Information Spreading and Levels of Representation in LFG

Avery D. Andrews
The Australian National University,
CSLI & Xerox PARC
Avery.Andrews@anu.edu.au

Christopher D. Manning
Stanford University
manning@csli.stanford.edu

A basic tenet of generative grammar is the existence of a distinction between the phonetic realization of utterances, concretely represented in terms of form, order and arrangement, and more abstract structures that play a central role in the formulation of syntactic constraints and the extraction of meaning. 'Classic' LFG (the theory and approach of the papers in Bresnan 1982) focused on two levels, c-structure, a fairly conventional form of superficial phrase-structure, and f-structure, a more abstract underlying level, similar in many respects to the S-structure of GB. These two levels are related by a correspondence function ϕ which maps each node of the c-structure to some substructure of the the f-structure. More recently, LFG workers have entertained proposals for some additional levels, especially 'argument structure' and 'semantic structure' (among others, Kaplan 1987 and Alsina in preparation). Here we will argue that the methods that have been proposed to integrate levels in LFG have substantial drawbacks, especially with regard to semantics. We will propose an alternative approach, which preserves the essential intuitions (and most of the actual analyses) of LFG, while expressing them in a manner that combines greater metatheoretical flexibility with less stipulation in the grammars of individual languages.

In the first section we will discuss some of the problems we find in the standard approach to relating levels in LFG, especially some that arise in connection with the semantics of 'scoping modifier' constructions that are, embarrassingly, handled quite straightforwardly in other 'unification-based' frameworks.[1]

[0]We would like to thank Farrell Ackerman, Joan Bresnan, Mary Dalrymple, Ron Kaplan, John Maxwell, Ivan Sag and Peter Sells for reading earlier drafts and discussion of the issues raised here. However, any remaining mistakes in approach or matters of fact are our responsibility.

[1]In spite of a recent movement to replace this term with 'constraint-based', we retain it because we think it captures the essential difference between LFG, GPSG, HPSG, etc., and other approaches such as RG or GB. Note in particular that virtually all current linguistic frameworks are based on the idea of simultaneous satisfaction of various different kinds of constraints, and are therefore appropriately described as 'constraint-based'.

In the second section we present our proposal, and show how it can be used to solve the scoping modifier problem. The essential idea is to generalize the way in which LFG spreads information. In classic LFG, different bar-levels in the tree structure (e.g., a phrasal node and its head) share information by being assigned the same correspondent in the f-structure, which automatically means that they share all f-structure information. On our approach, phrases and their heads are seen as functionally distinct, but certain kinds of information are shared in universally specified ways across certain grammatical functions, giving the effect of level-collapsing as before, but in a more flexible and (we argue) empirically more successful way. The approach can be seen as giving prime status to a structure similar to the annotated c-structure (introduced in Kaplan and Bresnan 1982 solely as an intermediate structure for use in generating f-structures). The annotated c-structure can be viewed as similar to the type of unified feature structure used in HPSG and other phrase structure grammar approaches and hence our resulting theory is a sort of hybrid of ideas from LFG and HPSG. In particular, we can maintain the strengths of traditional LFG analyses, based on grammatical functions, while implementing them in a constrained one-level syntactic theory akin to HPSG.

In the third section we move into new territory, showing how our approach lets us make some progress in analyzing 'reanalysis' and 'restructuring' constructions in Romance languages, which so far appear to have resisted insightful analysis in unification-based approaches. We will see that the combination of grammatical functions as normally used in LFG, together with our general approach to attribute-sharing provides a straightforward account of many of the basic properties of these constructions.

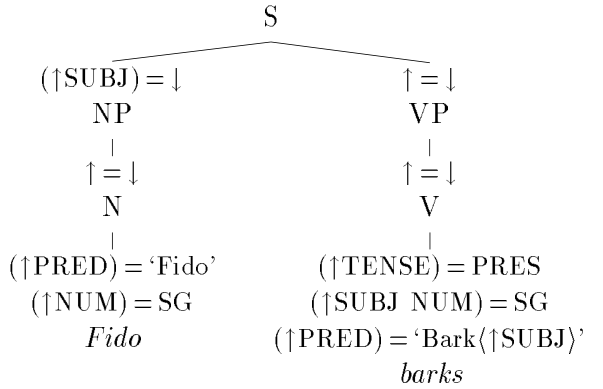
1 'Classic' LFG and some problems therein

1.1 ϕ and $\uparrow = \downarrow$

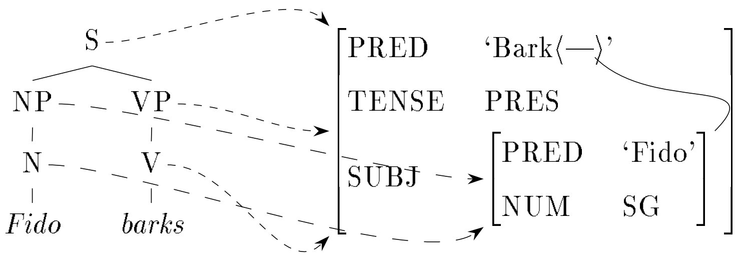
The correspondence between c- and f-structure in LFG is established in a surface-oriented fashion. Phrase-structure (PS) rules express the constraints applying to c-structure. These rules are supplemented by annotations, expressed as equations, that indicate the f-structure relationships between the correspondents of the c-structure nodes. The annotations are formulated as equations in the normal vocabulary of f-structures, together with the symbols '$\uparrow$' and '$\downarrow$', which mean 'the f-structure correspondent of the node above me' and 'my f-structure correspondent', respectively. Annotated c-structure rules produce annotated c-structures, and Kaplan and Bresnan (1982; henceforth KB) explain a method for solving the annotations that extends the c-structure to a full grammatical structure which is the minimal structure that satisfies the annotations. A variety of further constraints then apply to the f-structure portion of the full linguistic structure, thereby filtering the results of the c-structure rules (for example, agreement and subcategorization phenomena are handled in this way). Thus the sentence (1a) has the annotated phrase structure shown in (1b) which is solved to produce the full grammatical structure shown in (1c), where the dotted lines show ϕ mapping the c-structure to to the f-structure:

(1) a. *Fido barks.*

b.



c.



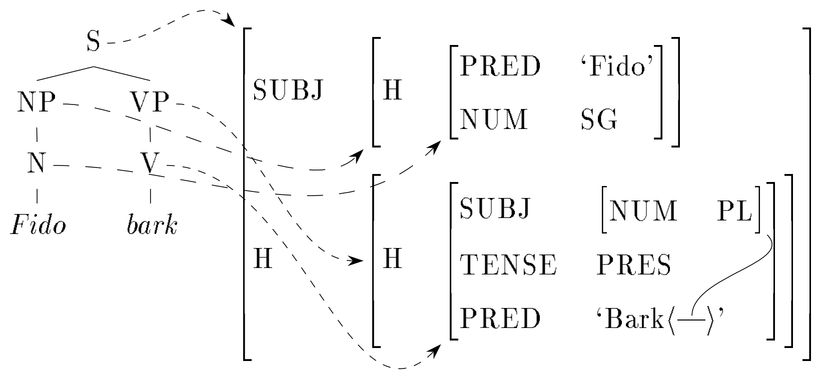
The annotation over the NP node in (1b) says that its f-structure correspondent bears the SUBJ relation to the f-structure correspondent of the node above it, which is the S. The ↑ = ↓ annotations over the remaining non-terminals assert that the nodes immediately above and below them have the same f-structure correspondent. All of this can be seen to be true in the complete sentence structure, (1c).

A particularly striking feature of ϕ is that heads and the constituents that they are heads of are typically identified in f-structure. In (1), for example, the ↑ = ↓ equations have the effect of 'collapsing levels' in the f-structure so that the S, VP and V, all have the same f-structure correspondent, as do the NP and the N. The role of ϕ in conjunction with ↑ = ↓ equations is to account for the fact that it is very common in language for certain members of a constituent, especially those that would traditionally be regarded as 'heads', to express or 'co-express' important properties of the whole constituent. Thus the gender, number and case of NPs are often marked on the head noun and the tense and mood of sentences on the verb. The properties of ϕ allow one to introduce these properties in the lexical entries of nouns and verbs and have them show up in the f-structure correspondents of the Ss and NPs they belong to.

Consider, for example, how the subject-verb agreement restriction on sentences of the form (1a) is accommodated (e.g., **Fido bark*, *dogs bark*, **dogs barks*). The lexical entry for *Fido* specifies that its f-structure correspondent has the value SG for the attribute NUM.[2] On the other hand the lexical entry for *barks* specifies that the NUM of the SUBJ is SG, while that for *bark* indicates that it is PL.[3]

Now if each individual c-structure node had its own f-structure correspondent, with, say, the traditional head of a constituent being the value of the H attribute of the whole, we wouldn't get any account of how agreement worked, because the grammatical properties of the SUBJ of the sentence would be different from those of the SUBJ of its verb (and the VP would have no SUBJ at all). This is shown in (2) for the sentence **Fido bark*. As in Chomsky (1992), feature "checking" (here unification)

(2)



is viewed as a local relation, and so something has to be done to produce a local relationship between the verb and its subject. The solution to this problem in 'Classic' LFG is to say that the f-structure correspondent of the head of a phrase is the same as the f-structure correspondent of the phrase itself. The concept of 'head' is treated as a (possibly epiphenomenal) aspect of c-structure, or maybe annotated c-structure, rather than of f-structure. When the H-values in (2) above are identified with their 'functional mothers', a conventional f-structure as in (1c) results, whereby

[2] This is not really true as agreement is semantically determined by the referent of an NP (Nunberg 1977). However, we will assume for simplicity that features of the conventional referent of an NP are marked in the lexicon. We will see later that agreement features also appear on the semantic projection, and there does not seem any difficulty in extending what is presented here to a more semantically faithful analysis.

[3] Actually, due to the nature of English morphology, things are more complicated than this, as discussed in Andrews (1990), but this description would be accurate for a language where singular and plural were distinguished by different endings, rather than presence of an ending versus its absence. Note also that for the purposes of exposition we are ignoring PERSon entirely and regarding the facts as involving purely a contrast in number.

the f-structure correspondent (ϕ-correspondent) of the head N of the subject NP is the same thing as the SUBJ-value of the ϕ-correspondent of the verb. And, by the principle that an attribute such as NUM must have a unique value, this value cannot be both SG and PL, and so agreement is checked.

This functional identification also plays an important role in semantic interpretation. In (2), for example, the lexical form 'Bark$\langle$—$\rangle$' is not connected with its argument, the subject NP *Fido*. Instead, the argument is an 'H-great-aunt'. But the identifications effected by ϕ cause the f-structure of the subject NP to be a functional sister of the f-structure correspondent of the verb.

Although this apparatus of mother-head identifications delivers an impressive amount of appropriate descriptive power for a rather small formal investment,[4] it is a bit too simple to be true. In particular, the f-structure identification of the correspondents of c-structure nodes and their heads via $\uparrow = \downarrow$ annotations erases too much grammatical information too quickly, and is both too inflexible and too stipulative. One of the deficiencies arises with agreement. Agreement of heads and determiners with the morphosyntactic features of their containing NP is normally accommodated by annotating the determiners and heads of nominal nodes (both maximal and submaximal, in X-bar theory) with an $\uparrow = \downarrow$ equation. This works out well for English, but is insufficient for most languages with NP-internal agreement phenomena, in which the agreeing features are normally spelled out on various kinds of adjectival constituents as well. These would normally be treated as members of the value of a 'set-valued' grammatical function ADJUNCT (KB:214–216), so that an NP such as the Icelandic one in (3a) would get an f-structure along the lines of (3b):

(3) a. stóran Sóvieskan kafbát
large.ACC.MASC.SG Soviet.ACC.MASC.SG sub.ACC.MASC.SG
'A large Soviet submarine'

b. $$\begin{bmatrix} \text{ADJUNCT} & \left\{ \begin{matrix} \begin{bmatrix} \text{PRED} & \text{`Large'} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{`Soviet'} \end{bmatrix} \end{matrix} \right\} \\ \text{PRED} & \text{`Submarine'} \\ \text{GEND} & \text{MASC} \\ \text{NUM} & \text{SG} \\ \text{CASE} & \text{ACC} \end{bmatrix}$$

[4] The very modest complexity of the formal ideas is revealed by the fact that, as far as we know, LFG is the only current syntactic theory for which educational grammar development systems have been produced that run on low-end personal computers.

But this structure is not good enough, because it doesn't contain anything representing the gender, number and case agreement between the NP as a whole and its prenominal attributive adjectives.

Although various solutions to this problem have been proposed, they all require some kind of special machinery to get agreement to happen, such as adding explicit annotations to adjective nodes that say $(\uparrow \text{NUM}) = (\downarrow \text{NUM})$, and so on. In the present architecture of LFG, this special machinery looks like some kind of marked excrescence on the grammar of a language, rather than a natural and expected property of the construction. But agreement is universally just as characteristic of attributive adjectives as it is of head nouns and determiners, so this kind of difference in treatment is not very well motivated.

Another problem is that the conventional treatment understates the degree of linkage between c- and f-structure. One of the original design goals of LFG was that f-structure would be a level abstracted from surface structure at which important universal properties of language could be stated without regard to the variation in c-structure across languages. However, it now seems clear that this separation of form and function was carried too far. F-structures are related to c-structures by equational annotations that are attached more or less *ad libitum* to phrase-structure rules, and although there are some observations about what kinds of combinations occur, there does not seem to be any clear-cut theory of what is possible. Why, for example, does sharing of a major category feature between mother and daughter always seem to imply f-structure sharing (i.e., functional identification) as well?[5] Standard LFG has not provided a good basis for tackling this kind of question in a productive way.

1.2 The introduction of semantics

In addition to the above metatheoretical difficulties, when one tries to introduce semantics into LFG, serious descriptive problems arise. Probably the most important task of syntax in conveying meaning is to associate the argument-positions of verbs with the linguistic material that characterizes the arguments. And the grammatical functions (GFs) defined in f-structure play an essential role in doing this. In particular, the lexical form of a verb or other predicator specifies the grammatical relations to be borne by the various arguments of the verb, and the rest of the syntax then specifies how these grammatical relations are expressed in terms of linear order, case-marking, agreement and whatever else might be relevant.

[5]There are at least apparent counterexamples to the converse, such as nominally headed sentences in Warlpiri (Simpson 1991), and verbally headed NPs in English (Schachter 1976), though recent work in X-bar theory has tried to remove some of these anomalies (Abney 1987).

Thus the original view in LFG was that semantic interpretation was done off f-structure: "the f-structure is the sole input to the semantic component" (KB:175, see also Halvorsen 1983). In later approaches to semantic interpretation in LFG, such as Kaplan (1987), semantic structure (σ-structure) is thought of as being a level of structure similar in nature to f-structure, but formed with different primitives. An argument-taking logical predicate is represented by the predicate being the value of the feature REL and the arguments values of indexed features ARGn. Just as each node in c-structure has an f-structure correspondent, so do most substructures in f-structure have σ-structure correspondents, which are furthermore specified by means of a correspondence σ from f- to σ-structure that has a lot in common with the c- to f-structure correspondence ϕ.

In one typical notation (Kaplan et al. 1989), parts of semantic structure corresponding to parts of f-structure are designated by prefixing the relevant f-structure designating expression with σ. Here, for example, is what might be added to the lexical entry for *barks* to get semantics for a sentence such as *Fido barks*:[6]

(4) $(\sigma{\uparrow}\mathrm{REL}) = \mathrm{bark}$
$(\sigma{\uparrow}\mathrm{ARG1}) = \sigma({\uparrow}\mathrm{SUBJ})$

With the aid of a $(\sigma{\uparrow}\mathrm{REL}) =$ "fido" equation in the lexical entry for *Fido*, we will get the following f- and σ-structures for *Fido barked* (ignoring grammatical features such as tense and number):

(5) $$\begin{bmatrix} \mathrm{SUBJ} & \begin{bmatrix} \mathrm{PRED} & \text{`Fido'} \end{bmatrix} \\ \mathrm{PRED} & \text{`Bark}\langle\text{---}\rangle\text{'} \end{bmatrix} \qquad \begin{bmatrix} \mathrm{ARG1} & \begin{bmatrix} \mathrm{REL} & \text{``fido''} \end{bmatrix} \\ \mathrm{REL} & \mathrm{bark} \end{bmatrix}$$

This may be regarded as an extension of the structure (1c) in the previous section (but we refrain from presenting everything in one place, as the result becomes rather cluttered).

There are however certain problems with this approach. A metatheoretical one is that it involves a lot of redundancy. For example, REL and ARG attributes can only be found in σ-structures, PRED and SUBJ attributes in f-structures. Nevertheless we are forced to provide in every equation stipulations of exactly which projection all attributes are attributed to. When everything is spelled out in full (as in Fenstad et al. 1987) this begins to get rather messy. We would like to be able to state such attributions only once, and the architecture we introduce in Section 2 will allow (indeed, require) us to do this.

Tough descriptive problems also appear. It was shown above that the f-structure 'flattens' various kinds of nesting relationships found in

[6] Note how on the left of these equations one applies σ to $\uparrow$ (which designates a piece of f-structure) to get the semantic correspondent of that f-structure (i.e., we calculate $((\sigma\uparrow)$ REL) etc.), while on the right of the second equation one first finds the SUBJ of a functional structure, and then the σ-correspondent of that.

c-structure. For example, the V, the VP and the S all have the same f-structure correspondent, and therefore the same σ-structure correspondent. Often this flattening doesn't matter, but sometimes it does. A case where it makes serious trouble is in the interpretation of 'concentrically scoped' modifiers in examples such as these (Andrews 1983):

(6) a. An alleged former racketeer was elected.

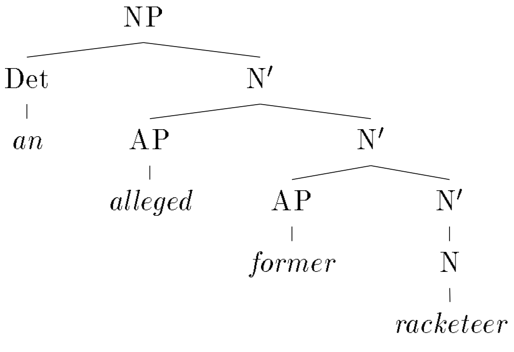
b. John knocked three times twice.

c. John knocked twice three times.

In most unification-based theories, these examples are completely unproblematical (see, for instance, Pollard and Sag forthcoming, Ch. 8). But the use of level-flattening and sets to represent adjuncts creates severe difficulties for LFG. The prenominal adjectives and postverbal adverbs would standardly be treated as members of a set-valued ADJUNCT GF, producing an f-structure such as this for (6a):

(7)
$$\begin{bmatrix} \text{SPEC} & \text{INDEF} \\ \text{ADJUNCT} & \left\{ \begin{array}{l} \begin{bmatrix} \text{PRED} & \text{`Alleged'} \end{bmatrix} \\ \begin{bmatrix} \text{PRED} & \text{`Former'} \end{bmatrix} \end{array} \right\} \\ \text{PRED} & \text{`Racketeer'} \end{bmatrix}$$

This f-structure would be the correspondent of all the nominal nodes in a c-structure like this:

(8)


But this f-structure is obviously a hopeless basis for semantic interpretation, since the semantic structure is clearly concentrically nested in the same way that the c-structure is: *former* can be regarded as an 'operator' that applies to a descriptive term, forming a new descriptive term that describes things that once satisfied the description provided by the operand (when *former* applies to *racketeer*, we get a term describing entities that at one time in the past satisfied the term *racketeer*). *Alleged* is likewise an operator applying to descriptions: it is applied to entities which someone has alleged to satisfy the operand description. Each N′

in the c-structure corresponds to a descriptive term in the semantics, and the order of adjectives determines the semantic interpretation correspondingly: an *alleged former racketeer* is different from a *former alleged racketeer* (this concept not being as pragmatically felicitous).

A similar story can be told about the frequency adverbs in examples (6b–c), which arguably involve a rule VP $\rightarrow$ VP AdvP, the semantics of the AdvP determining how the kind of action described by the lower VP is related to that of the higher one. In particular, if the AdvP is a frequency adverb, the kind of event designated by the higher VP is the kind designated by the lower, repeated the appropriate number of times. Hence the order of frequency adverbs determines the rhythm (note in particular that the semantic scope is not determined by linear order, but by c-structure nesting relations, as discussed more extensively in Andrews (1983)).

It should be clear that accounting for the semantics of these constructions requires a direct relationship from c-structure to σ-structure, or at least one not mediated by the usual kinds of f-structures. For on the latter view, once two of the nominal nodes in (8) have been functionally identified, they cannot have distinct semantic correspondents, which leads to the absurd view that the nominals in the completed sentence structure all have the same meaning. It is hard to see how a reasonable semantic structure could be obtained at all under these assumptions.

There are other indications that the collapsing of syntactic structure induced by the conventional use of $\uparrow = \downarrow$ equations goes too far. Relative clauses are presumably some kind of adjunct (so that the whole NP can have the same number as the head, for example). But multiple relative clauses are not entirely equivalent in function:

(9) a. The headway that they made on the project that was summarized in the report was insufficient to secure further funding.

b. *The headway that was summarized in the report that they made on the project was insufficient to secure further funding.

It seems that the 'idiom-chunk' *headway* must be licensed by a verb *make* in the inner relative clause, rather than the outer. This seems consonant with some sort of concentrically nested structure, but difficult to understand if the two relative clauses are just multiple values of an ADJUNCT function.

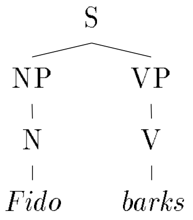
An intuitively appealing approach to these problems is to have the semantic correspondence σ come not off f-structure, but off c-structure. Then there could be semantic units that corresponded to constituents, but not necessarily to actual f-substructures. This approach has been advanced in Halvorsen and Kaplan (1988). Consider for example (8). In this structure, each nominal node will have a semantic correspondent, and the various N′s at least will have *different* correspondents, representing the senses of *racketeer*, *former racketeer*, and *alleged former racketeer*. It should be reasonably clear that we can get an analysis of concentrically scoped 'operator' modifiers on this kind of approach, without working through any particular analysis (one will be presented in the next section).

But projecting both ϕ and σ separately off the c-structure also leads to technical difficulties. To appreciate them, reconsider *Fido barks*, with lexical entries such as those below, which have been modified so that the up and down arrows refer to c-structure nodes rather than their f-structure correspondents, and the ϕ correspondence is mentioned explicitly in the same way that σ is:

(10) a. *Fido*, $(\phi\uparrow \mathrm{NUM}) = \mathrm{SG}$
$(\phi\uparrow \mathrm{PRED}) = \text{'Fido'}$
$(\sigma\uparrow \mathrm{REL}) = \text{“fido”}$

b. *barks*, $(\phi\uparrow \mathrm{TENSE}) = \mathrm{PRES}$
$(\phi\uparrow \mathrm{SUBJ\ NUM}) = \mathrm{SG}$
$(\phi\uparrow \mathrm{PRED}) = \text{'Bark}\langle\uparrow\mathrm{SUBJ}\rangle\text{'}$
$(\sigma\uparrow \mathrm{REL}) = \mathrm{bark}$
$(\sigma\uparrow \mathrm{ARG1}) = \sigma(\phi\uparrow \mathrm{SUBJ})$

We want our sentence to then have a full structure as represented below (correspondences from c- to f- and σ-structure omitted, to reduce clutter):

(11) 

f-s: $\begin{bmatrix} \mathrm{SUBJ} & \begin{bmatrix} \mathrm{PRED} & \text{'Fido'} \\ \mathrm{NUM} & \mathrm{SG} \end{bmatrix} \\ \mathrm{TENSE} & \mathrm{PRES} \\ \mathrm{PRED} & \text{'Bark}\langle\text{—}\rangle\text{'} \end{bmatrix}$

σ-s: $\begin{bmatrix} \mathrm{REL} & \mathrm{bark} \\ \mathrm{ARG1} & \begin{bmatrix} \mathrm{REL} & \text{“fido”} \end{bmatrix} \end{bmatrix}$

But it doesn't quite work.

The reason is that we haven't actually managed to connect the semantics of the argument NPs to the argument positions of the verb. The final equation in (10b) says that the value of ARG1 in the semantic correspondent of the (V dominating the) verb is the same as the semantic correspondent of the SUBJ. But, this doesn't work anymore, because σ is no longer defined from f-structure to σ-structure, but directly from c-structure to σ-structure.

To fix this, one might propose introducing an f-to-σ-structure correspondence alongside of the others, which we might call σ^*. The needed σ^* links could be established by associating with NPs an annotation such as this:

(12) $\sigma^*(\phi\uparrow) = \sigma\uparrow$

The effect of this little annotation would be to add σ^* lines from the subsidiary f-structures to the appropriate subsidiary σ-structures.[7] Note that in a structure like (8), it could be associated only with the NP (i.e., the maximal nominal): if it appeared on the N′'s, they would all get the same σ-structure, and we'd be back where we started.

However, adding a new projection with accompanying annotations like (12) is yet another stipulation in a situation that is already unacceptably laden with stipulations, and is best taken, we think, as an indication that there is a basic flaw in the entire conception. By the time σ^* is added, the correspondence relations are being represented and thought about in a manner that makes them indistinguishable from ordinary attributes. ϕ and σ are acting just like ordinary feature attributes, whose values contain further attributes that have similar functions and behavior (similar to the SYNTAX and SEMANTICS attributes in HPSG (Pollard and Sag 1987), or the distinguished attributes of FUG (Kay 1979)). But the present implementation is unwieldy, essentially because you often need to access one kind of attribute through the value of another kind of attribute, a move that this representation does not easily support.

Therefore in the next section we will describe an alternative approach in which different kinds of attributes are distinguished not by where they are placed, but in terms of how they behave with respect to certain universal principles of value-sharing. Subsequently, we will find that the basic idea of this alternative, to implement the idea of information spreading in terms of conventions for value-sharing between different levels of a structure, not only provides reasonable analyses of nominal and verbal adjuncts but also helps explain certain other sorts of phenomena that have long resisted insightful analysis in LFG and other unification-based frameworks. In addition, the resulting picture becomes interestingly similar to certain recent notions of feature percolation within the principles-and-parameters approach, such as Grimshaw (forthcoming).

2 Information spreading

Our approach depends on a generalization of the standard notion of information-spreading in LFG, which is most often achieved by having a node and one of more of its daughters share the same f-structure. Our notion on the other hand is more like the usual generative conception of feature percolation. To see what this concept amounts to, reconsider the f-structure of (2) that we briefly contemplated in the previous section for **Fido bark*:

[7] Note that σ^* is to be thought of as going directly from f-structure to semantics, rather than being defined as σ composed with the inverse of ϕ, since this latter approach has various problems, such as in handling unspecified (non-overt) subjects.

(13) $$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{H} & \begin{bmatrix} \text{PRED} & \text{`Fido'} \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix} \\ \\ \text{H} & \begin{bmatrix} \text{H} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{NUM} & \text{PL} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`Bark}\langle\text{—}\rangle\text{'} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The problem with this structure is that since the verb's subject information and the subject NP's information never get combined, the contradiction in NUM-values never gets discovered.

But suppose we had a convention to the effect that the traditional f-structure features such as SUBJ, NUM, etc., were 'shared across' H, that is, a convention that if x is the H-value of y, and F is an f-structure attribute, then $(x\ F) = (y\ F)$. Now, if the f-structure attributes include SUBJ, TENSE, NUM and PRED, but not H, the application of this convention will turn the f-structure corresponding to *Fido barks* (like (13) except that the verb specifies [NUM SG]) into something very similar to a conventional f-structure:

(14) $$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`Fido'} \\ \text{NUM} & \text{SG} \\ \text{H} & \ldots \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{PRED} & \text{`Barked}\langle\text{SUBJ}\rangle\text{'} \\ \text{H} & \ldots \end{bmatrix}$$

In particular, if we erase or ignore the H attributes and their values, as suggested by the ellipsis dots in (14), then it looks exactly like an f-structure.

Our approach will in fact be to use such conventions for value-sharing to do the work of the $\uparrow = \downarrow$ annotation. We shall say that a *projection* is defined by specifying two sets of attributes, the first set being those that are to be shared, the second the set that they are to be projected across. Now, although projections can achieve the effects of the $\uparrow = \downarrow$ annotation, they also offer promise of escaping some of its limitations, because although they do effect identifications, they don't do so by simply obliterating levels of structure, but rather by identifying some of the information found at different levels. This offers a prospect of getting useful effects by setting up a number of different projections.

On the other hand we retain one of the guiding intuitions of LFG, that certain sets of nodes are equivalent for certain purposes (by virtue of corresponding to the same f-structure). In fact we will see that we can reconstruct the conventional levels as sets of nodes (essentially, quotient sets) between which certain classes of attributes are shared, and that, for certain purposes, it is useful to do so. In the remainder of this section, we will first present the basic formal idea, then present our approach to semantic representation, and finally develop an analysis of 'scoping' adjectives such as *former* and *alleged*.

2.1 Expressions and resolution

Let us suppose that a c-structure node corresponds not to an f-structure node as currently conceived, but to a richer feature structure that specifies semantic and phrase-structure categorical properties as well, and thus resembles (13), but with even more information specified. Projections will then be used to effect the identifications achieved by $\uparrow = \downarrow$ in standard LFG, without erasing the constituency information needed to construct the semantic interpretation.

We will call these new feature-structures 'Expressions', and for the present we will be a bit vague about the exact nature of the mapping whereby c-structure nodes correspond to expressions. There are two main possibilities. One is that expressions are basically just like f-structures, objects distinct from c-structure nodes, but related to them by a correspondence relation. The other is that expressions are just c-structure nodes with additional properties beyond those of phonology and constituency. Under this conception, we can think of a representation like the annotated c-structure being the complete representation, as indicated in (15), where dotted lines indicate regions of sharing of the traditional f-structure features:

(15)

S

Left daughter of S:
$(\uparrow \text{SUBJ}) = \downarrow$
NP
N
$(\uparrow \text{REL}) =$ "fido"
$(\uparrow \text{NUM}) = \text{SG}$
Fido

Right daughter of S:
VP
V
$(\uparrow \text{TENSE}) = \text{PRES}$
$(\uparrow \text{SUBJ NUM}) = \text{SG}$
$(\uparrow \text{REL}) = \text{bark}$
$(\uparrow \text{SUBJ}) = (\uparrow \text{ARG1})$
barks

At the moment it is unclear to us whether anything empirical hinges on this and its attendant issues.[8]

[8] One relevant area is the treatment of discontinuous constituents. If we take the approach to discontinuity proposed in Simpson (1991), whereby distinct c-structure nodes can be annotated as, say, SUBJ of their clause, it would seem natural to distinguish an expression from the c-structure nodes corresponding to it, so that two distinct

Regardless of which approach is assumed, we can continue to use annotated c-structures as before. On the first (standard LFG) interpretation, the $\uparrow$ and $\downarrow$ arrows will be interpreted as desgnating the expression corresponding to the node above and below the arrow, respectively; on the second they will be interpreted as designating these nodes themselves. Although most (and perhaps all) linguistic issues don't seem to be affected by the choice, we will implicitly assume the former conception by generally drawing expressions in a feature structure format, as used in standard LFG.

Below are some lexical entries that might be proposed under this view:

(16) a. *Fido*, $(\uparrow\text{CAT}) = \text{N}$
$(\uparrow\text{NUM}) = \text{SG}$
$(\uparrow\text{REL}) = $ "fido"

b. *barks*, $(\uparrow\text{CAT}) = \text{V}$
$(\uparrow\text{TENSE}) = \text{PRES}$
$(\uparrow\text{SUBJ NUM}) = \text{SG}$
$(\uparrow\text{REL}) = \text{bark}$
$(\uparrow\text{SUBJ}) = (\uparrow\text{ARG1})$

c. *bark*, $(\uparrow\text{CAT}) = \text{V}$
$(\uparrow\text{TENSE}) = \text{PRES}$
$(\uparrow\text{SUBJ NUM}) = \text{PL}$
$(\uparrow\text{REL}) = \text{bark}$
$(\uparrow\text{SUBJ}) = (\uparrow\text{ARG1})$

They are basically like the ones we had before, but with the σ and ϕ correspondences left out. One other change is the omission of PRED-features. As is discussed later, once one has semantic structures and explicit linkages between grammatical functions and argument positions, there is no need to have PRED-features as well.

With these lexical entries and appropriate phrase-structure rules, *Fido barks* will have a full expression like this, before projections are formed, assuming a two bar-level X-bar theory, with S = V^2:

---

c-structure nodes could correspond to one sub-expression, bearing the GF SUBJ, just as in standard LFG.

(17)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{H} & \begin{bmatrix} \text{H} & \begin{bmatrix} \text{CAT} & \text{N} \\ \text{BAR} & 0 \\ \text{REL} & \text{“fido”} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{BAR} & 1 \end{bmatrix} \\ \text{BAR} & 2 \end{bmatrix} \\ \text{H} & \begin{bmatrix} \text{H} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{REL} & \text{bark} \\ \text{ARG1} & [\quad] \\ \text{CAT} & \text{V} \\ \text{BAR} & 0 \end{bmatrix} \\ \text{BAR} & 1 \end{bmatrix} \\ \text{BAR} & 2 \end{bmatrix}$$

Suppose now that CAT and all the familiar f-structure and semantic features, but not BAR, are shared along H.

If we display the results of this sharing, representing H-values with ellipsis dots, we get:

(18)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{REL} & \text{“fido”} \\ \text{NUM} & \text{SG} \\ \text{CAT} & \text{N} \\ \text{H} & \ldots \end{bmatrix} \\ \text{ARG1} & [\quad] \\ \text{TENSE} & \text{PRES} \\ \text{REL} & \text{bark} \\ \text{CAT} & \text{V} \\ \text{H} & \ldots \end{bmatrix}$$

The effecting of the required sharings will be called ‘resolution’.

The effect of resolution on (17) is to identify the SUBJ-values at the levels of the V, VP and S, as well as the feature values at the three nominal levels. So agreement is obviously checked, and the semantics of the subject is brought into the correct relationship to the REL of the verb.

In fact, if we suppress the nonsemantic attributes, and forget about tense and number, we clearly get the kind of semantic structure we want.

It should be apparent that this approach is very similar to standard LFG. In both cases, it is claimed that a central mechanism of grammar is (in effect) the sharing of attributes between different levels of the linguistic structure, especially, heads and the phrases that they are heads of. But while standard LFG locates the shared attributes in an f-structure, and then asserts that different nodes have the same f-structure, we relate the nodes by what are in effect 'head' relations, and then assert that certain attributes are shared across these head relations. Due to the central role of such 'vertical' attribute-sharing, standard LFG analyses can be transferred into this framework without significant change.[9] But, as we shall see, our approach confers a greater degree of flexibility, which provides analyses of further phenomena which have proved recalcitrant for standard LFG.

There are various approaches one might take to formalizing Resolution. For example, one possibility would be to implement it as meta-rules in the style of GPSG, or even as macro-expansions.[10] We do not presently have a good grip on either the empirical or mathematical issues relevant to the choice, so we will present one we find straightforward and linguistically appealing, based on some recent work by Johnson (1991). He proposes to replace the standard function application notation of LFG by the use of relations. An attribute is a relation between two f-structure objects, so that, for example, $(x \text{ SUBJ}) = y$ is taken as an alternative notation for $\text{SUBJ}(x, y)$, where SUBJ is taken as the name of a relation.[11]

An ordinary single-valued grammatical function F is then asserted to obey a single-valuedness postulate:

$$(\forall xyz)((\text{F}(x, y) \wedge \text{F}(x, z)) \rightarrow y = z)$$

'Set-valued' attributes such as ADJUNCT can easily be modelled by not subjecting them to an axiom of this form (note that for the strictly syntactic uses of set-valued attributes, the notions of nullity, union, etc. that Johnson provides are not in fact necessary). A functional description is assembled in the usual way, and can be treated either as a collection of

[9] In fact, so far, the required changes are the strictly mechanical ones of replacing $\uparrow = \downarrow$ annotations with $(\uparrow \text{H}) = \downarrow$ annotations, and replacing PRED equations with appropriate combinations of REL and ARG equations. This amounts to no substantive linguistic change at all.

[10] That is, the linguist writes 'Head($\uparrow$, $\downarrow$)', and the grammar compiler expands this to '$(\uparrow\text{H}) = \downarrow \;\&\; ((\neg(\uparrow \text{SUBJ}) \;\&\; \neg(\downarrow \text{SUBJ})) \vee (\uparrow\text{SUBJ}) = (\downarrow\text{SUBJ})) \;\&\; \ldots$'. For example, such an implementation scheme could be used in PATR (Shieber 1986).

[11] In even more recent, as yet unpublished work, it is proposed to replace this treatment with a 3-place 'arc' relation, relating two f-structures and a relation label. This move permits atomic values to function as grammatical relation labels, and makes the formulations of universal principles smoother, but it is not relevant for our present concerns.

assertions about the denotata of a collection of special ad-hoc constants (one for each f-structure object that the functional annotations assert to exist), or as an existentially quantified conjunction of assertions:

$$(\exists x_1 x_2 \ldots x_n)(\mathrm{SUBJ}(x_1, x_2) \wedge x_1 = x_3 \wedge \ldots)$$

This assertion is then conjoined with the assertions comprising the theory, and the structure of the sentence is the minimal model that satisfies the full set of assertions. The significant result is that as long as all of the conjoined formulae take the form of a collection of existential quantifiers followed by a collection of universal quantifiers followed by a quantifier-free formula, the consistency of the conjunctional is decidable (as an NP-complete problem). Such formulae constitute the 'Schönfinkel-Bernays' class (Johnson credits Harry Lewis with pointing out the relevance of the S-B class and its properties).

One might then propose to add projections as follows: if the set of attributes **A** is supposed to spread across the set of attributes **B**, we simply add an axiom of the following form for each **a** in **A** and **b** in **B**:

(19) $$(\forall xyz)(\mathbf{b}(x, y) \rightarrow (\mathbf{a}(x, z) \leftrightarrow \mathbf{a}(y, z)))$$

Since (19) is Schönfinkel-Bernays, we can add it to LFG without fear of creating undecidability.

But we aren't quite done yet, for the reason that LFG depends not just on the collection of assertions having a model, but on it having a unique minimal model, which is the structure of the sentence described by the assertions.[12] Restricting ourselves to S.-B. formulae clearly won't guarantee this, since such a formula might demand that some attribute, say, CASE, take either of two values, such as ERG or DAT. The resulting description will then not have a unique minimal model, but only two minimal ones.

This has always demanded a somewhat special treatment of disjunctive specifications, and has been dealt with by interpreting disjunction as a 'free-choice' operator between structures. That is, it is sufficient for there to be a minimal model for there to be one for every consistent model resulting from the selection of one disjunct in each disjunction. But this reveals a potential problem for the interpretation of conditionals in LFG. For if we regard $A \rightarrow B$ as logically equivalent to $\neg A \vee B$, the interpretation of disjunction will (ceteris paribus) lead to two possible structures for a conditional, one where A is false, and another where B is true. But that is not what we actually want, since we only want our conditional to cause B to be true[13] when something other than the conditional causes A to be

[12] Both this problem and the solution were pointed out to us by Ron Kaplan and John Maxwell.

[13] That is, to go into the description of which the sentence-structure is to be the minimal model.

true (though, of course, something entirely unrelated to the conditional might independently cause B to be true).

The solution to this problem is to interpret a conditional $\mathrm{A} \rightarrow \mathrm{B}$ as being equivalent to $\neg\mathrm{A} \vee (\mathrm{A}_c \,\&\, \mathrm{B})$, where the subscript c indicates that A is to be a constraining or 'nonconstructive' specification rather than a constituting or 'constructive' one.[14] Further, negations are assumed to always be nonconstructive, so $\neg\mathrm{A}$ is also a constraining rather than a constituting specification. Therefore, the conditional allows two structures, one where A is false; and another where some independent factor makes A true (since the conditional specifies A nonconstructively), and in addition, B is true (specified constructively).

Interpreting conditionals in this way allows them to be integrated into LFG processing with the desired semantics. Of course, various further conditions will presumably have to be imposed in order to assure reasonably efficient processing. For example, things will go much better if all of the substructures mentioned in the antecedent of the conditional are connected by the relations mentioned there, since this will greatly restrict the number of ways in which each conditional has to be instantiated in order to test it for applicability.

It is easy to see that (19) can be expressed as a pair of conditionals, interpreted in this manner. It therefore does not create conceptual or foundational problems for LFG. Introducing such 'constructive conditionals' (so-called because they can add information to the sentence-structure) does however raise a problem in linguistic metatheory: what is the role of these conditionals? Do they appear in language-particular grammars, or only as universal principles? In this paper we will only propose them for universal principles, though there are some intriguing possibilities for language-particular application as well. But to make serious proposals along these lines, one would need to have a restrictive theory of what kinds of conditionals are found in natural-language grammars, something we will not attempt to provide here.

2.2 Semantic representation

Since one of our major concerns is achieving a better integration between semantics and syntax, one of the things we have to do is adopt some definite ideas about semantic representation. We have so far followed the *de facto* unification-grammar standard of using attribute-value structures to mimic conventional logical forms, with attributes such as REL, ARG1, and so forth. But there is now a large and ever-growing body of work, in a variety of syntactic frameworks, which makes it quite clear that the traditional atomic REL-value with some blandly labelled arguments is

[14] Constructive specifications are part of the description that the sentence-structure is going to be the minimal model of; constraining ones serve as filters on the structure just chosen. See KB:209–210.

quite unable to support the kinds of principles that seem to be involved in the syntax-semantics interface.[15] What this work seems to indicate is that lexical items have some kind of internal structure in terms of semantically (more) primitive notions, similar to what was envisioned in Generative Semantics and the work of Gruber (1965).

Our approach will be to suppose that this internal structure is present within the values of the REL attributes. For instance, one could adopt the proposals of Jackendoff (1990). Then, in the representation of *Harry buttered the bread*, the REL value for the main predicate would be something like (Jackendoff 1990:129,164):

(20)
$$\left[\begin{array}{ll} & \mathrm{AFF}_{+\mathrm{vol}}([\quad]_i, [\quad]_j) \\ _{event} & \mathrm{CAUSE}(\ [\quad]_i, [\mathrm{INCH}\ [\mathrm{BE}([\mathrm{BUTTER}], [\mathrm{ON_d}\ [\quad]_j])]]) \end{array}\right]$$

Here the empty brackets with subscripts represent positions into which arguments will be inserted, the subscripts being Jackendoff's notation for indicating that positions are to be unified.

The syntactic structure supplies fillers for the argument-positions via unification: A lexical entry for this sense of *butter* with grammatical relations specified for the arguments would look like this:

(21)
$$\left[\begin{array}{ll} \mathrm{REL} & \left[\begin{array}{ll} & \mathrm{AFF}_{+\mathrm{vol}}([\ ], [\ ]) \\ _{event} & \mathrm{CAUSE}([\ ], [\mathrm{INCH}\ [\mathrm{BE}([\mathrm{BUTTER}], [\mathrm{ON_d}\ [\ ]])]]) \end{array}\right] \\ \mathrm{SUBJ} & [\ ] \\ \mathrm{OBJ} & [\ ] \end{array}\right]$$

REL is here functioning more or less as an equivalent to the traditional PRED feature, but the intended interpretation of our REL is hopefully clearer than that of PRED: its value is a semantic representation. We do not have an 'argument list' as in a PRED-value, but rather, certain positions within the semantic structure are equated with the values of grammatical relations. In what follows, we will abbreviate REL values by means of predicate symbols in the obvious way:[16]

(22)
$$\left[\begin{array}{ll} \mathrm{REL} & \mathrm{Butter}([\ ], [\ ]) \\ \mathrm{SUBJ} & [\ ] \\ \mathrm{OBJ} & [\ ] \end{array}\right]$$

The notational differences from standard LFG (no PRED-attribute, no quotes around the value of REL, no GFs appearing in argument-lists) are

[15] Among others, Jackendoff (1990), Grimshaw (1990), Van Valin (1990), Baker (1988), and Bresnan and Kanerva (1989).

[16] We continue to assume that semantic representations are typed (as indicated by the *event* notation, above) but we will generally omit the type.

intended to remind one of the theoretical difference that the value of REL is simply a semantic representation. Later on, we will sketch a Linking Theory whereby lexical items don't need to specify what grammatical functions express what arguments, but only specify argument positions that require linking.

It should be clear how the proposals of the previous section carry over to our modified framework: in (17), the SUBJ attribute introduced by the lexical entry of *bark* should be changed from being equated with the value of an ARG1 attribute, to being equated with a position within the REL value for *bark*. The effect of resolution will then be to propagate the SUBJ attribute up to the top level, causing its value to unify with that of the syntactically introduced subject. Meanwhile, the RELs will also propagate up to their 'maximal projections', so that the final result will be:

(23)
$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{REL} & \text{“fido”} \\ \text{NUM} & \text{SG} \\ \text{CAT} & \text{N} \\ \text{H} & \ldots \end{bmatrix} \\ \text{TENSE} & \text{PRES} \\ \text{REL} & \text{Barks([\ ])} \\ \text{CAT} & \text{V} \\ \text{H} & \ldots \end{bmatrix}$$

Not all of this is relevant for semantic interpretation. In fact, in this case, all we need is the REL-values (we are not yet purporting to have anything definite to say about the semantics of grammatical features), so the portion of (23) that is actually relevant for interpretation is:

(24) [REL Barks([REL "fido"])]

Having presented the rudiments of a semantic representation scheme, let us now return to analyzing scoping modifiers.

2.3 Scoping modifiers

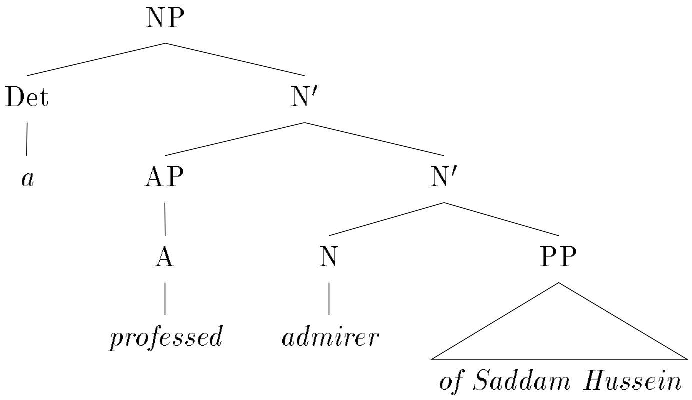
In a framework like the one we are developing here, the semantic structures of modifiers such as *professed* and *alleged* will have to be on the whole pretty much like those of the corresponding verbs, differing primarily in what grammatical functions are associated with the various arguments (we don't have lambda-abstraction or lexicon-internal transformations to provide multiple diatheses for single underlying lexical items). The lexical structure of *professed* will then be an elaboration of:

(25) $$\begin{bmatrix} \text{REL} & [_{state}\,\text{SAYS}([\ \ ],\ [_{state}\,\text{BE}\ ([\ \ ],\ [\text{AT}\ ([\ \ ])])])] \\ ?? & [\ \ ] \\ ??? & [\ \ ] \end{bmatrix}$$

Our problem is then to make a proposal about the mystery GFs that will deliver the results we want with minimal fuss.

Consider first the c-structure below:[17]

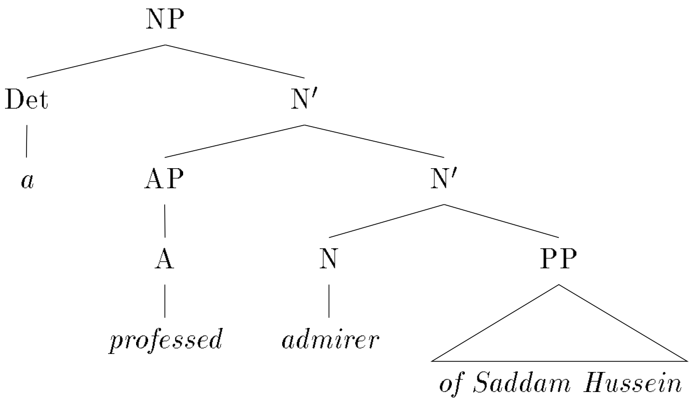
(26)



Now consider first the relation between the head N and the N′ immediately over it. We need to share between these nodes both the semantic and the f-structure attributes of *admirer*, which could be achieved by spreading both kinds of attributes along H. But this approach will not work at all for the two N′ nodes: we want them to share their f-structures, but they can't share their semantics. Rather, the upper N′ should share its semantics with the AP, so that it will have the same REL-value as the adjective *professed*. Anticipating the future somewhat, we will also suppose that Det and N′ are the semantic and syntactic heads, respectively, of the NP. We can depict these relationships on the tree by using solid lines to indicate regions of syntactic feature sharing, and dotted lines to indicate regions of semantic feature sharing, as in (27). These intuitions about the sharing of different types of information are completely shared by the similar analysis of noun phrases in Pollard and Sag (forthcoming:Ch. 8).

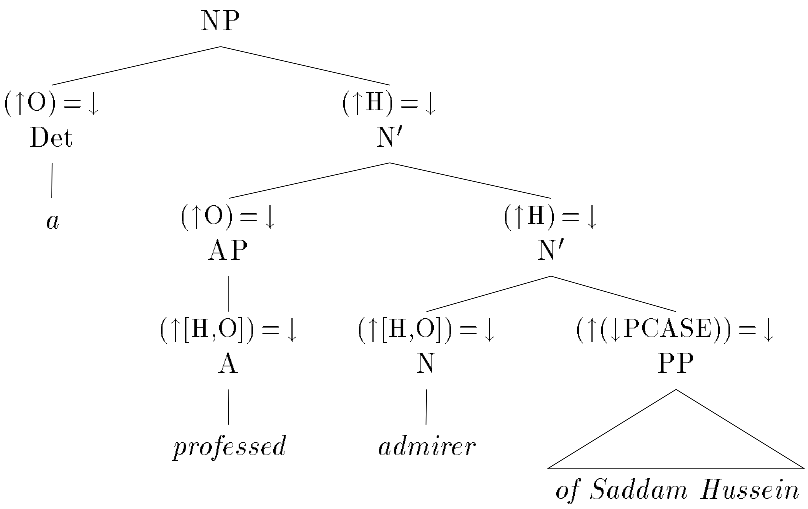
[17]Keep in mind that we are no longer committed to c-structure as a true level of representation: the tree-structure is real, but the category information that is conventionally presented might not be the most appropriate selection from what is provided by the expression.

(27)



We can formalize these notions by introducing what are in effect two different head functions, a 'syntactic head' H, and a 'semantic head' O (for 'operator'). The lexical head of an X′-construction will be both H and O, while in the N′ → AP N′ construction, the AP will be the O, and the N′ will be the H. Semantic attributes will be shared along O, f-structural syntactic ones along H. The use of H and O leads to (28) as the partially annotated (and still quite provisional) c-structure for (27). Lexical item specifications have been omitted, and we are using the notational convention '$[F, G]$' to indicate simultaneously applying F and G grammatical functions (i.e., $(\uparrow[\mathrm{H,O}]) = \downarrow$ abbreviates the two equations $(\uparrow\mathrm{H}) = \downarrow$ and $(\uparrow\mathrm{O}) = \downarrow$ and says that the value of both these attributes is the same).

(28)



In this and subsequent structures we omit the A′ level, in order to reduce clutter.[18]

To get a workable semantic structure out of this we will need to make some decisions about how to represent the arguments of the various pred-

[18] However, since complements are in fact forbidden with prenominal modifers and adverbs, it is conceivable that the A′ level is missing in fact.

icates involved. The Object of Attitude of *admirer* looks like a conventional oblique object and will be treated as one.

What, then, of the entity that is serving as the Experiencer of *admirer* and the Professor of *professed*? There are two basic possibilities. One is to follow the general lead of Bach (1968), introducing into the structure an attribute whose value serves in effect to represent an entity, to which apply predicates supplied by the nouns and other components of the NP. The other approach is that of Montague (also Keenan and Faltz 1985), wherein the senses of N′ nodes (functions from possible worlds to sets of individuals) are operated upon directly by the semantics.

The general design of our formalism favors the former approach. Providing a semantics for NPs on the latter approach would require extensive use of lambda-abstraction, which has a substantial overlap in functionality with unification. And, in general, it is undesirable to have two formal devices of similar functionality together in one theory, unless one has strong motivation for the necessity of both, and clear criteria indicating when each should be used.

So we will introduce a new attribute, IND, whose value will function as a logical variable, and which will spread along both H and O (that is, it will belong to both the semantic and the functional projections). Thus, for *admirer*, the structure we propose is:

(29) $$\begin{bmatrix} \text{REL} & \text{Admire}([\ ],[\ ]) \\ \text{IND} & [\ ] \\ \text{OBL}_{\text{OF}} & [\ ] \end{bmatrix}$$

It is not entirely clear how this ought to be specified in the language of f-descriptions. One possibility, along the lines that are conventional for PRED-values, would be:

(30) $(\uparrow\text{REL}) = \text{Admirer}((\uparrow\text{IND}), (\uparrow\text{OBL}_{\text{OF}}))$

But we don't believe that the formal issues involved here are central to this paper.

Moving on to the arguments of *professed*, here we will use IND again for the Professor argument, but we need to do something different for the Claim Professed. This is already specified as being the H, so one idea would be to say that this argument should be expressed by H, but this won't work, because it would require us to have H spreading along O, which would crunch together many of the things we need to keep distinct (in particular, we would wind up with a substructure where the meanings of *professed* and *admirer* are *both* supposed to be the value of REL).

What we will do is introduce a new attribute, ARG, which belongs to the semantic projection but not the functional one. The N′ appearing as 'complement' to *professed* will then be introduced as an [H,ARG] rather than just an H, as in structure (28) above. Given these decisions, the upper N′ in (28) will now receive (31) as its pre-resolution structure, where we indicate that something is the value of two attributes by stacking the attribute labels to the left of their shared value. Now, by the principles we

(31)
$$\left[\begin{array}{ll}
\text{CAT} & \text{N} \\
\text{BAR} & 1 \\
\text{O} & \left[\begin{array}{ll}
\text{CAT} & \text{A} \\
\text{BAR} & 1 \\
\begin{array}{l}\text{O}\\\text{H}\end{array} & \left[\begin{array}{ll}
\text{CAT} & \text{A} \\
\text{BAR} & 0 \\
\text{REL} & \text{Profess}([\ ],[\ ]) \\
\text{IND} & [\ ] \\
\text{ARG} & [\ ]
\end{array}\right]
\end{array}\right] \\
\begin{array}{l}\text{H}\\\text{ARG}\end{array} & \left[\begin{array}{ll}
\text{CAT} & \text{N} \\
\text{BAR} & 1 \\
\begin{array}{l}\text{O}\\\text{H}\end{array} & \left[\begin{array}{ll}
\text{CAT} & \text{N} \\
\text{BAR} & 0 \\
\text{REL} & \text{Admire}([\ ],[\ ]) \\
\text{IND} & [\ ] \\
\text{OBL}_{\text{OF}} & [\ ]
\end{array}\right] \\
\text{OBL}_{\text{OF}} & [\text{“Saddam Hussein”}]
\end{array}\right]
\end{array}\right]$$

have already discussed, all of the attributes of the N^0 except BAR become attributes of the lower N′. Consequently, the values of the two instances of OBL_{OF} get merged, so Saddam Hussein gets specified as being the object of admiration (but his internal structure is omitted as irrelevant to the present discussion). Furthermore, the semantic attributes of the adjective *professed* get propagated up two levels of O's so as to become attributes of the upper N′, as indicated informally by the dotted arrows in (32) (they also become attributes of the intervening AP, but that is left out to reduce clutter, as is the internal structure of the N^0 node, whose content is no longer relevant).

As a consequence of all this value-sharing, the ARG introduced by the lexical entry of *professed* winds up having the same value as the ARG introduced by the c-structure rules as one of the grammatical functions

(32)
$$
\begin{bmatrix}
\mathrm{CAT} & \mathrm{N} \\
\mathrm{BAR} & 1 \\
\mathrm{IND} & X \\
\mathrm{O} & \begin{bmatrix}
\mathrm{CAT} & \mathrm{A} \\
\mathrm{BAR} & 1 \\
\begin{array}{l}\mathrm{O}\\ \mathrm{H}\end{array} & \begin{bmatrix}
\mathrm{CAT} & \mathrm{A} \\
\mathrm{BAR} & 0 \\
\mathrm{REL} & \mathrm{Profess}([\ \ ],[\ \ ]) \\
\mathrm{IND} & [\ \ ]
\end{bmatrix} \\
\mathrm{ARG} & [\ \ ]
\end{bmatrix} \\
\begin{array}{l}\mathrm{H}\\ \mathrm{ARG}\end{array} & \begin{bmatrix}
\mathrm{CAT} & \mathrm{N} \\
\mathrm{BAR} & 1 \\
\mathrm{REL} & \mathrm{Admire}([\ \ ],[\ \ ]) \\
\mathrm{IND} & [\ \ ] \\
\mathrm{OBL_{OF}} & [\text{“Saddam Hussein”}] \\
\begin{array}{l}\mathrm{O}\\ \mathrm{H}\end{array} & [\ldots]
\end{bmatrix}
\end{bmatrix}
$$

borne by the lower N′. So, when we fuse these two values, and display only those attributes that appear on the semantic projection (currently REL, IND and ARG), we get the following:

(33)
$$\begin{bmatrix} \text{REL} & \text{Profess}([\ ], [\ ]) \\ \text{IND} & [\ ] \\ \text{ARG} & \begin{bmatrix} \text{REL} & \text{Admire}([\ ], \text{“Saddam Hussein”}) \\ \text{IND} & [\ ] \end{bmatrix} \end{bmatrix}$$

Furthermore, it is here only the REL values that are actually necessary to present what is relevant for the meaning; a maximally reduced semantic display is:

(34)
$$\begin{bmatrix} \text{REL} & \text{Profess}([\ ], \begin{bmatrix} \text{REL} & \text{Admire}([\ ], [\text{“Saddam Hussein”}]) \end{bmatrix}) \end{bmatrix}$$

At this point it is beginning to look as if the REL attributes are unnecessary and ought to be eliminated, but they will look more functional again when we get to ‘real’ adjuncts (nonscoping modifiers).

Summarizing the analysis so far, the PS-rule for N′ will be:

(35)
$$\text{N}' \quad \rightarrow \quad \left\{ \begin{array}{cc} \begin{array}{c}\text{AP}\\ (\uparrow\text{O}) = \downarrow\end{array} & \begin{array}{c}\text{N}'\\ (\uparrow[\text{H,ARG}]) = \downarrow\end{array} \\ \\ \begin{array}{c}\text{N}\\ (\uparrow[\text{H,O}]) = \downarrow\end{array} & \ldots \end{array} \right\}$$

As for lexical entries, adjectival *professed* will get:

(36) *professed*, $(\uparrow\text{CAT}) = \text{A}$
$(\uparrow\text{REL}) = \text{profess}((\uparrow\text{IND}), (\uparrow\text{ARG}))$

This entry is semantically identical to that of the verb, the differences being in the category and argument-bindings, although we really should introduce some sort of generic aspectual specification into the reading of the adjective.

Our present analysis clearly extends to prenominal adjectives that do not assign a semantic role to IND, such as *alleged* and *former*. *Alleged* is essentially two-place, but its ‘allegator’ argument is not linked to any syntactic attribute, and is understood existentially. *Former* on the other hand is one-place:

(37) a. *alleged*, $(\uparrow\text{CAT}) = \text{A}$
$(\uparrow\text{REL}) = \text{allege}([\text{ARB}], (\uparrow\text{ARG}))$

b. *former*, $(\uparrow\text{CAT}) = \text{A}$
$(\uparrow\text{REL}) = \text{Past}((\uparrow\text{ARG}))$

We will conclude by presenting an advantage of our style of treatment over a Montagovian one, in which the scoping adjectives would be taken as operators that apply to a common-noun sense to produce another one. On a Montagovian account, a scoping modifier could apply to a nominal in a very wide variety of ways. For example, there might be an adjective **allegig* [*sic*] such that *allegig murderer* meant 'a person who alleges that other people are murderers'. But in fact this sort of thing does not happen: the entity described by the whole nominal is always the bearer of the argument-role that the head noun alone would assign as the referent of an NP in which the noun appeared unmodified. While it is presumably possible to formulate some sort of constraint to enforce this regularity in an Montagovian analysis, in the present approach it comes out as a natural consequence of how the machinery works: given that IND spreads along H, the IND-value for the whole nominal will always be the same thing as the IND-value of the head nominal, thereby delivering this result.

2.4 Non-scoping modifiers

Although scoping modifiers are interesting, and under-analyzed in LFG, they constitute a rather small portion of what are traditionally called adjectives, most of which behave rather differently, and call for a significantly different analysis. Consider an NP such as:

(38) a greedy, rude, unscrupulous property-developer

There is no clear sense here that each adjective is taking what follows it as some kind of argument: they all seem to be applying in parallel, as it were. One might suppose that this kind of parallel application is a kind of coordinate AP. There may be something to this, but these adjective-series can't be ordinary coordinate structures because they are of quite restricted distribution. Main clauses can end in adjective sequences:

(39) The property developer was greedy, rude, unscrupulous.

But this requires a rather special intonation pattern, and is furthermore not possible in embedded clauses:

(40) *The fact that the property-developer was greedy, rude, unscrupulous did not bother his attorney.

This example can be saved by putting *and* in front of *unscrupulous*.

We shall suppose then that main clauses and NPs allow the appearance of adjective sequences (by means of rules that are probably unrelated). In the case of NPs, a rule such as (41) seems called for:

(41) $\mathrm{N}' \quad \rightarrow \quad \mathrm{AP}^* \quad \mathrm{N}'$

Having (41) recursive lets us have scoping adjectives before or after non-scoping ones, with what seem to be the appropriate meanings:

(42) a. A former greedy, unscrupulous property developer

b. A greedy, unscrupulous former property-developer

What of the grammatical functions and semantics? Since these non-scoping adjectives are interpreted intersectively, as if they were coordinated, the standard LFG analysis in terms of a set- or multiple-valued ADJUNCT function seems appropriate, and this is what we will in fact propose. Formally, we will employ a multiple-valued ADJUNCT feature, and, to oil the wheels of the semantics, we will introduce a third projection, the α-projection which will spread IND and the agreement features across ADJUNCT, as well as along H and O.[19] This expresses the idea that there is some sort of universal constraint whereby nominal adjuncts both modify and agree with the NP they are contained in, although it would of course be good to have it follow from the architecture of the theory that IND and the agreement features should behave alike.

What of the N′? If we think of O as being the semantic head function, we are led to expect that there should always be an O, and the daughter N′ is the only plausible candidate. It is then an [H,O]. The annotated version of (41) is thus:

(43) $$\mathrm{N}' \quad \rightarrow \quad \begin{array}{c} (\qquad \mathrm{AP} \qquad)^* \\ (\uparrow \mathrm{ADJUNCT}) = \downarrow \end{array} \qquad \begin{array}{c} \mathrm{N}' \\ (\uparrow [\mathrm{H,O}]) = \downarrow \end{array}$$

The N′ *greedy, unscrupulous property developer* will now give rise to a semantic projection like this:

(44) $$\left[\begin{array}{ll} \mathrm{REL} & \text{Property-developer}([\ \ ]) \\ \mathrm{IND} & [\ \ ] \\ \mathrm{ADJUNCT} & \left[\begin{array}{ll} \mathrm{REL} & \text{Greedy}([\ \ ]) \\ \mathrm{IND} & [\ \ ] \end{array}\right] \\ \mathrm{ADJUNCT} & \left[\begin{array}{ll} \mathrm{REL} & \text{Unscrupulous}([\ \ ]) \\ \mathrm{IND} & [\ \ ] \end{array}\right] \end{array}\right]$$

What is needed for the meaning can be recovered from the values of REL and ADJUNCT alone. At this point the REL attribute seems to look a bit less useless, since it contains the principal or 'core' description of the entity referred to by the nominal, while the ADJUNCT substructures are introducing secondary ones.

[19] This grouping of features corresponds to the 'Φ features' of recent work in GB, but we choose a different name since ϕ is already used within LFG.

There are some well-known facts about non-scoping adjectives which our analysis does not accommodate, in particular the fact that there is very little possible variation in the order of adjectives (this lack is shared by conventional LFG). It is in fact normally assumed that each adjective applies a restriction to the set of individuals designated by the head, the restrictions being applied in the order innermost first. This suggests that some kind of concentric structuring might be occurring with these adjectives as well. But it is difficult to devise any clear tests for the claims about the order in which non-scoping adjectives are 'applied'. Furthermore, they differ clearly in their behavior from scoping adjectives, which do not appear in series separated by pauses:

(45) *Nick is a former, future Tsar.

Note that there is nothing wrong with the conjunctive meaning that the scoping adjectives in this example are trying to have (they can for example be coordinated with *and*). The example just doesn't have that meaning. This indicates that there is a fundamental difference in nature between the two kinds of adjectives, which our analysis captures.

Our discussion so far has dealt solely with nominal modifiers, but if we assume that verbal projections also have an IND attribute, representing a Davidsonian event index, then our theory can be naturally extended to handle a wide range of verbal modifiers as well. The nature and coverage of the theory is very similar to that of Parsons (1990): an operator adverb such as *allegedly* receives an analysis parallel to the operator adjective *alleged*, while VP modifiers such as *slowly* or *violently* act as nonscoping modifiers of events, and are analyzed in the same way as the adjectives treated in this section.

2.5 Determiners

To indicate how a complete analysis of NPs would work, we would need to deal with at least determiners and quantity expressions as well. This is more than we would like to attempt here, so we will just add a few words about determiners and outline how a complete analysis might look. Following Kamp (1981) and Heim (1982), let us assume that there are basically two types of determiners, one requiring something equivalent to 'Quantifier Raising' (May 1985), the other not requiring this.

The former are determiners such as *no*, *every* and *few*, for which an analysis as logical operators with scope has frequently seemed quite appealing. We will assume that these are correctly treated as generalized quantification operators involving two open sentences plus bound variables (to keep things simple, we will assume one, but see Heim (1982)). For *every man loves Marilyn*, for example, our logical form would be along the lines of:

(46) (every, x, man(x), loves(x, m))

The latter sort includes items like the definite article for which a quantificational analysis has tended to seem less plausible. In our linguistic representation we will treat these as forming 'terms', which just sit in the semantic correspondent of whatever syntactic position they occupy. So for *the child bit the dog* we would have a logical form equivalent to:[20]

(47) bit((the, x, child(x)), (the, x, dog(x)))

For nonscoping determiners, we need to manufacture a semantic representation with three significant attributes, one representing the specifier itself, another for the free variable, and a third for the content of the NP. For these we suggest the new attributes SPEC, VAR, and COND, all on the semantic projection. SPEC and VAR will be atomic-valued, taking the determiner and what is in effect a bound variable as their values, while COND will have as its value the descriptive material provided by the nominal. One difficult question is whether COND should be introduced as such in PS annotations, or linked to another GF, the obvious candidate being ARG. We tentatively assume the latter.

We can now suppose that determiners are introduced by a rule similar to that for prenominal adjectives, along the lines of:

(48) $\text{NP} \rightarrow \underset{(\uparrow \text{O}) = \downarrow}{\text{Det}} \quad \underset{(\uparrow[\text{H,ARG}]) = \downarrow}{\text{N}'}$

The structure introduced by a nonscoping determiner such as *the* will then be:

(49) $\begin{bmatrix} \text{SPEC} & \text{DEF} \\ \text{VAR} & [\ ] \\ \text{IND} & [\ ] \\ \text{COND} & [\ ] \\ \text{ARG} & [\ ] \end{bmatrix}$

Now, since the determiner is introduced as an O, all the semantic attributes (including IND) will be shared between it and the NP. Hence the value of VAR will become identified with the value of IND for the NP, and the value of COND will be identified with the maximal N′ introduced by rule (48). And, since SPEC and VAR are semantic attributes, they will also be shared between the specifier and the NP, providing the NP with the kind of semantic representation that we want. In particular, *the professed admirer of Saddam Hussein* will get the semantic structure:

[20] To provide an analysis of indefinites along the lines of Heim (1982) we would need to say that they have scope in the same manner as quantifiers like *every*, but do not introduce their own quantificational operator. Rather they introduce just a variable (the VAR attribute) which is later bound by a quantifier at the level of discourse interpretation. Further discussion of the treatment of quantification in discourse is beyond the scope of this paper.

(50)
$$\left[\begin{array}{ll} \text{SPEC} & \text{DEF} \\ \text{VAR} & [\ ] \\ \text{COND} & \left[\text{REL}\quad \text{Profess}\ ([\ ], \left[\text{REL}\quad \text{Admire}([\ ], [\text{“S. H.”}])\right])\right] \end{array}\right]$$

This is the sort of structure we want. Note also that SPEC, VAR and COND are attributes we want to retain in the semantic display.

Scoping determiners submit to the same basic treatment, but require some additional machinery to handle scope (as quantifier scope is clearly completely independent of the precedence and dominance relations of c-structure). Indeed, it is perhaps worth considering whether scope effects ought to be integrated into the rest of the syntax at all. Someone might claim, for example, that an appropriate 'logical form' (abstract grammatical structure, over which inference rules are to be defined) for *every boy loves some girl* would be *loves((every,x,boy(x)), (some,x,girl(x))*, just as was proposed for nonscoping determiners. The appearance of 'ambiguity' would then be accommodated by appropriately adjusting the inference rules.

The difficulty with this, we think, is that it seems quite clear that when people say this sort of sentence, they mean one of the conventional readings or the other: therefore these sentences ought to be treated as ambiguous. A complete analysis would thus need some mechanism to resolve scope, so that the semantic structure of a quantifier can appear superordinate to the semantics of other parts of the sentence. There are various possibilities that could be adopted here, ranging from classical Cooper storage (Cooper 1983) to using inside-out functional uncertainty, as suggested in Halvorsen and Kaplan (1988), to using the assumptions mechanism of Dalrymple (1992). All of these produce similar and basically adequate results, and so we will not develop any of the options in detail here.

So, summing up the developments thus far, we have introduced a new approach to LFG organized around two key ideas:

- a single informational feature-structure, which we have called an Expression
- the use of *projections* as a means of describing the spread of information throughout this structure.

Let us briefly contrast the alternative. The current LFG rubric of 'levels of co-description' involves setting up a number of distinct places in the sentence-structure, and using subscripts on arrows in annotations to route different kinds of information to the appropriate place. However, as we tried to illustrate above, there are clear problems in getting at the parts of what we intuitively think of as the same object when they are separated onto different levels. We have provided a unified feature structure, considerably closer to the annotated c-structure than conventional f-structure is, and replaced level-collapsing with a notion of information spreading. Our lexical entries are no longer littered with stipulative and also redundant designators for the σ and ϕ correspondences, but rather this information is stated just once in the definition of the projections. Traditional LFG analyses that depend on the properties of f-structure will continue to work as before (since f-structure information is still spreading in roughly the same way) , but some of the difficulties of incorporating semantic interpretation have been overcome (since the technique employed to do it is different).

Formally, our notion of a projection is this:

(51) A projection is defined by:

(a) A set of attributes which are projected
(b) A set of attributes which are projectors

Resolution is a (declarative) relationship of equality of attribute values which is set up as follows:

For a projection P, with $f \in$ projected(P) and $g \in$ projector(P), and for a feature-structure node e,
if $(e\ g) = x$ then $(e\ f) = (x\ f)$.

A *subset display* is just the information relevant to a certain projection and is formed thus:

From the top, recursively, throw away all feature-value pairs for which the feature is not in the projected set.

Resolution is a declarative relationship that describes what is metaphorically information spreading (we can think of it as unifying the values of projected attributes linked across projector attributes). We can then produce the equivalent of conventional f-structures, σ-structures etc. by applying a subsetting operation which just discards all information irrelevant to a certain projection.[21] This subsetting both provides a display of related information that is helpful to the human reader, and a representation of a linguistic level, when only attributes and values on that level determine a constraint.

We have shown how various noun phrase phenomena can be elegantly described by defining f-, σ-, and α-projections. The syntactic GFs spread only across H and the purely semantic attributes (REL, ARG, SPEC, etc.) spread only across O links. IND, CASE and the agreement features (PERS,

[21] Normally the information relevant to a projection will be the attributes and their values for attributes in the shared set. However, we note that while ARG is in the shared set of the σ-projection, it is not needed in (i.e., relevant to) the final semantic display. This suggests a minor difference between σ-structure and LF, which we could trivially implement in the current formalism by adding another projection like σ, but which didn't include ARG.

projector(f) = {H, ARG}
projected(f) = {SUBJ, OBJ, OBJ_θ, ADJUNCT, COMP, ...}

projector(m) = {H}
projected(m) = {VFORM, PMARKING, ...}

projector(σ) = {O}
projected(σ) = {REL, ARG, SPEC, SCOPE, VAR, ...}

projector(α) = {H, O, ADJUNCT}
projected(α) = {IND, CASE, PERS, NUM, GEND}

Table 1: Definition of Projections

NUM, GEND) are shared across both H and O and across ADJUNCT. We have analyzed CAT as also spreading along the f-projection, while BAR is the only feature we have mentioned that does not spread along any projection. Note that the agreement of attributive adjectives now follows without stipulative PS-rule annotations (although the convention whereby this is achieved is not very interesting). The agreement features appear on both semantic and syntactic heads, and agreement is automatic (though overt only if a language possesses the appropriate variety of morphological forms).

We will need to extend these proposals slightly in the next section and our final proposal is shown in Table 1. The f-structure GFs are deemed to also spread over ARG (recall that up until now H and ARG have always co-occurred), while morphosyntactic features such as verbal form or prepositional marking are specified to spread only across H. One could imagine defining yet other projections, but for the moment we will move on and show how these ones can be used to develop a more satisfactory account of interesting verbal phenomena involving complex predicates.

3 Complex verbs in Romance

We have developed a notion of projections in order to deal with various aspects of noun-phrase syntax. Now we will apply it to a rather different kind of construction, 'light verbs'[22] in Romance languages. The problem is easily stated.

Romance languages have various infinitival constructions that behave in the general manner one would expect of complements. In particular, in these constructions, clitic object pronouns, (with which we assume that the reader has at least a passing acquaintance) appear attached to whatever verb they are semantically associated with. Here are some examples from Spanish:[23]

[22] Also called 'restructuring' and 'reanalysis' verbs.

[23] We will be using the standard orthographies for all Romance languages. In these orthographies, the proclitics are normally written as separate words while the enclitics

(52) a. **Sp** Luis comió las manzanas amarillas
Luis eat.3SG.PAST the apples yellow
'Luis ate the yellow apples.'

b. Luis **las** comió
Luis 3PL.FEM ate
'Luis ate them.'

c. Luis insistió en comer las manzanas amarillas
Luis insisted on eat.INF the apples yellow
'Luis insisted on eating the yellow apples.'

d. Luis insistió en comer-**las**
'Luis insisted on eating them.'

e. *Luis **las** insistió en comer

and here from Italian:

(53) a. **It** Piero conosce la signorina molto bene
Peter knows the young-lady very well

b. Piero **la** conosce molto bene.
Peter her knows very well

c. Piero affermava di conoscer-**la** molto bene.
Peter says 'of' to.know-her very well
'Peter says he knows her very well.'

d. *Piero **la** affermava di conoscere molto bene.

However there is also a class of verbs, the 'light verbs', which take what superficially appear to be complement structures, but the putative complement manifests various kinds of behavior quite unexpected of a complement. Perhaps the most obvious is the phenomenon of 'clitic climbing', wherein a clitic pronoun semantically associated with the supposed complement verb can appear on the light verb instead:

(54) a. **Sp** Luis trató de comer-**las**
Luis try of eat.INF-them
'Luis tried to eat them.'

b. Luis **las** trató de comer

(55) a. **It** Maria finisce di batter-**la** a macchina domani.
Maria finishes 'of' to.hit-it on machine tomorrow
'Mary finishes typing it tomorrow'

are written joined to their hosts. We will sometimes set off enclitics with a hyphen for clarity. The two bold letters before the first sentence in each group of examples indicate the language: Spanish, French, Italian or Catalan.

b. Maria **la** finisce di battere a macchina domani.

If we assume conventional complement-style phrase structures for these infinitives, it is quite unclear how to get the object clitics to match up with the verbs that they bear semantic roles to.

Manning (1992) surveys the prospects for analyzing these constructions in LFG and HPSG, and concludes that none of the obvious standard proposals are satisfactory (and finds little prospect of better in other unification-based frameworks). Some important points which we take him to have established are:

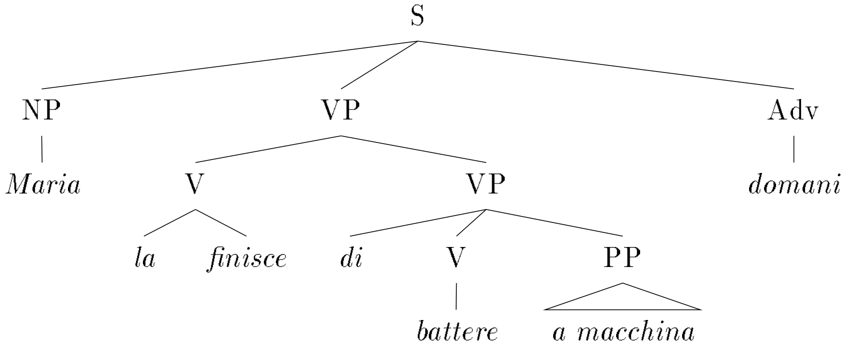
1. The sequence of verb forms cannot be treated as a unitary 'complex verb', whether formed in the lexicon and inserted under a V^0 node in the syntax or formed as a 'verbal complex' in the syntax.

2. The clitic pronoun(s) + (first) light verb sequence forms some sort of constituent, such as perhaps a V^0 node, excluding the heavy verb.

3. The overall structure is right-branching, with the heavy verb and its phrasal complements forming a constituent.[24]

Thus, the overall phrase-structure we have to work with (for (54b)) is:

(56)



Once this is established as the c-structure, the options within standard LFG become few. In fact, the only workable solution seems to be to:

1. treat the 'complement' (heavy) VP as bearing a new grammatical function, say XXCOMP

2. introduce the clitics with a functionally uncertain annotation along the lines of:
   $(\uparrow \text{XXCOMP* OBJ}) = \downarrow$

[24] The argument for right-branching hierarchical structure for French auxiliaries given in Manning (1992) is weakened by the existence in French of an independently needed process that gaps **avoir** and any attached clitics, but the argument retains its full force when applied to Spanish or Italian light verbs since these cannot otherwise be gapped.

But aside from being rather unattractive, this solution faces various difficulties.

One is a simple matter of typology and projection of grammars from data. If functional uncertainty annotations such as that above were the correct approach, one would expect to see different kinds of clitics appearing with different annotations: some having XXCOMP*, some not; some languages having XXCOMP*, others just (XXCOMP), etc. Plus there is the possibility of all sorts of other GFs showing up: some clitics might for example raise out of XCOMP, others out of XXCOMP, etc.

But nothing like this seems to happen. Instead, the possibility of climbing seems to be determined by the presence of a light verb, so that if one object marker can climb with these verbs, they all can.[25] Furthermore, if clitics can climb up one level, they can climb up through more than one level of light verbs:[26]

(57) a. **It** Maria avrebbe voluto andare a prender-**li** lei stessa
'Maria would have (**avere**) wanted to go to get them herself.'

b. Maria **li** avrebbe voluti andare a prendere lei stessa

(58) a. **Sp** quiero tratar de terminar de mostrar-**te-lo** mañana
'I want to try to finish showing them to you tomorrow.'

b. **te lo** quiero tratar de terminar de mostrar mañana

Furthermore, Rizzi (1982) shows that a number of syntactic phenomena show differences between the climbing and the regular complement structures, differences which in a transformational framework suggest that the verbs have combined into some sort of complex verb. We want an analysis in which these phenomena all emerge as natural consequences of something, rather than just being stipulations hung off a label. This is a challenging agenda, which we will not be completely successful in fulfilling, but we believe that the notion of projections presented here allows some progress to be made.

Our presentation will proceed in two stages. First, we will make our general proposal (3.1), and discuss a variety of data from Romance languages that support it (3.2). Then we will develop two more technical issues in greater detail: clitic placement (3.3), and linking theory (3.4).

[25]The ability of a verb to behave as a light verb has frequently been regarded as something to be listed (with the list varying between languages and ideolects), but this property can probably be predicted from the semantic interpretation of the verb (Napoli 1981). Regarding uniform clitic placement, there is a possible exception in Romanian (Mallinson 1986) where a feminine singular clitic remains on a participle when other clitics climb onto an auxiliary, but the Romanian situation seems quite different. There is little evidence that the auxiliary and verb are not a verbal complex here, and the behavior of the feminine singular clitic seems explicable in terms of morphological structure.

[26](57) is from Rizzi (1982:22); (58) was provided by a Spanish informant.

The discussion in these latter sections becomes rather difficult at times, and we hope that the principles we propose can be improved upon, but we think that it is important to make definite suggestions about difficult problems (which are hard in any framework, at the moment), rather than to leave them unacknowledged.

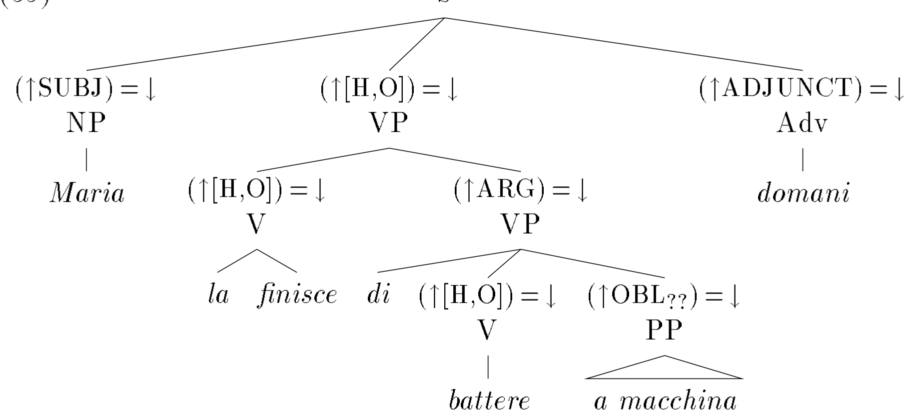
3.1 The basic proposal

The basic effect we obviously want to get is that the governable grammatical functions are shared between the light verb and its 'complement', while such properties as finiteness and complementizer features are not (so that each verb can select the appropriate feature values for its 'complement'). Not only will such a move provide a basic analysis of clitic climbing, but it will immediately address obvious issues in the syntax of causatives, such as why a causative+verb combination seems to act as is if the whole complex can only have at most one of each of the governable grammatical functions such as subject, object and indirect object.

This much is reasonably clear, but choosing a particular execution of the idea is more difficult. One might consider imitating the analysis of operator adjectives in the previous section: the light verb would be introduced as an O, the 'complement' as an [H,ARG]. But this is unsatisfactory, because we would not then get grammatical function spreading from both the light verb and the VP complement, which we want, and we would get morphological form features spreading up from the VP complement, whereas we want them to spread up from the light verb. To correctly analyze this situation, we need to tease apart three kinds of heads: morphological heads, functional heads and semantic heads (cf. Ackerman and Webelhuth 1991). Let us reinterpret H as indicating the morphological head: category and morphological form features will spread over H. Thus the sharing of morphological and semantic features between the light verb and its dominating VP argues that these two are connected by [H,O]. However, the grammatical relation sharing between the light verb, the sister VP and their mother suggests that both are functional heads. We could introduce another attribute to indicate functional heads, but at least for the moment it seems completely adequate for functional head to be a derived notion. Let us suggest that the VP 'complement' is introduced as a simple ARG and that anything connected by either H or ARG will be a functional head. In this way, the governable grammatical functions are shared between the light verb, the VP complement and their mother. This definition of the projections was summarized above in Table 1.

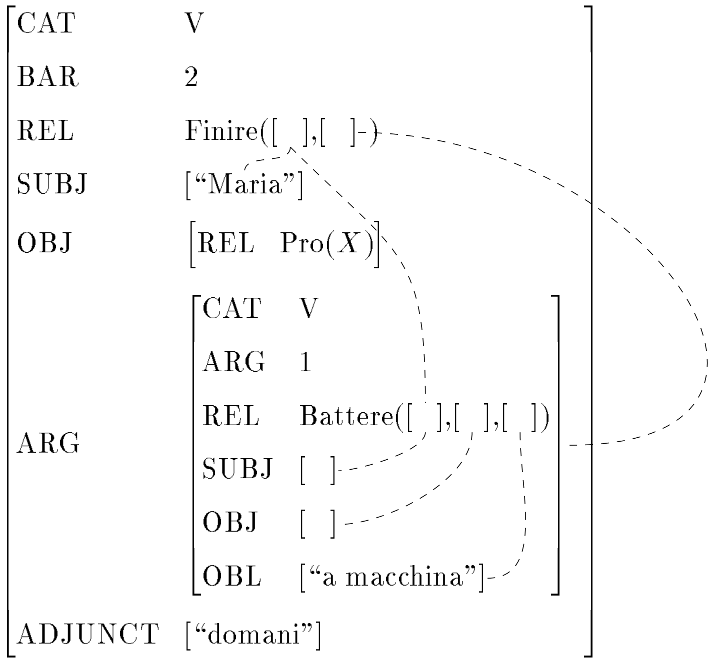
We can now present an annotated version of (55b):

(59)



When resolution has applied other than across ARG, we get a structure like (60). Here, we have specified only a REL value for pronouns, and

(60)



we are assuming that **domani** is a matrix adjunct while **a macchina** is bearing some sort of oblique relation to the verb (even if the combination **battere a macchina** is interpreted idiomatically, its components obviously have sensible and relevant interpretations too).

At this point, spreading the governable GFs across ARG obviously produces the effects we need of hooking up the object clitic in the upper level with the object attribute at the lower level, thus satisfying the Completeness and Coherence Conditions, and supporting an appropriate semantic interpretation. It furthermore does so in such a way as to provide an immediate and straightforward account of another basic property of these constructions, which is that they seem to function in many respects as a superficial morphological alternative to morphological causatives such as those found in Japanese or Chicheŵa. In these languages it is plausible to analyze causatives as involving incorporation (implemented by an actual movement rule) of the lower verb into a higher verb represented morphologically by the causative affix (Baker 1988). Romance causatives seem to have essentially the same kinds of syntactic properties as these apparently incorporational causatives, except that it is hard to see how any movement-based notion of incorporation could actually be used to produce them.

Baker (1988:202–294) proposes a rather obscure and problematic notion of 'abstract incorporation (at LF)' (for the difficulties, see especially fn. 37, p. 467), but our approach provides a much more straightforward account. In Romance, an ARG-substructure is introduced in the syntax. In languages with overtly incorporational causatives, on the other hand, this kind of substructure is introduced in the morphology: The verb-stem of a causative is introduced as bearing an ARG GF to the affix, so that a verb-stem such as **sak-ets** 'laugh-cause' in Chicheŵa would receive the following morphological structure:

(61)
```
            V
      ______|______
(↑ARG) = ↓       ets
    V
    |
   sak
```

Now in an actual sentence, such as

(62) Baluzi a-na-sak-ets-a ana
lizard SP-PRES-laugh-CAUSE-FV children
'The lizard made the children laugh.' (Baker 1988:162)

the feature-structure of the object NP **ana** will be shared (as the value of OBJ) between the upper and lower levels of the feature-structure of the verb, so that the overall form of the expression will be basically the same as in a Romance causative, in spite of the major difference in morphological technique. In fact, we will see later that our projections machinery in effect allows us to define a notion of 'abstract incorporation' of the kind that Baker needs, but it works better, because it is not based on movement, but on a kind of quotient-set formation.

We will now proceed to examine various properties of the Romance languages that support our approach, and will then consider the problems of clitic-placement and linking.

3.2 General considerations

Perhaps the basic observation about light verb constructions in the literature is that in many respects they act like a single clause-nucleus. Since there are also various respects in which they act like complement structures, they are commonly described with terms such as 'Restructuring', 'Reanalysis', 'Clause Union' or 'Clause Reduction'.

From the standpoint of a theory such as LFG, perhaps the most striking point is that they seem to have a single array of grammatical relations. This is not immediately obvious with the intransitive light verbs we have been considering, but is entirely well known with causatives. In French, for example, we have:

(63) a. **Fr** Nous avons fait partir Marie
we have made leave Mary(DO)
'We made Mary leave.'

b. Nous l' avons fait laver à Marie
we it have made wash to Marie(IO)
'We made Marie wash it.'

When the downstairs verb is intransitive, its 'logical subject' is rendered as an ordinary direct object; when it is transitive, as an **à**-marked object (a traditional indirect object), however one might want to formalize this in LFG (OBJ_{goal}, perhaps). Explaining why something like this has to happen in a theory like GB is difficult, although certainly not impossible; but it is immediately necessary in a framework such as LFG or Relational Grammar, if these constructions are (overtly) monoclausal. This general observation furthermore immediately shows why it would be unsatisfactory to adopt the analysis mentioned above where heavy and light verbs were distinguished by whether their complements were XCOMPs or XXCOMPs.

If complex predicate constructions take a single array of grammatical relations, it would seem that they must somehow undergo Linking (of semantic roles with grammatical relations) as a unit, and one would therefore expect that they would undergo 'lexical' operations such as Passive, etc., as units.

There is evidence that this actually happens, perhaps the most spectacular of which is the existence of 'long passives' (Aissen and Perlmutter 1983:389ff, Burzio 1986:376):[27]

(64) a. **Sp** Los obreros están terminando de pintar estas paredes
'The workers are finishing painting these walls.'

[27] These do not occur with all restructuring verbs, however, but only those in the semantic class of *begin*, *continue*, *finish*. It seems that this is a semantic restriction.

b. Estas paredes están siendo terminadas de pintar (por los obreros)
'These walls are being finished to paint (by the workers).'

c. **Les** pintan las paredes a los dueños
3PL.DAT paint.3PL.PRES the walls for the landlords
'They paint the walls for the landlords.'

d. Estas paredes **les** están siendo terminadas de pintar a los dueños
'These walls are being finished to paint for the landlords.'

e. *Estas paredes están siendo terminadas de pintar-**les** a los dueños

The (complex predicate) sentence (64a) can be passivized as shown in (64b). (64d) gives further confirmation of what is going on. As shown in (64c) an indirect object is doubled by a matching clitic appearing on the verb. In the passive (64d), the doubling clitic (necessarily) climbs so as to appear before the passive auxiliary, again indicating how all the verbs in this sentence appear to be acting (in some sense) like one large verb.

Another example is the 'Long Object Preposing' with **si**-impersonals discussed by Rizzi (1982).[28] (65d) versus (65c) indicates that long object preposing is good only when there has been complex predicate formation (since complex predicate formation would normally be accompanied by clitic climbing):

(65) a. **It** Si vuole vender-gli queste case a caro prezzo.
si wants sell-him these houses at a high price
They want to sell him these houses at a high price.

b. Gli si vuole vendere queste case a caro prezzo.

c. *Queste case si vogliono vendergli a caro prezzo.

d. Queste case gli si vogliono vendere a caro prezzo.

There is some dispute as to whether the 'preposed objects' become subjects, as argued by Rizzi, or merely topics, as claimed for European Portuguese by Raposo and Uriagereka (1990), but their proposal doesn't explain facts such as the near obligatoriness of clitic climbing with Long O.P. sentences (see further the discussion in Manning (1992)).

Further arguments are provided by *Tough*-Movement in Spanish (Aissen and Perlmutter 1983), which is normally strictly clause-bounded:

(66) **Sp** *Tales cosas son difíciles de insistir en hacer
'Such things are difficult to insist on doing.'

[28] These examples are given as cited by Rizzi, though informants seem to greatly prefer that the singular forms of the light verb **vuole** be switched to plural.

But, in the case of a restructuring verb, a pair of verbs again seems to act like a single complex verb, and 'long distance' *Tough*-Movement becomes possible:

(67) **Sp** Estos mapas serán difíciles de empezar a hacer
'These maps will be difficult to begin to make.'

This anomaly will fall out naturally if the object of the apparent VP-complement of the light verb actually belongs to the complex predicate as a whole, at least for the purposes of *Tough*-Movement.

The final phenomenon we will consider is perfective auxiliary selection in Italian. The basic facts are well known (see Rizzi (1982) and especially Burzio (1986)). It will be sufficient for present purposes to know that **essere** is used with unaccusative verbs, giving basic facts such as these:

(68) a. **It** Piero ha/*è mangiato con noi
'Piero has/*'is' eaten with us.'

b. Piero ha/*è voluto questo libro
'Piero has/*'is' wanted this book.'

c. Piero *ha/è venuto con noi
'Piero *has/'is' come with us.'

The facts of present interest are what happens when a restructuring verb that takes **avere** as its auxiliary, like **volere** in (68b), takes a verbal complement:

(69) a. **It** Piero ha/*è voluto mangiare con noi
'Piero has/*'is' wanted to eat with us.'

b. Piero ha/è voluto venire con noi
'Piero has/'is' wanted to come with us.'

Avere is always good, but if the downstairs verb would normally take **essere**, then **essere** is also possible.[29] Importantly, if other evidence shows that restructuring has occurred (for example, the occurrence of clitic-climbing),[30] then **essere** is necessarily used as the auxiliary, but conversely, if restructuring cannot occur, auxiliary change is impossible. For example, VP pied-piping accompanying *Wh*-movement, which prevents complex predicate formation (as discussed in Rizzi 1982) prevents the auxiliary change (70a) and clitic climbing is impossible without the auxiliary change having occurred (70b):

[29] Restructuring verbs that take **essere** as an auxiliary always maintain this auxiliary. We will not further discuss them in this section, but their existence in no way undermines the point that is being made.

[30] Recall that restructuring is optional with these modal and aspectual verbs.

(70) a. **It** La casa paterna, tornare alla quale Maria
the house paternal return.INF to the which Maria
{ avrebbe voluto / have.COND wanted | *sarebbe voluta / 'be'.COND wanted } già da molto tempo, ...
already since much time
'Her paternal home, to which Maria would have wanted to go back for a long time, ...'

b. *?Maria **ci** ha dovuto venire molte volte
'Maria has had to come there many times.'

We thus have the result that in cases of restructuring, it is the right hand verb that is determining whether the auxiliary is **avere** or **essere**. This result seems completely general. Rizzi (1982:22–23) demonstrates that in cases with complex predicates, no matter how many restructuring verbs occur between the auxiliary and the rightmost verb, it is still the rightmost verb that determines auxiliary selection:[31]

(71) a. **It** Maria li avrebbe voluti andare a prendere lei stessa
[avere] [essere] [avere]
'Maria would have (**avere**) wanted to go to get them herself.'

b. Maria ci sarebbe dovuta cominciare ad andare
[avere] [avere] [essere]
'Maria would have (**essere**) had to begin to go there.'

c. Maria li avrebbe potuti stare per andare a prendere
[avere] [essere] [essere] [avere]
lei stessa
'Maria would have been able to be on the point of going to get them herself.'

In cases of restructuring, if we think of the heavy verb (the one at the bottom of the pile) as being in effect *the* verb of the whole clause, it is understandable how its nature might affect auxiliary selection (we will present an actual proposal to explain this later).

We have now surveyed a range of evidence which suggests that massive grammatical relation sharing between light verbs and their 'complements' might well be the right general approach. We now turn to the more detailed task of constraining clitics to appear in the right positions.

[31]The auxiliaries that verbs would normally select are shown beneath them in these examples. Note that each example has a climbed clitic proving that restructuring has taken place and the auxiliary shown is in each case the only choice possible.

3.3 Placing clitics

The mainstream in Romance generative linguistics, starting from works such as Rizzi (1982) and Aissen and Perlmutter (1983), has presented a picture in which clitic climbing is a necessary and exceptionless result of complex predicate formation. In more recent work (such as Rosen 1989) the clitics are commonly presented as moving up to the I^0 node of the light verb, since the verbal complement lacks independent inflection. However, it is widely known that this picture is a simplification and there are numerous cases in which other diagnostics indicate complex predicate formation, but some or all of the clitics have failed to climb. Moore (1990) assembles many of these cases and suggests that clitic climbing is basically optional (even if it is sometimes subject to complicated preferences). However, on the other hand, one cannot put clitics just anywhere, and there remain strong constraints (especially when multiple clitics belong to one complex predicate). We will build on Moore's approach in developing our theory of clitic climbing. This will necessarily make the result somewhat more complicated than accounts using the standard simplification, but our account will be in much closer accord with the actual data. While we will not formally model everything, our theory will thus generate a constrained superset of the occurring data, rather than a convenient subset.

Our analysis involves three major ingredients. The first has already been provided. Since grammatical relations (such as OBJ and OBJ_θ) are shared throughout the levels of the complex verbal construction, clitics expressing these relations can in principle appear at any level, thereby accounting for the fact that a clitic semantically associated with a lower verb can appear either on it or the higher one (though there is certainly a tendency to prefer the higher position under a wide range of circumstances).

This basic optionality of clitic climbing will allow us to account for the following sorts of facts. Clitic climbing is a sufficient test for and usually occurs in cases of complex predicate formation (72b). However, with restructuring verbs, clitics can often remain downstairs despite the fact that other diagnostics such as auxiliary change (72c) or *tough*-movement (72d–e) indicate that complex predicate formation has taken place.[32] With causative verbs, downstairs cliticization of the embedded object yields mixed judgments (72f) (Moore 1990:326, Aissen and Perlmutter (1983) fn. 20, Alsina forthcoming) but putting this case and the last together, it seems that despite strong preferences, it is best to work forward from the position that clitic climbing is basically optional.

[32] (72e) shows 'copy Raising' where a clitic copies the raised argument (Moore 1990:324). Downstairs cliticization as in (72d) is not always very good (Moore 1990, footnote 9) and downstairs cliticization does not seem to be possible with the long passives that we saw earlier (64e). We have no explanation for these facts at the moment.

Moreover, clitic climbing is impossible or strongly disfavored in certain circumstances, and downstairs cliticization is then clearly acceptable. In cases of morphological overload when a higher verb cannot carry an extra clitic because there is already a clitic in the appropriate slot, the clitic appears downstairs (72g), and, in Spanish, a downstairs animate accusative clitic will not climb onto a higher verb which has its own dative clitic (72h) (Rivas 1977, Pizzini 1982).

(72) a. **Sp** **Me** permitieron comprar**la**.

b. **Me la** permitieron comprar.
'They permitted me to buy it.'

c. **It** Maria ha/**è** dovuta venir-**ci** molte volte.
'Marie has had to come there many times'
(Rizzi 1982 (84a), Ch. 1, fn. 26)

d. **Sp** Estos mapas van a ser difíciles de poder hacer**les** a los clientes.
'These maps are going to be difficult to be able to make for the clients.' (Moore 1990 (20a))

e. Esta carta es fácil de querer escribir**la**.
'This letter is easy to want to write.' (Moore 1990 (17a))

f. */?/%Juan **me** hizo vender-**la**.
'Juan made me sell it.' (Moore 1990 (22))

g. **Te** hicieron dár-**mela**.
***Te me la** hicieron dar.
'They made you give it to me.' (Moore 1990 (14a–b))

h. **Me** permitieron saludar-**la**.
*?**Me la** permitieron saludar.
'They permitted me to greet her.' (Pizzini 1982 (38a–b))

Simply allowing clitics to appear anywhere won't do, since there are additional constraints. In particular, given only the sharing of grammatical relations, the theory would allow an argument of an upper verb to be expressed as a clitic on a lower one, but this does not happen. An illustrative instance is the verb **permitir** in Spanish, which takes an indirect object for the person to whom permission is extended. When the downstairs object is inanimate, this object clitic can climb (72b), when animate, it doesn't (72h) (Pizzini 1982:64). However, our current concern is that it is inconceivable to 'lower' the clitic expressing the indirect object of **permitir** onto the lower verb:

(73) **Sp** *Permitieron comprár**mela**

To formulate this second ingredient, we will use a 'no-lowering' principle based on a notion of 'semantic dependence'. To motivate our definition of this notion, consider the proposal that (73) is bad because a clitic semantically associated with the upper verb is showing up on the lower one. This seems basically right, but there is a problem in specifying exactly what is meant by 'semantically associated with'. The clitic **me** is in fact associated with **comprar**, in that it expresses its Agent. The problem must then not be that the clitic doesn't have a semantic association with the lower verb **comprar** (since, in fact, it does), but that it does have one with respect to the higher verb **permitir**.

Given the way in which we have set our theory up, this means that if you were to travel through the whole Expression, from its root (i.e., top), and move only along REL, ADJUNCT and the other attributes that are relevant to semantic interpretation (including the ones that are found within the conceptual structures that are lexically associated with REL-values, that we haven't detailed), you couldn't get to the clitic **me** without going through (the portion of the Expression corresponding to) **permitieron**, but you wouldn't need to go through **comprar**, since you could get to it via the path that makes it the Recipient (of permission) of **permitieron**. This will be the basis of semantic dependence:

(74) x is semantically dependent on y iff you can't get from the root of the Expression to x along the semantically interpreted attributes without passing through y.

The general effect we want then is that a clitic can't appear on a verb unless it is semantically dependent on it. Note that this allows raising but not lowering of clitics, because a clitic functioning as an argument of a lower verb will still be semantically dependent on the higher ones.

To formulate the constraint a bit more precisely, we need to say a bit more about the structures in which clitics are introduced. We will suppose that they are introduced morphologically as part of the verb that they are attached to, and that they bear a relation CL to that verb. CL will be a 'c-structure' relation, one that is tightly tied to the actual overt arrangement of formatives, and not subject to spreading (as, indeed is also the case with O, H and ARG). So a form like **mandármela** will be associated in the lexicon with a structure like:

(75)
$$\left[\begin{array}{ll}
\text{REL} & \text{Mandar([ ],[ ],[ ])} \\
\text{CL} & \left[\begin{array}{ll}
\text{PERS} & \text{I} \\
\text{NUM} & \text{SG} \\
\text{REL} & \text{Pro([ ])} \\
\text{IND} & \text{[ ]}
\end{array}\right] \\
\text{CL} & \left[\begin{array}{ll}
\text{PERS} & \text{III} \\
\text{NUM} & \text{SG} \\
\text{GND} & \text{FEM} \\
\text{REL} & \text{Pro([ ])} \\
\text{IND} & \text{[ ]}
\end{array}\right]
\end{array}\right]$$

Note that the clitics are *not* introduced bearing grammatical relations. This is because clitics seem to appear in a fixed order, determined by their inherent feature content, regardless of what grammatical relations they bear (see Perlmutter 1971 and Bonet i Alsina 1991 for extensive discussion of this).

GFs such as OBJ, etc., must then be associated with the clitics by some other means. Recalling that conditional formulas may be added to an LFG grammar without creating undecidability, the following would seem to do the job:

(76) $\text{CL}(x, y) \rightarrow \{\text{OBJ}, \ldots\}(x, y)$

This rule is implicitly universally quantified, and says that if a structure y is a clitic of x, then it bears OBJ or one of various other GFs to x as well. Note that without something like (76), one faces the unattractive prospect of adding similar disjunctive GF-introducing annotations to each of the positions available for clitics, which is descriptively inadequate, due to a failure to capture generalizations. The form in which (76) is presented also makes it far more amenable to integration with, or derivation from, other principles than individual annotations sitting on nodes would be. Expressing multiple grammatical functions seems to be a typical characteristic of clitics, suggesting that the content of (76) is largely fixed by UG.

We will regard these clitics as introduced within the Expressions corresponding to the words that carry them, so the structure we get for the complex VP in (73) is (77) where rule (76) has provided the GFs, and Linking has connected them to argument positions, as indicated. Concurrently, Resolution will spread the SUBJ and two OBJ attributes through all levels of the structure, but this is not relevant for our present purposes. What is relevant is that the CL attributes remain associated

$$
(77)\quad \left[\begin{array}{ll}
\text{BAR} & 1 \\
\text{HO} & \left[\begin{array}{ll}
\text{CAT} & \text{V} \\
\text{BAR} & 0 \\
\text{SUBJ} & \left[\begin{array}{ll}
\text{PERS} & \text{III} \\
\text{REL} & \text{Pro}([\ ]) \\
\text{IND} & [\ ] \\
\text{NUM} & \text{PL}
\end{array}\right] \\
\text{REL} & \text{Permitir}([\ ], [\ ], [\ ]) \\
\text{ARG} & [\ ]
\end{array}\right] \\
\text{ARG} & \left[\begin{array}{ll}
\text{BAR} & 1 \\
\text{HO} & \left[\begin{array}{ll}
\text{CAT} & \text{V} \\
\text{BAR} & 0 \\
\text{REL} & \text{Comprar}([\ ], [\ ]) \\
\text{OBJ}_{\text{REC}}\ \text{CL} & \left[\begin{array}{ll}
\text{PERS} & \text{I} \\
\text{NUM} & \text{SG} \\
\text{REL} & \text{Pro}([\ ]) \\
\text{IND} & [\ ]
\end{array}\right] \\
\text{OBJ CL} & \left[\begin{array}{ll}
\text{PERS} & \text{III} \\
\text{NUM} & \text{SG} \\
\text{GND} & \text{FEM} \\
\text{REL} & \text{Pro}([\ ]) \\
\text{IND} & [\ ]
\end{array}\right]
\end{array}\right]
\end{array}\right]
\end{array}\right]
$$

with their original positions. Furthermore, the REL attributes will spread across O, so that the REL-value of the upper verb will become the REL of the entire sentence. A convenient (but not very principled) selection of information from the Resolved from of (77) is (78).

(78)
$$\left[\begin{array}{ll}
\text{SUBJ} & \left[\begin{array}{ll} \text{PERS} & \text{III} \\ \text{REL} & \text{Pro}([\ ]) \\ \text{IND} & [\ ] \\ \text{NUM} & \text{PL} \end{array}\right] \\
\text{REL} & \text{Permitir}([\ ], [\ ], [\ ]) \\
\text{ARG} & \left[\begin{array}{ll} \text{REL} & \text{Comprar}([\ ], [\ ]) \\ \text{CL} & [\ ] \\ \text{CL} & [\ ] \end{array}\right] \\
\text{OBJ}_{\text{REC}} & \left[\begin{array}{ll} \text{PERS} & \text{I} \\ \text{NUM} & \text{SG} \\ \text{REL} & \text{Pro}([\ ]) \\ \text{IND} & [\ ] \end{array}\right] \\
\text{OBJ} & \left[\begin{array}{ll} \text{PERS} & \text{III} \\ \text{NUM} & \text{SG} \\ \text{GND} & \text{FEM} \\ \text{REL} & \text{Pro}([\ ]) \\ \text{IND} & [\ ] \end{array}\right]
\end{array}\right]$$

There we can see that the first person clitic is not semantically dependent on the verb it is attached to, since we can get to it by starting at the whole Expression, going down into its REL (Permitir(...)), and then to the second argument position there, whose value is the same as that of the CL attribute of the lower verb (that this is equal to the first argument position of the lower verb is irrelevant: what matters is that we don't have to go *through* the lower verb). A somewhat more precise version of our constraint is then:

(79) No Lowering: If y bears the CL relation to x, then y must be semantically dependent on x.

It should now be clear that when clitics dependent on a lower verb are introduced on a higher one, then the constraint is not activated, since the clitics are also semantically dependent on a higher verb. The correct placement of clitics at different levels in cases of morphological overloading will also now follow without further stipulation.[33] However, we will leave unformalized the other restrictions that can cause clitic climbing to be disfavored in these **permitir** constructions (72h).

Formally, the no-lowering constraint is not Schönfinkel-Bernays, since its consequent involves a condition stated in terms of quantification over paths through the structure: the consequent states that every path from the root to y must contain (that is, go through) x. In order to preserve decidability, the constraint must therefore apply non-constructively, serving as a filter, in the same manner as the Completeness and Coherence Conditions. It would be desirable to investigate the possibility of generalizing the constraint against clitic-lowering to a more general anti-lowering prohibition, but we won't do this here.[34] Instead we will push on to formulate a 'no-splitting' constraint, which applies when two or more clitics appear within a complex predicate.

The most straightforward manifestation of no-splitting arises when the downstairs verb is ditransitive. It is then impossible for one argument to be expressed by a clitic on the downstairs verb, and another by one on the matrix:[35]

(80) a. **Sp** **Te lo** quiero decir.
'I want to tell it to you'

b. Quiero decír-**telo**.

c. ***Te** quiero decír-**lo**.

d. ***Lo** quiero decír-**lo**.

e. **It** Piero voleva dar-**melo**.
'Piero wanted to give it to me.' (Rizzi 1982, fn. 26)

f. *Piero **mi** voleva dar-**lo**.

[33]This is providing that in an example like *Te hicieron dármela* 'They made you give it to me', **te** is semantically dependent on the causative verb. Put the other way, the fact that this sentence is good provides further syntactic evidence for the ternary causative relation proposed by T. Mohanan (1990) and Alsina (in preparation).

[34]An apparent difficulty in doing this comes from Romance causatives. Given the phrase structure for causatives argued for in Alsina (in preparation) and Manning (1992), in which the causee appears among and on a level with the other complements of the lower VP, the causee argument does appear as a complement of a verb it is not semantically dependent on.

[35]This no-splitting constraint appears in all of Italian, Spanish, Catalan and French (excluding from consideration the reflexive clitic **se**, which, following Grimshaw (1982), we would not analyze as bearing a grammatical relation). However, as well as the Romanian case mentioned in footnote 25, there are some other apparent exceptions in various Romance dialects, such as Franco-Provençal; see the references in Kayne (1991:661,663).

Additionally, when there are two or more light verbs and a heavy verb, and two of the lower verbs have clitic dependents (81a), then there are further restrictions. One clitic can never hop over another clitic (81e) and it is impossible for two clitics to both shuffle up one verb (81f) (Aissen and Perlmutter 1983:366):

(81) a. **Sp** Juan quiere permitir**me** hacer**lo**.
'Juan wants to permit me to do it.'

b. Juan **me** quiere permitir hacer**lo**.

c. Juan quiere permitír**melo** hacer.

d. Juan **me lo** quiere permitir hacer.

e. *Juan **lo** quiere permitir**te** hacer.

f. *Juan **me** quiere permitir**lo** hacer.

These sorts of facts make old-timers nostalgic, since they are just the sorts of things that follow automatically from the idea that clitic climbing is a successive-cyclic transformation that takes all of the clitics currently associated with the verb. They also work out pretty well in GB, using relativized minimality (Moore 1990). But the resources behind these analyses are not available in unification-based frameworks. On the other hand, given the amount of optionality in clitic placement seen above, it would be hard to maintain the sort of analysis given to these data within Relational Grammar by Aissen and Perlmutter (1983).

There are various approaches that one might take to ruling out these cases. The proposal we will adopt is a variant of one suggested to us by John Maxwell. The idea is that we define a domain starting from the verb the clitic appears on and extending to the lowest verb that the clitic is a semantic dependent of (strictly, a domain not including the upper boundary, but including the lower boundary, if it is distinct from the upper boundary). This domain corresponds to what in the successive-cyclic transformational account was the set of verbs that the clitic had been attached to earlier in the derivation. Then our constraint is that there cannot be any clitics in that domain (regardless of which verbs they are semantically dependent on), for if there were, these clitics would also have had to have climbed.

We can express this notion easily using an off-path constraint on a functional uncertainty expression (see Dalrymple 1993).[36] Let Sem be the set of attributes in the semantic projection. Then we can use the

[36] Note, however, that since our use of off-path constraints is confined to a putative principle of UG, we are not making any claims about the appropriate use of off-path constraints in language-particular rules.

following conditional, which is again not S.-B., but which can apply as a nonconstructive constraint:[37]

(82) No-splitting: $\mathrm{CL}(x, y) \rightarrow (x\ \underset{\neg(\rightarrow \mathrm{CL})}{\mathrm{Sem}^*}) = y$

The mode of application of this constraint requires some explication. The antecedent condition is that y be a clitic attached to x. The consequent says that there must be a path along semantic attributes leading from y to x (because functional uncertainty expressions are interpreted as requiring the existence of some path), where each position along that path does not have a CL attribute.

This constraint rules out raising one clitic off a verb but not the other, since the other clitic will then appear within the prohibited domain, and furthermore extends to the more complex cases in (81). A further effect of the principle is that it prevents a single argument from being co-expressed by multiple clitics appearing on different verbs—a non-trivial effect, since Spanish clitics have a significant capacity for clitic-doubling with full phrasal complements, and therefore don't necessarily introduce unique indexes (traditionally on PRED features) that block merger with full NPs, or, presumably, each other.

It remains to generalize the no-splitting constraint, or get it to emerge as a consequence of other independently motivated principles. However, just the fact that we have managed to find a way to capture some of the real facts about clitic placement gives hope that somebody will be able to find a better way to realize this goal.

3.4 Linking

Previous work in LFG has assumed that the linking of semantic roles with grammatical relations happened in the lexicon. But our analysis of complex verbs does not permit us to maintain this. Our treatment of Romance causatives, for example, requires the combination of causative and embedded verb to be created in the syntax, and linked as a unit. Therefore linking also must happen in the syntax.

In this paper we will not attempt to develop a full-fledged theory of linking, but will merely attempt to show how the approach to linking developed by Alsina (in preparation) can be formalized and applied within our general approach. There are a variety of problems remaining to be solved (especially with the linking of adjuncts), and we can't be sure that our formalization is in accord with Alsina's intent on all points, but the discussion should suffice to show that there is probably no essential conflict between our analyses and Alsina's theory.

[37] No-splitting bears a tantalizing resemblance to no-lowering, but we have not found a way to combine them that doesn't run afoul of the properties of causatives and **permitir** which were discussed above.

First, some general background. The theory is a development of the 'Lexical Mapping Theory' presented in a series of publications by Alsina, Bresnan, Mchombo, Moshi and others, and shares with this theory the property of depending on the simultaneous effects of a classification of grammatical functions on the one hand, and of positions in the 'argument structure' of lexical items on the other, where 'argument structure' is understood as a presentation of some but not all of the information provided by the semantic representation of a lexical item (Grimshaw 1990). Although derived from the LMT, the theory is called the 'Functional Mapping Theory', for the reason that it no longer applies in the lexicon.

In the argument-structure, the argument positions appear in an ordering determined by their thematic roles:

(83) agent > benefic. > goal > instr. > theme/patient > locative

This ordering is relevant to various grammatical phenomena, and it is therefore convenient to list the argument positions of a predicate in accord with the ordering in (83).

The next primitive concept is that of 'internal' argument. The internal arguments are those receiving an 'object-like' grammatical function—normally the Patient/Theme, but possibly others as well. Being internal is treated as a diacritic feature of argument positions.[38] *Give* will have one argument structure where the Recipient is not marked internal, giving rise to the construction where the Recipient is oblique, and marked with the preposition *to*, and another where it is, giving rise to the double-object construction. Now, if an argument is not marked as internal it will be realized as an oblique, with one exception: a non-internal argument at the top of the list (having the highest role in the hierarchy) will be an 'external' argument, and will be the subject rather than an oblique.The two lexical forms for **give** can thus be represented as:

(84) a. $\mathrm{Give}([\ \ ], [\ \ ]_{\mathrm{int}}, [\ \ ]_{\mathrm{int}})$

b. $\mathrm{Give}([\ \ ], [\ \ ], [\ \ ]_{\mathrm{int}})$

The maximal non-internal argument will be obligatorily realized as subject, unless it is 'suppressed' in the sense of Grimshaw (1990), in which case it is expressed as an oblique (a PP in English).[39] Extensionally, maximal non-internal arguments seem to be the same as 'first arguments' on the 'action tier' of Jackendoff (1990).

[38] We take it that at the moment there is no actual theory of internal and external arguments, but just some devices for notating them, and observations about what kinds of arguments are or can be external. In particular, we regard Grimshaw's (1990) resort to double parentheses around unaccusative subjects as a clear indication that her attempt to develop a hierarchy-based theory did not fully succeed.

[39] Neither Grimshaw nor Alsina offers a clear formal interpretation of what 'suppression' really amounts to. A diacritic feature which makes it both unnecessary and impossible for the argument to be expressed as the subject would seem to deliver the desired behavior, though not in a very satisfying way.

At the level of grammatical functions, the classification is between functions that are [±subj(ect)] and [±obl(ique)]. The SUBJ GF is [+subj −obl]; the various object GFs, [−subj −obl]. In a substantial departure from the original assumptions of LFG, Alsina suggests that there can be an indefinite number of distinct OBJ-values, suggesting that this might be formalized by indexing the OBJ attributes in some manner.[40] Obliques are [+obl], and can appear multiply. Finally, there is a principle that [+obl] implies [−subj].

The mapping between argument-structure and grammatical functions is then controlled by three main principles:

(85) a. An (unsupressed) external argument must be [+subj −obl].

b. An internal argument is [−obl].

c. A clause has one and only one [+subj] GF (but can have more than one [−subj] ones, or none).

For the two argument-structures in (84), the first principle will cause the first item on the argument-lists to be subjects, while the remaining internal items will be objects. But if passivization applies, the external argument is suppressed, and one of the internal arguments can be subject. According to what we have seen so far, either of the two internal arguments in the second entry could be subject, while in English only the higher one on the hierarchy actually can:

(86) a. John was given the book.

b. *The book was given John.

But there are other languages, such as the 'symmetric' Bantu languages (for example, Kichaga and Kinyarwanda (Bresnan and Moshi 1990)) where passives corresponding to both (a) and (b) above are acceptable. We must therefore provide for additional principles that will allow some languages to impose further limits on what internal arguments can be subjects. We will also need principles relating linear order to relative order on the argument list. Alsina provides an account of these phenomena, motivated primarily by data from the Bantu languages, but they are beyond the scope of this paper.

In the past, LFG linking theories have been formulated to apply in the lexicon, but the present theory needs to apply in the syntax, so that it can work for syntactically composite complex verbs. Consider for example a 'faire-infinitive' causative sentence such as:

[40] This amounts to accepting the original proposal of Kimenyi (1980) that 'object' need not be a single-valued grammatical function, and explaining the differences in behavior between various kinds of objects in multiple-object constructions in terms other than their grammatical functions.

(87) **Ca** He fet llegir la carta al nen
I-have made read the map to-the boy

We will have two REL-values, a causative one for **fet**, and an ordinary verbal one for **llegir**.[41] But our linking theory has to apply in such a way as to give us only one subject and one accusative object, the same restrictions as apply in ordinary simple clauses with only one verbal REL-value. The pre-Resolution form of the Expression for (87) will be:

(88)
$$\left[\begin{array}{ll}
\mathrm{BAR} & 2 \\
\begin{array}{l}\mathrm{H}\\ \mathrm{O}\end{array} & \left[\begin{array}{ll}
\mathrm{BAR} & 1 \\
\begin{array}{l}\mathrm{H}\\ \mathrm{O}\end{array} & \left[\begin{array}{ll}
\mathrm{CAT} & \mathrm{V} \\
\mathrm{BAR} & 0 \\
\mathrm{REL} & \mathrm{Far}([\ \ ],[\ \ ]_{\mathrm{int}},[\ \ ]) \\
\mathrm{ARG} & \left[\begin{array}{ll}\mathrm{REL} & *(\ldots,[\ \ ],\ldots)\end{array}\right] \\
\mathrm{SUBJ} & \left[\begin{array}{ll}\mathrm{PERS} & \mathrm{I}\\ \mathrm{NUM} & \mathrm{SG}\\ \mathrm{REL} & \mathrm{PRO}\end{array}\right]
\end{array}\right] \\
\mathrm{ARG} & \left[\begin{array}{ll}
\mathrm{BAR} & 1 \\
\begin{array}{l}\mathrm{H}\\ \mathrm{O}\end{array} & \left[\begin{array}{ll}
\mathrm{CAT} & \mathrm{V} \\
\mathrm{BAR} & 0 \\
\mathrm{REL} & \mathrm{Llegir}([\ \ ],[\ \ ]_{\mathrm{int}})
\end{array}\right] \\
\mathrm{OBJ} & [\textit{la carta}] \\
\mathrm{OBJ} & [\textit{al nen}]
\end{array}\right]
\end{array}\right]
\end{array}\right]$$

Here the '*' represents an underspecified semantic structure that will merge with any REL-value. The material in the argument-list of '*' indicates that some argument of the base predicate is to unify with the second argument of the cause predicate (this unification is subject to complex conditions which will not be discussed here).

Resolution will now cause the GF values to be shared between structures connected by H and ARG, and the semantic function values ARG and REL to be shared between structures connected by O. A more compact presentation of the relevant information in (88) after resolution will therefore be the following, where the ellipsis dots stand for [H,O] attributes whose relevant attributes have spread up to the higher levels of the structure:

[41] The causative relation is here assumed to be ternary, following the proposals of Mohanan (1990) and Alsina (in preparation).

$$(89)\quad \begin{bmatrix} \text{CAT} & \text{V} \\ \text{BAR} & 2 \\ \text{REL} & \text{Far}([\ \ ],[\ \ ]_{\text{int}},[\ \ ]) \\ \text{ARG} & \begin{bmatrix}\text{CAT} & \text{V}\\ \text{BAR} & 1 \\ \text{REL} & \text{Llegir}([\ \ ],[\ \ ]_{\text{int}}) \\ \ldots & \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PERS} & \text{I}\\ \text{NUM} & \text{SG} \\ \text{REL} & \text{PRO}\end{bmatrix} \\ \text{OBJ} & [\textit{la carta}] \\ \text{OBJ} & [\textit{al nen}] \\ \ldots & \end{bmatrix}$$

Now, since linking is not applying in the lexicon, this structure contains argument-positions and grammatical function values, but these are not yet connected to each other: the argument-positions mostly have nothing in them, while the grammatical function values are not characterized as the arguments of anything.

The linking principles are now going to have to achieve the effect that even though (89) contains two REL-values, it constitutes only a single *argument domain* for the purposes of linking: there can be only one subject, and one accusative object, just as in a simple clause with only one REL-value. We can achieve this effect with the following idea: the linking principles apply to structures that are either the top level of whole expressions, or the value of one of the standard f-structural GFs, that is, SUBJ, OBJ, ADJUNCT, etc., but, crucially, not ARG. Linking will then apply to all of (89), and would apply internally to the subjects and objects if these had complex internal structure, but it won't apply to the value of ARG.[42]

What the linking rules will thus be aware of is not all of the information presented in (89), but rather the following:

[42] Equivalently, we can think of the expression being partitioned into domains through which f-projection attributes spread (all things connected by H or ARG links), and linking applying (once) in each such partition. The existence of these partitions show how the notion of f-structure levels remains latent in our representation.

(90)
$$\begin{bmatrix} \text{CAT} & \text{V} \\ \text{BAR} & 2 \\ \text{REL} & \text{Far}([\ \ ], [\ \ ]_{\text{int}}, [\text{REL Llegir}([\ \ ],[\ \ ]_{\text{int}})]) \\ \text{SUBJ} & \begin{bmatrix} \text{PERS} & \text{I} \\ \text{NUM} & \text{SG} \\ \text{REL} & \text{PRO} \end{bmatrix} \\ \text{OBJ} & [\textit{la carta}] \\ \text{OBJ} & [\textit{al nen}] \end{bmatrix}$$

Here the ARG attribute has been ignored, and its value has been inserted into the appropriate position in the argument-list of the causative predicate.

We then need to show that linking principles such as those in (85) can be expressed in our formalism. We will formulate them as conditionals, that equate certain substructures (argument-positions and GF-values) under certain conditions, or else impose constraints that the effects of other equations must satisfy. It seems virtually certain that these conditionals will take us outside the S.-B. class. Consider for example the second linking principle, which should allow internal argument positions to be unified with the value of OBJ grammatical functions. We want to say something like 'if x is an internal argument position w.r.t. y (y being the whole expression, in the present case), and z is the value of an OBJ-GF w.r.t. y, then x and z can be equated'. Now, x is an argument-position of y if you can get from y to x going down along REL and the attributes that the internal structures of predicates are made of, and x has whatever further properties characterize an argument position (that is, it can actually unify with an argument). But this notion cannot be described by a quantifier-free predicate-calculus formula over the primitive attributes that expressions are made of, since such an argument position might be indefinitely far down in the internal structure of the expression. We need to use the equivalent of a functional uncertainty expression, of the general form (y Sem* ArgPos) $= x$.

But in spite of not being S.-B., conditionals with such antecedents can in fact be added to LFG without losing decidability. This should be intuitively obvious from the fact that they do not create additional sub-feature-structures, but merely require that pre-existing structures unify, and can be seen more formally by translating the conditionals into disjunctions as was discussed in Section 2.

We can therefore make free use of relations such as 'x is an argument-position of y', taking this to hold, for example, when y is the entire structure in (90), and x any of the empty positions within the REL-value of y, but not when, say, x is the value of one of the GFs in (90). Alsina's external argument mapping principle (85a) can then be stated semi-formally as follows:

(91) If x is the external argument position of y, and z is the SUBJ of y, then x can be equated to z.

and the second principle could be translated similarly. Note that this principle is formulated so as to apply optionally, but, typically, if it does not apply, the structure will be Incomplete, Incoherent or both (more on these notions shortly). The third linking principle then applies as a filter, constraining the output of the others. It prevents internal arguments from becoming subjects when there is an external argument (since the external must be subject, and there can only be one subject). Conversely, when there is no external argument, the third principle requires that one of the internal arguments become the subject.

Continuing with the analysis of the causative, Alsina proposes a case-marking principle that requires that only the lowest internal argument on the argument-list can be accusative; all the others must be dative. The principle says that if x is an internal argument superior on the argument-list to another internal argument, then x must be dative case. This is clearly within the class of conditional rules we can allow. Therefore the 'causee subject' comes out dative, marked with the preposition **a** or **à** (**al** when combined with the article in (87)) while the object of the embedded verb comes out accusative.

The previous example does not motivate everything we have been assuming, since because the causee subject of **Llegir** is identified with the second argument of the causative predicate, which is internal, it is actually not clear whether the causee should otherwise be considered an external argument or not. If our definition of external argument is that it is the maximal non-internal argument in an argument domain, and we maintain that a causative complex predicate constitutes a single argument domain, then our prediction is that it isn't an external argument. As we will see, this seems to be correct, because even when the causee isn't unified with an argument position of the causative verb, it is still the case that only the causer can be expressed as the subject.

Catalan and the other Romance languages also have causatives where the causee subject is not expressed. Alsina analyzes these 'faire-par' causatives as involving a binary causative relation, as opposed to the ternary causative relation above, that resulted in the causee appearing as an internal argument. This construction is possible only when the highest argument of the embedded verb isn't marked as internal:

(92) a.	**Ca**	El	mestre	ha	fet	treballar	molt	aquest	curs.
		the	teacher	has	made	work	a lot	this	term
		The teacher has made people work a lot this term.							

b. *El fum farà sortir de la casa.
the smoke will make come out of the house
The smoke will make people come out of the house.

When the lower verb is 'unaccusative', realization of its highest (internal) argument is obligatory.

The relevant structure of these complex predicate constructions, after resolution, and looking only at the root node, would be:

(93)
$$\begin{bmatrix} \text{REL} & \text{`Far([ ],[ ]-)'} \\ \text{SUBJ} & [\cdots\ ] \\ \text{ARG} & \begin{bmatrix} \text{REL} & \text{`Treballar([ ])} \\ \ldots & \end{bmatrix} \end{bmatrix}$$

The first arguments of **far** and **treballar** are both 'external' in the conventional use of the term. However, we wish to maintain that a light verb construction like this constitutes a single argument domain, and in such a domain it is only the highest non-internal argument, that is, the causer of the causative verb, that actually counts as an external argument for the purposes of linking. Our principles thus make expression of the causer obligatory, not permitting the causee argument to become a SUBJ-value while the causer argument of **Far** is left unexpressed. Indeed, none of the linking rules in (85) are applicable to the highest argument of **treballar**; hence it must either remain unexpressed or be expressed by an adjunct. But why, then, cannot the unaccusative causee subject be left unexpressed? We will suppose that it is a principle of UG that unsuppressed internal arguments must be expressed.[43] There are many further aspects of causative linking considered by Alsina, but this should be sufficient to illustrate that his general approach is consistent with our style of formalization.

Above we promised to say something about Completeness and Coherence. Our current proposal for Completeness is not very interesting: it is that certain positions in the semantic structure of a verb require a syntactic filler. For example, the Recipient of *hand* cannot be omitted, while that of *give* may be:

(94) a. John handed some money *(to Mary).

b. John gave some money (to Mary)

[43]An alternative would be to follow Alsina's formulations a bit more closely, and divide linking into two stages. In the first stage, argument positions would be freely associated with GF-values, in the second, the linking principles would be applied as filters to the results of the free association. The second principle can then say that all internal argument positions *must* be associated with a [−obl] GF. Further information about the typology of unexpressed arguments would be desirable before trying to decide which style of analysis is best.

Although it would be highly desirable to be able to say something interesting about what is going on here, we do not have anything to add, other than to point out that the literature still tends to neglect the problem of obligatory adjuncts.

For Coherence, we have something more useful to offer. Take a resolved expression, and now remove all the relations (i.e., arcs) connecting subexpressions except the ones that are semantically interpreted. An expression containing a governable GF whose value has not been equated with an argument-position will now disintegrate into at least two disconnected pieces, since nothing will connect the semantic structure of the GF's value with that of the rest of the expression. So the Coherence principle can be stated literally as the requirement that the semantic projection be coherent.

The last task we will attempt is to give an explanation for Italian auxiliary-selection. Essentially, what we need is to provide a way of capturing Picallo's (1990) intuition that in cases of restructuring, it is the last heavy verb that acts as the main verb, determining argument structure and auxiliary choice. But this is quite straightforwardly achieved by saying that restructuring light verbs (unlike the homophonous non-restructuring complement-taking verbs and causative verbs) do not contain any argument positions at all. Although a 'control' light verb semantically links its controller with the logical subject of the embedded verb, it does not introduce any argument positions for linking.

The rule is then the same as for simplex clauses: the auxiliary is **avere** iff the verb's subject is an external argument.[44] Consider the structures of the examples in (69), repeated here for convenience:

(69) a. **It** Piero ha/*è voluto mangiare con noi
'Piero has/*'is' wanted to eat with us.'

b. Piero ha/è voluto venire con noi
'Piero has/'is' wanted to come with us.'

When **volere** is taken as 'light', the post-Resolution structure for (b) that is relevant for linking will be:[45]

(95)
$$\begin{bmatrix} \text{REL} & \text{Perf}([\ \ ]\text{-}) \\ \text{ARG} & \begin{bmatrix} \text{REL} & \text{Volere}(\text{—},[\ \ ]\text{-}) \\ \text{ARG} & \begin{bmatrix} \text{REL} & \text{Venire}([\ \ ]_{\text{int-A}}) \end{bmatrix} \end{bmatrix} \\ \text{SUBJ} & [\textit{Piero}] \\ \text{ADJUNCT} & [\textit{with us}] \end{bmatrix}$$

[44] To deal with reflexives in the same manner as Alsina (in preparation), we should strictly say, "only an external argument".

[45] Here, we are glossing over the representation of the comitative adjunct.

The lexical form of **volere** specifies that the Wanter is to be identified with the maximal unsupressed argument of the ARG-verb. The specification 'maximal unsupressed' is required to accommodate passivized verbs in the ARG-value, in examples such as:

(96) **It** Quei prigionieri vorebbero essere già stati liberati
'Those prisoners would like to already be freed.'

In the structure in (95), the Wanter is not an argument position, so there is only one argument position to be linked, and it is internal (w.r.t. **venire**), so the auxiliary **essere** is selected. In contrast, in the non-restructured variant, the Wanter serves as a linkable argument, and the auxiliary is **avere** because the Wanter is external (w.r.t. **volere**). Of course, in this case, the Wanter is also functionally identified with the argument of **venire**, but in this second, 'non-restructuring' configuration, to get from the REL of **volere** to the internal argument of **venire**, one must go through an expression that bears a GF, since the second argument of **volere** will be the value of XCOMP. Using the notion of *argument domain* that we introduced earlier, the condition for auxiliary selection is that **essere** is used if the value of SUBJ is an internal argument in the same argument domain as the auxiliary; otherwise **avere** is used. Thus the domains of grammatical function spreading, linking and auxiliary assignment are all correctly predicted to coincide.

There is, we think, considerably more to be done in sorting out how these notions of linking and argument domain are related to each other and to yet other notions. In particular we have said nothing about the contrasts in causatives between Italian on the one hand and French and Spanish on the other, that were observed by Zubizarreta (1985) (see also Rosen 1989). However, it does seem that the notions we have introduced can be used to give coherent and formally precise analyses to a substantial body of material that has resisted such analysis in the past.

4 Conclusion

Generative grammar traditionally sets great store on setting up 'levels of representation': places where certain kinds of information are placed, each with their own distinctive format. While much can be achieved with this approach, it often becomes unclear how the information distributed across different levels can be utilized in an integrated fashion. In this paper we have pursued a somewhat different approach, supposing that all the different kinds of information are in effect present in the same place, but kinds of information are distinguished by how they spread and interact in accord with universal principles of value sharing. The most important technique has been the postulation of principles for automatic sharing of certain attributes across others, a special case of a more general notion of 'conditional rule', a formal device that greatly increases what can be done with LFG-like syntactic theories. Thus we do not employ multiple levels of syntactic representation, but the levels of Classic LFG can be recaptured as derived notions within our new system.

This technique has allowed us to analyze some complex phenomena that have previously resisted analysis within LFG, and to do so with a favorable ratio of principles to stipulations. Our account provides predictive analyses of agreement and both scoping and non-scoping modifiers, and provides a reasonable basis for a formally rigorous and linguistically adequate account of verbal complex predicates.

Bibliography

Abney, S. P. 1987. *The English Noun Phrase in its Sentential Aspect.* PhD thesis, MIT.

Ackerman, F., and G. Webelhuth. 1991. Topicalization and German complex predicates. MS, UCSD and UNC.

Aissen, J. L., and D. M. Perlmutter. 1983. Clause reduction in Spanish. In D. M. Perlmutter (Ed.), *Studies in Relational Grammar 1*, 360–403. Chicago: University of Chicago Press. [Earlier version published in 1976].

Alsina, A. in preparation. *Predicate Composition: A Theory of Syntactic Function Alternations.* PhD thesis, Stanford.

Andrews, A. D. 1983. A note on the constituent structure of modifiers. *Linguistic Inquiry* 14:695–697.

Andrews, A. D. 1990. Unification and morphological blocking. *Natural Language and Linguistic Theory* 8:507–558.

Bach, E. 1968. Nouns and noun phrases. In E. Bach and R. Harms (Eds.), *Universals in Linguistic Theory*, 91–124. New York: Holt, Rinehart and Winston.

Baker, M. C. 1988. *Incorporation: A Theory of Grammatical Function Changing.* Chicago, IL: University of Chicago Press.

Bonet i Alsina, M. E. 1991. *Morphology after Syntax: Pronominal Clitics in Romance.* PhD thesis, MIT.

Bresnan, J., and J. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20:1–50.

Bresnan, J., and L. Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2):147–186.

Burzio, L. 1986. *Italian Syntax: a Government-Binding Approach.* Dordrecht: D. Reidel.

Chomsky, N. 1992. A minimalist program for linguistic theory. MIT Occasional Papers in Linguistics, Number 1.

Cooper, R. 1983. *Quantification and Syntactic Theory.* Dordrecht: D. Reidel.

Dalrymple, M. E. 1992. Categorial semantics for LFG. In *Proceedings of COLING-92*, Nantes, France.

Dalrymple, M. E. 1993. *The Syntax of Anaphoric Binding.* Stanford, CA: CSLI.

Fenstad, J. E., P.-K. Halvorsen, T. Langholm, and J. van Benthem. 1987. *Situations, Language and Logic.* Dordrecht: D. Reidel.

Grimshaw, J. 1982. On the lexical representation of Romance reflexive clitics. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 87–148. Cambridge, MA: MIT Press.

Grimshaw, J. 1990. *Argument Structure.* Cambridge, MA: MIT Press.

Grimshaw, J. forthcoming. Extended projection. MS, Brandeis University.

Gruber, J. 1965. *Studies in lexical relations.* PhD thesis, MIT.

Halvorsen, P.-K. 1983. Semantics for Lexical-Functional Grammar. *Linguistic Inquiry* 14:567–615.

Halvorsen, P.-K., and R. M. Kaplan. 1988. Projections and semantic description in Lexical-Functional Grammar. In *Proceedings of the International Conference on Fifth Generation Computer Systems, FGCS-88*, 1116–1122, Tokyo. Institute for New Generation Systems.

Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* PhD thesis, University of Massachusetts, Amherst.

Jackendoff, R. 1990. *Semantic Structures.* Cambridge, MA: MIT Press.

Johnson, M. 1991. Features and formulae. *Computational Linguistics* 17:131–151.

Kamp, H. 1981. A theory of truth and semantic representation. In *Formal Methods in the Study of Language.* Amsterdam: Mathematical Center. Reprinted in J. Groenendijk et al, eds. (1984) *Truth, Interpretation and Information*, Dordrecht: Foris.

Kaplan, R. M. 1987. Three seductions of computational psycholinguistics. In P. Whitelock, H. Somers, P. Bennett, R. Johnson, and M. M. Wood (Eds.), *Linguistic Theory and Computer Applications*, 149–181. London: Academic Press.

Kaplan, R. M., and J. Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 173–281. Cambridge, MA: MIT Press.

Kaplan, R. M., K. Netter, J. Wedekind, and A. Zaenen. 1989. Translation by structural correspondences. In *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*, 272–281, University of Manchester.

Kay, M. 1979. Functional grammar. In *Proceedings of the Berkeley Linguistics Society 5*, 142–158.

Kayne, R. S. 1991. Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22:647–686.

Keenan, E. L., and L. M. Faltz. 1985. *Boolean Semantics for Natural Language*. Dordrecht: D. Reidel.

Kimenyi, A. 1980. *A Relational Grammar of Kinyarwanda*. Berkeley: University of California Press.

Mallinson, G. 1986. *Rumanian*. Dover, NH: Croom Helm.

Manning, C. D. 1992. Romance is so complex. Technical Report CSLI-92-168, Stanford University, Stanford CA.

May, R. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

Mohanan, T. 1990. *Arguments in Hindi*. PhD thesis, Stanford University.

Moore, J. 1990. Spanish clause reduction with downstairs cliticization. In K. Dziwirek, P. Farrell, and E. Mejías-Bikandi (Eds.), *Grammatical Relations: A Cross-Theoretical Perspective*, 319–333. Stanford, CA: Stanford Linguistics Association.

Napoli, D. J. 1981. Semantic interpretation vs. lexical governance. *Language* 57:841–887.

Nunberg, G. 1977. *The Pragmatics of Reference*. PhD thesis, CUNY Graduate Center.

Parsons, T. 1990. *Events in the semantics of English: a study in subatomic semantics*. Cambridge, MA: MIT Press.

Perlmutter, D. M. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Picallo, M. C. 1990. Modal verbs in Catalan. *Natural Language & Linguistic Theory* 8:285–312.

Pizzini, Q. A. 1982. The positioning of clitic pronouns in Spanish. *Lingua* 57:47–69.

Pollard, C., and I. A. Sag. 1987. *Information-Based Syntax and Semantics*. Vol. 1. Stanford, CA: Center for the Study of Language and Information.

Pollard, C., and I. A. Sag. forthcoming. *Information-Based Syntax and Semantics*. Vol. 2. Stanford, CA: Center for the Study of Language and Information.

Raposo, E., and J. Uriagereka. 1990. Object agreement in the impersonal *-se* passive construction in European Portugese. In K. Dziwirek, P. Farrell, and E. Mejías-Bikandi (Eds.), *Grammatical Relations: A Cross-Theoretical Perspective*, 387–399. Stanford, CA: Stanford Linguistics Association.

Rivas, A. 1977. *A Theory of Clitics*. PhD thesis, MIT.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris. [Chapter 1 published in 1978 as "A Restructuring Rule in Italian Syntax"].

Rosen, S. T. 1989. *Argument Structure and Complex Predicates*. PhD thesis, Brandeis University.

Schachter, P. 1976. A non-transformational account of gerundive nominalizations in English. *Linguistic Inquiry* 7:205–242.

Shieber, S. M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Chicago: Chicago University Press.

Simpson, J. H. 1991. *Warlpiri Morpho-Syntax*. Dordrecht: Kluwer.

Van Valin, R. 1990. Semantic parameters of split intransitivity. *Language* 221–260.

Zubizarreta, M. L. 1985. The relation between morphophonology and morphosyntax: The case of Romance causatives. *Linguistic Inquiry* 16:247–290.