. 1993a. Lexico-Logical Form. Ms., University College London.
- Brody, Michael. 1993b. θ -theory and arguments. *Linguistic Inquiry* 24:1–23.
- Brody, Michael. 1994. Phrase structure and dependence. In *UCL working papers in linguistics* 6, 1–33. Department of Phonetics and Linguistics, University College London.
- Brody, Michael. 1995a. *Lexico-Logical Form: A radically minimalist theory*. Cambridge, Mass.: MIT Press.
- Brody, Michael. 1995b. Towards perfect chains. Ms., University College London. [Published in *Handbook of syntax*, ed. Liliane Haegeman. Dordrecht: Kluwer.]
- Brody, Michael. 1997. Mirror Theory. Ms., University College London.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1994. Bare phrase structure. Ms., MIT, Cambridge, Mass. [Appeared as MIT Occasional Papers in Linguistics 5. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. Published in *Government and Binding Theory and the Minimalist Program*, ed. Gert Webelhuth, 383–439. Oxford: Blackwell (1995).]
- Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*, 219–394. Cambridge, Mass.: MIT Press.
- Cinque, Guglielmo. 1993. On the evidence for partial N movement in the Romance DP. Ms., University of Venice.
- Cinque, Guglielmo. 1995. Adverbs and the universal hierarchy of functional projections. Abstract of paper presented at the 18th GLOW Colloquium, Tromsø.
- Collins, Chris, and Höskuldur Thráinsson. 1993. Object shift in double object constructions and the theory of Case. In *MIT working papers in linguistics 19: Papers on Case and agreement II*, 131–174. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Hale, Kenneth, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 53–109. Cambridge, Mass.: MIT Press.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Koizumi, Masatoshi. 1993. Object agreement phrases and the split VP hypothesis. In *MIT working papers in linguistics 18: Papers on Case and agreement I*, 99–148. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Koopman, Hilda. 1996. The spec head configuration. Ms., UCLA, Los Angeles, Calif.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.
- Lebeaux, David. 1989. Relative clauses, licensing and the nature of the derivation. Ms., University of Maryland, College Park.
- Manzini, Maria Rita. 1992. *Locality*. Cambridge, Mass.: MIT Press.

- Nunes, Jairo. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation, University of Maryland, College Park.
- Sportiche, Dominique. 1994. Adjuncts and adjunction. Abstract of paper presented at the 17th GLOW Colloquium, Vienna.
- Watanabe, Akira. 1991. *Wh*-in-situ, Subjacency and chain formation. Ms., MIT, Cambridge, Mass.
- Williams, Edwin. 1994. *Thematic structure in syntax*. Cambridge, Mass.: MIT Press.

Department of Phonetics and Linguistics
University College London
Gower Street
London WC1E 6BT
England
m.brody@ling.ucl.ac.uk