3

Long-distance Dependencies, Constituent Structure, and Functional Uncertainty

RONALD M. KAPLAN AND ANNIE ZAENEN

1 Introduction

Tree representations are used in generative grammar to represent very different types of information. Whereas in structuralist practice (at least as reconstructed by early transformationalists), Phrase Structure Markers were used to represent surface cooccurrence patterns, transformational grammar extended their use to more abstract underlying structures where they represent, for example, 'grammatical relations'. The claim embodied in this extension is that the primitives of a tree representation, namely, linear order, dominance (but not multi-dominance) relations and syntactic category labels, are adequate to represent several types of information that seem quite dissimilar in nature. They have been used, for example, to represent the dependencies between predicates and arguments needed for semantic interpretation and also the organization of phrases that supports phonological interpretation.

Lexical-Functional Grammar (like Relational Grammar) rejects this claim[1] and proposes to represent information about predicate argument

This paper originally appeared in *Alternative Conceptions of Phrase Structure*, ed. Mark Baltin and Anthony Kroch (Chicago: Chicago University Press, 1989), 17–42.

[1] In recent work in phrase-structure-based frameworks there has been some weakening of this claim. For example, almost all proposals now separate out linear order from dominance relations and represent grammatical functions mainly in terms of the latter and not the former. See Pullum (1982) for an early proposal separating these two aspects.

Formal Issues in Lexical-Functional Grammar
edited by
Mary Dalrymple
Ronald M. Kaplan
John T. Maxwell III
Annie Zaenen

dependencies in structures that allow multi-dominance and ignore linear order. Moreover these frameworks claim that the primitives in these representations are not categories like noun or sentence. Rather they are of a different nature that approximates the more traditional functional notions of subject, object, etc. In a certain sense LFG formalizes a more traditional approach than the one found in transformational grammar. The use of tree representations (called constituent structures or c-structures in LFG) is restricted to the surface structure, which is assumed to be the input to the phonological component; information about predicate argument dependencies and the like is represented in the functional structure (f-structure).

Given this view on the use of phrase structure representations, it is a bit of an anomaly that the original formulation of LFG (Kaplan and Bresnan 1982) used c-structures to state generalizations about so-called long-distance dependencies of the type illustrated in (1):

(1) Who did Bill claim that Mary had seen?

Most previous accounts of long-distance phenomena, done in generative frameworks where no other explanatory devices are available, were stated in phrase structure terms. Early LFG proposals (Kaplan and Bresnan 1982, Zaenen 1980, 1983) in effect incorporated and developed such c-structural notions without seriously examining the assumptions underlying them. But given that LFG makes a clear distinction between the functional and phrasal properties of an utterance and encodes predicate-argument relations specifically in functional structure (f-structure), this approach embodies the claim that these relations are not directly relevant to long-distance dependencies. This is a surprising consequence of this approach, given that so many other syntactic phenomena are more sensitive to properties and relations of f-structure than to those of c-structure. Indeed, a deeper investigation of long distance dependencies reveals that they too obey functional rather than phrase structure constraints. This motivates the revision to the LFG treatment of long distance dependencies that we propose in this paper. This treatment depends on a new formal device for characterizing systematic uncertainties in functional assignments.

The organization of the paper is as follows: in the first section we give an argument based on data from Icelandic that functional notions are necessary to account for generalizations about islands in that language. In the second section we sketch the mechanism of functional uncertainty that is needed to formalize these generalizations (for a more extensive discussion of the mathematical and computational aspects of this mechanism, see Kaplan and Maxwell 1988). In the third section we show how the system handles some rather recalcitrant data from English, and in the last section we discuss a case in which multi-dominance (or a similar many-to-one mechanism) is needed to get the right result.

2 The Relevance of Functional Information: Icelandic Island Constraints

It is well known that long distance dependencies involving adjuncts are more restricted than those involving arguments. To give an example from English, we can contrast example (1), where the initial *Who* is interpreted as an argument of the predicate *see* within the sentential complement of *claim*, with the following:

(2) *Which picture did they all blush when they saw?

In (1) the embedded *see*-clause is an argument of the matrix verb *claim*, whereas in (2) the embedded clause is an adjunct to the main proposition. This contrast cannot be accounted for simply in terms of node labels, because in both (1) and (2) S and/or S′ appear in the 'syntactic binding domain' (as defined, for example, in Zaenen 1983). In English, it can be plausibly claimed that these sentences differ in the configurations in which the nodes appear, so that a c-structure account of the contrast is not implausible. A similar contrast in acceptability is found in Icelandic. In the Icelandic case, however, it can be shown that no difference in surface phrase structure configuration can plausibly support an account of this kind of contrast.

To show this we will first quickly summarize the arguments given for surface structure in Thráinsson (1986) and then consider how they bear on the issue of extraction out of sentences dominated by PP's. Thráinsson (1986)[2] shows that sentences with an auxiliary or a modal have a surface structure that is different from those that have no auxiliary or modal. Both types are illustrated in (3a) and (3b) respectively:

(3) a. Hann mun stinga smjörinu í vasann.
He will put butter-the in pocket-the.
'He will put the butter in his pocket.'

b. Hann stingur smjörinu í vasann.
He puts butter-the in pocket-the.
'He puts the butter in his pocket.'

[2] Thráinsson's paper is written in a transformational framework but his generalizations translate in an obvious way into the framework used here. We use his analysis because it gives a very intuitive account of the data, but of course our remarks apply to all phrase structure accounts that hypothesize that the two types of PP's discussed below have the same attachment at some moment of the derivation.

A first place where the difference shows up is when a so-called wandering adverb is added to either of these sentences: whereas for (3a) there are only two possible positions for such an adverb as illustrated in (4), for (3b) there are the additional possibilities illustrated in (5):

(4) a. Hann mun **sjaldan** stinga smjörinu í vasann.
He will seldom put butter-the in pocket-the.
'He will seldom put the butter in his pocket.'

b. *Hann mun stinga **sjaldan** smjörinu í vasann.
He will put seldom butter-the in pocket-the.

c. *Hann mun stinga smjörinu **sjaldan** í vasann.
He will put butter-the seldom in pocket-the.

d. Hann mun stinga smjörinu í vasann **sjaldan**.
He will put butter-the in pocket-the seldom.

(5) a. Hann stingur **sjaldan** smjörinu í vasann.
He puts seldom butter-the in pocket-the.
'He seldom puts the butter in his pocket.'

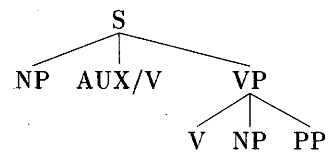
b. Hann stingur smjörinu **sjaldan** í vasann.
He puts butter-the seldom in pocket-the.

c. Hann stingur smjörinu í vasann **sjaldan**.
He puts butter-the in pocket-the seldom.

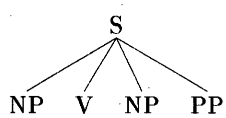
This is not the only contrast between the two types of sentences; indefinite subjects and 'floating' quantifiers show the same placement contrasts. We refer to Thráinsson (1986) for examples of these two latter phenomena.

Rather than proposing that these three types of elements are introduced by different rules in sentences with and without auxiliaries, Thráinsson proposes that it is the constituent structure of the clause that differs while the constraints on the distribution of the adverbs, indefinite subjects and quantifiers remain the same. The generalization is that the adverbs, indefinite subjects and quantifiers are daughters of S but can appear in any linear position. Thus they can be placed between each pair of their sister constituents (modulo the verb second constraint, which prohibits them from coming between the first constituent of the S and the tensed verb). This will give the right results if we assume that the c-structure for sentences with an auxiliary is as in (6) whereas sentences without an auxiliary have the structure in (7):

(6)



(7)



To be a bit more concrete, we propose to capture this insight in the following partial dominance and order constraints; these account for word order and adverb distribution in the sentences above:

(8) Dominance Constraints:
- S can immediately dominate {V, VP, NP, PP, ADV}
- VP can immediately dominate {V, VP, NP, PP}
- V is obligatory both in S and in VP.

(9) Ordering Constraints:
- for both S and VP: V<NP<PP<VP
- for S: XP immediately precedes V[+tense] (verb-second constraint)

These constraints (given here in a partial and informal formulation), together with LFG's coherence, completeness, and consistency requirements, provide the surface structures embodying the generalization proposed by Thráinsson.

Given this independently motivated difference in c-structure, let us now return to the difference between arguments and adjuncts. Icelandic differs from English in allowing Ss in PPs, as shown in (10) to (13):[3]

(10) Hann fór eftir að ég lauk verkinu.
He went after that I finished work-the.
'He left after I finished the work.'

[3]These constructions are analyzed as PP's in Icelandic because in all these cases the S′ alternates with a simple NP:

(i) Jón kom eftir kvöldmatinn.
'Jon came after dinner.'

(ii) Jón var að hugsa um Maríu.
'Jon was thinking about Maria.'

In general, the simplest hypothesis about Icelandic phrase structure rules is that an S′ is permitted wherever an NP can appear (if the meaning allows it).

(11) Jón var að þvo golfið eftir að María hafði skrifað bréfið.
John was at wash floor-the after that Mary had written letter-the.
'John was washing the floor after Mary had written the letter.'

(12) þú vonaðist til að hann fengi bíl.
You hoped for that he will-get car.
'You hope that he will get a car.'

(13) Jón var að hugsa um að María hefði líklega skrifað bréfið.
John was at think about that Mary had probably written letter-the.
'John was thinking that Mary had probably written the letter.'

(10) and (11) illustrate cases in which the PP clause is an adjunct, whereas (12) and (13) are examples in which the PP clause is an argument. We will use these complex embedded structures because they allow a straightforward illustration of the patterns of long-distance dependencies: we find cases that exhibit the same local categorial configurations (PP over S′), but differ in their long-distance possibilities:

(14) *Hvaða verki fór hann eftir að ég lauk?
Which job went he after that I finished?
'Which job did he go after I finished?'

(15) *þessi bréf var Jón að þvo golfið eftir að María hafði skrifað.
This letter was John at wash floor-the after that Mary had written.
'This letter John was washing the floor after Mary had written.'

(16) Hvaða bíl vonaðist þú til að hann fengi?
Which car hoped you for that he will-get?
'Which car did you hope that he would get?'

(17) þessi bréf var Jón að hugsa um að María hefði líklega skrifað.
This letter was John at think about that Mary had probably written.
'This letter John was thinking that Mary had probably written.'

What these examples illustrate is that extractions are allowed from the PP-S′ configuration only when it is an argument; it forms a wh-island when it functions as an adjunct.[4]

In defining the original c-structure formalization for long-distance dependencies, Kaplan and Bresnan (1982) noted that the correlation of ex-

[4] This is true for tensed clauses, but we have not yet investigated infinitives. It is well known that they tend to be less strong as islands, but further studies are needed to understand fully the influence of tense on island constraints.

traction constraints with categorial configurations is far less than perfect. They allowed bounding-node specifications in individual phrase-structure rules to characterize the variations of long-distance dependency restrictions across languages and across different nodes of the same category in a particular language. Indeed, the formal devices they introduced are sufficient to accurately describe these Icelandic facts: The argument and adjunct PP's can be introduced in separate phrase-structure expansions, with only the PP receiving the ADJunct function assignment boxed as a bounding node. But it is clear that the boxing device is used to import functional distinctions into the c-structure. Looking back at the discussion in Kaplan and Bresnan (1982) one realizes that it is always the case that when one instance of a given category is boxed as a bounding node and another is not, those instances also have different functional schemata attached (ADJ vs. one of the oblique argument functions in the Icelandic example, or the COMP vs. RELMOD functions that distinguish English *that*-complement S′s from relative clauses.). Kaplan and Bresnan, while realizing that extraction domains cannot be defined in terms of obvious natural classes of c-structure categories or configurations, did not then recognize that natural classes do exist at the functional level.

They actually considered but quickly rejected the possibility of defining long-distance dependencies in terms of f-structure configurations partly because no rigorous functional formalization was at hand and partly because examples like (18) (Kaplan and Bresnan 1982, 134) seemed to indicate the long-distance relevance of at least some categorial information that would not be available in f-structure:

(18) a. She'll grow that tall/*height.
b. She'll reach that *tall/height.
c. The girl wondered how tall she would grow/*reach.
d. The girl wondered what height she would *grow/reach.

These examples suggest that adjective phrases can only be extracted from AP positions and noun phrases only from NP positions, and, more generally, that fillers and gaps must have matching categories. Thus, they ignored the apparently functional constraints on long-distance extractions and defined special formal mechanisms for encoding those constraints in c-structure terms. In this they remained similar to other structure-oriented theories of the day.

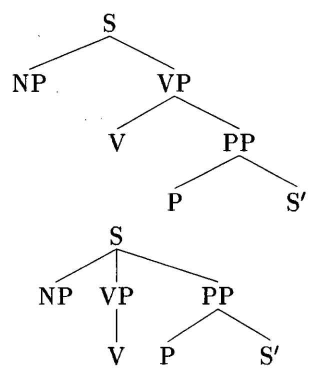
The Icelandic data given above, however, suggest that a more functional approach would capture the facts more directly. In section 3, we will show that the data in (18) can also be naturally analyzed in functional terms. In fact constraints on extraction that in LFG terms are functional in nature have also been proposed by syntacticians working in a completely structure-oriented theory. The Icelandic data discussed above can be seen as a case of the Condition on Extraction Domain proposed in Huang (1982), which can be interpreted as an emerging functional perspective formulated in structural terms. It states that

(19) No element can be extracted from a domain that is not properly governed.

Intuitively the distinction between governed and nongoverned corresponds to the difference between argument and nonargument. But it is clear from Thráinsson's arguments for the difference in structure between sentences with and without an auxiliary that the correct notion of government cannot be simply defined over c-structures. To represent the difference between the two types of PP's as in (20) would go against Thráinsson's generalization.

(20) a.

b.



Indeed, adverb placement shows that both adjunct and argument PP's are sisters of S when there is no auxiliary but are both in the VP when an auxiliary is present:[5]

(21) a. Ég vonaðist **alltaf** til að hann fengi bíl.
I hoped always for that he will-get car.
'I always hoped that he would get a car.'

b. Ég hef **alltaf** vonast til að hann fengi bíl.
I have always hoped for that he will-get car.
'I have always hoped that he would get a car.'

[5]Zaenen (1980) proposes that the extractability from an S is determined by a lexical property of the complementizer that introduces it. Under that hypothesis the adjunct/argument contrast discussed here would be unstatable, since the same complementizer appears in both constructions.

c. *Ég hef vonast **alltaf** til að hann fengi bíl.
I have hoped always for that he will-get car.

(22) a. Hann fór **alltaf** eftir að ég lauk verkinu.
He went always after that I finished work-the.
'He always went after I finished the work.'

b. Hann hefur **alltaf** farið eftir að ég lyk verkinu.
He has always gone after that I finished work-the.
'He has always gone after I finished the work.'

c. *Hann hefur farid **alltaf** eftir að ég lyk verkinu.
He has gone always after that I finished work-the.

This pattern does not change when in the context of a long-distance dependency, as the following contrast illustrates:

(23) a. *Hvaða verki fór hann **alltaf** eftir að ég lauk?
Which job went he always after that I finished?
'Which job did he always go after I finished?'

b. Hvaða bíl vonaðist þú **alltaf** til að hann fengi?
Which car hoped you always for that he will-get?
'Which car did you always hope he would get?'

Thus in Icelandic the same c-structure configuration allows for extraction when the PP is an argument but not when the PP is an adjunct. Netter (1987) draws a similar conclusion from data concerning extraposition of relative clauses in German. Given these facts, an adequate structurally-based account will have to appeal to stages in a derivation[6] and assume different tree structures for these sentences at the moment the relevant movement takes place. Whether this is feasible or not will depend on one's view on principles like cyclicity and the like and we leave it to practitioners of structural approaches to elaborate these accounts. From our nonderivational perspective the most straightforward approach seems also the most reasonable one: we will assume that long distance dependencies are sensitive to functional information and investigate further how such constraints can be formulated in functional terms.[7]

[6] A reanalysis of the verb and the prepositions as one unit would not obviously account for this contrast, and in any event, such an analysis has no independent motivation. Maling and Zaenen (1980) argue explicitly that there is no such reanalysis in Icelandic, and the fact that an adverb cannot be placed between the preposition and the following clause is further evidence against such a proposal:

(i) *Ég vonaðist til **alltaf** að hann fengi bíl.
I hoped for always that he will-get car.

[7] As far as we can see, the Icelandic data also do not allow for a syntactic account in frameworks like GPSG which define 'government' solely on the surface structure.

3 The Formal Account: Functional uncertainty

Standing back from the details of particular constructions or particular languages, long-distance dependencies seem difficult to characterize because they involve rather loose and uncertain connections between the superficial properties of local regions of a string and its more abstract functional and predicate-argument relations. For many sentences this connection is very direct and unambiguous. If, for example, the first few words of an English sentence have the internal organization of an NP, it is often the case that those words also function as the subject of the sentence. Of course, there are uncertainties and ambiguities even in simple sentences: in a garden-path sentence such as (24), it is not clear only from the local evidence which words make up the initial NP, and those words thus are compatible with two different functional configurations. This local ambiguity is resolved only when information about the later words is also taken into account.

(24) The cherry blossoms in the spring.

Local uncertainties of this sort have never seemed difficult to describe, since all grammatical theories admit alternative rules and lexical entries to account for all the local possibilities and provide some method of composition that may reject some of them on the basis of more global contextual information. What distinguishes the uncertainties in long-distance dependencies is that the superficial string properties local to, say, a fronted English topic are compatible with an unbounded number of within-clause functional or predicate-argument relations. The infinite set of possibilities cannot be specified in any finite number of alternatives in basic rules or lexical entries, and which of these possibilities is admissible depends on information that may be available arbitrarily far away in the string.

Structural approaches typically handle this kind of unbounded uncertainty through conspiracies of transformations that introduce empty nodes and prune other nodes and thereby destroy the simple connection between the surface and underlying tree structures. Our solution to the uncertainty problem is much more direct: we utilize a formal device that permits an infinite set of functionally constrained possibilities to be finitely specified in individual rules and lexical entries.

Kaplan and Bresnan (1982) observed that each of the possible underlying positions of an initial phrase could be specified in a simple equation locally associated with that phrase. In the topicalized sentence (25):

(25) Mary John telephoned yesterday.

the equation (in LFG notation) $(\uparrow \text{TOPIC}) = (\uparrow \text{OBJ})$ specifies that *Mary* is to be interpreted as the object of the predicate *telephoned*. In (26):

(26) Mary John claimed that Bill telephoned yesterday.

the appropriate equation is ($\uparrow$ TOPIC) = ($\uparrow$ COMP OBJ), indicating that *Mary* is still the object of *telephoned*, which because of subsequent words in the string is itself the complement (indicated by the function name COMP) of the top-level predicate *claim*. The sentence can obviously be extended by introducing additional complement predicates (*Mary John claimed that Bill said that .... that Henry telephoned yesterday*), for each of which some equation of the general form

($\uparrow$ TOPIC) = ($\uparrow$ COMP COMP ... OBJ)

would be appropriate. The problem, of course, is that this is an infinite family of equations, and hence impossible to enumerate in a finite disjunction appearing on a particular rule of grammar. For this technical reason, Kaplan and Bresnan abandoned the possibility of specifying unbounded uncertainty directly in functional terms.

Instead of formulating uncertainty by an explicit disjunctive enumeration, however, a formal specification can be provided that characterizes the family of equations as a whole. A characterization of a family of equations may be finitely represented in a grammar even though the family itself has an infinite number of members. This can be accomplished by a simple extension of the elementary descriptive device in LFG, the functional-application expression. In the original formalism function-application expressions were given the following interpretation:

(27) $(f\ s) = v$ holds if and only if f is an f-structure, s is a symbol, and the pair $<s, v> \in f$.

This notation was straightforwardly extended to allow for strings of symbols, as illustrated in expressions such as ($\uparrow$ COMP OBJ) above. If $x = sy$ is a string composed of an initial symbol s followed by a (possibly empty) suffix string y, then

(28) $(f\ x) \equiv ((f\ s)\ y)$
$(f\ \epsilon) \equiv f$, where ϵ is the empty string.

The crucial extension to handle unbounded uncertainty is to allow the argument position in these expressions to denote a set of strings. The interpretation of expressions involving sets of strings is derived in the following way from the interpretation (28) for individual strings. Suppose α is a (possibly infinite) set of strings. Then we say

(29) $(f\ \alpha) = v$ holds if and only if
$((f\ s)\ \mathrm{Suff}(s, \alpha)) = v$ for some symbol s,
where $\mathrm{Suff}(s, \alpha)$ is the set of suffix strings y such that $sy \in \alpha$.

In effect, an equation with a string-set argument holds if it would hold for a string in the set that results from a sequence of left-to-right symbol choices. For the case in which α is a finite set this formulation is equivalent to a finite disjunction of equations over the strings in α. Passing from finite disjunction to existential quantification captures the intuition that unbounded uncertainties involve an underspecification of exactly which choice of strings in α will be compatible with the functional information carried by the surrounding surface environment.

We of course impose the requirement that the membership of α be characterized in finite specifications. More particularly, it seems linguistically, mathematically, and computationally advantageous to require that α in fact be drawn from the class of *regular languages*. The characterization of uncertainty in a particular grammatical equation can then be stated as a regular expression over the vocabulary of grammatical function names. The infinite uncertainty for the topicalization example above can now be specified by the equation given in (30):

(30) $(\uparrow \text{TOPIC}) = (\uparrow \text{COMP}^{*}\ \text{OBJ})$

involving the Kleene closure operator. One remarkable consequence of our functional approach is that appropriate predicate-argument relations can be defined without relying on empty nodes or traces in phrase-structure trees. This allows us to make the phrase-structure representations much more faithful to the sentence's superficial organization. Note that a particular within-clause grammatical function can be assigned by a long-distance dependency only if the phrase-structure rules optionally introduce the nodes that would normally carry that function in simple clauses.[8] This formulation is possible only because subcategorization in LFG is defined on f-structure via the Completeness and Coherence conditions and is independent of phrase-structure configurations.

The mathematical and computational properties of functional uncertainty are discussed further in Kaplan and Maxwell (1988). Here we summarize the mathematical characteristics briefly: it is clearly decidable whether a given f-structure satisfies a functional description that includes uncertainty specifications. Since a given f-structure contains only a finite number of function-application sequences, it contains only a finite number of strings that might satisfy an uncertainty equation. The membership problem for the regular sets is decidable and each of those strings can therefore be tested to see if it makes the equation hold.

[8] Thus a constraint like the one proposed by Perlmutter (1971) that (tensed) clauses must have local surface subjects (and hence that question movement of the subject is not allowed) would follow in a straightforward way from making the NP constituent bearing the subject equation obligatory in the phrase structure rule.

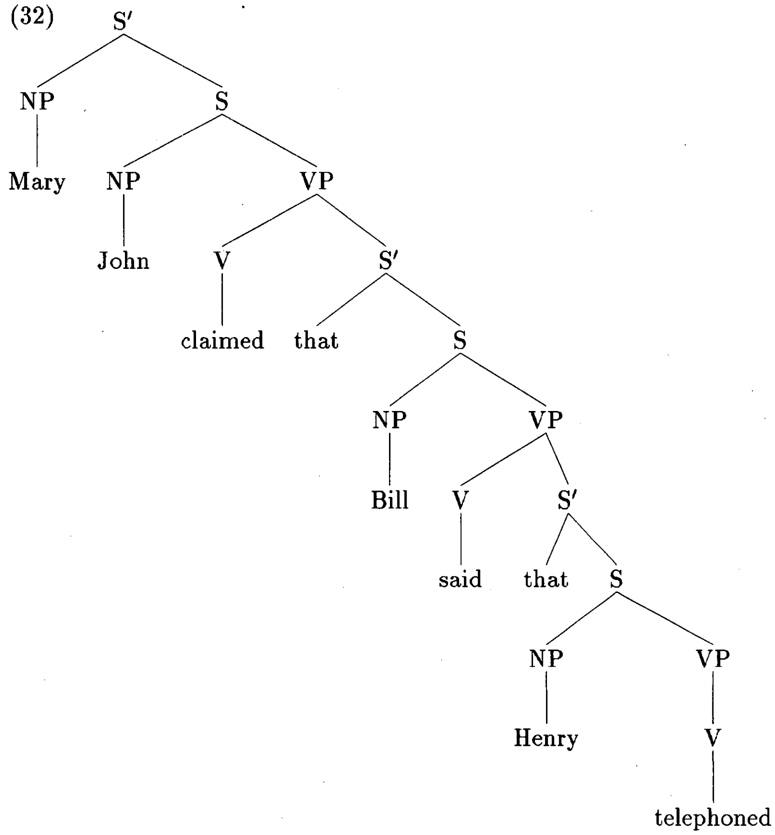
It is less obvious that the satisfiability problem is decidable. Given a set of equations describing a functional structure for a sentence, can it be determined that a structure satisfying all the equations does in fact exist? For a trivial description with a single equation, the question is easy to answer. If the equation has an empty uncertainty language, containing no strings whatsoever, the description is unsatisfiable. Otherwise, it is satisfied by the f-structure that meets the requirements of any string in the language, say the shortest one. The difficult case arises when the functional description has two uncertainty equations, say $(f\,\alpha) = v_\alpha$ and $(f\,\beta) = v_\beta$. If α contains (perhaps infinitely many) strings that are initial prefixes of strings in β, then the strings that will be mutually satisfiable cannot be chosen independently from the two languages. For example, the choice of x from α and xy from β implies a further constraint on the values v_α and v_β: for this particular choice we have $(f\,x) = v_\alpha$ and $(f\,xy) = ((f\,x)\ y) = v_\beta$, which can hold only if $(v_\alpha\ y) = v_\beta$. Kaplan and Maxwell (1988) show, based on a state-decomposition of the finite-state machines that represent the regular languages, that there are only a finite number of ways in which the choice of strings from two uncertainty expressions can interact. The original equations can therefore be transformed into an equivalent finite disjunction of derived equations whose remaining uncertainty expressions are guaranteed to be independent. The original functional description is thus reducible to a description without uncertainty when each of the remaining regular languages is replaced by a freely chosen member string. The satisfiability of descriptions of this sort is well-established. A similar proof of satisfiability has been developed by Mark Johnson (p. c.).

If the residual uncertainties include an infinite number of strings, then an infinite number of possible f-structures will satisfy the original description and are thus candidates for the f-structure that the grammar assigns to the sentence. This situation closely resembles the general case that arises for descriptions without uncertainties. As Kaplan and Bresnan (1982) noted, if a description is consistent then an infinite number of f-structures will satisfy it. These f-structures are ordered by a subsumption relation and Kaplan and Bresnan defined the subsumption-minimal satisfying structure to be the grammatically relevant one. The family of f-structures that satisfy the residual uncertainties is also ordered, not just according to subsumption but also according to the lengths of the strings that are chosen from the regular set. We extend the minimality condition of LFG by requiring that the f-structure assigned to a sentence include only the shortest strings realizing a particular uncertainty. In this way we follow the general LFG strategy of excluding from consideration structures that involve arbitrarily redundant information. See Kaplan and Maxwell (1988) for further discussion.

This is a general formalism that may apply to phenomena that are traditionally not thought of as falling into the same class as long-distance dependencies but that nevertheless seem to involve some degree of uncertainty. Johnson (1986) and Netter (1986) have used it in the analysis of Germanic infinitival complements and Karttunen (1989) discusses how similar extensions to Categorial Unification Grammar can account for related facts in Finnish that would otherwise require type-raising. Halvorsen (1988) has extended its use to the semantic domain, where it offers a simple characterization of various kinds of quantifier scope ambiguities. In this paper we illustrate the formalism by showing how it can be used to represent different conditions on long-distance dependencies. Consider the multi-complement sentence (31) whose c-structure and f-structure are given in (32) and (33):

(31) Mary John claimed that Bill said that Henry telephoned.

(32)



(33)

$$\left[\begin{array}{ll}
\text{TOPIC} & \left[\text{PRED} \ \text{`Mary'}\right] \\
\text{PRED} & \text{`claim}\,\langle(\uparrow \text{SUBJ}),\ (\uparrow \text{COMP})\rangle\text{'} \\
\text{SUBJ} & \left[\text{PRED} \ \text{`John'}\right] \\
\text{COMP} & \left[\begin{array}{ll}
\text{PRED} & \text{`say}\,\langle(\uparrow \text{SUBJ}),\ (\uparrow \text{COMP})\rangle\text{'} \\
\text{SUBJ} & \left[\text{PRED} \ \text{`Bill'}\right] \\
\text{COMP} & \left[\begin{array}{ll}
\text{PRED} & \text{`telephone}\,\langle(\uparrow \text{SUBJ}), (\uparrow \text{OBJ})\rangle\text{'} \\
\text{SUBJ} & \left[\text{PRED} \ \text{`Henry'}\right] \\
\text{OBJ} &
\end{array}\right]
\end{array}\right]
\end{array}\right]$$

Notice that the tree in (32) has no empty NP node in the embedded clause. The link in the functional structure (33) indicates that the relation between the topic and the object of the most deeply embedded complement is one of functional identity, just like the relation between a functional controller in a raising or equi construction and its controllee (see the discussion of functional control in Kaplan and Bresnan 1982). Thus the same subsidiary f-structure serves as the value of both the TOPIC function and the OBJ function in the complement. The linguistic conditions on the linkages in functional control and long-distance dependencies are quite different, however. The conditions on functional uncertainty in long-distance dependencies can be subdivided into conditions on the potential functions at the end of the uncertainty path (the *bottom*, OBJ in this example) and conditions on the functions in the middle of the path (the *body*, here COMP*).

In the example above, the bottom is the function OBJ. Of course, there are a variety of other within-clause functions that the topic can have, and the equation might be generalized to

(34) $(\uparrow \text{TOPIC}) = (\uparrow \text{COMP}^* \text{ GF})$

where GF denotes the set of primitive grammatical functions. As we discuss in Section 3, this is too general for English since the topic cannot serve as a within-clause complement. A more accurate specification is

(35) $(\uparrow \text{TOPIC}) = (\uparrow \text{COMP}^* \text{ (GF}-\text{COMP))}$

where GF−COMP denotes the set of grammatical functions other than COMP. This might appear to be still much too general, in that it permits a great number of possible bottom functions most of which would be unacceptable in any particular sentence. But whatever bottom function is chosen will have to be compatible with all other requirements that are imposed on it, not only case-marking and agreement etc. but also the general principles of consistency, completeness, and coherence. Although phrase-structure rules no longer play a role in insuring that the topicalized constituent will be linked to the 'right' place within the sentence, these functional conditions will rule out unacceptable sentences like (35):

(36) *Mary, he said that John claimed that Bill saw Peter.

(This sentence does have an interpretation with *Mary* as a vocative but we ignore that possibility here.) If OBJ is chosen as the bottom function and the body reaches down to the lowest clause, the features of *Mary* will be inconsistent with the features of the local object *Peter*. An inconsistency would arise even if *Peter* were replaced by a repetition of the word *Mary* because of the instantiation property of LFG's semantic forms (Kaplan and Bresnan 1982, 225). If the body does not reach down to the lowest clause, then one of the intermediate f-structures will be incoherent: neither of the predicates *claim* nor *say* take objects. The f-structure would also be incoherent if some other function, say OBJ2, were chosen as the bottom or if the body were extended below the lowest clause.

The following sentence has the same c-structure as (36) but is grammatical and even ambiguous, because the function ADJ can be chosen as the bottom:

(37) Yesterday, he said that Mary claimed that Bill telephoned Peter.

This is acceptable because ADJ is in GF but is not one of the *governable* grammatical functions, one that can serve as an argument to a lexical predicate, and thus is not subject to the coherence condition as defined by Kaplan and Bresnan (1982).

Similarly, restrictions on the sequence of functions forming the body can be stated in terms of regular predicates. The restriction for Icelandic that adjunct clauses are islands might be expressed with the equation:

(38) $(\uparrow \text{TOPIC}) = (\uparrow (\text{GF}-\text{ADJ})^* \ \text{GF})$

This asserts that the body of the path to the clause-internal function can be any sequence of non-adjunct grammatical functions, with the bottom being any grammatical function that may or may not be ADJ. For English the body restriction is even more severe, allowing only closed and open complements (COMP and XCOMP in LFG terms) on the path, as indicated in (39):

(39) $(\uparrow \text{TOPIC}) = (\uparrow \{\text{COMP}, \text{XCOMP}\}^* \ (\text{GF}-\text{COMP}))$

Given this formalism, the theory of island constraints becomes a theory of the generalizations about the body of possible functional paths, expressible as regular predicates on the set of uncertainty strings. For example, if RELMOD is the function assigned to relative-clause modifiers of noun-phrases, that function would be excluded from the body in languages that obey the complex-NP constraint.

Other conditions can be stated in the phrase-structure rules that introduce the uncertainty expression. These rules are of the general form indicated in (39):

(40)
$$\text{S}' \longrightarrow \underset{\substack{(\uparrow \text{DF}) = \downarrow \\ (\uparrow \text{DF}) = (\uparrow \text{ body bottom})}}{\Omega} \qquad \Sigma$$

where Ω is to be realized as a maximal phrasal category, Σ is some sentential category, and DF is taken from the set of discourse functions (TOPIC, FOCUS, etc.). This schema expresses the common observation that constituents introducing long-distance dependencies are maximal projections and are sisters of sentential nodes. Restricting the introduction of discourse functions to rules of this sort also accounts for the observation that discourse functions need to be linked to within-clause functions (see Fassi-Fehri 1988 for further discussion). The rule in (41) for English topicalization is an instance of this general schema:

(41) $$\mathrm{S}' \quad \rightarrow \quad \begin{array}{c} \text{XP or S}' \\ (\uparrow \textsc{topic}) = \downarrow \\ (\uparrow \textsc{topic}) = (\uparrow \{\textsc{comp}, \textsc{xcomp}\}^* (\textsc{gf}-\textsc{comp})) \end{array} \quad \mathrm{S}$$

In English, S′ and any XP can occur in topic position. Spanish, on the other hand, seems to be a language in which some topic constructions allow NP's but not PP's (Grimshaw 1982).

4 Illustrations from English

As mentioned above, Kaplan and Bresnan (1982) noticed an apparent category-matching requirement in sentences like (42)–(43) (Kaplan and Bresnan 1982, example 134):

(42) a. The girl wondered how tall she would grow.
b. *The girl wondered how tall she would reach.

(43) a. The girl wondered what height she would reach.
b. *The girl wondered what height she would grow.

Grow seems to subcategorize for an AP and *reach* for an NP. But subcategorization in LFG is done in functional terms, and it turns out that independently motivated functional constraints also provide an account of these facts. First observe that *reach* but not *grow* governs the OBJ function, as indicated by the contrast in (44):

(44) a. *That tall has been grown.
b. That height has been reached.

Grimshaw (1982) shows that passivization is not dependent on syntactic category but on whether the verb takes an OBJ.[9] The verb *grow*, on the other hand, establishes a predicational relationship between its subject and its adjectival complement and thus governs the XCOMP function. The relevant lexical entries for *reach* and *grow* are as follows:

(45) a. reach: $(\uparrow \textsc{pred}) = \text{`reach}\langle(\uparrow \textsc{subj})(\uparrow \textsc{obj})\rangle\text{'}$
b. grow: $(\uparrow \textsc{pred}) = \text{`grow}\langle(\uparrow \textsc{subj})(\uparrow \textsc{xcomp})\rangle\text{'}$
$(\uparrow \textsc{subj}) = (\uparrow \textsc{xcomp subj})$

[9] Jacobson (1982) points out that the verbs *ask* and *hope* are not susceptible to this analysis.

Sentence (42a) is acceptable if XCOMP is chosen as the bottom function: XCOMP makes the local f-structure for *grow* complete. *Tall*, being a predicative adjective, also requires a local subject, and that requirement is satisfied by virtue of the control equation $(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$. The choice of XCOMP in (42b) is unacceptable because it makes the local f-structure for *reach* be incoherent. Choosing OBJ satisfies the requirements of *reach*, but the sentence is still ungrammatical because the f-structure for *tall*, in the absence of a control equation, does not satisfy the completeness condition. In (43a) the choice of OBJ at the bottom satisfies all grammaticality conditions. If OBJ is chosen for (43b), however, the f-structure for *grow* is incoherent. If XCOMP is chosen the f-structure for *grow* is complete and coherent, and the sentence would be acceptable if *what height* could take the controlled subject. Although some noun-phrases can be used as predicate nominals (*She became a doctor*, *She seems a fool*), others, in particular *what height*, cannot (**She became that/a height*, **She seems that height*, **I wonder what height she became/seemed*). Whether or not the restrictions ultimately turn out to be functional or semantic in nature, it is clear from the contrasts with *become* and *seem* that they have nothing to do with syntactic categories.

Not only is category-matching unnecessary, it does not always yield the correct results. Kaplan and Bresnan (1982) discussed the examples in (46) (their 136) where a simple category-matching approach fails:

(46) a. That he might be wrong he didn't think of.
b. *That he might be wrong he didn't think.
c. *He didn't think of that he might be wrong.
d. He didn't think that he might be wrong.

In these examples the category of a fronted S′ can only be linked to a within-clause position that is normally associated with an NP. Kaplan and Bresnan complicated the treatment of constituent control to account for these cases by allowing the categories of the controller and controllee both to be specified in the topicalization rule. A closer look at the lexical requirements of the verbs involved, however, gives a more insightful account. Bresnan (1982) proposes association principles between syntactic categories and grammatical functions. These principles lead to the following VP rule for English:

(47) VP $\rightarrow$

$$\begin{array}{ccccc} \text{V} & \text{(NP)} & \text{(NP)} & \text{PP*} & \ldots \quad (\text{S}') \\ & (\uparrow \text{OBJ}) = \downarrow & (\uparrow \text{OBJ2}) = \downarrow & (\uparrow (\downarrow \text{PCASE})) = \downarrow & (\uparrow \text{COMP}) = \downarrow \end{array}$$

This rule embodies the claim that in English the OBJ function is only associated with NP's and the COMP function only with S′. Adopting these principles, we propose the following partial lexical entries for *think* and *think-of*:[10]

(48) a. think: (↑ PRED) = 'think<(↑ SUBJ) (↑ COMP)>'
b. think: (↑ PRED) = 'think<(↑ SUBJ) (↑ OBL_{OF})>'

The difference between the grammatical and ungrammatical sentences in (46) follows if COMPs cannot be the bottom of an uncertainty in English (whereas OBJ, OBJ2, and obliques such as OBL_{OF} can). For (46a) the choice of OBL_{OF} for the bottom is compatible with the semantic form in (48b), so the sentence is acceptable. Since COMP cannot be the bottom, OBL_{OF} and (48b) are also the only possible choices for (46b), but with this string the requirement that the preposition *of* be present is violated (this requirement is similar to the conditions on idiosyncratic case-marking, the details of which do not concern us here).[11]

It is true that the OBL_{OF} slot in (46a) is filled in a way that would be impossible in sentence internal position (46c), but this follows simply from the phrase-structure rules of English. There is no rule that expands PP as a preposition followed by an S′, no matter what functional annotations might be provided; as we have seen in Section 1, this is a very language-specific restriction. But as far as the functional requirements of *think-of* go, nothing in the f-structure corresponding to an S′ prevents it from serving as the OBL_{OF}.

Under this account of long-distance dependencies, then, there is no need to parameterize them in terms of particular phrase-structure categories. This proposal also easily handles the following contrasts, discussed in Stowell (1981):

(49) Kevin persuaded Roger that these hamburgers were worth buying.

(50) *That these hamburgers were worth buying, Kevin persuaded Roger.

(51) Louise told me that Denny was mean to her.

[10]This analysis assumes an unlayered f-structure representation of oblique objects related to the proposal of Bresnan (1982) and slightly different from the two-level approach discussed by Kaplan and Bresnan (1982) and Levin (1986). The only change necessary to accommodate the two-level representation would be to allow the bottom to be a two-element sequence such as OBL_{OF} OBJ, the same sequence that *think-of* would subcategorize for under that approach.

[11]One kind of sentence that is not ruled out on syntactic grounds is:

That John saw Mary Bill kissed.

We assume that this is out for semantic reasons: that-clauses, regardless of their grammatical function, correspond to semantic propositions and propositions are not kissable.

(52) That Denny was mean to her Louise told me (already).

They can be compared to

(53) *Kevin persuaded Roger the news.

(54) Louise told me the story.

(53) shows that *persuade* does not subcategorize for an OBJ2, while (54) shows that *tell* does take an OBJ2 as an alternative to the COMP assignment for (51). The relevant lexical information is given in (55).[12]

(55) tell: (↑ PRED) = 'tell<(↑ SUBJ) (↑ OBJ) (↑ OBJ2) >'
or
(↑ PRED) = 'tell<(↑ SUBJ) (↑ OBJ) (↑ COMP)>'

persuade: (↑ PRED) = 'persuade<(↑ SUBJ) (↑ OBJ) (↑ COMP)>'

The ungrammaticality of (50) follows again from the fact that the bottom cannot be a COMP, whereas (52) is acceptable because an OBJ2 is permitted.

Our proposal is different from the one made in Stowell (1981) in that adjacency plays no role for us, so we do not need incorporation rules to account for (51–52). This is in keeping with our view that phrase-structure rules and functional structure are in a much looser relation to each other than in the theory that Stowell assumes. The fact that the incorporation analysis of (51) is not independently motivated is in turn a confirmation for this view.

Both the present proposal and the one elaborated in Stowell (1981) can be seen as accounts of a generalization made in phrase-structure terms by Higgins (1973), namely, that S′ topicalization is only possible from an NP position. Indeed, the present functional approach covers the cases

[12]Notice that according to our proposal the grammaticality of (i) does not license (ii):

(i) John persuaded Roger.

(ii) *That these hamburgers were worth buying, John persuaded.

Arguments slots in LFG are reserved for certain semantically restricted types, as the following unacceptable string illustrates:

(iii) *John persuaded the fact.

One way to achieve this is to assume that each GF is associated with a thematic role and that lexical rules do not change these associations. For instance, a verb like *give* takes a goal and a theme, and in the OBJ, OBL_{GOAL} realization the theme is linked to the OBJ and the goal to the OBL_{GOAL}. In the OBJ, OBJ2 construction, however, it is the goal that is linked to the OBJ and the theme to the OBJ2. For different ways to formulate this correspondence that preserve thematic role assignments, see Bresnan (1982) and Levin (1986). With *persuade*, the goal argument is obligatory and the prepositional argument is optional, as is shown by (iv):

(iv) *John persuaded that Bill had left.

Higgins himself discusses. These include contrasts involving extraposition like those in (56):

(56) a. That Susan would be late John didn't think was very likely.
b. *That Susan would be late John didn't think it was very likely.

Extraposition is a lexical rule that for each extraposable entry of the form in (57a) adds a lexical entry of the form shown in (57b):

(57) a. $(\uparrow \text{PRED}) = \text{`R}\langle(\uparrow \text{SUBJ}) \ldots \rangle\text{'}$
b. $(\uparrow \text{PRED}) = \text{`R}\langle(\uparrow \text{COMP}) \ldots\rangle(\uparrow \text{SUBJ})\text{'}$
$(\uparrow \text{SUBJ PERS}) = 3$
$(\uparrow \text{SUBJ NUM}) = \text{SG}$
$(\uparrow \text{SUBJ GEND}) = \text{NEUT}$

This rule applied to the lexical entry for *likely* yields (58) and accounts for the alternation in (59):

(58) likely: $(\uparrow \text{PRED}) = \text{`likely}<(\uparrow \text{COMP})>(\uparrow \text{SUBJ})\text{'}$
$(\uparrow \text{SUBJ PERS}) = 3$
$(\uparrow \text{SUBJ NUM}) = \text{SG}$
$(\uparrow \text{SUBJ GEND}) = \text{NEUT}$

(59) a. That Susan will be late is likely
b. It is likely that Susan will be late.

Since a PRED value must be linked to a thematic function, either directly or by a chain of functional control, expletive *it* as in (59) is the only possible realization of the nonthematic SUBJ in (59b):

(60) it: $(\uparrow \text{PERS}) = 3$
$(\uparrow \text{NUM}) = \text{SG}$
$(\uparrow \text{GEND}) = \text{NEUT}$
$\neg(\uparrow \text{PRED})$

With the extraposition entry in (58) the ungrammaticality of (56b) easily follows. The function COMP is not a legal uncertainty bottom, so that with this entry a complete functional structure cannot be assigned. Choosing SUBJ as the uncertainty bottom would be compatible with the entry corresponding to (57a), but this choice would result in the subject having a sentential PRED value, which the features for expletive *it* do not allow.

The lexical extraposition rule also interacts with the phrase structure rule that introduces sentential subjects to exclude (60):

(61) *John didn't think (that) that Susan would be late was very likely.

Whereas the phrase-structure rule for embedded clauses is as given in (62a), main clauses also allow the one given in (62b):

(62) a. S $\longrightarrow$ (NP) VP
$(\uparrow \text{SUBJ}) = \downarrow$

b. E $\longrightarrow$ (XP) VP
$(\uparrow \text{SUBJ}) = \downarrow$

E is the category for a root-node expression and XP can be any phrase that can bear the subject function, namely S′, PP (as in *Into the room jumped a rabbit*; see Levin 1986 for discussion) and NP. In embedded position, however, we only find NP.[13]

Our discussion ignores embedded questions but it is clear that some contrasts like the one exemplified in (62) can be treated along the same lines:

(63) a. *Whether John would come early she didn't wonder.
b. Whether John would come early she didn't know.

Translating the observations of Grimshaw (1979) into our framework, we would hypothesize that *wonder* takes a COMP only whereas *know* allows for a COMP and an OBJ. But the general problem of embedded questions needs further study: it is well known that in some cases they are more OBJ-like than *that*-clauses. We have not studied their behavior in enough detail to propose a general treatment.

We have shown in this section that a functional approach can account for the basic correspondences that characterize long-distance dependencies as well as previous category-matching approaches do, and also for a variety of additional facts that have seemed rather puzzling under a categorial analysis.

5 Interactions with Functional Control: Japanese Relatives

There are no multiply dominated substructures in phrase-structure trees, and, hence, any two nodes are connected by just one path. This is not the case with paths in functional structure. The following example shows such a multiple-path configuration:

(64) Mary John expected to walk.

[13]These rules also allow us to account for the ungrammaticality of (i) and (ii):

(i) *That John will be late seems.

(ii) *That John will be late Bill doesn't think seems.

We simply assume that *seem* has only the 'derived' lexical entry in (57b) and not the one in (57a). Thus the thematic argument with *seem* is always a COMP and never a SUBJ, and indeed there are no sentences like (iii) that might lead to (ii):

(iii) *John/The fact seems.

(65)
$$\left[\begin{array}{ll} \textsc{topic} & \left[\textsc{pred} \ \text{`Mary'}\right] \\ \textsc{pred} & \text{`expect}\langle\ (\uparrow \textsc{subj})\ (\uparrow \textsc{xcomp})\rangle\ (\uparrow \textsc{obj})\text{'} \\ \textsc{subj} & \left[\textsc{pred} \ \text{`John'}\right] \\ \textsc{obj} & \\ \textsc{xcomp} & \left[\begin{array}{ll} \textsc{pred} & \text{`walk}\langle\ (\uparrow \textsc{subj})\ \rangle\text{'} \\ \textsc{subj} & \end{array}\right] \end{array}\right]$$

The matrix OBJ and the XCOMP SUBJ in this example are identified by an equation of functional control. This means that there are two equivalent ways of resolving the topic uncertainty in this construction, if XCOMP is allowed in the body and OBJ and SUBJ are both allowed at the bottom. Although there appears to be no need for both of these uncertainty paths in English, this formal possibility offers a simple account for certain interactions between coordination and long-distance dependencies in Japanese.

Saiki (1985) observes that some relative clauses in Japanese are constrained so that in a coordinate structure, when a SUBJ is bound in one conjunct, a SUBJ must also be bound in the other.[14] When there is a nonsubject in the one there has to be a nonsubject in the other conjunct too. The pattern is illustrated by the following examples:

(66) Takashi ga kat-te Reiko ga tabeta ringo.
Takashi SUBJ bought Reiko SUBJ ate apple.
'the apple which Takashi bought and Reiko ate.'

(67) Hon o yon-de rekoodo o kiita gakusei.
Book OBJ read record OBJ listened-to student.
'the student who read a book and listened to a record.'

(68) *Ookiku-te Reiko ga katta suika.
Big Reiko SUBJ buy watermelon.
'the watermelon which was big and which Reiko bought.'

(69) *Takashi ga nagut-te Reiko o ketobashita otoko.
Takashi SUBJ hit Reiko OBJ kicked man.
'the man whom Takashi hit and who kicked Reiko.'

Bresnan, Kaplan, and Peterson (1985) present a functionally-based theory of coordination within the LFG framework. According to this theory, coordinate structures are represented formally as a set in f-structure,

[14] Native speakers of Japanese seem to differ about the exact generalizations here. Our analysis is meant to illustrate the interaction between different components of the grammar, but as we are not experts in Japanese grammar, we remain agnostic about the exact analysis of Japanese.

with the elements of the set being the f-structures corresponding to the individual conjuncts. LFG's function-application primitive is extended in a natural way to apply to sets of f-structures: a set is treated as if it were a function with the properties that are common to all its f-structure elements. As Bresnan, Kaplan, and Peterson show, this simple extension, which is orthogonal to the extension (27) that we are proposing here, is sufficient to provide elegant accounts for the wide variety of facts that coordinate reduction rules and across-the-board conventions attempt to handle. The theory of coordination also interacts properly with the present theory of long-distance dependencies: a path of functional uncertainty that passes into a set will be resolved independently for each of the set's elements. Thus, for sentence (70a) the topic uncertainty will be resolved as XCOMP OBJ for the first conjunct and as XCOMP OBL_{TO} for the second.

(70) a. Mary John expected to see and give the book to.
b. *Mary John expected to see Bill and give the book to.

But even though the paths are allowed to differ from one conjunct to the other, it must be the case that if an uncertainty is resolved inside one of the functions it must also be resolved inside the other, as illustrated by (70b).

The fact that uncertainties are resolved independently for each conjunct, as required for the English example (70a), may seem incompatible with the Japanese pattern in (66–69). Indeed, if the within-clause role of the relative clause head is specified by a single uncertainty expression whose bottom allows either SUBJ or non-SUBJ functions, the constraint against mixing functions would not be satisfied. There is an obvious way of describing these facts, however, by specifying the within-clause function as a choice between two uncertainties, one with a SUBJ bottom and one with GF−SUBJ, as in the following rule for Japanese relative modifiers, adapted from Saiki (1985):

(71)
$$\text{NP} \rightarrow \underset{(\uparrow \text{ RELMOD}) = \downarrow}{\text{S}'} \quad \underset{\substack{(\uparrow \text{ RELMOD} \left\{\begin{matrix}\text{XCOMP}\\ \text{COMP}\end{matrix}\right\}^* \text{ SUBJ}) = \downarrow \\ \text{or} \\ (\uparrow \text{ RELMOD} \left\{\begin{matrix}\text{XCOMP}\\ \text{COMP}\end{matrix}\right\}^* (\text{GF}-\text{SUBJ})) = \downarrow}}{\text{NP}}$$

The analysis of these examples does not depend on the fact that f-structures can contain separate but equivalent paths. But there are other Japanese examples that contain two equivalent paths, one of which ends in a SUBJ and the other in a non-SUBJ. This situation arises in causatives, which, following Ishikawa (1985), are assumed to have the following lexical schemata:

(72) $(\uparrow \text{PRED}) = \text{'cause}<(\uparrow \text{SUBJ})(\uparrow \text{OBJ2})(\uparrow \text{XCOMP})>\text{'}$
$(\uparrow \text{XCOMP SUBJ}) = (\uparrow \text{OBJ2})$

The functional control equation identifies the XCOMP's SUBJ with the OBJ2 of the matrix. Saiki (1985) noticed that in this situation our formalization predicts that either of the uncertainties in (71) can lead to the common element, so that causative phrases ought to be conjoinable with other clauses in which either a SUBJ or non-SUBJ is relativized. That this prediction is correct is shown by the acceptability of the following phrases (Saiki 1985):

(73) Takashi o nagutte, Reiko ga Satoru o ketobas-ase-ta otoko
Takashi OBJ hit, Reiko SUBJ Satoru OBJ kick CAUS man
'the man who hit Takashi and who Reiko caused to kick Satoru.'

(74) Takashi ga nagutte, Reiko ga Satoru o ketobas-ase-ta otoko
Takashi SUBJ hit, Reiko SUBJ Satoru OBJ kick CAUS man
'the man who Takashi hit and who Reiko caused to kick Satoru.'

Within a classical transformational framework, the causative could be analyzed as a raising or equi construction, but at the moment of wh-movement, the information about the 'deep structure' subjecthood of the noun phrase would be unavailable. It would thus be expected to behave only as an object. With trace theory and other enrichments of phrase structure approaches, one can imagine stating the right conditions on the long distance dependency. Again, however, there is no convergence of surface structure configuration and the configuration that must be postulated to account for these cases.

6 Conclusion

LFG proposes a distinction between functionally conditioned and c-structure dependent phenomena. We have argued that long-distance wh-constructions are in fact functionally conditioned, contrary to what was previously assumed, and hence should be accounted for in the f-structure. The Icelandic facts show that c-structure dominance relations are not always relevant, the English facts show that node labels alone do not allow the proper distinctions to be made, and the Japanese causative illustrates a case in which multi-dominance is necessary. In short, the primitives of phrase-structure representation are much less adequate than those of functional structure.

Of course phrase-structure accounts of these phenomena are possible if several (successive) tree structures are admitted to encode different types of information and if traces and/or reconstruction are introduced to give the effect of multi-dominance. It is clear, though, that these accounts are not more economical than the LFG approach: besides the succession of tree structures and abstract traces, further principles must be defined to govern the mapping from one tree representation to another (such as the pruning convention proposed in Thráinsson (1986) and distinctions between casemarked and non-casemarked positions as in Stowell (1981)). We are not suggesting that such representations and principles are incapable of yielding the right empirical results. But for the claim that functional generalizations can be stated in terms of structural primitives to be interesting, it has to be shown that the postulated phrase structures are independently motivated. As the Icelandic case illustrates, there are clear cases where they are not. Given this lack of convergence, we conclude that phrase-structure accounts obscure the basically functional nature of long-distance dependencies. In part this is because they do not formally distinguish them from purely distributional generalizations such as those concerning the ordering of adverbs in Icelandic.

Acknowledgments

Joan Bresnan, Per-Kristian Halvorsen, Lauri Karttunen, John Maxwell and an anonymous reviewer are thanked for comments on earlier versions of this paper, Joan Maling and Höskuldur Thráinsson for help with the Icelandic data and the students of the LFG course (Stanford, winter 1987) and the audience of the conference on Alternative Conceptions of Phrase Structure (NYU, July 1986) for reactions to oral presentations.

References

Bresnan, Joan. 1982. Control and Complementation. In *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 282–390. Cambridge, MA: The MIT Press.

Bresnan, Joan, Ronald M. Kaplan, and Peter Peterson. 1985. Coordination and the Flow of Information through Phrase Structure. Unpublished MS.

Fassi-Fehri, Abdulkader. 1988. Agreement in Arabic, Binding, and Coherence. In *Agreement in natural language: Approaches, theories, descriptions*, ed. Michael Barlow and Charles A. Ferguson. Stanford University: CSLI/The University of Chicago Press.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279–326.

Grimshaw, Jane. 1982. Subcategorization and grammatical relations. In *Subjects and other subjects*, ed. Annie Zaenen. 35–56. Bloomington: Indiana University Linguistics Club.

Halvorsen, Per-Kristian. 1988. Situation Semantics and Semantic Interpretation in Constraint–Based Grammars. In *Proceedings of the International Conference on Fifth Generation Computer Systems, FGCS-88*, 471–478. Tokyo, Japan, November. Also published as CSLI Technical Report CSLI-TR-101, Stanford University, 1987. Reprinted in Part IV of this volume.

Higgins, F. R. 1973. On J. Emonds's analysis of extraposition. In *Syntax and Semantics*, ed. John P. Kimball, Vol. 2, 149–195. New York: Academic Press.

Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar.* Doctoral dissertation, MIT.

Ishikawa, Akira. 1985. *Complex predicates and lexical operations in Japanese.* Doctoral dissertation, Stanford University.

Jacobson, Pauline. 1982. Comments on 'Subcategorization and Grammatical Relations'. In *Subjects and other subjects*, ed. Annie Zaenen. 57–70. Bloomington: Indiana University Linguistics Club.

Johnson, Mark. 1986. The LFG treatment of discontinuity and the double infinitive construction in Dutch. In *Proceedings of the Fifth West Coast Conference on Formal Linguistics*, ed. Mary Dalrymple, Jeffrey Goldberg, Kristin Hanson, Michael Inman, Chris Piñon, and Stephen Wechsler, 102–118. Stanford University. Stanford Linguistics Association.

Kaplan, Ronald M., and Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan, 173–281. Cambridge, MA: The MIT Press. Reprinted in Part I of this volume.

Kaplan, Ronald M., and John T. Maxwell. 1988. An Algorithm for Functional Uncertainty. In *Proceedings of COLING-88*, Vol. 1, 297–302. Budapest. Reprinted in Part II of this volume.

Karttunen, Lauri. 1989. Radical Lexicalism. In *Alternative Conceptions of Phrase Structure*, ed. Mark Baltin and Anthony Kroch. 43–65. Chicago University Press.

Levin, Lori S. 1986. *Operations on lexical forms: Unaccusative rules in Germanic languages.* Doctoral dissertation, MIT.

Maling, Joan, and Annie Zaenen. 1985. Preposition stranding and passive. *Nordic Journal of Linguistics* 8.

Netter, Klaus. 1986. Getting things out of order. In *Proceedings of COLING-86*, 494–496. Bonn.

Netter, Klaus. 1987. Nonlocal dependencies and infinitival constructions in German. Unpublished MS, University of Stuttgart.

Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax.* New York: Holt, Reinhart, and Winston.

Pullum, Geoffrey K. 1982. Free Word Order and Phrase Structure Rules. In *Proceedings of the Twelfth Annual Meeting of the North Eastern Linguistic Society*, ed. James Pustejovsky and Peter Sells, 209–220. University of Massachusetts at Amherst.

Saiki, Mariko. 1985. On the coordination of gapped constituents in Japanese. In *Papers from the Twenty-First Regional Meeting of the Chicago Linguistic Society*, ed. William H. Eilfort, Paul D. Kroeber, and Karen L. Peterson, 371–387. University of Chicago. Chicago Linguistic Society.

Stowell, Timothy. 1981. *The Origins of Phrase Structure*. Doctoral dissertation, MIT.

Thráinsson, Höskuldur. 1986. On auxiliaries, AUX and VP's in Icelandic. In *Topics in Scandinavian Syntax*, ed. Kirsti Koch Christiansen and Lars Hellan. Dordrecht: D. Reidel.

Zaenen, Annie. 1980. *Extraction rules in Icelandic*. Doctoral dissertation, Harvard University. Reprinted by Garland Press, New York, 1985.

Zaenen, Annie. 1983. On syntactic binding. *Linguistic Inquiry* 14:469–504.