Merge and Mirrors
Edwin Williams
June 2007; revised June 2008

There are five or so grammatical processes which target heads, in that they attract them, or attach things to them: Affix Hopping, Morphological Lowering, Verb or Head Raising, and less obviously, Adverb Placement and the basic Merge operation. Here I will propose a system in which these are manifestations of a single process--COMBINE--which operates in a single cycle, deriving phrasal syntactic structures and the morphosyntactic marking of them simultaneously. 'Mirror effects' arise from the operation of COMBINE, but not always--mirrors are size-relative, or 'fractal'[1] in a sense made precise in sections 2 and 5.

In some common recent formulations of the grammatical system, these five processes are parts of different components, as shown here:

(1)

Phrasal Syntax	===>	**Derivational and Inflectional Morphology**
Merge		Affix Hopping
Raising		Morphological Lowering
Adverb placement		

As I have argued elsewhere (Williams (2006)) derivational morphology is modularly separated from phrasal syntax, the arrangement grammatical system advanced here can be drawn up as the following:

(2)

Derivational Morphology	===>	**Phrasal Syntax, Inflectional Morphology**
COMBINE		COMBINE

Mirror effects are generally understood to arise from the operation of Head Movement regimented by the Head Movement Constraint—if Head Movement applies successively in a structure like (98a) to adjoin the verb to successively higher affixes, the resulting derived word will 'mirror' the underlying phrasal structure:

(3) a [af1 [af2 [af3 V]]]
b V+af3+af2+af1 [...]

COMBINE too will derive mirrors, but not always, and not under the Head Movement regime. In fact, non-mirroring structures are found everywhere; here are two, and many more will follow:

(4) a. [un grammatical] ity ==> un[grammatical+ity]
b. [[under stand] past] ==> under [stand + past] ==> under stood

There are several reasons for packing into one operation all of the head-targeting operations, which I will understand to be the manifestations of COMBINE. First, theoretical redundancy is reduced, with the consequence that any adjustment to one of the manifestations will affect all the others, clearly the best situation on general grounds. Also, uniting them simplifies their interaction

1 to use a term suggested to me by Klaus Abels.

with other operations, and even with themselves and each other; that is, the manifestations cannot interact intricately. As an example of a simple interaction, consider the morphosyntax of the verb in a verb raising language--the raising of the verb is sometimes followed by a lowering of Tense onto the verb:

(5) T ...[be kommen]] => [T[be kommen]] ==> [be [kommen T]] => bekam (German. 'get')

Here, *bekommen* is first raised to T, but then T is then lowered onto the root of *bekommen*, to get the strong form of the past tense *bekam,* parallel to *kommen~kam (*German 'come'). In the theory with COMBINE these operations happen simultaneously.

An argument has been given by Embick and Noyer (2001) for an architecture of the grammatical system that would make COMBINE impossible. They argue specifically that Raising and Affix Hopping must be in different components, components that are linearly sequenced with no feedback. Head Raising and other phrasal syntactic operations are all completed before any morphological operations such as Affix Hopping (called "Lowering" by them). Their argument for such architecture is based on the following kinds of sentences, in which Tense attachment and VP fronting are seen to interact in a particular way:

(82) a. ...and [see Mary[[he will t]
b. ...and [see Mary [he did t]
c *...and [saw Mary] [he t]

In their account, the reason that (c) is bad is that its derivation would require the application of Affix Hopping (of T onto V) before VP fronting, and the organization of the grammar they wish to support forbids this, as Affix Hopping belongs to the morphological component, and VP fronting to the phrasal syntactic component.

But their predictions are too wide. Of the English Hopping affixes, it is only Tense that shows this behavior, as in both of the following the affix is clearly lowered before VP fronting:

(44) a. and [talk-ing to Bill] [she was t]
b. and [eat-en her dinner] [she has t]

We know that *-ing* and *-en* are Lowered affixes, and not affixes that trigger Verb Raising, because they are exactly like Tense in the familiar V-Adverb-Direct-Object paradigms:

(8) a. *I was seeing recently Bill
b. *I have seen recently Bill
c. *I saw recently Bill

So the fact in (82c) needs an explanation, but the explanation given by Embick and Noyer says that all Lowering affixes will behave similarly, whereas it is really only Tense that interacts with VP Fronting in the way that it does. I believe that there is an account of this difference in behavior in Representation Theory (Williams (2003)): VP fronting precedes the Merging in of Tense, but not the other affixes, and the ordering reduces to the Merge order of the trigger for Fronting versus Tense; in other words, it reduces to a fact about F-structure. But be that as it may, the Embick and Noyer

argument cannot be taken as an obstacle to combining Hopping and Raising into a single rule.

The operations subsumed by COMBINE have the following properties: they are all head-seeking; they are all bounded by the clause boundary in an absolute way; and they are 'mirror-creating', at least sometimes. Each of these properties will get detailed treatment in the sections that follow.

1. The formulation of COMBINE.

If we write a simple rule to carry out the operations just discussed it is already interesting that the result is ambiguous between Raising and Lowering:

(56) $X > [Y\ Z]_{YP} ==> [X+Y\ Z]_{X+Y}$[2]

If X is Tense and Y is V then (56) is instantiated by either Tense raising in French, or Tense Lowering in English. In more complex cases, X itself might have a complement already; then the question arises what to do with X's complement. If we assume that time adverbial clauses are complements of Tense, then the following suggests the general pattern:

(10) [past [when I was there]] > [saw Bill] ==>
a. [[saw+past Bill] [when I was there]]
b. *[[when I was there][saw+past Bill]]
c. *[[saw+[past [when I was there]] Bill]

The complement of X is added after the complement of Y, giving the 'write wrap' operation familiar from Categorial Grammar. The general rule is then

(11) $[X_{affix}\ Y]x > [W\ Z]w ==> [[X\ W]\ Z]\ Y]$

Next we drop the requirement that X be an affix, and we arrive at the final formulation of COMBINE:

(89) a. $[X\ Y]_x > [W\ Z]_w ==> [[X\ W]\ Z]\ Y]$ 'COMBINE'
b. $X > [W\ Z]_W ==> [[X\ W]\ Z]$ Affix Hopping and Raising
c. $X > W ==> [X\ W]$ simple Merge

As (b) and (c) show, not only are Affix Hopping and Raising special cases of COMBINE (where the complement to X is absent; but simple Merge itself is a special case as well, where both the complements of both X and W are absent. This is intentional. COMBINE is not simply the replacement for head-movement rules; it is the rule of syntactic formation itself. So the expression 'X>Y' is always to be interpreted as, 'X and Y are in the "workspace", and if X and Y are put together with Y the complement of X, then the spellout of that combination is given by (89a)."

So COMBINE is not an operation on a tree, but the assembly of two units into one. As such, it solves the "extension" problem of Head Movement. In recent treatments, the effects of the "strict cycle" have been obtained by a requirement that every operation "extend" the tree, and movements like

2 For the time being, "α>β" means α takes β as a complement.

WH movement clearly do extend the tree. But Verb Raising does not:

(46) [T [V NP] ==> [T+V [t NP]]

This singular failure of extension led to the possibility that Verb Raising is not part of the syntactic cycle, but is a phonological rule (Chomsky (2001)). The problem does not arise with COMBINE: COMBINE adds T and triggers the raising of the verb at the same time, so the left-hand structure in (46) never exists.

The right-wrap property of combine directly explains a generalization about complement clauses in Williams (1974)--that when a clause associated with a scoping item is extraposed, it extraposes to the end of its scope. To such scopal items and their associated clauses are given here:

(194) a. [must [in order to S]]
b. [more [than S]]

The generalization about extraposition and scope is illustrated in the following:

(15) a. John must win more than I do in order to get a prize
must (win more than I do) (in order to get a prize)
b. John must win more in order to get a prize than I do
more (John must win x in order to win a prize) (I must win x in order to get a prize)

The only surface difference is the order of the *than* and *in order to* clauses. The relative scopes of *must* and *more* are given beneath each example, and it is clear that the scopally higher item has its associated clause in the rightmost position.

COMBINE directly derives this generalization; these are the derivations of the two examples

(16) a. [must [in order to get a prize]] > [more [than I do]] > [John win] ==>
[must [in order to get a prize]] >John win more than I do ==>
John must win more than I do in order to get a prize
b. [more [than I do]] > [must [in order to win a prize]] > [John win] ==>
[more [than I do]] > John must win in order to get a prize ==>
John must win more in order to get a prize than I do

The first lines of (a) and (b) are not syntactic structures, they are rather charts of derivations; in (a), *more* is merged before *must;* in (b), the reverse; if COMBINE effects the merging, then the correct ordering of the complement clauses automatically happens.

Importantly, the relative scopes of items is *determined by the derivation*, *and not by the structures*. In the examples just considered, the presence of the complement clauses gives a telltale trace of what the derivation order was, but if those clauses were not present *must* and *more* nevertheless interact scopally (that is, "John must win more" is ambiguous), and the relative scope of *must* and *more* is determined by the derivational history. It is not structures that will mirror meaning, but rather derivations, with COMBINE determining exactly in what way the structure will fail to mirror meaning.

COMBINE thus solves two problems so far, the extension problem and the problem posed by the scope/complement clause correlation. A number of problems remain: if COMBINE is both raising and lowering, how are these different constructions differentiated?; since 'head of X' is generally ambiguous, which head is referred to in the formulation of COMBINE?; and finally how is COMBINE bounded? These problems are in fact connected to one another.

2. Bounding Combine.

As a quick example to illustrate how COMBINE can go wrong, I will implement Larson's (1989) analysis of Dative constructions in the system under development here. In Larson's analysis, the Dative argument of a ditransitive verb is taken by the verb as a first argument, and the theme argument is a specifier, as in (a). To get the right word order, Raising moves the verb (*give* here) to a higher abstract verb position:

(17) a. $[V[[\text{my book}]_{NP}\ [\text{give}\ [\text{to Fred}]]_{VP}\ ]]$
b. $[\text{give}\ [\text{my book}]_{NP}\ [t\ [\text{to Fred}]]_{VP}]$

This has the familiar problem of failing the extension condition. We may get Larson's intuition with COMBINE, as follows:

(18) $[\text{give}\ [\text{to Fred}]]_{V} > [\text{my book}]_{NP} \Longrightarrow [[\text{give}\ [\text{my book}]_{NP}]_{V}\ \text{to Fred}]$

Here, *give* has attached to the direct object *my book*. This is permitted by COMBINE, but not forced by it. It is permitted because in an X-bar projection, every phrase that intervenes between X and XP counts as a head of XP, and for maximum simplicity, so should X and XP themselves. The attachment of the verb to the direct object gives us a constituent structure that is not consistent with the binding facts that motivated the Larsonian analysis: by many tests, the direct object asymmetrically c-commands the dative. So in order to calculate the c-command relations, we must consult not the derived structure, but the derivation itself, in which the direct object does c-command the dative argument in an obvious sense. This is consonant with our previous conclusion about the relation of semantics (scope) to structure.

But the attachment of *give* to the direct object is not forced; in fact, COMBINE could as well attach *give* to the ultimate lexical head of the direct object, which I will assume to be *book*[3]*:*

(19) $[\text{give}\ [\text{to Fred}]]_{V} > [\text{my book}]_{NP} \Longrightarrow *[[\text{my}\ [\text{give book}]_{NP}]_{V}\ \text{to Fred}]$

Clearly give has gone too far into the direct object. But what is the principled limitation that governs COMBINE?

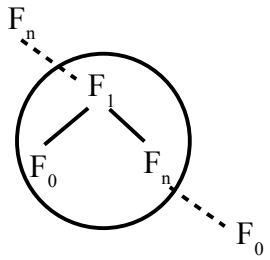
I will suggest that it is Functional Structure, that COMBINE cannot breach the Fn/F0 boundary, and that it is unique in being limited in this way.

The first step is to define the Fn/F0 boundary. It has something to do with the transition from a lexical item at the bottom of one F-structure to the top Functional element of an different embedded F-

3 The head could be *the*, but the point would be the same.

structure:

(20)



I don't think that is possible at this point to say what an F-structure is, or to say what the difference between a lexical item and a functional item is, and in the concluding section I expose some of the reasons for my skepticism. I will simply assume that there is a linearly ordered set of elements $F_n...F_0$ for each category that is the functional structure of that category, and I will assume that we know an Fn/F0 boundary when we see one—we know, for example, that "try to swim" is two F-structures, and "can swim" is one F-structure. These two cases, incidentally, mean that we cannot us a shared feature ("+V" for sentences for example) to mark off F-structures, as then that feature would occur on every node of both of the F-structures "try" and "to swim", and we could then not distinguish the two-F-structure case "try to swim" from the single-F-structure case, "can swim"--in both cases, every node would be "+V".

A systematic review of syntactic processes reveals that none of them respect the Fn/F0 boundary in an absolute way:

(21) a. Not WH: Who do you think [$_{Fn}$ t will win]
b. Not NP movement: John is expected [$_{Fn}$ t to win]
c. Not case assignment: John expects [$_{Fn}$ him to win]
d. Not agreement (Tsez, Polinsky and Potsdam (2001)):
enir uza magalu b-acruli r-/b-ixyo
Mother [boy bread-ABS(III) III ate] IV/III-know
'The mother knows that the boy ate the bread
e. Not phonological dependence:
1. Who does he think['s coming]
2. Who did he say ['z coming]

In the face of such a catalog one must wonder whether the Fn/F0 boundary is meaningful at all—maybe it has no properties that distinguish it from any $F_i.F_{i+1}$ boundary. In some theories, some Fs do play a fundamental role; so CP, IP, and NP are bounding nodes for Subjacency; and CP and vP are (strong) phases in "Derivation by Phase", but neither of these privileges the Fn/F0 boundary. What we are looking for are operations or processes that are absolutely blocked by Fn/F0 *for any n,* not just for some particular ones like CP and IP and vP. Phases, furthermore, are not absolute boundaries anyway, because the Phase Impenetrability Condition allows the embedding phrase to communicate with the left edge of the embedded phase. So there is a high possibility that nothing at all respects the Fn/F0 boundary in an absolute way. In that case, it is all simply embedding, with some nodes (vP, etc) picked out for special roles, but no special standing for Fn/F0 boundaries, and no special role in phrasal syntax for the lexical/functional distinction.

But I will proceed with the idea that COMBINE is bounded by Fn/F0. If this is so, and if COMBINE is the only operation that is bounded in this way, then we have a fundamental connection between F-structure and the basic rule for assembling linguistic units. Why there should be such a connection remains unclear, but I think it is interesting enough for the moment to identify such a connection.

The following definition of COMBINE builds in the Fn/F0 bounding:

(22) COMBINE: in $[X\ Y]_x > [W\ Z]_w$ (that is, if F is to take W as a complement)
X attaches to the head of W if W meets certain conditions to be specified shortly; if the head of W fails, then X COMBINES with Z if Z is the Functional complement of W

We will give the caret the following meaning: the two constituents on either side of the caret are to be thought of as "in the workspace", and the caret signifies the intent to combine them into a single syntactic object; the definition of COMBINE tells how to spell out that object. The definition is recursive, in that if the immediate head of W fails, then COMBINE continues searching for an attachment target for X in the functional complement of W (Z). I will refer to this behavior as "default to the complement", and it is the limitation to *Functional* complements in the recursive part of the definition that gives the Fn/F0 bounding. There are other definitions that give approximately the same results but this one best fits the range of cases that I will discuss in what follows.

Two things remain to make COMBINE a plausible candidate for the job I have in mind for it. First, it must be specified what it means for W to fail to be an appropriate target for X; and second, some parametric variation must be introduced into it to give rise to the variety of constructions which are subsumed by COMBINE. I will introduce now a pair of lexical parameters on X which will fill in both these missing pieces. The are both 'size' parameters that X imposes on W, which together determine whether X can target W in the first place, and then how X attaches to W.

The first is the Phrasal Size Parameter (PSP), which determines whether X attaches to the whole of W, or to the (phrasal) head of W. A simple example of the distinction that this parameter draws is the difference between a preposition (*at,* for example) and a case (locative, for example)--the preposition attaches as a prefix to DP, whereas the case-marking attaches to some head of DP, possibly N. Each of these is a "head" of DP—that is, N is the ultimate lexical head of DP, and DP is a head of itself trivially. Accordingly, the two values for the Phrasal Size Parameter are X^0 and XP. It will be useful to refer to these as the head-attaching and the prefixing values of the PSP, respectively.

The other parameter is the Morphological Size Parameter (MSP). It does two things. First, it determines whether or not there is a morphological union of X and W; and second, it determines what size of morphological head in W that X must attach to. If X specifies a morphological union, there are two possibilities: "root" and "stem", where root is smaller than stem. A simple illustration of the difference in the settings of these two parameters is the difference between derivational prefixes in English and Tense in English. Prefixes like *re-* attaches to a complex form like *under-take* prefixing to the whole: *re-under-take*. Tense on the other had does not affix to complex forms, but rather targets the ultimate verbal root, so we get T_{past}>re-under-take ==> re-under-[take+T_{past}] ==> re-under-took. Tense is COMBINED later than *re-*and so strictly takes scope over *re-,* but is nevertheless realized in a position subordinate to *re-* in the surface because it targets a smaller morphological unit than *re-,* namely *root.*

The MSP plays an additional role. If it is present, then morphological union is mandated, and in the right circumstances (namely, if the PSP is prefixal), that will force what looks like syntactic movement; and if the MSP is absent, that (appearance of) movement does not take place. The presence vs absence of the MSP will play in present system something like the role of the "strong"/"weak" feature distinction in checking theories, and will allow us to recover some of the difference between raising and lowering constructions. An important difference between MSP and feature strength in checking theories is that when A is COMBINED with B, only the MSP of A, the top merged element, governs COMBINE; any MSP values of the element in B that is targeted are not relevant at the time that A is merged. In checking theories, on the other hand, feature strength on both the 'probe' (A here) and the 'goal' (B here) are relevant. So by comparison COMBINE reduces analyses, always a welcome result unless empirically inadequate.

The role of the parameters will be illustrated for some simple well-known systems. We begin with the raising/lowering difference between French and English:

(23) a. English T:
1. T > [NP VP]
2. T: Phrasal Size: X^0
Morphological Size: root
3. T_{past} > [NP VP] ==> [NP [[V+T] ...]$_{VP}$

b. [under [stand+T_{past}]] => understood

c. French T:
1. T>[NP VP]
2. T: Phrasal Size: XP
Morphological Size: root
3. T > [NP VP] ==> [NP [[V+T][...]$_{VP}$]
4. [NP [V+T [Adv [...NP]$_{VP}$]]

As demonstrated in Emonds (1978) and Pollock (1989) , French verbs raise over VP- attached adverbs and English verbs do not:

(24) a. John kisses [often t Mary]
b. Jean embrasse [souvent t Marie]

In the present system the difference is achieved by assigning different Phrasal Size Parameters to Tense in the two languages—in English the value is X^0, meaning that T attaches to the head of the phrase that it COMBINES with (a.3); in French the value is XP, which means that T prefixes to the phrase that it combines with (as in (c.3)). Never mind for the moment that in French T prefixes to the VP, and not the entire [NP VP] constituent that it COMBINES with; this discrepancy will be dealt with in a later section. In French, something further than prefixation occurs—the head of the VP moves to the prefix position of T, and this happens because of the Morphological Size Parameter value of T in French, which is "root". In order to satisfy the requirement of this parameter, COMBINE begins a search inside of VP for something to combine T with morphologically, by the algorithm given earlier; when it finds such a thing, it COMBINES T with that thing, and puts the result in the position designated by the PSP, that is prefixed to the VP. It can be seen from this example that one of the functions of the MSP is to trigger V "movement", though of course it is not really movement in the classical sense (or is it?).

The MSP value for French Tense (root) is a stand-in for a more nuanced account. D. Sportiche (pc) informs me that in fact sometimes French Tense targets the root, as in English, and sometimes it targets something larger:

(25) a. Future: venir-viendra;
pre[venir] -previendra
b. Present: mouvoir-meut;
pro[mouvoir]; pro[meut]

In the Future and the Present, derivative verbs show the same allomorphy as their roots, as shown in (a,b); but for other tenses, such as the Past, it looks like something higher is targeted, something like the "derived stem", and so regularization occurs:

(26) Past: voir-verra
pre[voir]
*pre[verra]
pre[voir]a

So a first approximation to such facts is that the MSP in French is not constant across tenses, but is "root" for Future and Present, and "stem" for Past.

It is not necessary for an item to have a MSP. If it is prefixal (that is PSP = XP), then we will get simple prefixation without movement, as happens with the English complementizer *that:*

(27) English *that:*
a. C>TP
b. phrasal size: XP
c. Morphological size: none (i.e. free form)

(a) indicates that F-structure permits C and TP to be removed from the workspace and COMBINED. (b) is the prefixal PSP, and there is no MSP. Since there is no MSP, there is no search in TP for a morphological target, and since the PSP is prefixal, the COMBINE result is simply [that TP].

Subject-Aux-Inversion sentences require the postulation of a morphologically invisible C, call it Q:

(28) English SAI complementizer Q:
a. Q>TP
b. PSP: XP
c: MSP: aux-V^0

The presence of the MSP value forces COMBINE to search in TP for something satisfying the parameter, and the result is joined with Q and prefixed to the TP: [aux-V+Q TP]. The V+Q combination is not morphologically distinguishable from V itself in English, but we know that in some languages it is: in varieties of Dutch for example, the complementizer agrees with the subject under certain conditions; and when the V is fronted under conditions similar to English, the fronted verb takes

the morphology of complementizer agreement, not the usual morphology of verbal agreement (see Zwarts (2006) for discussion and examples).

Because of the MSP values, it is necessary that COMBINE operate with actual morphemes, not with unrealized feature complexes. Put another way, COMBINE must know what feature complexes are realized as morphemes, and what ones are not. In this respect, COMBINE differs from theories in which lexical insertion is post-(phrasal)syntactic. So, for example, we may suppose that there is a F-structure position for the progressive aspect, with value either +progressive or -progressive. In English, only the +progressive is spelled out as an actual morpheme, the -progressive value is not realized as a morpheme. This will make a difference in how COMBINE with T>VP as input searches for the MSP target—it will skip over -progressive values of the progressive node and continue searching lower, but it will not skip over a +progressive node, as such a node will contain the morpheme 'be', an appropriate MSP target, and the search ends there:

(29) a. T > [$_{Fi}$ be [$_{F0}$ seeing]] ==> [$_{Fi}$ [[T+be] [$_{F0}$ seeing]]
b. T [-Prog [...[V ...]] ==> T+-Prog V
c. T [+Prog [...[V ...]] ==> * ...+Prog..T+V... (e.g. *be saw)

In (a) T correctly COMBINES with be; in (b) it incorrectly combines with -progressive, and in (c) it incorrectly skips over an appropriate target (+progressive is realized as the auxiliary *be)* and combines with V. It is the fact that +progressive has a morpheme and -progressive does not that regulates these outcomes.[4]

3. F-free Zones.

According to the formulation given in the previous section, COMBINE will recursively search through a single F-structure for its target, but not beyond a single F-structure. The formulation does not mean that COMBINE will not operate in the absence of F-structure; rather, COMBINE will operate in a more limited way, as there will be no recursive search, no "default to the complement". There are at least two realms of syntactical objects with no F-structure—derivational morphology, and the domain of pure theta-role assignment in phrasal syntax. The latter domain can be thought of as the bottom VP shell in standard F-structure syntax, or as the initial level in Representation Theory (Williams (1994)). In both, we find that the absence of F-structure gives rise to the characteristic strictures of the constructions that arise in these 'F-free zones'.

3.1 Derivational Morphology.

I have written elsewhere (Williams (2006)) that derivational morphology, in the traditional sense, is a separate module from phrasal syntax. Inflectional morphology, on the other hand, is part and parcel of phrasal syntax; hence the diagram in (2). The alternative of course is to integrate all of morphology into phrasal syntax, making one big undifferentiated syntactical system (OBUSS), as one finds for example in Distributed Morphology. One of the arguments against OBUSS is that it is impossible to define the target attachment conditions of derivational affixes without the modularization

4 Auxiliaries exist in English because in fact the English verb is restricted to a single inflectional affix, and these are few in number (-0, *-s, -ed, -ing, -en*); so compound aspects and tenses necessarily involve multiple verbs in a single F-structure

of derivational vs. phrasal morphology. An example is the English prefix *re-*, which only attaches to "process-result" (or roughly "telic") predicates:

(30) a. The butter remelted
b. *John relaughed.

Significantly though, *re-* distinguishes between telic *words* and telic *phrases:*

(31) a. *John re-made Mary sad
b. John re-saddened Mary

Make by itself is not a telic predicate, only when combined with a result predicate is a telic predicate achieved. *Sadden,* on the other hand, is a telic predicate all on its own. Now *sadden* derives from "make sad" in any theory, mine included. The problem for OBUSS is that *sadden* and *make sad* will have essentially the same underlying structure, with incorporation applying to derive *sadden.* The problem for the attachment of *re-* is, how is it to know that the meaning-neutral application of incorporation has happened (or will happen)? The more general problem for OBUSS is that it has no notion of "telic word", which seems to be the governing concept for the attachment of *re-*. So in Williams (2006) I conclude that *re-* is attached in a derivational system that feeds, but is not part of, the phrasal syntactic system, an arrangement in which "telic word" makes perfect sense. The conclusion is a broad one, since any affix which can appear interior to *re-* must also be part of the derivational system, and that includes all of the derivational prefixes of English (*re-mis-align*, *re-de-bone*, etc.,), and also all of the incorporations of the *sadden* kind, as well as derivational suffixes. See Williams (2006) for full discussion.

So there is a derivational system consisting of stems, roots, prefixes, suffixes, etc., but no functional elements. The big question is, what is the system of assembly in this domain. The answer is COMBINE of course, and in fact some important details of the character of derivational morphology follow from this fact.

At first blush it would appear that derivational morphology 'violates' boundaries, as in the following derivation:

(32) -er_N >[drive_V >truck_N]==> -er >[truck drive]$_V$ ==>[truck drive-er]

First *drive* and *truck* join in a theta relation; then a nominal element, *-er,* moves into a verbal complex, which some versions of affix bounding would block, as violating a category-change boundary. But in fact, COMBINE can be understood as exactly the rule that is performing the operation: COMBINE attaches the left term to the head of the (appropriate) head of the right term, and that is exactly what is happening in (32). What COMBINE could *not* do is attach -er to the *non*head of the right term, in case the head of the right term was not an appropriate target (i.e. **trucker drive*). The reason is, since there is no F-structure here, there can be no "default to the complement"; for that, we need a single F-structure.

COMBINE in fact solves some "bracketing paradoxes" in morphology. It has been observed by Pesetsky (1985) that there is a conflict between the scope and the morphological requirements of *-er* and *un-* in *unhappier:*

(33) a. -er [unhappy]
b. un- [happier]

(a) represents the correct structure from the point of view of meaning, but as *-er* normally cannot attach to trisyllable adjectives, (b) is the correct form for that restriction. But COMBINE can get both. The order of derivation is

(34) a. -er > un- > happy ==>
b. -er > [un[happy]] ==>
c. un [happier]
d. -er: MSP: short stem

The reason (c) is possible is that *un-* has the MSP value "short stem", and *unhappy* has the short stem *happy* in its head position, and so *-er* can target it. The order of the derivation gives the scope order, and COMBINE gives the spell-out, which in this case does not mirror the derivation. And note that no use of "default to the complement" was used--*un-* is not the head in any sense, and so *happy* is the first, and only available, head.

Another kind of paradox in morphology is "level ordering" paradoxes. These are contested, but it is at least worth mentioning how COMBINE would treat them. Under the assumption that *-ity* is a root-attaching affix, and un- is a stem attaching affix, and that roots are smaller than stems, we get again a representation that does not mirror the meaning:

(35) a. meaning: [un grammatical] ity
b. root/stem ordering: un [grammatical ity]

But with the specifications in (a), we can get the derivation in (b) directly with COMBINE:

(36) a. ity > un > grammatical
-ity: MSP = root
un-: MSP = stem
grammatical = root, stem

b. ity > un > grammatical ==>
ity > [un [grammatical]] ==>
[un [grammaticality]]

Again, no use of "default to the complement" can occur here, but nothing blocks the derivation in (b). Importantly, here and throughout, scope is determined by the merge order, not by the resulting structure. COMBINE only gives the spellout; the derivation determines the scope relations. So *unhappier* is the comparative of *unhappy*, not the negative of *happier;* and *-ity* nominalizes *ungrammatical,* rather than un- negating *grammaticality.* This is consistent with the conclusion about Williams' generalization in section 1.

3.2 Theta Structure.

There are a number of constructions which might or might not be derived in derivational morphology. These include the extended family of constructions that Baker (1988) treats as cases of incorporation. They all involve movement of a head to another head, and thus are candidates for COMBINE. The question then arises whether the details of the constructions are compatible with the bounding of COMBINE, and in particular whether there is F-structure involved. I will consider the French causative and the noun-incorporating constructions as representative examples.

Of course if these constructions are part of derivational morphology, then they are reconciled with COMBINE as in the previous section, as there is no F-structure in derivational morphology. But if they are derived in the system of phrasal syntax, as is usually assumed, then the role of F-structure becomes problematic, and in fact most analyses of these constructions are incompatible with COMBINE as it is formulated. I will suggest though that there are adequate phrasal syntactic analyses which are fully compatible with COMBINE.

3.2.1 Clause-Union Causativization.

The thoroughly "clause union" style of causative construction, including the well studied Romance causative, presents a sharp challenge to the ideas presented here. Generally the construction is taken to involve a movement of the embedded verb to the governing causative verb, as illustrated here:

(37) Jean a fait manger$_i$ [le pomme t_i a Pierre]$_{vP}$ (French)
Jean has made eat the apple to Pierre
"Jean has made Pierre eat the apple"

Here the complement by general assumption is at least as big as vP, and hence contains functional structure. But *fait* and *manger* are not part of the same functional structure, and so this movement is not possible under the formulation of COMBINE arrived at in section 2. To reconcile such constructions with COMBINE it is necessary to assume that the embedded complement contains no functional structure at all, not even "little v"; then COMBINE can target the unique lexical head of the complement and move it as indicated, under the values for *faire* "MSP=word, PSP=XP."

This conclusion is incompatible with the assumption that the external argument is assigned not by the V, but by a functional element, "little v". But in fact "little v" wreaks havoc in causative constructions under any analysis. Since a causative construction has two agents, (*Jean* and *Pierre* in (37)), it should have two "little v"s. But there is in fact strong independent evidence that the embedded complement is a pure V projection, with no "little v"--in addition to assigning the external argument role, "little v" has another property, the capacity to assign accusative case. So if one reasons that two agents implies two "little v"s, then one must reason further that two "little v"s implies two accusative cases, but that is characteristically unavailable in causative constructions. To see why consider the following:

(38) *...[$_{vP}$ Jean [$_{v'}$ v fait [$_{vP}$ Pierre [$_{v'}$ v [manger le pomme]]]]

Acc 1 Acc 2

If the higher "little v" is available for case assignment, then it remains unexplained why the embedded

direct object and the embedded agent cannot both be assigned accusative case. Instead, the embedded agent must get dative case, and in fact the reason this is called "clause union" is because the resulting configuration has exactly the case frame of a single clause: V-Acc-Dat.

If we eliminate "little v", we solve the problem of the extra accusative case, and we also enable COMBINE to perform the raising of the embedded verb. We must conclude that the verb itself assigns the agent theta role.

The conclusion is that COMBINE is compatible with the kind of clause union that is found in causatives, but only under radical circumstances: the embedded clause must contain no functional structure. If it is correct that clause union occurs in the phrasal syntactic component, then there must be at least one “F-free zone” in phrasal syntax. In Representation Theory this is the initial level, Theta Structure, in the series of levels that constitute the architecture of that theory. In a standard theory, it will correspond to the lowest VP structure of the clause, before any functional elements are introduced.

3.2.2 Noun Incorporation.

Certain implementations of Noun Incorporation involve Head-to-Head movement, but are nevertheless incompatible with COMBINE. For example, Baker in early work (Baker (1988)) analyses Noun Incorporation as in (a), where a head noun is removed from a DP over the intervening F-structure of the DP and the into domain of the governing verb; and he gave examples like (b) to support this characterization of the process:

(39) a. N_i+V [$_{DP}$ ... t_i...]
b. Kanekwarunyu wa'-k-akyatawi'tsher-u:ni
it.dotted.DIST PAST-I-dress-make
"I dress-made a polka-dotted one"

In (39b), only the head *dress* has been removed from the complement DP, the rest is a DP remnant. COMBINE cannot effect such a relation, as the two positions related are in different F-structures.

But In Baker's recent work on the same topic (Baker (1995)), the analysis of Noun Incorporation is entirely what one would expect if COMBINE were the operation behind Incorporation. In particular, Baker has proposed that free NPs like the "remnant" NP in (39) are not really arguments of the verb like English direct objects are, but are more like adjuncts, and he has presented compelling arguments based on binding, quantification, and other considerations. Under his Polysynthesis Parameter hypothesis, when free NPs are not arguments of the verb the verb must incorporate a Noun or be overtly case-marked. This means in our terms that the apparent direct object is not the complement of the verb; rather just the bare Noun *dress* is the complement, and so we have another F-free zone in which COMBINE can operate:

(40) V > N ==> [N V]

In Representation Theory of Williams (2003) (40) would take place at "Theta Structure", the only F-free zone in Phrasal syntax. The phrase "polka-dotted one" in (39) is added later. This explains why incorporation never includes any higher material from the DP, like Demonstrative marking, etc., but only the bare noun itself.

3.3 Breaching Fn/F0 After all

Having advertised COMBINE as bounded by Fn/F0 boundaries, I must point out that in fact in a limited way it can breach such boundaries. Specifically, it can access the very top head of a lower F-structure, but if that element fails to meet the target conditions of the governing head, there is no further search via 'default to the functional complement', according to the formulation I have given. So in fact COMBINE allows such forms as "think-that", and in general F_0+F_n combinations are allowed, where n is maximal.

There are several examples of such constructions. In French and German, Romanian, and other languages, there are constructions in which prepositions combine with the determiners of the DPs they govern:

(41) a. zu [dem Bahnhof]$_{DP}$ ==> zum Bahnhof (German)
to the train-station
P >D ==> P+D (no default to the complement required)
b. *[zum N Relative$_{restrictive}$] (Prinzhorn p.c.)
c. de [le vin] ==> du vin (French)
of the wine
d. Mă îndrept către parc / *către parc-ul (Romanian)
me head towards park towards park-the
'I'm heading towards the park' (A. Mardale (ms))

Curiously the contraction is impossible in German when there is a relative clause in the DP (Martin Prinzhorn, p.c.); COMBINE does not explain this restriction. In Romanian, the P+D combination does not morphologically reflect the D, presumably a defect of morphology, not of phrasal syntax.

Another kind of example of F_0+F_n is *wanna* contraction in English. Under the assumption that the complement is a *to*P, so that to is the top head of the complement, *wanna* contraction is entirely compatible with COMBINE. All that is required is to give *want* an optional MSP value, and it will then attract *to*:

(42) want > [to VP]$_{toP}$ ==> wanna [VP]$_{toP}$
MSP=stem

The well-known absence of *wanna* contraction in the presence of a wh-trace in the subject position of the complement follows from COMBINE plus the supposition that the subject requires a higher projection than toP, presumably TP or CP. Since *wanna* contraction breaches an Fn/F0 boundary there can be no default to the complement, so *to* remains inaccessible in the larger structure:

(43) Who do you want [t [to win]$_{toP}$]$_{CP}$

So in this account, it is not the linear intervention of the trace that blocks contraction, but rather the extra structure that the subject trace requires the complement to have. Unfortunately I have no independent evidence that it is extra structure and not linear intervention that the contraction is blocked by, but the difference in the character of the explanation is worth noting.

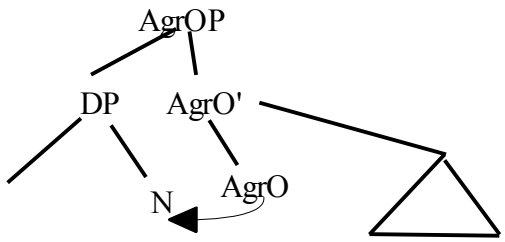
4. Co-generation and the Fn/F0 boundary.

A number of syntactic relations which might be viewed as head-targeting do not seem to obey Fn/F0 bounding. I will review some of these cases, and suggest that the grammatical architecture of Representation Theory (Williams (2003)) together with COMBINE provides correct analyses fully consistent with the bounding of COMBINE already in place.

4.1 Co-generation and Case/Agreement Markers

In the wrong theoretical setting, case and agreement marking provides a rich trove of difficulties for COMBINE. Consider for example a story like the the following about how the Direct Object gets its case: First the direct object is moved from its base position and adjoined to AgrO. Then AgrO is lowered, or copied, onto the head N of the Direct object:

(44)



COMBINE cannot perform this kind of operation, as an element in one F-structure is targeting an element deep inside of another. And yet, the operation looks like a head-targeting operation, and we are supposing that COMBINE is the only such. Case marking and agreement will have to be rethought from the ground up.

The rethinking that is required is the story of NP embedding in the first place. In standard views, clausal and NP embedding occur in the same way: in both cases, the item to be embedded is built up in the workspace, and finally is merged into the matrix structure. In Representation Theory (RT; Williams (2003)) things go differently. Clauses are indeed embedded in the standard way, all at once. But NPs are "stitched" into the clause that embeds them in a succession of steps; to use the terminology of Williams (2003), NPs are "co-generated" with the clause that embeds them, through the succession of levels that are the architecture of RT.

A comparison of a partial schematization of clausal F-structure with a partial schematization of nominal F-structure gives a hint about how co-generation might proceed, and why co-generation is the method of nominal embedding:

(45) Clause: Universal Quantif... > Existential Closure > .... > theta domain (VP)

b. DP: [Q_{univ} . . . [Dem . . . [Numord . . . [RC . . . [Numcard . . [Cl . . . [A . . . NP]]]]]]]

(Cinque 2005)

Both lines are skeletons of their respective F-structures. The bottom line is taken directly from Cinque (2005). The top line is my own construction, but consistent with a consensus view of the constitution

of the clause.

Arrows connect what I believe to be related parts of the two functional structures. I take the related points not to just be similar or parallel to each other, but to be derivationally connected. The first one, the connection between the natural site in CP of Universal Quantification, and the site in NP where Qs such as *every* occurs, represents the fact that the scope position of Universal quantification is high in the clause *and* the marking of the quantifier (by the morpheme *every)* is high in the DP.

The second arrow is more subtle--I have drawn a correspondence between the site of Existential Closure in the Clause, and the position of the demonstrative in the DP. This arrow represents again represents a correspondence between the structure of the clause and the structure of DP. To begin with DPs, *indefinite markers* occur strictly inside of Universal quantifiers, as seen in examples like "Every 3 people", and indefinite markers, being in complementary distribution with demonstratives, are best understood as instantiations of Dem. And from the point of view of clausal structure, it seems no accident that DemP is the complement of *Every: i*n clausal structure, The universal quantifier *every* requires a "variable" to bind, and that variable seems to correspond exactly to the [Dem(onstrative)...N] substructure of the NP, and in fact [Dem...N] can be used as quasi-grammatical variables in such examples as the following:

(46) Every man thinks that [that man]$_{\text{DemP}}$ will win.

We conclude then that "[Dem...NP]" seems a likely locus for the linguistic notion of "variable", so that *within* a DP, "[Every [Dem...N]]" instantiates a "Quantifier-Variable" structure directly (all that is missing from the picture is the scope). According to the second arrow, DemP in DP structure corresponds to the position of "Existential Closure" of the Direct Object, on the assumption that Existential Closure is the lowest scope position in the clause, and that DemP is the smallest nominal projection that can be a variable:

(47) [Bill [$_{\text{Existential Closure}}$ Ex [...saw [$_{\text{DemP}}$ a man]$_{x}$]

So the same notion of variable, namely [Dem...N] is used both within DP, for "[Every [Dem N]]" and outside of DP, for "Ex...[(Dem) N]" structures. This I think rationalizes a correspondence between the Dem position with DP and the Existential Closure site in the clause. It remains to give the implementation of course (co-generation).

As for the last correspondence in (45), it instantiates the lowest level of RT, the point at which the N is first introduced to the V, and selection by the V is imposed. Selection by the V is restricted to the core N and perhaps some of its AP modifiers, as no verb selects for higher nominal material—that is, no verb selects for a demonstrative NP, and no verb selects for a universal quantifier. So there is a real correspondence between the theta domain of the clause, and the NP nucleus of a DP. As emphasized in Williams (2003), an embedded *clause* bears a completely different relation to the VP it is embedded in—it does not exhibit the pointwise correspondence to its embedding matrix illustrated for DP in (45), and in fact the only relation is "selection", and the relation has access only to the outer shell of the clause (CP say, and thereby C), not to the inner core of the clause, in that no verb that takes a *that* clause ever selects the verb that will appear in the *that* clause. Compared to clauses, NPs are 'inside out'. Also, clauses do not enter into case relations or agreement relations with their embedding clause, nor do they enter into quantificational relations, and in fact they are systematically excluded

from NP positions on this account. The conclusion is that CPs are embedded in the classical sense, but NPs are "co-generated".

On this final point, it is worth noting that in Baker's work on Polysynthetic languages (Baker (1996)), NPs and CPs are sharply different in their embedding: overt NPs in these languages do not occupy argument positions as they do in English; rather, they are adjuncts, and Baker presents a number of arguments based on binding, quantification, and other considerations for this conclusion. On the other hand, CPs in the same languages do occupy argument positions, just as they do in English. In the theory outlined here and given more fully in Williams (2003) this aligns strikingly with the conclusion that NP arguments, but not CP arguments, are co-generated. In a language in which NP co-generation is trivial, as in the Polysynthetic languages, CPs nevertheless occupy the same structural argument positions as English embedded CPs. This follows if it is co-generation that is curtailed in these languages, not arguments in general.

Now for implementation of the notion "co-generation" that makes sense of these correspondences. The essence is this: at the beginning of derivation, the N and V enter into a primitive "selection" relation at the start, and then the N and the V both "grow" through the application of a series of binary relations, where each binary relation has some semantic or structural content, and each relation can morphologically "mark" either the N projection or the V projection or both. The following is just a selection of such relations; see Williams (2003) for more:

(48) Co-generation:

a. selection: V>N	==>	[V N]	
b. case/agreement: [V N]	==>	[[V+PHI] [N+ACC]]	
c. telicity/partitivity: Part: [[V+...] [N+...]]	==>	[[V+...+Part] [N+...]]	(Russian)
	==>	[[V+...] [N+...+Part]]	(Finnish)
d. Every: [[V+...+PHI...+...][N+...+ACC]]	==>	[[V+...+PHI...] [Every +...N+...+ACC]]	

In these representations, things like '[V+...+Part]' stand for partially built-up projections, this one with V at its base and 'Part(itive)' at its top. Each of the binary relations (a-d) is a level of Representation Theory, a general theory of derivation, but some aspects of that theory not crucial to the present point will be omitted. As inspection shows, the binary relations are applied successively, building up the structure of both the NP and the VP it is embedded in simultaneously, hence the term "co-generation". The first is the core "selection" relation between N and V. This relation marks neither term in English, but in languages with "lexical case", that case might be marked here on the noun; and in a language with classifiers, the classifier might mark the verb. Next is the case/agreement relation, which in general marks both N and V, the N with structural case, and the V with agreement features. Making case-assignment and agreement each part of a single binary relation expresses the general fact that case and agreement go together; that is, a verb (or Tense) only agrees with a N which it (potentially) case-marks. Another binary relation is telicity/atelicity determination. I include this relation because it is well-known that the "telicity" of the verb-object combination can be marked either on the verb, as in the Russian verbal prefix system, or on the Noun, as in Finnish partitive construction, in which partitive case on the Direct Object can be interpreted as imparting an "atelic" interpretation to the VP even when the Direct Object is interpreted as definite, just as in the English connative "eat at the apple" construction. And finally, there is the "Every" relation, which relates a Scope (the first term, a verbally based projection: "[V+...+PHI...]") to a variable (the second term, "[Dem...N+...+ACC]").

To return to the original problem, there is no such operation as illustrated in (44), in which an element in the verbal domain is lowered into the nominal domain, breaching Fn/F0. Rather, each binary operation is capable of adding a morphological mark to either of the two terms, and in the case of case/agreement, a "case" mark can be added to the nominal term. And importantly, the two terms of the binary relations are not subject to the Fn/F0 boundary, as they are not involved in head-to-head relations, *but the marks they introduce are*, as these marks are head-seeking. So it is never the case that an F from the verbal domain is lowered into the nominal domain; rather, any relation can add an F to either or both of the nominal and verbal domains that it applies to, and the spellout of those marks is effected by COMBINE, the only head-targeting operation.

Subjects are introduced in a fashion parallel to Direct Objects, except that the "Selection" step holds between N and an already partially developed verbal projection, say vP; so parallel to (48) we would have:

(49) a. selection: N>vP ==> [N vP]
b. Tense/case/agreement: [N vP] ==> [[N+NOM] [vP+PHI+T]]

If Tense is a binary operator as indicated in (b), we now have an answer to the question raised earlier of why French is not a VSO language. Recall that French Tense has the prefixal value for PSP, which would lead to VSO if the derivation was as follows:

(50) T>[NP VP] ==> T+V [NP VP]

But in fact, Tense is a binary operator bundled with case/agreement, and so tense "distributes" over NP and VP, giving

(51) [T>NP T>VP]

T>NP[5] is realized as nominative case on NP (see Williams 1994; Pesetsky and Torrego (2002)), and T>VP is realized as tense and agreement marking on the verb. T>NP (and also T>VP) is realized morphologically by COMBINE, based on the MSP and PSP values of T. In sum then, since Tense is a binary operator, its realization involves two steps: first, it applies a mark ("T" here) to one or the other or both of the two terms; then COMBINE spells out that mark on each term. So in French, the PSP of T will realize T>VP by fronting the verb to the front of VP, but not at the front of the entire clause, and so French is not VSO.

4.2 Quantifiers

The *Every* relation in English and lots of languages marks only the second term (the variable), and this might lead us to wonder whether *Every* is treated correctly as a binary ([V+], [N+]) relation; but in fact there are some languages that mark the verbal projection instead of the nominal projection. In the 80's and 90's linguistic semanticists studied a number of languages in which universal quantification is marked on the verbal projection (dubbed 'A-quantifiers'). In most of these languages, the quantifier turned out to be an unselective binder (that is, it could be understood as a binder of any

5 Here I have again used the caret (">") to indicate the "complement of" relation. Later in this section I will systematize the relation between the caret and the binary operators; in brief, each binary operator applies a caret to (potentially) each of its terms.

unbound NP in its scope), but in some cases, such as the one shown below, the scope markers are specific to particular grammatical relations, so one for the subject, a different one for the object, etc.,

(52) a. Actor/subject:
Garri-djarrk-dulubom duruk
we.pl-together-shootPSTPF dog
we all shot the dogs
b. Absolutive:
Guluban garri-djagged-na-ng
flying.fox we-mob-see-PSTPF
we saw a mob of flying foxes (Mayali (Nick Evans, as reported in Partee(1995)))
c. Q: [V N] ==> [Q+V N] (Mayali)
==> [V Q+N] (English)

And so in fact we conclude that Quantifiers also are binary relations, in the spirit of the theory of Generalized Quantifiers, capable of morphologically "marking" either term, as (c) indicates.

4.3 Bounding of Binary relations

The point that the terms of the relations themselves are not subject to Fn/F0 bounding can be made most easily with quantifiers. The two terms are the scope and the NP itself. Clearly sometimes quantifiers take scope wider than a single clause. In such cases the first term spans more than one clause (or more generally, more than one Functional structure):

(53) a. Someone or other considers [each of the candidates a bore], each > someone
b. Each: $F_{i \ldots} F_0 F_n \ldots V$, NP

Here, the scope of *each* is some F_i in the matrix clause, and the second term is an NP in the lower F-structure. In principle, *each* could apply a mark to either term, and Fn/F0 would bound the realization of that term under COMBINE. If the language were Mayali, it would mean that the mark could not be realized lower than F_0 of the matrix. But since it is English, it would mean that the mark could not be realized lower than N_0, which of course it is not. So the lesson is, the binary relations themselves do not involve COMBINE, but the marks they apply do.

So, if the terms of the binary relations themselves are not bounded by Fn/F0, but only the morphological marks applied to those terms, then the obvious question is, is there any bounding of the terms themselves?

In fact there is a natural bounding provided by Representation Theory, the Level Embedding Conjecture (LEC). The LEC requires a number of assumptions, some straightforward, and one not. First, F-structure exists; that is, there is a sequence $F_n > \ldots F_0$ that defines the maximal architecture of the clause. Second, any substructure of the maximal structure can serve as an embedded clause, if the embedding verb will allow it—in other words, there are not only small clauses, but clauses of all sizes. Third, clauses are built up in a series of steps, the "levels" of Representation Theory, and at any given level all of the items (and in particular, clauses) in the workspace are built up only to the size permitted by that level. For example, there is a level that corresponds to TP; at that level, clauses are built up to the size of TPs, but not to the size of CPs. If there is a clause with a verb in it that takes TP

complements, then the TP can be embedded under that verb at that level, and at that point, it ceases to grow; but if there is a verb which takes a CP complement, then that verb will have to wait until a later level, when the clauses in the workspace are built up to the size of CPs. This third supposition is the LEC, and it is the novel part of the theory. In Williams (2003) it is shown that some diverse effects follow from the LEC: the opacity of non-bridge verbs, constraints on possible remnant movements, controls on reconstruction, differential locality of processes targeting different Fi 's, and a generalized constraint on improper movement.

One way to think of the LEC is as a particular relation between F-structure and the "workspace": The F_n>...F_0 sequence is a "clock" that ticks from 0 to n, and at each tick everything in the workspace is built up another level. NPs and the verbal projections they are embedded in are put together from the start as per the the co-generation regime described earlier in this section. Embedded clauses and their embedding clauses on the other hand are not co-generated; they are generated separately in the workspace up to the point when the size of the clause to be embedded matches the size selected for by the verb in the embedding clause. Thought of this way, the LEC is not a separate principle, rather it is simply the "timing" built into the architecture of the theory, it is the relation of F-structure to the workspace.[6]

This architecture provides a natural bounding on processes. Simply put, you cannot move something out of a clause until that clause is embedded (how could you?). So if the target of a movement is F_i, then that movement will happen at the level corresponding to F_i, and at that level the maximum size embedded clause that there can be is F_i, so there can be no movement from embedded Spec of F_j, j>i, because Spec F_j *doesn't exist yet.* This yields improper movement effects; i.e., there can be no movement from embedded Spec CP to embedding Spec TP, because at the point where Spec TP can be targeted, there are no CPs in the workspace.

LEC also limits the terms of binary relations in two different ways. First, consider "scope reconstruction" cases like the following:

(54) a. Someone needs to be there (ambiguous; need>someone and someone> need)
b. Everyone needs to be there (non-ambiguous; everyone>need)

Some quantifiers can scope beneath *need*, others not, at least according to a widespread judgment. Let us suppose that there are F-structure positions corresponding to the scope of *someone* ($F_{someone}$) and *everyone (*$F_{everyone}$), and suppose that $F_{everyone} > F_{someone}$ (consistent with our conclusion earlier). Then we may suppose that the complement of *need* is F_{need}, where $F_{need} < F_{everyone}$. With this configuration of Fs, it will be impossible for *everyone* to have scope beneath *need*, because that complement has no scope position for *everyone.* The earliest scope position for *everyone* will be in the matrix clause. Because the scope position for *someone* is smaller, it can scope beneath *need;* schematically,

(55) Theta<-- Existential-Closure <--Case-licensing<--- Universal Quantification

complement of *need*

6 In the present theory it is not necessary—in fact not possible—to have separate workspaces for separate clauses or phases, as in Minimalist theories of the kind in Chomsky (1995). Everything is in the same workspace.

For this sort of case, the size of complement clauses automatically bounds terms of binary relations. For other cases, we need something further, but something quite natural in the setting provided by LEC. Consider, for example, the fact that *everyone* can scope out of the complement of need, but *not* out of the complement of *say:*

(56) a. John needs to talk to everyone (everyone > needs[7])
b. John said that he had talked to everyone (*everyone> says)

The problem is that there are two scope positions for universal objects, one inside of the complement of *say* (since that complement is a bigger than the complement of *need* and the complement of *need* has one) and another in the matrix clause. We may bound quantification by saying that only the lower one is available. This means that a quantifier must scope to the lowest F that allows it. This is not the same as saying that a quantifier is clause-bounded absolutely; only that if a clause is big enough, the quantifier will be bounded by that clause.

Returning to *each:* when the first term of *each* (the *scope* term) contains more than one F-structure, as here,

(57) Each: $[F_{i \ldots}F_0\ F_j \ldots F_0]$, NP

then it must be that j<i.

This last law we will apply generally:

(58) Term Bounding:
No term in a binary relation can have the form $F_{i \ldots}F_j \ldots$ unless i>j.

Term Bounding is independent of LEC, in the sense that we could keep the LEC but reject Term Bounding, but they are clearly akin. In terms of timing, the LEC says, you cannot do something until it is time (in terms of the F-clock), whereas Term Bounding would say, you must do something when it is time, so it is an 'earliness' principle, but plays such a different role in the present context that I have not adopted Pesetsky's (1989) term for it. Term bounding plays the role of principles that limit probing play in Minimalist theories, such as in "Derivation by Phase" (Chomsky (1999)).

There are two kinds of binary relations we have discussed--ones where the two terms are sisters (as with Tense), and others where one of the terms is inside of the other (as with quantification). Term Bounding only applies to the latter; or more accurately, it applies nonvacuously only to the latter.

First consider the case where the terms are sisters: at the level at which T is applied, there will be something like "$[[\text{John}]\ [\text{leave}]]_P$" in the workspace (a). Tense is binary, and it takes as its two arguments the two sisters that make up this phrase P, and it marks both of them (b). Its arguments are the two sisters themselves, not copies of them, so anything that it does to either argument is registered in the phrase P itself (c):

(59) a. workspace: $\{[[\text{John}]\ [\text{leave}]]_P\}$

7 *need>everyone* is also possible, so there must be a scope position for universal objects low enough to be in the complement of *need.*

b. T_{past} ([John], [leave]) => T_{past}>John, T_{past}>leave => $John_{Nom,}$ $leave_{past}$
c. new workspace: {[[$John_{Nom}$] [$leave_{past}$]]$_P$}

Next consider the case where one of the terms is contained in the other, as in is the case for quantifiers like *every.* In the workspace will be something like [[John [saw boy]]$_P$ (a). Every is binary. P itself is the first term. Some NP in P, *boy* here, is the second term (b). The result is (c):

(60) a. workspace: {[[John [saw boy]]$_P$}
b. every([John saw boy], boy) => every> [John saw boy], every>boy =>
[John saw every boy], [every boy]
c. new workspace: {[John saw [every boy]]}

As it happens, "every> [John saw boy]" (*every* applied to the scope argument) has no morphological consequences under COMBINE, so only the variable argument is marked with *every.*

The two cases are parallel in a several respects, but different on one point: in the second sort of case, the second argument occurs within the first, and so the question arises, how deeply within. Term Bounding is the answer. Again, Term Bounding plays the role of conditions bounding probing in some Minimalist theories.

So far, it seems like there is an intrinsic difference between *every* and *Tense:* the first has its second term inside the first, but the second has its two arguments as sisters. Actually though there is no deep difference between the two. The two possibilities correspond to the difference between weak and strong features in checking theories; the cases where the two terms are sisters are the strong cases, and the cases where the second term is inside of the first are the weak cases, and I will adopt the strong/weak terminology, though obviously the implementation is different from what it is in a checking theory.

First of all, Tense is not necessarily strong—there are cases where Tense is weak, as in, for example, (non) raising cases in Italian:

(61) T: {[arriv- Gianni]} => Tense ([arriv- Gianni], Gianni) ==> [arriv+T_{past} $Gianni_{Nom}$] ==>
arrivano Gianni

"Tense:{ ....}" is short for, "we are about to apply Tense to workspace, and it applies to *everything* in the workspace, though I here indicate only the item of interest. Next , "Tense(..., ...)" is Tense applying to its two arguments, since it is binary; the next form shows the marking imposed on the two terms and incorporated by COMBINE, and the last line shows the final form.

In this example Tense has 'probed' for the second argument; such probing is bounded by Term Bounding.

And on the other hand, it is possible that quantifiers are sometimes strong. In such a case, the variable would have to be sister to the scope; such might be the case with floated quantifiers:

(62) Both: { [[The men] [t left]]} ==> Both: ([The men], [t left]) ==> [[The men] [both t left]]

There is the further question, how is movement, such in WH movement, effected in the present theory. I will reserve this topic for a subsequent paper, making only a few preliminary remarks here. First, COMBINE effects a kind of movement, for an item having PSP=XP and any value at all for MSP. But this movement is Fn/F0 bounded, and so cannot be extended to WH movement. WH movement is similar to *every,* in that it is a binary relation (scope, variable) and in that it probes for the second argument, and in that it is governed by Term Bounding. The two problems to be further explored are the difference in the output (the second term is prefixed to the first) and pied-piping. Pied-piping is a particularly thorny problem in the present context, as the relation between the wh-marking and the second term cannot be effected by COMBINE in cases like the following:

(63) John, [[a picture of who's mother] I saw in the paper]

The marking is the *wh* on *who's*, and the second (moved) term is [a picture of X's mother], and COMBINE cannot apply the mark here because of Fn/F0 bounding—as many as 3 Fn/F0 boundaries intervene between wh and X, each marked by a pair of brackets in the following:

(64) wh > [a picture [of [[X]'s mother]]]

4.4 Does Fn/F0 bound Case/Agreement?

Bobaljik and Wurmbrand (ms) have raised the possibility that Agreement itself is clause-bounded (bounded by Fn/F0), even for small clauses. But in our terms, Case/Agreement should not be Fn/F0 bounded--only the marks it applies should be, because Agreement is one of the binary relations that stitches an NP into its VP, and we have already seen that the terms of these relations are not Fn/F0 bounded, because they are not head-seeking, only COMBINE is.

Bobaljik and Wurmbrand base their case partly on an analysis of the "long passive" in German, in which the direct object of an infinitive becomes Nominative, with only the matrix verb taking on passive morphology and agreeing with the Nominative:

(65) weil die Traktoren [t zu reparieren] versucht wurden
since the tractors (NOM) to repair tried were
'since they tried to repair the tractors'
Bobaljik and Wurmbrand (2003)

If *die Traktoren* remained in situ, then the Case/Agreement relation with the matrix Tense would breach Fn/F0. Significantly, Fn must be extremely small here; the embedded verb is does not have passive morphology, and by common accounts, that would indicate that it does not have the accusative case-assigning potential which is located in "little v"; so the embedded complement is a VP without a containing vP shell.

Their main evidence that *die Traktoren* has not remained *in situ* is based on scope considerations; if die Traktoren is replaced with a quantified NP, it can be shown that it is strictly exterior to the scope of *versucht.* That fact by itself is not enough to secure the prediction however because of the well-known phenomenon of scope reconstruction, so Bobaljik and Wurmbrand posit a principle, the Agreement-Scope Correlation, which says that an item cannot take scope lower than the

position at which it was licensed by Agreement. So, in (65) *die Traktoren* has moved out of the embedded infinitive, and it is Agreement itself which has forced the movement, not some late process of remnant-creating scrambling.

But there are grounds to question both the generality of their conclusions, and in fact the Agreement-Scope Correlation as well. First, I will put aside cases of what might be called "secondary agreement"--there are cases in which an NP in an embedded clause agrees with both its own verb and the matrix verb, and clearly remains in situ--Bobaljik and Wurmbrand cite such cases themselves. The more narrow question is, is licensing agreement Fn/F0 bound? In Icelandic, it seems not to be, as the following example suggests:

(66) a. Honum mundu sennilega virðast [þeir (vera) hæfir].
him. DAT would.3 PL probably seem [they.NOM (be) competent]
'They would probably seem competent to him.'
Icelandic, Holmberg and Sigurðsson (ms)

b. I consider [every boy sick]$_{SC}$ every > consider, *consider> every

In (66) the embedded subject has been assigned Nominative case and agrees with (only) the matrix verb, and is clearly *in situ.* So even if Movement has taken place in German in (65) it has not taken place in (66), so Fn/F0 is not a general bound on agreement, even if it is one in German. And the existence of (66) of course raises the possibility that the movement in German is not required by Agreement even in German, but is rather an optional late scrambling. In this connection it would be interesting to know whether the scope facts in the Icelandic construction were the same as the German one, but I do not. In some comparable English cases (b) though, the scope is clearly higher than the position which is licensed. I assume that the small clause in (b) is simply too small to house the universal quantifier.

Bobaljik and Wurmbrand are able to allow cases like the Icelandic one, but still exclude the German long passive in situ, by appeal to an unappealing phasal notion; specifically, Fn/F0 can be breached if the target is on the left edge of the embedded Fn. But by left edge they mean, Spec of Fn or head. The reason they cannot mean literally *left edge* is because *die Traktoren* in the German long passive *is* on the left edge, but it is a complement, not a spec or head. So there is no unitary designation that picks out "head and spec", and so their formulation contains an irreducible disjunction, always a warning to look for a better account.

Second, there are grounds to question the Agreement Scope Correlation. It is clearly violated in the following English case:

(67) &Someone seems [t to be there]

Example (67) is ambiguous, with *someone* taking scope in the matrix or in the embedded trace position, the latter being forbidden by the Agreement Scope correlation, because *someone* is clearly licensed by the matrix Tense. Bobaljik and Wurmbrand cite the example themselves, pointing out that the base position of someone is the specifier of the lower clause. But since that position is accessible, it raises the question why movement is required for this case—in the German case, movement is required exactly because the dirrect object is not accessible, but that would not apply here. Unanswered, this question challenges the Agreement Scope Correlation.

Third, I showed in section 4.3 in this section that different quantifiers behave differently with respect to Scope Reconstruction into complement infinitives—that *every* could not scope under *need,* but *someone* could. If this is so, then the Agreement Scope Correlation cannot be right, as there is no reason to expect different quantifiers to enter into different Agreement relations, and so they should uniformly show only matrix scope. Under the account I gave of *every, need,* and *someone* in raising constructions (example (54)), scope and agreement are correlated, but in a more nuanced way than under the Agreement Scope Correlation. And as a by-product of course Agreement is freed of the Fn/F0 bound, the point of primary interest here.

5. Mirrors are Fractal[8]

There are two sorts of mirror principles that one might entertain:

(68) a. Surface syntactic organization of morphemes mirrors derivation
b. Verbally attached morphemes mirror F-structure

The first says that surface syntax will reflect perfectly the order in which morphemes were entered into the derivation, and to the extent that that reflects semantics, that the surface order will transparently reflect semantic relations. This is obviously false of natural languages, though it is true of invented languages like logical languages and computer languages.

The second principle, the one that actually goes by the name "Mirror Principle", is also false, and only slightly less obviously. It says that the order of morphemes in a word will reflect the order in which those morphemes entered into the derivation. The second mirror principle is generally understood to follow from the Head Movement Constraint, which says that the agglomeration of affixes on a word must proceed in an incremental fashion, thus guaranteeing that the result will reflect the order of introduction into F-structure.

A simple example that shows both wrong is *understood.* The semantic, and derivational, relation of Past and *under* is clearly Past $>^{9}$ *under*, by any account. And yet, in the word *understood* Past is realized beneath *under.* In the context of the Head Movement Condition special rules of morphology can be introduced to derive the found form from the expected form:

(69) [[under stand] Past] ==> [under stood]

Whatever we call such rules, they clearly undermine the mirror relation between derivation and form.

In the present context there are no such rules, there is only COMBINE. So it is worthwhile to take a step back and catalog mirror effects de novo. Clearly, language exhibits mirror effects. This is why we can meaningfully say that semantics is compositional, not just with respect to derivation, but with respect to form. But in the face of examples like *understood* we must reconsider how mirrors arise. I will suggest that mirrors arise, but strictly relative to the size parameters of COMBINE.

8 The title of this section was “Mirrors are Size-Relative” until Klaus Abels suggested to me the term “fractal”, which expresses perfectly the idea that mirrors occur at different scales, but there is no over-all mirror.

9 here the caret means "has scope over", not "takes as a complement

5.1 Types of Mirrors, Intact and Broken[10]

In the spirit of Relativized Minimality, which in fact incorporates the Head Movement Constraint itself, it is worthwhile to distinguish mirrors among elements of like type (intra-mirrors), and mirrors that hold among elements of different types (inter-mirrors). The types we will consider are adverbs, Functional elements, and NPs.

To assess the presencex of inter-mirrors, we would ask, to what extent is the semantic or derivational relation between NPs and Fs, or NPs and Adverbs, or Adverbs and Fs, mirrored in their surface relation? The answer is, somewhat. In (a) for example, we see that *completely* is ungrammatical when it precedes the tensed auxiliary, presumably because *completely* modifies something smaller than TP, maybe VP. So (a) demonstrates a mirror relation between Adverb and Tense. But (b) illustrates a broken mirror: *evidently* is derivationally and semantically higher than Tense, but nevertheless can appear beneath it the surface.

(70) a. F(Tense) and Adverb: *John completely has finished.
b. F(Tense) and Adverb: John has evidently finished. *evidently* > *T*

When we look at the other inter-mirrors, we find the same thing. In the following, NP is higher than Tense in the surface, but Tense has higher scope than everybody, and for that matter, so does the negation contracted with it:

(71) F(Tense) and NP: Everybody wasn't there

And in the following, we see that the mirror of NP and Adverb is not observed, at least in the main interpretation (suddenly [everybody left]):

(72) Adverb and NP: Everybody suddenly left

So inter-mirrors are broadly not observed.

Intra-mirrors are more often observed, but in fact in each case we find broken mirrors.

For Fs, the Head Movement Condition guarantees that surface Fs will mirror one another, and elementary facts like (a) are expected:

(73) a.*Have John might t left
b. Procel$_i$ sum t_i knigata (Bulgarian)
Read have I t the book
c. *Sum procel knigata

But exactly the kind of movement shown in (a) is found, in the the phenomenon of "Long Head Movement", illustrated in (b,c); the surface order has the participle preceding the underlying order, and in fact the underlying order is ungrammatical. I will discuss Long Head Movement in more detail shortly.

10 If you are ready to agree that it is the intact mirrors and not the broken mirrors that need explanation then you may skip this section.

For Adverbs, we find a robust intra-mirror for free adverbs, as shown here:

(74) a. *Slowly John probably t left
b. Slowly John left
c. John probably left

The nonmirroring preverbal order *slowly > probably* is ungrammatical, even though (b,c) show that each of the adverbs can independently appear in the positions in which they appear in (a); clearly a mirror effect.

But again, we find broken intra-mirrors with adverbs. In Greek, the prefixal adverb *ksana-* ("again") appears attached to the main verb, and yet is understood as having the free adverb *entelos* ("completely") in its scope:

(75) O Janis ksana-moline to potami entelos *ksana> entelos*
John again-polluted the river completely

So no intra-mirror is absolute.

Within morphology proper, we find everywhere non-mirroring structures of the *understood* type. For example, in Lithuanian, Embick and Noyer (2001) finds a prefixal reflexive morpheme outside of the Tense morpheme, clearly an inversion of the underlying order T> reflexive:

(76) is-si-laikau
preverb-self-hold.hand
si + lenk-iu -> lenkiuo-si
-> *lenki-si-u
self bend (Embick and Noyer 2001)

The reflexive really is an affix, and not a clitic, as it appears inside of other doubtless affixes, the preverbal affixes. In Embick and Noyer's theory, Tense is added first, and then the problem is to prevent the prefix, which is being lowered, from attaching to the root of the verb itself, and not to the [root+T] combination. Another example is found in Malagasy, where Keenan and Polinsky (1998) finds a non-mirroring order.

(77) Tsy m+amp+atory anao ny kafe
not pres+cause+sleep you.acc the coffee
"Coffee makes you not sleep" not the expected "Coffee doesn't make you sleep"
Keenan and Polinsky (1998)

Here, the surface order cause>not>sleep is expressed on the surface as not>cause>sleep. I classify this case as a broken intra-mirror on the grounds that both *not* and *cause* are F-structure elements; *cause* has moved past *not* and attached to the root *sleep.*

In sum, the failure of different sorts of mirrors is so widespread that one wonders whether it is the failure of mirrors or the presence of them that needs explanation. Of course mirroring is a kind of

faithfulness, to use OT terminology, and will be expected to compete with other things. But what things? If the phrasal syntax of natural language was "Alphabetize the words", we would expect no mirroring, or simply minor accidental mirroring. Clearly syntax is not alphabetization, but if it were I am sure it would be rationalized as an aid to lexical retrieval. In the present context, where COMBINE is the only head-seeking operation, mirroring is competing with size parameters of morphemes and classes of morphemes, via their MSPs and PSPs. So the view I will outline in what follows is that mirroring is *size-relative;* that is, when two things are attached to the same third thing, they will attach in a mirroring order.

5.2 Long Head Movement

Long Head Movement LHM) seems to violate the Head Movement Condition, which itself is responsible for deriving mirror effects insofar as they involve F-structure:

(78) a. Procel$_i$ sum t_i knigata
Read have I t the book (Bulgarian, Rivero (1992))
b. Lennet$_i$ en deus Yann t_i al levr (Breton)
Read 3sgM has Yann the book (Breton, Roberts (2000))

In both examples, a participle has moved above its governing auxiliary. Roberts (2000) documents that the movement really is head movement, and not degenerate ("remnant") VP movement, on the grounds that these languages do have VP movement, but the movement illustrated in (78) differs from VP movement in being clause-bound, and in being blocked by certain intervening heads.

Although some (e.g. Roberts (2000)) have sought to square LHM with the Head Movement Constraint by further differentiating types by introducing new features, I think that Rivero put her finger on a more interesting feature of the construction: the auxiliaries that are moved over are always enclitic, and can never stand as the first element in a sentence. This sounds exactly like a size parameter, perhaps a MSP, the PSP here determining that the auxiliaries attracts the auxiliary to it, rather then being moved down onto it; so the specifications for the auxiliary *sum* in Bulgarian for example would be:

(79) en: PSP: _XP
MSP: $X_{word+}\_$ (i.e., clitic)

$X_{word+}\_$ designates something at least big enough to be eligible to host a clitic. Some of the auxiliaries do not trigger LHM; these will simply lack a MSP and not trigger. More than one clitic in a sentence can have the MSP shown in (79), in which case stacking will occur:

(80) Predstavili *smo mu je* t juce (Serbo-Croatian; Boscovic (1995))
Introduced are him.dat her.acc yesterday

Boscovic (see also Ackema and Čamdzić, (ms)) has argued that this is clustering, not long head movement, but from our perspective the distinction doesn't matter, as in any case the order is derived by COMBINE, and it is non-mirroring.

A further indication that the trigger for long head movement is the enclitic status of the first

auxiliary is that it is inhibited in the presence of fronted topics and WH phrases:

(81) a. *Koga poljubio je Ivan?
whom kissed is Ivan
b. Koga je poljubio Ivan

je normally triggers LHM, but not only is it not necessary (b), it is impossible (a); obviously, the fronted phrase itself satisfies the clitic's need; furthermore, LHM is a 'last resort' process, applying only when a clitic would otherwise go unsupported.

Evidence that LHM arises from COMBINE is the fact that it is clause-bounded; as mentioned before, Roberts (2000) shows that remnant VP topicalization is not an available analysis exactly on this account.

The inhibition of LHM in the presence of negation is telling:

(82) a. *Ne prorel sfim knigata. [*[correct forms?]]
not read I-have book-the
b. *Prorel ne sfim knigata.
read not I-have book-the
c. Ne sfim prorel knigata.
not I-have read book-the
I have not read the book.

The failure of LHM in (b) has led some to conclude that LHM is a case of head movement, as head movement in general is supposedly inhibited from crossing negation (no such general conclusion is possible with COMBINE). But strikingly, LHM is not possible *beneath* negation (a); furthermore, the form with no movement (c) is *grammatical. (*a) shows that it is not crossing negation which causes the problem, and (c) shows that whatever drives LHM in parallel cases without negation is neutralized with negation. The answer is simple, and is the same as for the fronted topics and WH phrases: *ne* itself satisfies the need of *sfim* to lean left, and since LHM is last resort, it does not happen. All that is required is to not give *ne* a MSP; that is, it is not itself enclitic.

5.3 Wide scope prefixes.

In Greek, verbal prefixes that in English would take very narrow scope take broad sentence-level scope, even scope over free adverbs; the Greek prefix *ksana-* ("re-": or again") illustrates:

(83) a. O Janis ksana-moline to potami entelos *ksana> entelos*
The John again-polluted the rover completely
b. John re-polluted the river completely ***re-** > completely*

The translation of the Greek example cannot be done with *re-,* because *re-* gives rise to narrow scope. Another systematically related difference between *re-* and *ksana-* is that re- has wide scope with respect to existential closure of the direct object and *re-* does not.[11]

11 See Williams (in review) for details.

This is clearly a broken intra-mirror, as *ksana-* and *entelos* are the same kind of thing--adverbs. But to preserve the mirror, one might imagine that *ksana-* was different in some way from other adverbs, and that they were invisible to its placement. But that is not so: when the other adverb is itself prefixal, the once more a mirror arises:

(84) a. Dhen tha ksana-kalo-fai edo *ksana>kalo*
not Fut again-well-eat-Pres-3rd-sing here (Rivero (1992))
b. ksana-kalo-fai = again [well [eat + T]] *again > well, T > again*

Here both *ksana-* and *kalo-* attach to the verb but in mirror fashion--*ksana-* has scope over *kalo-*. From this we conclude that *ksana-* and *kalo-* are targeting the same size thing--say, the (derived) verb stem, and that the mirror arises because of this:

(85) a. ksana- MSP: stem
b. kalo- MSP: stem

Since F-structure (and meaning) dictate that ksana- will be added to the verb after kalo, and since theey target the same size, they will stack up in a mirroring order.

Significantly though, verb-internal mirroring is not complete in Greek; elements that target something smaller than the verb stem will end up in a nonmirror relation with respect to these prefixes. The clear case of this is Tense in Greek, which is attached to the verbal root, and so winds up subordinate to these affixes, despite being higher than they are semantically and derivationally:

(86) a. Tense: MSP=root
b. T > ksana > kalo > fai ==> [ksana [kalo [fai + T]]]

The lefthand side of (b) represents the order in which things will be joined together as dictated by F-structure; and the righthand side is the resulting structure.

5.4 COMBINE positions Adverbs

Without explicitly remarking on it I have extended COMBINE to the positioning of adverbs, at least adverbs which are realized prefixally like Greek *ksana-*.[12] But let's consider the possibility that adverbs in general are positioned by COMBINE and explore the consequences.

A first consequence is that adverbs will respect the Fn/F0 boundary. And it seems true that adverbs are never lowered into clauses smaller than the clause in which they take scope; putting this in reverse, an adverb in a lower clause is never taken to have a higher clause as its scope (contrary to DPs, for which this does happen of course).

(87) [John wants [to bother Bill very much] =/=
John wants [to bother Bill] very much

12 *re-* is another story; see Williams (in review) for discussion.

But there is in fact an interesting class of cases in which an Adverb does seem to be lowered; however the circumstances in which the lowering occur are exactly in accord with COMBINE, and with the co-generation treatment of DP embedding detailed in section 4 and Williams (2003). The cases involve Adverbs like *probably* when they attach to quantifiers:

(88) a. John talked to probably every linguist
b. &John wants to talk to every linguist
c.~&John wants to talk to probably every linguist

Cinque (2001) dismisses such adverbs, which he calls Focusing Adverbs, from inclusion in the F-structure of the clause, but the considerations about to be exposed probably argue against this. The fact that *probably* occurs between a preposition and its object suggest that it is not dominated by the matrix clause, but is rather attached to *every* (as does **to probably Bill*). The (b) and (c) examples show first that its placement possibly contradicts the claim that adverbs cannot be lowered, and second, that its placement is sensitive to the scope of the adverb. Ordinarily, *every* in an embedded infinitive is ambiguous in scope, taking either clause as its scope. However, when *probably* has attached to it, it is forced to take wide scope. This is undoubtedly because *probably* itself cannot take scope beneath *want*, presumably because the F-structure of the complement of *want* is too small for the scope of *probably.* The question then is, how can this finding be squared with the idea that adverbs cannot be lowered from the clause in which they take scope?

What we have said about adverbs in the previous section, combined with what we said about the embedding of DPs, gives an answer. First, for the unambiguous reading of (c), we have an F-structure for the matrix like the following:

(89) probably > Every > ...want

Now recall that *every* itself is a binary operator, the first term of which is its scope (and corresponds to its position in the F-structure in (89), and its second term is an NP. The NP in this case is in the complement of want. Now, suppose that *probably* combines with *every* by a rule which prefixes adverbs to elements in the same F-structure, according to COMBINE, to give the binary operator "probably every"; then this binary operator relates the matrix scope, its first term, to the embedded NP, its second term, and marks the second term by prefixing it with "probably every". In such a derivation, COMBINE has never lowered the adverb; rather, the binary relation [probably every] has marked second term with "probably every", and COMBINE then will spell-out the consequences of that marking *within that term.* The net result is that COMBINE does not lower the adverb into the lower clause, despite the fact that COMBINE positions adverbs, and despite the fact that the adverb winds up in the lower clause.

So we now have two cases in which COMBINE attaches adverbs in positions that do not mirror their introduction into F-structure--the Greek prefix cases, and the attachment of adverbs to quantifiers. There is a third. When I introduced the parameters governing COMBINE, I relied on the the well-known French/English difference in Adverb/Direct Object ordering to motivate the values for the PSP, assuming that French verbs raise and English verbs do not:

(90) a. Jean T+embrasse souvent [t Marie]
"John kiss often Mary"

b. John t often [[T+kissed] Mary]

But we now have an embarrassing overlap: if COMBINE is positioning both verbs and adverbs, how can we tell whether a given nonmirroring order is the result of the positioning of the adverb or the positioning of the verb? I do not know how to give a definitive answer to this analytic question, but I do think that the facts of French are sufficient to show that both must be operative in French. In (90a) we see Tense in French attracting the verb to position higher than the adverb *souvent*, which is beneath Tense in F-structure. But we find that the same surface order holds (b) even when the adverb is *above* Tense in F-structure, for example *probablement* ("probably"), and in fact the expected result (c) is ungrammatical:

(91) a. probablement > Tense > souvent ...V
b. John T+embrasse probablement [t Marie]
c. *Jean probablement T+embrasse [t Marie]

Neither (b) nor (*c) is expected under V raising to T, given the F-structure in (a). There is in fact an absolute prohibition against any adverb between the subject and Tense, regardless of how high the adverb is. One might imagine that there are higher attractors than Tense that propel the verb to even higher positions, with the subject also moving higher; but no matter how high the verb and the subject move, there is no explanation of why *no* adverb can intervene that is based on simply the movement of the verb and the subject.

Rather, it appears that the prohibition arises from properties of the adverbs themselves. Suppose that adverbs as a class had a PSP of X^0_; that is, adverbs are suffixed to the X^0 head of the phrase they modify, by COMBINE. Then, irrespective of whether the verb raises or not, we expect the order "V+T Adv NP"[13]

(92) a. T > [souvent> embrasse Marie] ==> T>[[embrasse souvent] Marie] ==>
[[embrasse T] souvent] Marie
b. probablement > [T>embrasse Mary] ==> probablement > [embrasse+T Mary] ==>
[[embrasse T] probablement] Mary]

English adverbs will of course have PSP=XP.

The Search Rule for COMBINE also predicts that it will be the *first* X^0 host that captures the adverb; so assuming the functional structure in (a) below, and assuming that *avoir* lexicalizes perfect aspect, we expect (c), (d), and (e):

(93) a. Probablement > T > Aspect > rapidement ("rapidly") > V
b. Jean a probablement choici ses cours
John has probably chosen his courses
c. *Jean a choici probablement ses cours
John has chosen probably his courses
d. Jean a choici rapidement ses cours

13 There is more complexity here having to do with infinitives, in which a prefixal value for PSP for adverbs seem possible. See Pollock (1989) p. 377ff. for full discussion.

John has chosen rapidly his courses

Strictly speaking, Verb-raising accounts fail to predict these outcomes. Under those accounts, there must be raising for both the auxiliary and the main verb, because of (b) and (c). Given that, and given whatever triggers both raisings, there is nothing to prevent both verbs from raising above the position of the high adverb, giving (c). Again, it seems a property of adverbs themselves that settles the issue.

5.5 Mirrors in Morphology

We return now to the original case, *under[stand T],* which shows that the structure of the word is not a reliable guide to the derivation or the semantics. It seems to me that it is appropriate to use here the same strategy as in the other cases: the mirror fails because small sizes are targeted. In English, Tense targets the verb root; this is not universal--see for example the French Future discussed in section 3. For the Lithuanian reflexive case discussed earlier (example (76)), Tense targets the verb root. And for the Malagasy causative (77) the causative morpheme targets the X^0 head of its complement. See also the discussion in section 2 of *un[grammatical ity]]*.

In sum, Mirrors will arise where successive applications of COMBINE target the same size. The sizes can be various. To rehearse the cases discussed:

(94) French adverbs: PSP = X^0_ (to the right of the lexical head)
English adverbs: PSP = XP (to the right or left of XP)
Greek prefixal adverbs: MSP :_X^0 (prefix to stem)
French Present and Past Tense MSP: root_ (suffix to root)
English Tense: MSP: root_ (suffix to root)
French Future: MSP = stem, PSP=X^0
Lithuanian Tense: MSP = root, PSP=X^0
Lithuanian Reflexive: MSP =stem, PSP=X^0
French Tense MSP: root_ (suffix to root)
English Tense: MSP: root_ (suffix to root)

6. Conclusion.

If COMBINE is the sole rule which respects the Fn/F0 boundary, does it tell us something about clausehood, and other constituent types, and their F-structures? I am not sure that it does, and in addition, I think F-structure is less well understood than is generally appreciated. I have assumed that we know where the Fn/F0 boundaries are in any given sentence, but actually there is little to go on that is part of a general understanding.

A general program in linguistics for some time now is to map the $F_n...F_0$ sequence ("Cartography") and to explain the sequence. There are two broad subprograms of explanation. One is to identify intrinsic properties of elements in the sequence that account for their positioning in the sequence in terms of laws not specifically designed to regulate the sequence. Nilsen (2004) has proposed that polarity is such a property. We know for example that in a single clause, *not* must precede *yet*: *"John has yet seen nothing"/"No one has yet seen John". But this follows not from the theory of clause structure, but from a broader law, which holds across the clause boundary: "No/*some one thinks that John has yet arrived". So the theory of clause structure, and in particular of the F_n-F_0

sequence, is relieved of the duty to explain this ordering. And likewise for other cases that Nilsen discusses.

The other program of explanation proceeds by identifying substructures of F-structure with elements in an ontology of some kind, consisting of "events", "facts", "speech acts" etc., where the ontology has at least enough structure to say that events are parts of facts, and facts are part of speech acts, and not vice-versa; then, given that *completely* modifies events, and *actually* modifies facts, and *frankly* modifies speech acts, we can predict the relative ordering of these adverbs:

(95) a. John frankly actually completely startled Bill
b. *John completely frankly startled Bill
c. *John completely actually startled Bill
d. *John actually frankly startled Bill
e. But: John actually said that frankly he completely startled Bill

A problem for this program is illustrated in (e)--if we have 2 clauses, then we can get a "speech act" (modified by *frankly*) embedded under a "fact" (modified by *actually).* So the ontology must allow this. But it remains that this ordering is not possible within a single clause. So it appears that we cannot use the structure of the ontology to explain the ordering within a single clause without saying something like, the ordering in the ontology must hold within a single clause, but not across clauses. But with this qualification, we cannot really be said to be explaining something about language purely in terms of something outside of language, and we are still stuck with the clause, and the Fn/F0 boundary, as a mysterious entity. Further problems of the same kind are illustrated by "probably S", which is one clause, versus "It is probable that S" which is two; and "needn't V" which is one clause, versus "need to V", which is two, and "can V" versus "be able to V". These last examples also suggest that there is not an obvious semantic criterion for distinguishing functional verbs (F_is) from lexical verbs (F_0s).

Under what circumstances would Cartography be impossible? Consider the phrases NP and PP —and ask, what order do they occur in? There is no answer of course—either of them can embed the other. Now imagine that the F_i's of clausal structure were the same—they could occur in every possible order. Then there would be no mapping, or explaining, their order. There would at most be the project of identifying them, just as there is the project of identifying the basic parts of speech. So we might begin to understand F-structure by identifying what properties of it allow Cartography to be done on it.

It seems to me that F-structure has two properties that enable Cartography. First, it is nonrecursive, and therefore finite, and second, elements occur in a fixed linear order. In fact, it would be hard to have one of these properties without having the other, but I will treat them as separate properties. They enable Cartography in that if there were recursion within a clause, the clause itself would be infinite in its structure, and could not be mapped; and if elements could occur in any order, there would be no Cartography to do.

But in fact it seems that F-structure actually lacks both of these crucial properties. First, there is recursion within a single F-structure:

(96) ...[[[[[John eats black eyed peas] on the first day] in the second week] in every third month] in every other year] in every fifth century]....

Time modifiers of ever-increasing size can be added to (96) ad libitum, and if we ever run out of named units we use made-up ones ("every other unit of 100 million eons"). And (96) does have the stacking structure illustrated in (96), although it possibly also has a structure in which there is one very complex PP ("on the first day [in the second week [...]]". The existence of the stacking structure is confirmed by the fact that we can front a middle term of the sequence:

(97) In every other week John eats black eyed peas on Tuesday, in every third year.

This fronting would be impossible if the recursion occurred *within* a single PP. So recursion within a single F-structure exists, since on all accounts these stacked time modifiers occur in a single clause, and so F-structure is not finite.

Likewise, linearity fails to hold:

(98) John could (not) have (not) been (not) swimming

not can occur in any position in the auxiliaries except before the tensed one. It makes a different contribution to the meaning of the sentence in each position, but of course--in general, *any* difference in hierarchical structure will correspond to a difference in meaning, quite apart from the particular properties of F-structure. The point is, negation has no single privileged position of occurrence, and likewise for other elements like modals, as has been often noted.

The upshot is, the two properties that at first glance would seem to be characteristic of F-structure, and which are necessary to make Cartography possible, aren't even true of the F-structure of clauses. But of course they are roughly true, or else Cartography would indeed be impossible. But if they aren't completely true, then they are not *definitional;* that is, they do not tell us why the Fn/F0 boundary is a distinguished boundary, and they do not tell us what a clause is, as distinguished from other $F_i...F_j$ sequences, including ones which might include the Fn/F0 boundary in the middle—these too will show rough but not complete finiteness and linearity.

It was suggested in section 2 that only COMBINE is Fn/F0 bounded, and if this is so, then this the only general thing that we know about F-structure, beyond the facts of the form $F_i > F_j$ for particular i and j.

Bibliography

Ackema P. and A. Čamdzić (ms) "LF complex predicate formation: the case of participle fronting in Serbo-Croatian" University of Edinburgh

Bach E. (1979) ″Control in Montague grammar″. *Linguistic Inquiry* 10.515-531

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

Baker M. (1996). *The Polysynthesis Parameter*. New York [etc.]: Oxford University Press

Beghelli, F. and T. Stowell (1997) "The syntax of Distributivity and negation"*, in A.* Szabolcsi (ed) *Ways of Scope Taking*, Kluwer.

Bobaljik J. and S. Wurmbrand (ms) "Agreement Domains", UConn

Boskovic Z. (1995) "Participle Movement and Second Position Cliticization in Serbo-Croatian" *Lingua* 96, 245-266

Chomsky N. (1999) Derivation by Phase", MIT Occasional Papers in Linguistics, no. 18, Cambridge, MA: *MIT Working Papers in Linguistics*, Department of Linguistic and Philosophy, 1999. (Unpublished revision 1999; reprinted in *Festschrift for Kenneth Hale. Publication pending*)

Chomsky N. (1957) *Syntactic Structures*. The Hague: Mouton

Chomsky, N. (2001) "Derivation by Phase" In *Ken Hale: A life in language*, ed. By Michael Kenstowicz, pp. 1–52, Cambridge, Mass.: MIT Press

Cinque G. (2001) *Adverbs and Functional Heads*, Oxford University Press

Cinque G. (2005) "Deriving Greenberg's Universal 20 and Its Exceptions" *Linguistic Inquiry*, Volume 36, Number 3, 315–332

Di Sciullo A. M. and E. Williams (1986) "Noun Incorporation" *NELS 1985 Proceedings*

Embick D. and R. Noyer (2001) "Movement after Syntax" Linguistic Inquiry 32.4, 555-595

Emonds J. (1978) "The Verbal Complex V' – V in French" *Linguistic Inquiry* 9: 151-175

Fox D. and J. Nissenbaum (1999) "Extraposition and Scope: A case for overt QR", *Proceedings of the West Coast Conference on Formal Linguistics 18*

Holmberg A. and H. Sigurðsson (ms) "Icelandic Dative Intervention" handout, Lund University and University of Newcastle upon Tyne

Larson R. (1988). "On the Double Object Construction". *Linguistic Inquiry* 19, 335–391

Keenan E. and M. Polinsky (1998) "Malagasy" in Handbook of Morphology, A. Spencer and A. Zwicky eds., Blackwell, 563-623

Mardale, A. (ms) "Why on table is on the table?: Notes on the definites's article incorporation with prepositions in Romanian" CNRS/LLF – Université Paris 7 & Universitatea din Bucureşt

Nilsen O (2004) *Eliminating Positions: Syntax and semantics of sentence modification* Utrecht NL: LOT

Partee B. (1995) "Quantificational Structures and Compositionality" in Bach, E. Jelinek, E., Kratzer, A., and Partee, B. (eds) *Quantification in Natural Language*, Dordrecht: Kluwer., 541-601

Pesetsky D. (1985) "Morphology and Logical Form" *Linguistic Inquiry* 16. 193-248

Pesetsky D. (1989) "The **Earliness** Principle". http://web.mit.edu/linguistics/www/**pesetsky**/**earliness**.pdf

Pesetsky D. and E. Torrego (2002) "Tense, Case, and the nature of Syntactic Categories" to appear in *The Syntax of Time* Jacqueline Guéron and Jacqueline Lecarme (eds.), MIT Press.

Polinsky M. & E. Potsdam (2001) "Long distance agreement and topic in Tsez", *Natural Language and Linguistic Theory* 19.3, 583-646

Pollock Jean-Yves. (1989) "Verb Movement, Universal Grammar, and the Structure of IP" Linguistics Inquiry 20: 365-424

Riemsdijk H van. (1998) "Head Movement and Adjacency" *Natural Language and Linguistic Theory* 16:633-678

Rivero M. L. (1992) "Adverb Incorporation and the syntax of Adverbs in Modern Greek" *Linguistics and Philosophy* 15.3, 289-331

Roberts I. (2000) "Head Movement," in Mark Baltin and Chris Collins (eds) Handbook of Syntactic Theory. Oxford: Blackwells, pp. 113-147.

Williams E. (1974) *Rule Ordering in Syntax* MIT dissertation

Williams E. (1981) "Argument Structure and Morphology" *The Linguistic Review* 1.1, pp. 81-114.

Williams, E. (2003) *Representation Theory.* Cambridge, Mass: MIT Press

Williams E (2006) "Dumping Lexicalism" in *The Interfaces* G. Ramchand, ed, Oxford University Press

Williams E. (ms in review) "Derivational Morphology is Projective, not Realizational"

Wurmbrand S. (2001) *Infinitives: Restructuring and Clause Structure* Berlin/New York: Mouton de Gruyter

Zwart J-W (2006) "Complementizer agreement and dependency marking typology" in Marjo van Koppen, Frank Landsbergen, Mika Poss & Jenneke van der Wal (eds.), *Special issue of Leiden Working Papers in Linguistics 3.2,* 53-72.