

Type-driven semantic interpretation and feature dependencies in R-LFG

Mark Johnson

Revision of 23rd August, 1997

1 Introduction

This paper describes a new formalization of Lexical-Functional Grammar called R-LFG (where the “R” stands for “Resource-based”). The formal details of R-LFG are presented in Johnson (1997); the present work concentrates on motivating R-LFG and explaining to linguists how it differs from the “classical” LFG framework presented in Kaplan and Bresnan (1982).

This work is largely a reaction to the linear logic semantics for LFG developed by Dalrymple and colleagues (Dalrymple et al., 1995; Dalrymple et al., 1996a; Dalrymple et al., 1996b; Dalrymple et al., 1996c). As explained below, it seems to me that their “glue language” approach bears a striking resemblance to semantic interpretation in those versions of categorial grammar which exploit the Curry-Howard correspondence (Girard, Lafont, and Taylor, 1989; van Benthem, 1995), such as Lambek Categorical Grammar and its descendants. A primary goal of this work is to develop a version of LFG in which this connection is made explicit, and in which semantic interpretation falls out as a by-product of the Curry-Howard correspondence rather than needing to be stipulated via semantic interpretation rules.

Once one has enriched LFG’s formal machinery with the linear logic mechanisms needed for semantic interpretation, it is natural to ask whether these make any existing components of LFG redundant. As Dalrymple and her colleagues note, LFG’s f-structure completeness and coherence constraints fall out as a by-product of the linear logic machinery they propose for semantic interpretation, thus making those f-structure mechanisms redundant. Given that linear logic machinery or something like it is independently needed for semantic interpretation, it seems reasonable to explore the extent to which it is capable of handling feature structure constraints as well.

R-LFG represents the extreme position that *all* linguistically required feature structure dependencies can be captured by the resource-accounting machinery of a linear or similar logic independently needed for semantic interpretation. The goal is to show that LFG linguistic analyses can be expressed as clearly and perspicuously using the smaller set of mechanisms of R-LFG as they can using the much larger set of mechanisms in LFG: if this is the case then we will have shown that positing these extra f-structure mechanisms are not linguistically warranted. One way to show this would be to present a translation procedure which reduces LFGs to equivalent R-LFGs, but currently no such procedure is known. Thus we proceed on a case by case basis, demonstrating that particular LFG analyses can be expressed at least as well in R-LFG.

R-LFG is also of interest because it proposes a radically different basis for feature structure interaction. In “unification-based” theories of grammar feature structures are typically viewed as static objects, which are the solutions to systems of feature structure constraints (called *f-descriptions* in LFG) (Kaplan and Bresnan, 1982; Rounds, 1997; Shieber, 1986). However, linguists often talk informally of “feature assignment” and “feature checking”; notions which cannot be expressed in a pure unification grammar. As discussed below, LFG does contain formal devices which can express these notions indirectly, viz., the non-monotonic devices such as existential constraints and constraint equations. On the other hand, the resource oriented nature of R-LFG provides a direct and natural formalization of the intuitions behind feature assignment and feature checking.

The rest of this paper is structured as follows. The next section introduces type-driven semantic interpretation from f-structures, and the one after that sketches the architecture of R-LFG and compares it to that of standard LFG. The following section introduces the reader to the idea that features are resources by demonstrating that one method of describing agreement relationships in standard LFG already possesses a resource-oriented character. The section following that describes how very simple agreement relationships can be described in R-LFG, and the final substantive section shows how Andrews (1982) analysis of Icelandic Quirky Case marking can be re-expressed in R-LFG.

2 Type-driven semantic interpretation from f-structures

This section develops type-driven semantic interpretation from graph structured resources used in R-LFG, motivating it by considering type-driven semantic interpretation from linearly ordered or sequential structures of resources in categorial grammar.

As has been often observed, the types of semantic objects constrain how they can combine, and hence the interpretations that can be possibly constructed from a bag of semantic objects. For example, suppose the words *Sandy* and *snores* are given the semantic interpretations in (1) and (2) with the types as shown.

$$Sandy' : e \tag{1}$$

$$\lambda x.snores'(x) : e \multimap t \tag{2}$$

(The symbol ‘ \multimap ’ is the implication symbol of Linear Logic, so the type $e \multimap t$ would be written $e \rightarrow t$ in a Montagovian notation for types). Now, there is only one way of combining these semantic objects to form a saturated proposition of type t , namely by applying the semantic interpretation of the verb *snores* to the interpretation of *Sandy* as its argument, so this is the only possible interpretation of the intransitive clause *Sandy snores*. This combination can be depicted as a proof (shown in natural deduction format here), where the two input semantic forms constitute the assumptions, and the single saturated proposition produced by the combination constitutes the conclusion.¹

$$\frac{\lambda x.snores'(x) : e \multimap t \quad Sandy' : e}{snores'(Sandy') : t}$$

It is worth reflecting on what is going on here. The types alone determine whether a particular way of combining lexical meanings is possible or not. The λ -terms are purely decorative: they are determined (up to reduction and renaming of variables) by the meanings of the lexical inputs and the structure of the combination.

The idea that a logic can be used to describe the possible modes of combination of a collection of objects underlies the Curry-Howard correspondence, and is at the root of much recent work in Categorical Grammar (van Benthem, 1995). The formulae of such a logic are the types of the objects being manipulated, and a proof in this logic corresponds to a particular way of

¹The resulting semantic form has been simplified via β -reduction.

combining the objects. The λ -terms are purely decorative: they are images of the structure of this proof, and play no role in determining whether a combination is possible or not.

Unfortunately, in more complex sentences semantic type constraints alone are not sufficiently restrictive to provide just the available interpretations. For example, if the semantic interpretations of the three words in the sentence *Sandy likes Kim* are as given in (1), (3) and (4)

$$\lambda y \lambda x. \text{likes}'(x, y) : e \multimap e \multimap t \quad (3)$$

$$\text{Kim}' : e \quad (4)$$

(where '*likes*' associates to the right) then besides permitting a combination corresponding to the available interpretation

$$\frac{\text{Sandy}' : e \quad \frac{\lambda y \lambda x. \text{likes}'(x, y) : e \multimap e \multimap t \quad \text{Kim}' : e}{\lambda x. \text{likes}'(x, \text{Kim}') : e \multimap t}}{\text{likes}'(\text{Sandy}', \text{Kim}') : t} \quad (5)$$

the semantic type constraints alone also permit an interpretation in which subject and object are exchanged.

$$\frac{\text{Sandy}' : e \quad \lambda y \lambda x. \text{likes}'(x, y) : e \multimap e \multimap t}{\lambda x. \text{likes}'(x, \text{Sandy}') : e \multimap t} \quad \text{Kim}' : e}{\text{likes}'(\text{Kim}', \text{Sandy}') : t} \quad (6)$$

It is obvious why the unintended interpretation was obtained. The semantic types do not reflect any information about the syntactic structure of the sentence: merely requiring semantic type compatibility amounts to treating a sentence as a bag of words, ignoring all other structural relationships between the words. Clearly this is incorrect for a language like English (as this example shows).

Categorial grammar deals with this problem by refining the structural sensitivity of the system: the elements manipulated are taken to be a linearly ordered sequence of categories, rather than just a bag). Correspondingly, the types must be refined to make them sensitive to this additional structural information. The single implication ' \multimap ' used above is specialized into a rightward-looking implication '/' and a leftward-looking implication '\', respectively.

The types associated with intransitive and transitive verbs are refined from (2) and (3) to (7) and (8), which specify the directions in which their

arguments are to be found.

$$\lambda x. \textit{snores}'(x) : e \setminus t \quad (7)$$

$$\lambda y \lambda x. \textit{likes}'(x, y) : (e \setminus t) / e \quad (8)$$

This directional sensitivity rules out the unattested combination (6), only permitting a combination that corresponds to the available interpretation.

$$\frac{\textit{Sandy}' : e \quad \frac{\lambda y \lambda x. \textit{likes}'(x, y) : (e \setminus t) / e \quad \textit{Kim}' : e}{\lambda x. \textit{likes}'(x, \textit{Kim}') : e \setminus t}}{\textit{likes}'(\textit{Sandy}', \textit{Kim}') : t}$$

Categorial grammarians have developed many insightful linguistic analyses within this framework. The treatment of the syntax-semantics interface within a framework, such as Lambek Categorial Grammar and its descendants, is especially appealing: once the lexical types and modes of syntactic combination are specified, semantic interpretation comes “for free” via the Curry-Howard correspondence between proofs of type well-formedness and λ -terms.

However, the focus on linear order in categorial grammar goes against one of the central intuitions of Lexical-Function Grammar: that the level of word order and surface syntactic structure is not an appropriate one at which to state many cross-linguistic generalizations. Rather, many interesting cross-linguistic generalizations are more appropriately stated at the level of function-argument or f-structure.

For example, as Bresnan (1982) argues, the relationship between a verb and its direct object NP argument may be manifest in many different surface syntactic relationships:

- it may be indicated by an agreement marker on the verb, or by
- a case marker on the NP direct object, or by
- a syntactic configuration, where the object immediately precedes or follows the verb as is appropriate, or by
- some combination of the above.

At the level of function argument structure the cross-linguistic uniformity of grammatical relation changing operations such as Passive becomes apparent. Crucially for the analysis presented here, as far as is known function-argument relationships are interpreted uniformly cross-linguistically, despite their variation in surface syntactic realization.

Thus from the LFG perspective, the appropriate response to the unattested combination (6) is to make the types sensitive to function-argument structure rather than word order directly. That is, the input to the combinatory process of semantic interpretation should be f-structures, rather than strings of lexical items.

To some extent this is achieved in the work of Dalrymple and her colleagues. In their approach, semantic interpretation starts with an f-structure decorated with formulae from what they call a “glue language.” Semantic interpretation is obtained via a combinatory process sensitive to function-argument structure. Moreover, Dalrymple and colleagues have achieved an impressive empirical coverage using their glue language approach.

However, the glue language approach seems to suffer from a number of conceptual drawbacks:

- The formulae manipulated during the course of a derivation are an amalgam of linear logic terms, standard first-order terms and the “glue” relation ‘ \rightsquigarrow ’. No interpretation (model-theoretic or otherwise) for such amalgams has been offered, so the manipulations performed in the course of interpretation can be described as uninterpreted symbol pushing.
- The semantic combinatory operations in the glue language approach are formulated in terms of (first-order?) term unification, rather than the function application and abstraction operations familiar from model-theoretic semantics. It is known from the computational linguistics literature that first-order term unification can be used to simulate β -reduction of λ -terms in function application (Pereira and Shieber, 1987), but it is also known that this simulation only approximately captures the properties of function application. It would be interesting to see if a system where resources have a function-argument structure organization can be made to operate with the more standard function application and abstraction mechanisms, or if term unification is essential here.
- Semantic forms are explicitly constructed in the glue language approach, rather than merely reflecting the structure of the proof, as they do in a Lambek Categorical Grammar. In principle, the glue language formalism allows semantic interpretation rules to be written in which a rule fails to apply not because of a type incompatibility, but because of unification failure of semantic terms (i.e., terms on the

right of the ‘ \rightsquigarrow ’ relation). Thus these terms need not be restricted to the purely decorative role that semantic forms play in Lambek Categorical Grammar, but may determine the well-formedness of a proof. Again, it would be interesting to know if this is an essential property of semantic interpretation of f-structures, or if a system exploiting a Curry-Howard correspondence can be developed.

Thus the system developed here, R-LFG, is explicitly modelled on categorial grammars where semantic interpretation is obtained by a Curry-Howard correspondence. It differs from them in that the inputs to the derivational process have the graph structure of an f-structure, rather than the linear structure of a string. Borrowing the idea that features in feature structures can be described by modal operators in a multi-modal language (Kasper and Rounds, 1990; Rounds, 1997), grammatical relations can be formalized as propositional modal operators.

Returning to the earlier example, the NP *Sandy* and the transitive verb *likes* would be associated with the lexical entries (9) and (10).

$$Sandy' : e \tag{9}$$

$$\lambda y \lambda x. likes'(x, y) : OBJ e \multimap SUBJ e \multimap t \tag{10}$$

(The modal operators ‘SUBJ’, ‘OBJ’, etc., are semantically vacuous, i.e., always semantically interpreted by identity functions, and bind more tightly than the implication symbol ‘ \multimap ’). This entry indicates that the verb *likes* first applies to an object of type e (embedded within the OBJ grammatical relation), yielding a function which in turn applies to a subject of type e to yield a saturated proposition of type t .

Assuming that in a transitive clause such as *Sandy likes Kim* the NP *Sandy* can be identified as subject and *Kim* as object (in English, this occurs by virtue of their c-structure locations), the following derivation yields the one available interpretation for this sentence.

$$\frac{Sandy' : SUBJ e \quad \frac{\lambda y \lambda x. likes'(x, y) : OBJ e \multimap SUBJ e \multimap t \quad Kim' : OBJ e}{\lambda x. likes'(x, Kim') : SUBJ e \multimap t}}{likes'(Sandy', Kim') : t}$$

Following standard treatments of feature structures, re-entrancies are described by path equations $f_1 \dots f_m = g_1 \dots g_n$, which permit a resource structure $f_1 \dots f_m \alpha$ to be transformed to $g_1 \dots g_n \alpha$. For example, Subject Raising in LFG is described in terms of a re-entrancy between the matrix

subject position and the complement’s subject position, licensed by a path equation associated with the Subject Raising verb. The lexical items in the sentence *Sandy seems happy* would be associated with the lexical entries (9), (11) and (12).

$$\lambda P.seems'(P) : XCOMP t \multimap t, SUBJ = XCOMP SUBJ \quad (11)$$

$$\lambda x.happy'(x) : SUBJ e \multimap t \quad (12)$$

Again, assuming that *Sandy* and *happy* are identified as filling the SUBJ and XCOMP grammatical functions respectively, the following deduction shows how the available interpretation for *Sandy seems happy* can be obtained.

$$\frac{\lambda P.seems'(P) : XCOMP t \multimap t \quad \frac{\lambda x.happy'(x) : XCOMP(SUBJ e \multimap t) \quad \lambda x.happy'(x) : XCOMP SUBJ e \multimap XCOMP t}{happy'(Sandy') : XCOMP t} * \quad \frac{Sandy' : SUBJ e \quad SUBJ e = XCOMP SUBJ e}{Sandy' : XCOMP SUBJ e}}{seems'(happy'(Sandy')) : t}$$

The inference labelled ‘*’ requires the grammatical relation XCOMP to distribute over the implication operator ‘ \multimap ’.

3 R-LFG: a simplification of LFG

The architectural simplification of R-LFG is best appreciated when compared with that of standard LFG together with the linear logic semantics augmentation of Dalrymple et. al. This section starts by sketching the architecture of standard LFG, and then presents the revised architecture of R-LFG.

3.1 The architecture of standard LFG

Figure 1 shows the architecture of this “standard” LFG. The components of LFG as presented by Kaplan and Bresnan (1982) are shown inside the dotted box in this figure, and the linear logic machinery for semantic interpretation posited by Dalrymple et. al. is depicted outside this box: see these references for further details.

In LFG, a syntactic description of an utterance is taken to be a pair constituting of a c-structure and an f-structure.² The yield of the c-structure

²There are proposals for additional structures, which for simplicity are ignored here.

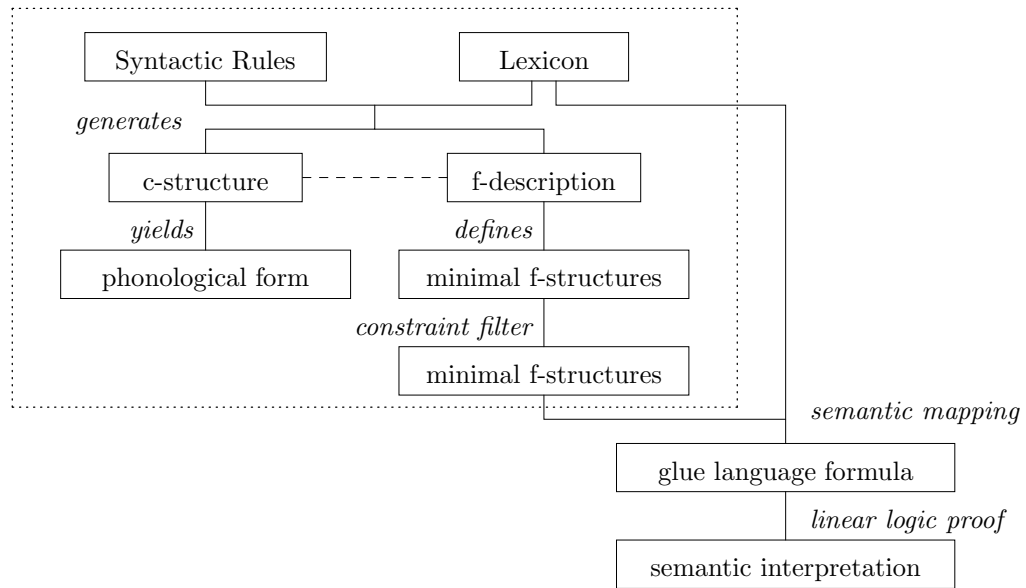


Figure 1: The architecture of standard LFG. The linear logic semantics component is shown outside the dotted box.

tree determines the phonological form of the sentence it describes.

The c-structure/f-structure pairs generated by an LFG are determined by the following procedure. The syntactic rules and lexical entries of an LFG together generate a set of c-structure trees, each of which is paired with a formula called an f-description which identifies which (if any) f-structures this c-structure can be paired with. The f-descriptions are boolean combinations of equations. These equations come in two kinds: *defining* and *constraining* equations.

The simplest account of the relationship between f-descriptions and the f-structures they describe seems to be procedural, following Kaplan and Bresnan (1982). First, the f-description is expanded into Disjunctive Normal Form (DNF) and the f-structure solution to each conjunct is determined as follows. The constraining equations are temporarily ignored (i.e., replaced with *true*) and if the resulting formula is satisfiable and has a unique minimal satisfying f-structure, that f-structure is a candidate solution to the conjunct. This candidate solution is a (true) solution to the conjunct just in case it also satisfies the formula obtained by replacing each constraining equation in the conjunct with corresponding defining equations. The set of solutions to an f-description is the union of the set of solutions to each conjunct of its DNF, so the f-description determines a finite number of f-structures.³

³ To appreciate some of the difficulties in giving a declarative treatment of LFG's constraint equations, consider a treatment of Case marking in which subject NPs are optionally assigned a nominative Case feature *NOM*, such as the Andrews (1982) analysis of Icelandic quirky case marking discussed in section 5.2, using the following LFG syntactic rule.

$$S \rightarrow \begin{array}{ccc} & \text{NP} & \text{VP} \\ & (\uparrow \text{SUBJ}) = \downarrow & \uparrow = \downarrow \\ & ((\uparrow \text{SUBJ CASE}) = \text{NOM}) & \end{array}$$

The parentheses surrounding the lower equation annotating the NP indicates that this defining equation is optional, reflecting the fact that the subject NP is only optionally assigned nominative case (as it may be assigned a ‘quirky’ non-nominative case by the verb, as explained below). This annotation presumably abbreviates the following disjunction:

$$(\uparrow \text{SUBJ CASE}) = \text{NOM} \vee \textit{true}$$

Clearly replacing this disjunction with *true* does not change the set of minimal models for any f-description which contains it, so the equation itself has no effect on the minimal models, and hence cannot result in the satisfaction of any constraint equations. Clearly this is not the intended interpretation: the “purpose” of this equation is to provide a Case feature to satisfy the requirements of the subject NP.

Kaplan and Bresnan (1982) do not discuss disjunction, but it appears they intend disjunctions to be interpreted as an abbreviatory convention, i.e., that their process applies only to individual conjunctions after expansion to a Disjunctive Normal Form (DNF). Thus

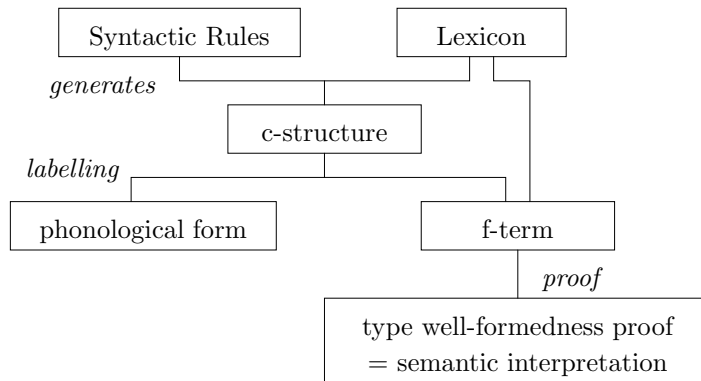


Figure 2: The architecture of R-LFG.

Dalrymple et. al. use these f-structures as the input to their semantic interpretation procedure. Certain elements in an f-structure are associated with formula in a *glue language*, which is an amalgam of linear logic and classical first-order logic, in effect mapping each f-structure into a formula of the glue language. For semantic interpretation to succeed this glue language formula must derive a term with the type of a saturated proposition: the argument of this term is the semantic interpretation of the sentence.

3.2 The architecture of R-LFG

The architecture of R-LFG is depicted in Figure 2. The most striking difference between LFG and R-LFG is that R-LFG does not contain an independent level of f-structure representation, since the same mechanisms used for semantic interpretation are also used to account for syntactic feature dependencies. Given that it is a simpler architecture, it should be preferred on grounds of parsimony.

The lexical entries and syntactic rules of R-LFG generate c-structure/f-term pairs in the same way that they generate c-structure/f-description pairs in LFG. In LFG several steps are required to obtain the f-structures that serve as the input to semantic interpretation from the f-descriptions. How-

their treatment, while not falling foul of the problem just noted, involves a rather curious mixture of proof-theoretic devices (e.g., DNF expansion) and model-theoretic devices (e.g., focussing on minimal models).

ever, in R-LFG the f-term serves as the input to semantic interpretation directly. Thus in R-LFG the linguistic effects of f-structure constraints must be obtained by other means, viz., the same logical mechanisms used for semantic interpretation.

As explained below, these logical mechanisms enforce a *resource accounting* which ensures that every predicate combines with an appropriate number of arguments and that every non-root semantic unit appears as the argument of some predicate. The semantic interpretation itself is determined by the pattern of predicate-argument combination via a Curry-Howard correspondence, as explained in more detail in Johnson (1997).

This same resource accounting mechanism is also used to describe feature dependencies. Purely syntactic features with no semantic content differ from semantically interpreted elements only in that they are semantically vacuous, i.e., given trivial interpretations which are systematically ignored by any functors which take them as arguments.

The resource logic used here differs considerably from the glue language used by Dalrymple et. al. That language includes first-order terms with equality, which can be used to encode feature structure unification in the manner of e.g., Definite Clause Grammars (see Shieber (1986) for a description of the relationship between the first-order terms of Definite Clause Grammars and attribute-value “unification” grammars) and hence directly simulate f-structure attribute-value constraints. While this would provide a straightforward way to encode f-structure constraints in the glue language, it is not clear that such an approach would constitute a real simplification of LFG, rather than just a reshuffling of its complexity.

For this reason, R-LFG uses a much simpler resource logic than the glue language of Dalrymple et. al. Inspired by recent work in Categorical Grammar such as Morrill (1994), the resource logic is propositional modal logic that encodes the types of the semantic objects being manipulated, and the semantic interpretation itself is provided by a Curry-Howard correspondence between proofs and λ -terms (Girard, Lafont, and Taylor, 1989). As van Benthem (1995) demonstrates, a wide variety of substructural logics possess a Curry-Howard correspondence, so the requirement that semantic interpretation is obtained in this way does not identify a particular logic. Rather, the precise logic used should be chosen to best fit the linguistic phenomena described by the theory. Moortgat (1997) develops the theory of propositional multimodal logics used here. The reader is referred to Johnson (1997) for the full details of R-LFG.

4 Describing agreement relationships with LFG

This section argues that Lexical-Functional grammarians typically use the formal devices of LFG to manipulate features as resources that are assigned and checked. It introduces two methods often used for describing agreement relationships in LFGs. It turns out that one method, which crucially relies on “constraining equations”, can be viewed as describing agreement in terms of resource dependencies. Thus resource based accounts of agreement are not a new innovation of R-LFG, but are already a familiar part of LFG. The principal claim behind R-LFG is that *all* linguistic dependencies can be expressed in this manner, and that the explicit resource-orientation of R-LFG simplifies and clarifies the nature of the linguistic dependencies concerned.

As sketched above and explained in more detail in Kaplan and Bresnan (1982), LFG’s f-descriptions contain two different kinds of equations. A defining equation instantiates the value of an attribute, while a constraining equation checks that a value is instantiated by a defining equation elsewhere in the f-description. The linguistic dependencies involved in simple agreement can be described using defining equations alone, or by using a mixture of defining and constraining equations. This latter method has a natural resource interpretation.

To keep things clear, the two methods for describing agreement relationships are explained using the same examples (13).

- (13) a. Sandy snores.
b. Professors snore.

Both methods of describing agreement relationships require that the agreeing items (in (13a), *Sandy* and *snores*) are capable of constraining the value of the same f-structure element; this is usually achieved by defining equations associated with syntactic rules. The agreeing items both impose constraints the value of that f-structure element, thus ensuring that only compatible items can appear simultaneously in a syntactic structure.

4.1 Agreement using defining equations alone

In this method, both agreeing items constrain the shared f-structure element using defining equations. For example, the grammar fragment in (14–18) generates exactly the two sentences in (13). The c-structure and f-structure generated by this fragment for (13a) is depicted in Figure 3.

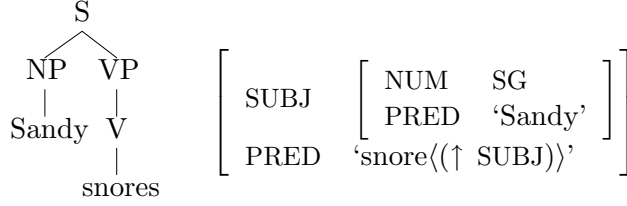


Figure 3: The c-structure and f-structure for *Sandy snores* generated by the fragment (14–18).

$$\begin{array}{lcl}
 \textit{Sandy} & \text{NP} & (\uparrow \text{ PRED}) = \text{'Sandy'} \\
 & & (\uparrow \text{ NUM}) = \text{SG}
 \end{array} \tag{14}$$

$$\begin{array}{lcl}
 \textit{Professors} & \text{NP} & (\uparrow \text{ PRED}) = \text{'professor'} \\
 & & (\uparrow \text{ NUM}) = \text{PL}
 \end{array} \tag{15}$$

$$\begin{array}{lcl}
 \textit{snores} & \text{VP} & \uparrow \text{ PRED}) = \text{'snore}(\uparrow \text{ SUBJ})\text{' } \\
 & & \underline{(\uparrow \text{ SUBJ NUM}) = \text{SG}}
 \end{array} \tag{16}$$

$$\begin{array}{lcl}
 \textit{snore} & \text{VP} & \uparrow \text{ PRED}) = \text{'snore}(\uparrow \text{ SUBJ})\text{' } \\
 & & \underline{(\uparrow \text{ SUBJ NUM}) = \text{PL}}
 \end{array} \tag{17}$$

$$\begin{array}{lcl}
 \text{S} & \longrightarrow & \text{NP} \quad \text{VP} \\
 & & (\uparrow \text{ SUBJ}) = \downarrow \quad \uparrow = \downarrow
 \end{array} \tag{18}$$

The lexical entries for subject NPs require that the value of their NUM attribute is SG or PL as appropriate. In addition, the underlined equation in each verb’s lexical entry also requires that this value is appropriate for the verb’s inflection. If the subject and the verb require different values for this f-structure element (as in the ungrammatical **Professors snores*), the corresponding f-description will require this element to be equal to two different values (e.g., SG and PL). However, the well-formedness conditions on f-structures do not permit this (Kaplan and Bresnan, 1982; Johnson, 1995) so the f-descriptions associated with such sentences are inconsistent, and the sentences themselves are correctly predicted to be ungrammatical.

Thus this method functions by arranging for ungrammatical sentences to be associated with an inconsistent f-description. This observation is in fact quite general: if all grammatical relationships are described using defining

equations (i.e., if we restrict attention to the monotonic constraints) then the only way such an equation can have a grammatical “effect” is by being inconsistent with other equations, i.e., by “causing” ungrammaticality.

More precisely, suppose we identify a subset of the elements of a f-structure as follows. The *semantically interpreted elements* are those which serve as the input to the semantic interpretation procedure (in the framework of Dalrymple et. al. these elements are associated with glue language formulae at some stage during the interpretation process). The idea is the semantically uninterpreted elements can be deleted from an f-structure without changing its semantic interpretation. In a typical LFG, the values of attributes such as PRED, SUBJ, OBJ, etc., are semantically interpreted, while the values of CASE and GENDER (in a grammatical gender language) are not semantically interpreted.

Now consider a “pure unification” grammar without non-monotonic devices such as “constraining equations”, e.g., in which all equations are defining equations, such as the PATR grammars of Shieber (1986). These are grammars in which all linguistic relationships are expressed with defining equations. It is possible to show that in such a grammar, if an equation which equates only non-semantic values is not inconsistent with other equations on some input, then deleting it from the grammar does not affect the language generated or the interpretations assigned. (A similar observation holds in monotonic grammars such as HSPG).

This means that if all grammatical relationships are described using defining equations, a nonsemantic feature defining equation only has an effect on the language generated if somewhere else in the grammar there are defining equations that are inconsistent with this one. For example, there is no point in adding a defining equation that introduces an attribute that does not appear elsewhere in the grammar, such as

$$(\uparrow \text{ HISTORICAL-ORIGIN}) = \text{ROMANCE} \quad (19)$$

unless other defining equations that can possibly be inconsistent with it are also introduced. But in order to be inconsistent with (19) these other equations must require the attribute’s value to be *different* to the value specified in the former equation, e.g.,

$$(\uparrow \text{ HISTORICAL-ORIGIN}) = \text{GERMANIC} .$$

Thus with defining equations alone, different grammatical properties are based on feature *oppositions* or contrasts. The formal machinery of these

$$\left[\begin{array}{l} \text{SUBJ} \left[\begin{array}{l} \text{PRED} \text{ 'Sandy'} \\ \text{NUM} \text{ SG} \end{array} \right] \\ \text{SUBJ} \left[\begin{array}{l} \text{NUM} =_c \text{ SG} \end{array} \right] \\ \text{PRED} \text{ 'snore}(\langle \uparrow \text{SUBJ} \rangle)' \end{array} \right]$$

Figure 4: A alternative minimal f-structure solution to the f-description for (13a) obtained by relaxing the functionality requirements on f-structures. Note that this f-structure never the less does not satisfy the constraining equations expressing subject-verb agreement.

way that the defining equation method does. For example, the constraint equation (22) requiring that the NP receive an ACC case value does not rely on the existence of other Case values besides ACC; it functions just as well if ACC is the only Case value used in the grammar. That is, while a defining equation ensures that an attribute has one value rather than another, a constraining equation ensures in addition that the feature has in fact been given a value independently. Thus this method more fully supports privative features than the defining equation method does.

Further, the constraining equation method does not rely on the functionality axiom or the constant-constant clash axioms in the same way that the defining equation method does. For example, even if the functionality requirement on f-structures were relaxed so that the defining equations in the f-description for (13a) could have the second minimal f-structure solution depicted in Figure 4 besides the one depicted in Figure 3, that f-structure would fail to satisfy the constraining equation expressing subject-verb agreement, and so would be ill-formed for independent reasons.

In fact, feature structures in R-LFG behave very much in this way. While attributes are permitted to be single-valued, no feature structure axiom forces them to be so. But since grammatical relationships are described in a way very similiar to the constraining equation method, in general the grammatical requirements of predicates will require that attributes are single-valued.

4.3 Resource management in LFG

The constraining equation method of describing agreement relationships can be described in terms of *resources*, where the resource is the feature value of the shared f-structure entity. Each such feature value is *produced* by *one or more* defining equations, and is *consumed* by *zero or more* constraining equations. This pattern of resource management is formalized by Intuitionistic Logic.

Interestingly, the special properties LFG endows the values of PRED attributes with provides them with special resource management properties also. The values of PRED attributes must be *produced* by *exactly one* argument, and must be *consumed* by *one or more* predicates. The logic LPC developed by van Benthem (1995) formalizes this resource management.

Thus LFG already incorporates a number of mechanisms which can be seen as performing resource management. R-LFG attempts to describe all syntactic relationships in terms of such resource management. Identifying the appropriate resource management mode for a particular grammatical relationship is a key step in developing its R-LFG description.

5 Resource accounting in R-LFG

Johnson (1997) formally defines R-LFG's f-terms and presents a Gentzen sequent calculus that describes the resource management relationships between features. It also presents labelled deduction systems for describing the mappings from c-structures to f-terms, and semantic interpretation from f-terms. That paper should be consulted for the technical details of R-LFG; this section presents that material in an informal (and hopefully more accessible) manner.

An f-formula is an expression that indicates the type of a constituent, or more generally, a single resource. The semantic type of a constituent can be determined from its f-formula, but just as in the categorial grammar example above, f-formulae also specify additional syntactic constraints.

Following Morrill (1994), we distinguish semantically contentful types from semantically impotent types. The basic semantically contentful types e , t , etc., are f-formulae (these are the types of individuals and truth values respectively), as are the basic semantically impotent types NOM, ACC, etc., (which are interpreted by constants, and whose value is systematically ignored by any function that takes them as an argument). The full set of f-formulae are obtained by closing these under the following operations.

If ϕ is an f-formula then $f\phi$ is also an f-formula, where f is an attribute; it denotes the result of embedding ϕ under the attribute f .

If ϕ_1, ϕ_2 are f-formulae then $\phi_1 \multimap \phi_2$ is also an f-formula; it is a linear implication which consumes ϕ_1 to produce ϕ_2 .

The f-formulae are related to the more usual types of model-theoretic semantics is defined by the mapping $(\cdot)^\natural$ and a new type constant \emptyset for the semantically impotent f-formulae.

$(\phi)^\natural = \phi$ if ϕ is a semantically contentful basic type,

$(\phi)^\natural = \emptyset$ if ϕ is a semantically impotent basic type,

$(f\phi)^\natural = (\phi)^\natural$ where f is an attribute, and

$(\phi_1 \multimap \phi_2)^\natural = (\phi_2)^\natural$ if $(\phi_1)^\natural = \emptyset$, and $(\phi_1)^\natural \rightarrow (\phi_2)^\natural$ otherwise.

For example, the natural type of an f-formula for an NP requiring a nominative case marking is $(\text{NOM} \multimap e)^\natural = e$. In general, it is required that the type of λ -term labelling an f-formulae ϕ (i.e., giving the constituent's semantic interpretation) be of type $(\phi)^\natural$. (Semantically impotent f-formulae are not labelled with λ -terms, as they have no natural semantic interpretation).

F-formulae are the building blocks of f-terms. Informally, a f-term is a graph-structured configuration of one or more constituents, or more generally, resources. F-formulae are f-terms, and if $\alpha, \alpha_1, \dots, \alpha_n$ are f-terms then:

$\alpha_1, \dots, \alpha_n$ is the *multiset* of resource structures $\{\alpha_1, \dots, \alpha_n\}$ (order is unimportant in a multiset, but the number of times an element appears is important),

$f \alpha$ is the result of *embedding* the structure α under the attribute f ,⁴

$f_1 \dots f_m = g_1 \dots g_n$ is a path equation which *restructures* an f-term by moving a resource structure embedded under the sequence of attributes $f_1 \dots f_m$ so that it is located under the sequence of attributes $g_1 \dots g_n$, and

(α) is an *optional* occurrence of the structure α .

⁴Johnson (1997) follows Moortgat (1997) in introducing a separate punctuation symbol to distinguish modal structures in f-terms from modal operators in f-formulae, but here we rely on context to distinguish these two usages.

An f-term describes a graph structure of constituents, or more generally, resources. The f-term associated with a sentence is required to simplify to a single resource of type t in order for the sentence to be grammatical. (This single requirement subsumes both the requirement that the f-description be satisfiable and the requirement that the Linear Logic glue formula simplify to an expression of type t in standard LFG). An f-term simplifies by applying linear implications, restructuring using path equations, distributing attributes over multisets, and either deleting optional elements or replacing them with their non-optional counterpart.

Attributes are permitted, but not required, to distribute and factor over multisets. That is, the following biimplication holds, where f is an attribute and α_1 and α_2 are f-terms:

$$f(\alpha_1, \alpha_2) \Leftrightarrow (f \alpha_1), (f \alpha_2).$$

Unlike LFG, R-LFG does not require that attributes are single-valued, nor does it enforce a constant-constant clash. Every f-term is “satisfiable” in that it represents some configuration of resources; grammaticality is determined by whether those resources can combine to produce a single element of type t (the type of a saturated proposition).

5.1 Nominative Case marking in English

A simple R-LFG fragment which describes structural nominative case assignment to subject NPs is presented below. The lexical entry for the nominative Case marked subject NP *Sandy* in (23) requires it to consume a NOM case resource in order to produce a resource of type e , and the lexical entry for the verb *snores* in (24) requires it to consume a resource of type e embedded within a SUBJ attribute in order to produce a resource of type t . The syntactic rule (25) specifies how the f-terms associated with the NP and VP (referred to by the meta-variable ‘ \downarrow ’ just as in LFG) are to be combined to produce the f-term for the S. In this case, a multiset consisting of the NP’s f-term and a NOM case resource is embedded within a SUBJ attribute, which together with the f-term associated with the VP yields the multiset f-term associated with the S.

$$\begin{array}{ll} \textit{Sandy} & \text{NP } \textit{Sandy}' : \text{NOM } \multimap e \end{array} \quad (23)$$

$$\begin{array}{ll} \textit{snores} & \text{VP } \lambda x. \textit{snores}'(x) : \text{SUBJ } e \multimap t \end{array} \quad (24)$$

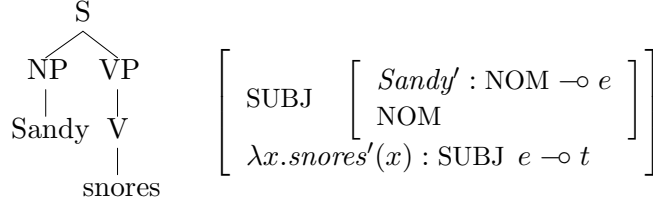


Figure 5: The c-structure and f-term for *She snores* generated by the fragment (23–25). The f-term simplifies straightforwardly to type t , yielding the semantic labelling $snores'(Sandy')$.

$$\begin{array}{ccc}
 \text{S} & \longrightarrow & \text{NP} \quad \text{VP} \\
 & & \text{SUBJ(NOM, } \downarrow) \quad \downarrow
 \end{array} \quad (25)$$

This fragment generates the c-structure and f-term depicted in Figure 5. The f-term simplifies to type t in the following steps:

$$\frac{\frac{\frac{Sandy' : \text{SUBJ(NOM } -o e)}{Sandy' : \text{SUBJ NOM } -o \text{ SUBJ } e} \quad \text{SUBJ NOM}}{Sandy' : e} \quad \lambda x.snores'(x) : \text{SUBJ } e -o t}{snores'(Sandy') : t}$$

5.2 Icelandic Quirky Case Marking

Quirky Case marking in Icelandic presents a more complex array of linguistic data which exercises a wider range of f-term machinery. This construction has proven difficult to encode in unification-based grammars, and has motivated several non-monotonic extensions to the basic unification grammar machinery, such as LFG’s constraint equations and a complex inheritance system in HPSG (Sag, 1995). The analysis presented here demonstrates how the resource sensitivity of R-LFG provides a simple way to encode the LFG analysis of Andrews (1982) without requiring recourse to complex extensions to the basic machinery of R-LFG.

In Icelandic, subject NPs are usually case marked nominative, as in (26a). However, a few verbs, such as *vantar* ‘lacks’ exceptionally case mark their subject NPs with accusative or some other non-nominative “quirky” case (26b). The subjects of subject raising verbs, such as *vir’ist* ‘seems’,

usually appear in nominative case (26c), but if the embedded verb is a quirky case assigning verb then the matrix subject is assigned the quirky case, rather than nominative (26d).

- (26) a. *drengurinn kyssti stúlkuna*
 the-boy.nom kissed the-girl.acc
 ‘The boy kissed the girl’
- b. *drengina vantar mat*
 the-boys.acc lacks food.acc
 ‘The boys lack food’
- c. *hann vir’ist elska hana*
 he.nom seems love her.acc
 ‘He seems to love her’
- d. *hana vir’ist vanta peninga*
 her.acc seems lack money.acc
 ‘She seems to lack money’

This pattern of data receives a straightforward informal account in terms of case assignment if we make the following assumptions:

- All NPs must receive exactly one case,
- Quirky case marking verbs always assign a quirky case,
- Case is preserved in Raising and other grammatical operations, and
- Structural nominative case is only optionally assigned.

Thus if a subject NP receives a quirky case, then that must be the case that it appears in. On the other hand, if the subject NP is not assigned a quirky case, then the only case available is structural nominative case.

This account can be formalized in R-LFG as follows. The phrase structure rules for this Icelandic fragment are the following.

$$S \longrightarrow \begin{array}{cc} \text{NP} & \text{VP} \\ \text{SUBJ}((\text{NOM}), \downarrow) & \downarrow \end{array} \quad (27)$$

$$\text{VP} \longrightarrow \begin{array}{c} \text{V} \\ \downarrow \end{array} \left(\begin{array}{c} \text{NP} \\ \text{OBJ} \downarrow \end{array} \right) \left(\begin{array}{c} \text{VP} \\ \text{XCOMP} \downarrow \end{array} \right) \quad (28)$$

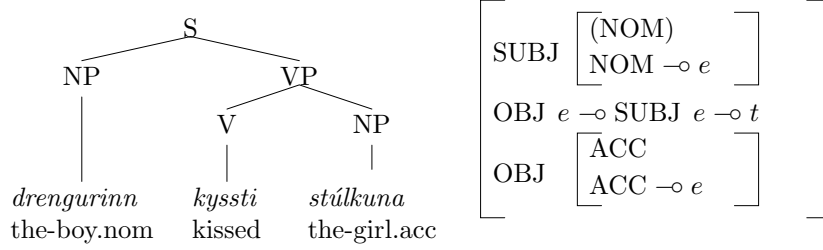


Figure 6: The c-structure and f-term for the single clause non-quirky Icelandic example (26a) generated by (27–31).

The phrase structure rule (27) differs from the corresponding English rule (25) in that it optionally embeds a NOM case under the SUBJ attribute. The phrase structure rule (28) introduces a verb, an optional direct object NP and an optional VP. It embeds the direct object NP’s f-term under the OBJ attribute and the VP’s f-term under the XCOMP attribute, as is standard in LFG.

The lexical entries (29–31) are required to generate the non-quirky single clause example (26a). The c-structure and f-term associated with this example are shown in Figure 6. It is straightforward to check that this f-term reduces to t , labelled with $kissed'(boy', girl')$.

$$drengurinn \quad \text{NP} \quad boy' : \text{NOM} \multimap e \quad (29)$$

$$stúlkuna \quad \text{NP} \quad girl' : \text{ACC} \multimap e \quad (30)$$

$$kyssti \quad \text{V} \quad \lambda y \lambda x.kissed'(x, y) : \text{OBJ } e \multimap \text{SUBJ } e \multimap t, \quad (31)$$

OBJ ACC

The single clause quirky case marked example is only slightly more complex. It can be described with the three additional lexical entries (32–34).

$$drengina \quad \text{NP} \quad boys' : \text{ACC} \multimap e \quad (32)$$

$$mat \quad \text{NP} \quad food' : \text{ACC} \multimap e \quad (33)$$

$$vantar \quad \text{V} \quad \lambda y \lambda x.lacks'(x, y) : \text{OBJ } e \multimap \text{SUBJ } e \multimap t, \quad (34)$$

OBJ ACC, SUBJ ACC

The lexical entry for the quirky case marking verb *vantar* ‘lacks’ in (34) differs from that for the non-quirky verb *kyssti* ‘kissed’ in that it assigns

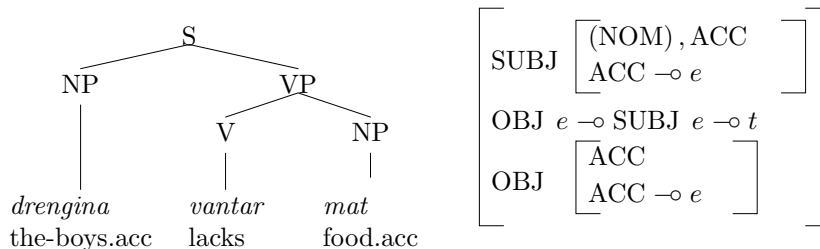


Figure 7: The c-structure and f-term for the single clause quirky case example (26b) generated by (27–34).

an accusative case to its subject (in the underlined part of the f-term) as well as to its object. The c-structure and f-term for (26b) are depicted in Figure 7. Again, it is straightforward to check that the f-term reduces to t , and is labelled with the λ -term $lacks'(boys', food')$. Note that if the subject were replaced with a nominative NP the f-term would no longer reduce to t , since the ACC case feature embedded under the SUBJ attribute could not be consumed.

The formalization of the non-quirky case Subject Raising example (26c) is very similar to the standard LFG account of Subject Raising (Bresnan, 1982). The lexical entry (35) for the Raising verb *vir'ist* ‘seems’ contains the path equation $SUBJ = XCOMP SUBJ$ which permits resources embedded under the SUBJ attribute to be restructured under the XCOMP SUBJ attributes. In this example, a resource of type e is lowered into the embedded clause. The f-term associated with this example is depicted in Figure 8. (Here we ignore the complexities of pronominal binding, and treat the pronouns simply as NPs that consume a nominative or accusative case resource). It is straightforward to check that this reduces to t , and is labelled with the λ -term $seems'(loves'(he', her'))$.

$$\begin{array}{lll}
 \textit{vir'ist} & V & \lambda P.seems'(P) : XCOMP \ t \ \rightarrow \ t, \\
 & & SUBJ = XCOMP SUBJ
 \end{array} \tag{35}$$

$$\begin{array}{lll}
 \textit{elska} & V & \lambda y \lambda x.love(x, y) : OBJ \ e \ \rightarrow \ SUBJ \ e \ \rightarrow \ t, \\
 & & OBJ \ ACC
 \end{array} \tag{36}$$

The syntactic rules and lexical entries introduced above that are independently needed to account for quirky case marking in single clause con-

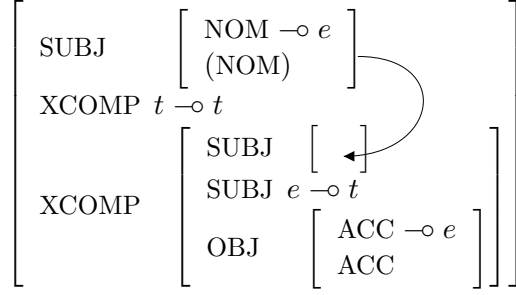


Figure 8: The f-term for the non-quirky Subject Raising example (26c) generated by (27–36).

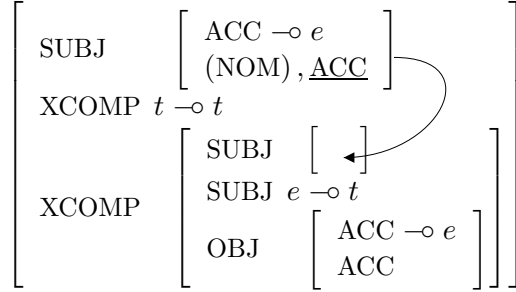


Figure 9: The f-term for the quirky case marked Subject Raising example (26d) generated by (27–36).

structions and for Subject Raising without quirky case also correctly account for the interaction of those two constructions, which was presented in (26d) on page 22. The f-term for this example is shown in Figure 9.

Just as in the single clause quirky case marking example (26b), the subject NP is assigned both an accusative case and an optional nominative case, so only an accusative subject NP can appear. If a nominative subject were inserted in matrix subject position it could consume the optional nominative case resource, but the accusative case resource assigned by the quirky verb to the subject would not be consumed, and so an f-term of type t could not be derived. It is straight-forward to check that the f-term depicted in Figure 9 simplifies to t , and that it is labelled with the λ -term

seems'(lack'(she', money')).

6 Conclusion

This paper has introduced a simplified version of LFG called R-LFG in which a single representation called an f-term plays the role of both f-description and f-structure. LFG's f-structure well-formedness constraints are re-expressed in terms of feature resource dependencies, which permits them to be checked by the same mechanism that performs semantic interpretation. It is not implausible that this can be done for many, if not most, LFG analyses, as many standard LFG analyses already have a resource oriented character, and it seems that the “core” LFG analyses of Raising, Control, etc., can be straightforwardly reexpressed in R-LFG.

Even if it turns out that the R-LFG project is ultimately untenable—perhaps it will be possible to demonstrate that some linguistically necessary properties of f-structures simply cannot be adequately captured using the resource logic machinery utilized for semantic interpretation—this research may still contribute by providing an alternative perspective on feature interactions in grammar and suggesting modifications or extensions to the standard LFG framework.

References

- Andrews, Avery D. 1982. The representation of Case in modern Icelandic. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*. The MIT Press, Cambridge, Massachusetts, pages 427–502.
- Bresnan, Joan. 1982. Control and complementation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*. The MIT Press, Cambridge, Massachusetts, pages 282–390.
- Dalrymple, Mary, John Lamping, Fernando C. N. Pereira, and Vijay Saraswat. 1995. Linear logic for meaning assembly. In *Proceedings of CLNLP*, Edinburgh.
- Dalrymple, Mary, John Lamping, Fernando C. N. Pereira, and Vijay Saraswat. 1996a. A deductive account of quantification in LFG. In Makoto Kanazawa, Christopher J. Piñón, and Henriette de Swart, edi-

tors, *Quantifiers, Deduction, and Context*. CSLI Publications, Stanford, CA.

- Dalrymple, Mary, John Lamping, Fernando C. N. Pereira, and Vijay Saraswat. 1996b. Intensional verbs without type-raising or lexical ambiguity. In Jerry Seligman and Dag Westerståhl, editors, *Logic, Language and Computation*. Center for the Study of Language and Information, Stanford, California, pages 167–182. Also in Proceedings of the Conference on Information-Oriented Approaches to Logic, Language and Computation/Fourth Conference on Situation Theory and its Applications, Saint Mary's College of California, Moraga, California. June 1994.
- Dalrymple, Mary, John Lamping, Fernando C. N. Pereira, and Vijay Saraswat. 1996c. Quantifiers, anaphora, and intensionality. *Journal of Logic, Language, and Information*, to appear.
- Girard, Jean-Yves, Yves Lafont, and Paul Taylor. 1989. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, England.
- Johnson, Mark. 1995. Logic and feature structures. In Mary Dalrymple and Ronald M. Kaplan, editors, *Formal Properties of Lexical-Functional Grammar*. CSLI Lecture Notes Series.
- Johnson, Mark. 1997. A resource sensitive reinterpretation of lexical-functional grammar. available via anonymous ftp from [lx.cog.brown.edu](ftp://lx.cog.brown.edu).
- Kaplan, Ronald M. and Joan Bresnan. 1982. Lexical-Functional grammar: A formal system for grammatical representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*. The MIT Press, chapter 4, pages 173–281.
- Kasper, Robert T. and William C. Rounds. 1990. The logic of unification in grammar. *Linguistics and Philosophy*, 13(1):35–58.
- Moortgat, Michael. 1997. Categorical type logics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*. The MIT Press, Cambridge, Massachusetts, pages 93–178.
- Morrill, Glyn V. 1994. *Type-logical Grammar: Categorical Logic of Signs*. Kluwer Academic Publishers, Dordrecht.

- Pereira, Fernando C.N. and Stuart M. Shieber. 1987. *Prolog and Natural Language Analysis*. Number 10 in CSLI Lecture Notes Series. Chicago University Press, Chicago.
- Rounds, William C. 1997. Feature logics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*. The MIT Press, Cambridge, Massachusetts, pages 475–533.
- Sag, Ivan A. 1995. HPSG problem set 4: Icelandic case. Technical report, The Center for the Study of Language and Information. Available as <http://hpsg.stanford.edu/hpsg/lecture-materials/pset4-icelandic.ps>.
- Shieber, Stuart M. 1986. *An Introduction to Unification-based Approaches to Grammar*. CSLI Lecture Notes Series. Chicago University Press, Chicago.
- van Benthem, Johan. 1995. *Language in Action: Categories, Lambdas and Dynamic Logic*. The MIT Press, Cambridge, Massachusetts.