Anaphoric Binding in Lexical-Functional Grammar

Dan W. Higinbotham

Follow this and additional works at: https://scholarsarchive.byu.edu/dlls

BYU ScholarsArchive Citation
Higinbotham, Dan W. (1986) "Anaphoric Binding in Lexical-Functional Grammar," Deseret Language and Linguistic Society Symposium: Vol. 12 : Iss. 1 , Article 8. Available at: https://scholarsarchive.byu.edu/dlls/vol12/iss1/8

This Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Deseret Language and Linguistic Society Symposium by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Recent work by Ron Kaplan within the framework of Lexical-Functional Grammar recasts long-distance dependencies (such as relativization or topicalization) in terms of permissable paths within functional structures. This paper proposes a slight addition to the new formalism which enables Lexical-Functional Grammar to state analogues to Chomsky's Binding Principles; to describe simply the conditions on the occurrence of parasitic gaps in English; and to state principles of anaphoric reference in nonconfigurational languages, which have been resistant to description within Chomsky's Government and Binding Theory.

The first section of this paper will be a short summary of some of the formalisms of LFG; both the traditional (see Bresnan and Kaplan 1982, section 4.7) and more recent treatments of long-distance dependencies will be considered. A modest extension to the more recent treatment will next be proposed in order to handle multiple wh-gaps. Its application to other cases of anaphoric binding will then be explored, including the phenomena of reflexives, pronouns, parasitic gaps, and the Japanese reflexive 'jibun'.

LEXICAL-FUNCTIONAL GRAMMAR

The first subsection will be a short review of the basic formalism of Lexical-Functional Grammar; the second will outline the traditional treatment of long-distance dependencies, and the third will discuss some recent developments in their treatment within LFG. The discussion and most of the examples of the first two sections will be based on chapter 4 of The Mental Representation of Grammatical Relations (Bresnan and Kaplan, 1982), particularly section 4.7.

Basic LFG Formalism

In LFG, the structure of a sentence is represented by two structures.
This constituent structure (c-structure) is related to the following functional structure (f-structure):

A functional structure (the material inside any pair of brackets in (2)) is a list of labels, each with an associated value. The labels either express names of features, such as NUMBER or TENSE, or functional roles, such as SUBJ (subject) or XCOMP (open complement). The values of feature labels are atomic, such as PLURAL or PAST. The values of role labels are subordinate functional structures. Each node of the c-structure tree is associated with a functional structure. Functional structures are flatter than tree structures, because the head of a phrase is associated with the same functional structure as the phrase itself. For example, the c-structure nodes NP and N (1) share the same functional structure in each of three cases in (2); S, VP, and V in (1) also share the same f-structure in (2).

C-structures are built with the usual context-free rewrite rules. Functional equations attached to elements in these rules specify relationships and constraints on the associated f-structures.
For example, in

\[(3) \text{VP} \rightarrow \text{V} \quad \text{S'} \quad i = 1 \quad (i \text{COMP}) = 1\]

the \(i\)'s refer to the f-structure associated with the VP-node, and the \(i\)'s refer to the f-structures associated with the elements under which they are found. \(i = 1\) means that the f-structure of the VP is the same as the f-structure of the V. \((i \text{COMP}) = 1\) means that in the f-structure of the VP, there is a role label COMP whose value is the f-structure of the S'.

There is also a special label PRED which is used in lexical entries, which encodes a word's predicate argument structure. Following is the entry for 'wondered'.

\[(4) \text{wondered: V, } (i \text{TENSE}) = \text{PAST} \quad (i \text{PRED}) = '\text{WONDER<(iSUBJ)(iCOMP)>}'\]

The PRED label's value includes the semantic relation name of the word, and a list of role labels which must be found in the functional structure which includes PRED as a label, in order for that f-structure to be well-formed. If any of these role labels are missing, the structure is said to be incomplete. If any of a set of other role labels which are not listed in PRED appear in the functional structure, it is said to be incoherent (this set is called the set of governable designators, which is the set of all role labels which appear in a PRED feature for some entry in the lexicon). During a parse, the implicit rule

\[(5) \text{V} \rightarrow \text{wondered} \quad (i \text{TENSE}) = \text{PAST} \quad (i \text{PRED}) = '\text{WONDER<(iSUBJ)(iCOMP)>}'\]

creates a node labelled V in the constituent structure (c-structure) tree, and associates it with an f-structure which has the two labels TENSE and PRED and the given values.

The Uniqueness condition requires that every label in an f-structure have exactly one value. After a sentence is parsed, the \(i\)-variables and \(!\)-variables can be replaced by numbers, which are indexes to f-structures. All functional equations which appear on elements of rules which were actually matched to parse the sentence are collected into a list. The \(i\)-variables and \(!\)-variables are then replaced by indexes to f-structures; the resulting list of equations is called the functional description (or f-description), because taken together, they constitute a description of the f-structure of the sentence. In order for the structure to be well-formed (and thus for the sentence to be grammatical), the equations must be compatible; that is, the equations taken together must not specify that a single label have more
than one value. There are cases where the equations specify that two f-structures are equal. In these cases, the two f-structures are merged, or unified, into a single structure. Each label in the two f-structures must therefore have a value compatible with the value of the same label in the other f-structure, if it exists. Values can be compatible if they are atomic and exactly the same, or if both are f-structures and can be unified as well. Thus, the unification process is recursive through all sub-f-structures.

In the case of optional modifiers, however, more than one phrase may bear the same relationship to the head, and so these presumably ought to have the same kind of label. Because of such cases, the value of some labels (which are not among the governable ones mentioned above) is allowed to be a set of f-structures, rather than a single f-structure.

One rule for VP's might therefore look like

(6) VP -> V NP pp*  
\[ T = 1 \quad (\uparrow \text{OBJ}) = 1 \quad ! \in \langle \uparrow \text{ADJUNCTS} \rangle \]

The $F_1 \in F_2$ means that f-structure $F_1$ belongs to the set of f-structures $F_2$.

Since some PP's have governable functions, and their function is basically expressed by the preposition itself, the following notation is also used:

(7a) VP -> V NP pp*  
\[ T = 1 \quad (\uparrow \text{OBJ}) = 1 \quad (\uparrow(\uparrow \text{PCASE}) = 1) \]

(b) PP -> P NP  
\[ T = 1 \quad I = 1 \]

(c) to: P, (\uparrow \text{PCASE}) = \text{OBL}_T O

When a prepositional phrase with 'to' is encountered, since the f-structure of the P is the same as the f-structure of the PP, the value of PCASE in the PP is $\text{OBL}_T O$. Consequently, in (7a), the schema $(\uparrow(\uparrow \text{PCASE})) = 1$ turns out to be the same as $(1 \text{OBL}_T O) = 1$.

The Uniqueness condition and the conditions of completeness and coherence serve to constrain the types of f-structure that are allowed. Certain other constraints are also necessary, such co-occurrence constraints as agreement, for example. For this purpose, LFG has constraint equations. The following is an example:

(8) NP -> NP N  
\[ (\uparrow \text{POSS}) = 1 \quad I = 1 \quad (\uparrow \text{CASE}) \in \text{GEN} \]
The symbol $\equiv$ will be used for constraint equations in this paper. The last equation is equivalent to saying that the NP must have a feature label CASE with the value GEN in order for the rule to match. Therefore, only genitive noun phrases can satisfy the rule. The following also express constraints on the f-structure.

(9a) $S \rightarrow$ NP $\rightarrow$ VP

$\downarrow\text{(SUBJ)} = 1$ $\downarrow\text{TENSE} = 1$

$\uparrow\text{(CASE)} = \text{NOM}$ $\downarrow\text{TENSE}$

(b) $VP' \rightarrow$ (to) $VP$

$\neg\downarrow\text{(TENSE)}$ $\downarrow\text{TENSE}$

The $\downarrow\text{(TENSE)}$ schema of (9a) means that the f-structure associated with the $S$ must have the feature TENSE with some value, though that value can be any legal value of TENSE. The $\neg\downarrow\text{(TENSE)}$ schema of (9b) means that the f-structure of the $VP'$ is not allowed to have a value for the feature TENSE.

**Traditional Long-distance Dependencies in LFG**

Long-distance dependencies are relationships across an uncertain number of levels in the parse tree. This kind of relation goes beyond the elements that are found in a single rule, and consequently such dependencies are unexpressible in the formalism discussed in the last subsection.

LFG therefore uses what are called bounded domination metavariables. These are symbolized by double arrows, which will be represented in this paper by $\bowtie$ for the double up-arrow, and $\bowtie$ for the double down-arrow.

These are used to connect fillers and gaps, and also to relate a feature of a phrase with an element deeply embedded in it. One such case is the following

(10) $S' \rightarrow$ NP $\rightarrow$ S$^*$

$\downarrow\text{(Q)} = \bowtie_{[\text{wh}]^{NP}}$ $\downarrow\text{TENSE} = 1$

$\uparrow\text{(FOCUS)} = 1$

$\downarrow\text{N}_{NP}$

The subscript [wh] of $\bowtie$ in the first equation line tells the type of element required somewhere within the scope of the (superscript) NP. In effect, it requires a wh-word somewhere in the NP. The subscript NP of the $\bowtie$ in the last line requires an NP gap somewhere within the (superscript) S, the second element of the rule. The $\bowtie$ is called the controller metavariable of this dependency; the $S$ is called the domain root, and all of the nodes under the $S$ which could contain a matching $\bowtie$ are said to belong to the control domain of the controller. The $\bowtie$ notation on the $S$ is discussed below.
The lexical entry for a wh-word will include a specification that it satisfies a $\psi_{(+wh)}$ condition.

(11) who: N, (\| PRED) = 'WHO'
    $\downarrow = \bullet_{(+wh)}$

The $\bullet$ is called the controller metavariable of the dependency. The functional structure of the N(oun) c-structure node created when this word is parsed will be associated via the $\bullet$ and $\psi$ metavariables, and could be assigned as the value of the role label $Q$ because of rule (11).

Gaps are introduced by rules like

(12) NP $\rightarrow$ e
    $\downarrow = \bullet_{NP}$

which says that an NP may be expanded as a gap which will be associated with an NP filler.

Because of the necessity to handle island conditions, the notation $\bullet$ in (10) above specifies a node as a bounding node. A pair $\psi$ and $\bullet$ may not be associated with each other if the path between the c-structure nodes on which the two are introduced (not including those two nodes) contains a bounding node. In (10) above, the $\bullet$ on the S keeps any gaps within the wh-clause from being associated with anything above the S in the tree. Bounding nodes may be relevant to some kinds of dependencies, but not others. The $\bullet$ notation implies bounding of all kinds of long-distance dependencies, but can be overridden with notations like

(13) NP $\rightarrow$ NP$\bullet$
    N
    (\| POSS) = $\downarrow$
    $\downarrow = I$
    (\| CASE) $\equiv$ GEN
    $\psi_{(+wh)} = \psi_{(+wh)}^{NP}$

where the last line specifically allows a $\psi$ above the NP and a $\bullet$ below the NP to be associated with each other if both are of the $[+wh]$ type. The part of the tree above such a notation is considered to be one control domain, and the part below, another control domain. This rule (revised from (10)), allows wh-words in possessive phrases, like the sentence

(14) The girl wondered whose playmate's nurse the baby saw.

The phrase "whose playmate's nurse" is associated via the $\psi$ in (10) with the $\bullet$ in the lexical entry for 'whose'; the functional structure assigned to 'whose' becomes the $Q$ element in (11).
The "proper instantiation" of the • and ♦ metavariables includes the following conditions (as in Bresnan and Kaplan 1982, p. 246):

(15a) No node is a domain root for more than one controller,
(b) Every controller metavariable has at least one control domain,
(c) Every controller metavariable corresponds to one and only one controllee in each of its control domains
(d) Every controllee metavariable corresponds to one and only one controller,
(e) All metavariable correspondences are nearly nested,
(f) Every domain root has a lexical signature.

Condition (15f) simply means that a domain root corresponds to a phrase in the string which is at least one word long. Nesting corresponds to the diagram in (16a) rather than that in (16b).

The crossing degree of a correspondence is the number of lines it crosses. The crossing degrees of the correspondences AD, BE, and CH are each two in (16b). A • and ♦ correspondence is nearly nested if the crossing degree is not above a crossing limit, which is a parameter of a particular grammar (and of a particular language). The crossing limit given (in Bresnan and Kaplan 1982, p. 262) for English is 0, and for Icelandic is 1.

Notice that the machinery developed here does not handle parasitic gaps like the following:

(17a) Which articles did John file ___ without reading ___?
(b) This is the kind of food you must cook ___ before you eat ___.

(Engdahl 1983, p. 5)

Even though the conditions of proper instantiation allow a controller metavariable to be associated with more than one controllee metavariable (in different control domains), that is, a filler with more than one gap, there is no way to show that one gap is dependent on another. With the conditions as
given, if the second gap in each sentence of (17) can be associated with the respective fillers in those sentences, they should be equally grammatical in sentences in which the first gap of each sentence is filled by a full NP; but such sentences are ungrammatical.

This completes the summary of the treatment of long-distance dependencies in LFG. The examples and definitions given, perhaps with slight modification, have been mostly from Bresnan and Kaplan (1982, chap. 4).

New Directions in the Treatment of Long-distance Dependencies

Recently, Kaplan and associates have discussed replacing the bounded domination metavariables with functional equations of the sort below:

$$(18) \ S' \rightarrow NP \ \ S$$

\[ (\downarrow Q) = (\downarrow \ldots) \]

\[ (\downarrow Q \ WH) \equiv + \]

\[ (\downarrow FOCUS) = \downarrow \]

\[ \downarrow = (\downarrow XCOMP^* \ OBJ) \]

Paths in f-structures between the controller and controllee are taken to contain all of the relevant bounding information, rather than paths in the c-structures.

The "..." notation is one possible notation for the inclusion relationship, defined as follows:

$$(19) (f \ldots) = g \text{ iff there exists an } a \text{ such that } (f \ a) = g \text{ or } ((f \ a) \ldots) = g.$$  

(Saiki 1985, p. 3)

The "..." can therefore be replaced by any sequence of labels. The schema therefore represents possibly more than one relationship. In (18) above, $$(\downarrow Q) = (\downarrow \ldots)$$ means that the f-structure labelled Q in the f-structure associated with the S' must be the same as (one of) the (sub-)f-structure(s) associated with the NP. The constraint equation below it ensures that whichever (sub-)f-structure that is, it must have a feature WH with the value +. Such a feature and value can come only from the lexical entry of a wh-word.

$$(20) \ \text{who: N, } (\downarrow \text{PRED}) = '\text{WHO}'$$

\[ (\downarrow \text{WH}) = + \]

Consequently, the two schema together ensure that the NP contains a wh-word.

The $$\downarrow = (\downarrow XCOMP^* \ OBJ)$$ schema is similar, except that the path is more constrained. XCOMP* means any number of labels XCOMP, so the schema means that the $$\downarrow$$ f-structure will be
unified with some f-structure which appears in an f-structure like

(21a)  \[ \text{XCOMP} \[ \text{XCOMP} \[ \ldots \text{XCOMP} \[ \text{OBJ} \] ] ] ] ] ]

or

(2)  \[ \text{OBJ} \] ] ]]

Once again, the schema is indeterminate. It may be that several of the XCOMP's can have an OBJ. The ! f-structure could be unified with any of them. Each such possibility leads to a different parse. Only those parses which have a PRED in the same f-structure as the OBJ can be coherent, and of those, only those which have a PRED expecting an OBJ. It is likely that the ! f-structure already has a PRED feature, and so it will not be able to unify with the OBJ f-structure of any object which has already been added to the tree and has a conflicting PRED feature. The schema shown here is just an example, and is not likely to be part of an actual rule of English. Where several possibilities exist, notations like \{A,B,C\}* may be used, denoting any sequence of labels on the alphabet \{A,B,C\}.

Rules like the following may use the $\in$ notation:

(22)  $S' \rightarrow \text{PP} \quad S$

\[ \langle \text{TOPIC} \rangle = ! \quad ! = ! \\
\quad \langle \in \text{XCOMP}\text{* ADJUNCTS} \rangle \]

This indicates that the topicalized PP is to be added to the ADJUNCTS of one of the XCOMP's in an XCOMP* path starting at the f-structure of the $S'$.

Notice that even though there is some indeterminacy as to where the $=$ unification or the $\in$ addition is to take place, in any one parse, the unification or addition takes place at only one f-structure. The "proper instantiation" conditions above which reference the domain root or control domain are no longer applicable, but controllers and controlees are still in a one to one correspondence. This means that parasitic gaps are still outside the realm of the theory. Furthermore, the nearly nested conditions are uninterpretable, since no ordering is defined on f-structures. Perhaps these issues will be clarified in the forthcoming paper which will outline the new theory (Kaplan and Zaenan, in preparation).
UTILIZING AND EXTENDING THE NEW FORMALISM

Consider the following sentences:

(23a) I wonder who George believes Mary persuaded Bill to claim that Jehoshaphat loved ___.

(b) Who did you try to get Matilda to persuade Sue to tell ___ that you love him?

(c) * Who should I ask whether I offended ___?

(d) * Who does the fact that ___ cheated bother Bill?

It appears that grammatical sentences connect the f-structure of the root of the control domain to the f-structure of the gap with a sequence of labels that can begin with any number of labels XCOMP or COMP, followed by one label from the set \{SUBJ,OBJ,OBJ2\}. The label WHCOMP, used as the label for wh-clauses, may not be found in the path at all.

Consider next the following sentences:

(24a) Who did you go to the store with ___?

(b) Which month do we always fly kites in ___?

(c) What did you write the song on the bottom of ___?

(d) Beethoven is the composer John sent Mildred a bust of ___.

(e) Who is this a picture of a caricature of ___?

(f) I wonder who he said the song was easy to sing ___ with ___.

(g) * Phineas wondered which football player critics of ___ have never played football.

We see from these sentences that the labels OBJ and OBJ2 may also be followed by any number of labels representing the oblique functions of prepositional phrases, which can be symbolized collectively by OBL\(_e\). These conditions together can be summarized by associating an equation of the following form with the filler:

\[
(25) \downarrow = (1 \{\text{COMP}, \text{XCOMP}\}^* \{\{\text{SUBJ}, \text{OBJ}, \text{OBJ2}\}^* (\{\text{OBJ}, \text{OBJ2}\})^+ \})
\]

\(<\text{A} \text{B}>\) means B must follow A in the path designation. OBL\(_e\) is shorthand for the set of oblique functions, and OBL\(_e^+\) means at least one occurrence of one of the oblique functions. This single schema expresses many of the island constraints.
The wh-island constraint is taken care of by the fact that no WHCOMP can appear in the path. The sentential subject constraint is taken care of by the fact that if SUBJ appears in the path, it must be the last label. The coordinate structure constraint is taken care of by the fact that no COORD is allowed in the path (see a discussion of coordination in LFG by Andrews, 1983).

Now notice that for each of the following sentences, the equation in (25) would allow either of the two gaps to be associated with the filler.

(26a) Which administrator did you relay the employee's complaints about ___ to ___?

(26b) Which daytime TV actress did you convince the admirer's of ___ to carve soap statuettes of ___?

In each of these sentences, either of the two gaps could be filled with a full noun phrase, and the sentence would remain grammatical; neither gap is dependent on the existence of the other. In each sentence, both gaps satisfy the necessary constraints for long-distance binding by the wh-phrase. These two sentences show that more than one gap can be associated with the same filler.

The equation in (25) is therefore inadequate to express the situation in the sentences of (26). The equation as given implies that the f-structure referred to by 1, namely the f-structure of the filler, is the same as the f-structure at the end of exactly one path which satisfies the given path schema. If long-distance equations were to allow any number of paths compatible with the schema, the (1 Q) = (1 ...) and (1 Q WH) Ξ + equations in (18) above would allow more than one wh-element in a filler, which is clearly incorrect (except possibly in the case of a coordination). Rather than change the interpretation of equations to allow more than one path compatible with the schema, a simple solution would be to associate the equation with the gaps rather than the filler; then the filler would be fulfilling two equations, one for each gap, and it would no longer necessary to allow a single equation associated with a filler to be fulfilled by two different gaps.

Implementing this solution would require the ability of paths to look not only downward and inward toward more embedded f-structures, but also upward and outward toward less embedded structures. For this purpose, we use a / in a schema to represent an upward search, and \ to represent a return to the normal downward search.
For example, in

\[(27a)\] NP $\rightarrow$ e
\[\text{(i / OBJ} \sim \text{TENSE}) = \text{PRES} \]
\[\text{(i GAP) = +} \]

\[(b)\] VP $\rightarrow$ V NP VP'
\[\text{i = i (i OBJ) = i (i XCOMP) = i} \]

\[(c)\]
\[
\text{TENSE } \text{PRES} \\
\text{OBJ} \quad \text{GAP } + \\
\text{XCOMP} \quad \text{SUBJ} \\
\]

the i in (27a) refers to the f-structure in (27c) which contains the GAP feature. In the schema (i / OBJ) refers to the outermost f-structure shown, and the schema (i / OBJ \sim \text{TENSE}) refers to the value of the label TENSE in the outermost f-structure; that value in (27c) is PRES. Paths specified by this kind of schema are not allowed to double back on themselves. If i refers to the f-structure containing the GAP feature in (27c), the schema (i / OBJ \sim OBJ) would not successfully refer to any f-structure. Specifically, a schema may not represent any path in which the last label of the path upward is the same as the first label in the path downward. Even though the values of OBJ and XCOMP SUBJ are the same, however, the schema (i / OBJ \sim \text{XCOMP SUBJ GAP}) is perfectly legal, and has the value +; (i / OBJ \sim \ldots \text{GAP}) could only be instantiated by the path (i / OBJ \sim \text{XCOMP SUBJ GAP}), and would therefore have the value + as well. If the last label of an upward path is set valued, however, any member of the set other than the one which is the last in the upward path, may be used as the first of the downward path, even though they belong to the same label.

Functional equations involving long-distance schemata are not evaluated until all others except constraint equations. Consequently, each gap will have been assigned a (possibly empty) f-structure within the f-structure of the sentence by whichever rule incorporates the gap. In the example above, the NP in (27a) is used in (27b), and becomes part of the f-structure of the sentence via the equation (i OBJ) = i. In this way, every gap will have been assigned an f-structure which is subordinate to the f-structure of the sentence by the time the long-distance equations are evaluated. Therefore, upward searching schema, such as the unlikely (i / OBJ \sim \text{TENSE}) used in this example, will be well-defined.
The / and \- notation of the above simplified example can be used to specify exactly the constraint needed for wh-binding:

(28a) \[ S' \rightarrow NP \quad S \]
\[ (\uparrow \text{FOCUS}) = ! \quad I = ! \]
\[ (\uparrow \text{FOCUS GAPTYP}) = Q \]
\[ (\uparrow \text{GAP}) \equiv + \]
\[ (\uparrow Q) = (\downarrow \ldots) \]
\[ (\uparrow Q \text{ WH}) \equiv + \]

(b) \[ NP \rightarrow e \]
\[ \downarrow = (\downarrow / \{ \text{OBL}^+ (\{\text{OBJ,OBJ2}\}) \} > \{\text{SUBJ,OBJ,OBJ2}\} ) \]
\[ (\uparrow \text{GAP}) = + \]
\[ (\uparrow \text{GAPTYP}) \equiv Q \]

(28b) reverses the schema given before in (25). It looks outward and upward from the f-structure which was assigned to the gap, allowing the same kinds of paths as discussed above (although traversed in reverse order).

In (28a), the constraint \((\uparrow \text{GAP}) \equiv +\) ensures that at least one gap has been associated with the filler. The constraint \((\uparrow \text{GAPTYP}) \equiv Q\) in (28b) ensures that the gap has been associated with a filler of the proper type.

Together, the equations of (28) allow just the kinds of gaps that can be associated with wh-fillers, including the multiple gaps in sentences like (26).

OTHER CASES OF ANAPHORIC BINDING

In comparing the theory of control in LFG to that in Government and Binding Theory (GB), Bresnan argues that "government is a functional, not a phrase-structural, relation." (Bresnan 1982, p. 318). In particular, "a word may govern the functions in the f-structure that immediately contains the word's features." (Bresnan 1982, p. 312). She also argues that it is f-command rather than c-command that is relevant to anaphoric control (see Chomsky 1981). F-command is defined as follows:

(29) For any occurrences of the functions \(\alpha, \beta\) in an f-structure \(F\), \(\alpha\) f-commands \(\beta\) if and only if \(\alpha\) does not contain \(\beta\) and every f-structure of \(F\) that contains \(\alpha\) contains \(\beta\).

In English, the pair of functional equations \((\uparrow \text{FIN}) \equiv +\) and \((\uparrow \text{SUBJ PRED}) = 'PRO'\) may be assigned to lexical entries. PRO is like a personal pronoun but has no phonetic realization. One of the conditions on PRO is that if it is anaphorically bound by an element in the same sentence, it
must be f-commanded by that element. Its controller cannot be an element of its minimal clause. Functional control, as opposed to anaphoric control, entails that all functional features of the controller and controllee be identical, so the controller and controllee share the same f-structure. In cases of anaphoric control, the controller and controllee simply have identical reference (Bresnan 1982, pp. 321,326,333,340). If we symbolize anaphoric control by labelling the controller with the label ANA within the f-structure of the controllee, the following optional equation may therefore be added to the lexical item which bears the equation that introduces PRO:

\[(30) \{1 \text{SUBJ ANA}\} = (1 / X \sim \{\text{SUBJ,OBJ,OBJ2,OBL}_e\})\]

where X is a label variable and OBL\_e is shorthand for a list of oblique functions. In cases where this equation is not assigned, PRO receives the arbitrary interpretation as in (31a), or can be assigned a discourse anaphor as in (31b):

(31a) Baking a unicorn could be difficult.
(b) Richard was in a panic. Baking a unicorn could be difficult. He should never have volunteered to do it.

PRO is anaphorically controlled within the sentence in

(32a) Baking a unicorn could be difficult for Richard.
(b)
In this case, the \( t \) of \( (i / X ^{\text{\{SUBJ,OBJ,OBJ2,OBL\}}} ) \) refers to the f-structure labelled 1. Since f-structure 1 is the subject of f-structures 2, 3 and 4, \( (i / X) \) can refer to any of these three f-structures. Although f-structures 2, 3 and 4, all have a subject, the restriction on doubling back on itself prevents the schema from referring to them. Only OBL\textsubscript{FOR} in f-structure 4 satisfies the schema. Consequently, the anaphoric controller of 'baking a unicorn' must be 'Richard'.

**Reflexives**

In GB, PRO is both an anaphor and a pronominal. The binding theory states that an anaphor is bound in its governing category, and a pronominal is free in its governing category. (Chomsky 1982, p. 20). Just as there are functional analogues of the conditions on PRO in LFG, there are also analogues of the binding principles for anaphors and proninals. Since government in LFG amounts to a word's governing the functions in the f-structure that immediately contains the word's features, as stated above, the governing category of an anaphor or pronominal corresponds to the f-structure one functional level above the one immediately containing its features. Therefore, in the lexical entry of a reflexive pronoun (an anaphor in GB) or a persona pronoun (a pronominal in GB), \( t \) corresponds to the f-structure immediately containing its features, and \( (i / X) \), where X is a label variable, to its governing f-structure.

The equation attached to a reflexive pronoun will therefore be something like

\[
(33) \quad (\uparrow \text{ANA}) = (i / X ^{\text{\{SUBJ,OBJ\}}} )
\]

This requires the reflexive to be bound to the subject or object in its governing f-structure. Some simple examples are

(34a) Mr. C. constantly praises himself.
(b) We questioned Bill about himself.
(c) Bill asked Sue to stifle herself.

In the first two, the reflexives are bound to the SUBJ and OBJ, respectively, of their governing f-structures. In the third, the SUBJ of the infinitive is functionally controlled by 'Sue', so the reflexive is bound to the SUBJ of the infinitive, which is its governing f-structure.

**Pronouns**

Since personal pronouns may or may not be coreferential to other elements in the sentence, a group of equations like the following may optionally be attached:
(35a) \( \neg ((\text{ANA}) = (\text{SUBJ,OBJ,OBJ2}) \equiv (\text{ANA}, X \rightarrow \text{SUBJ,OBJ,OBJ2})) \)

(b) \( (\text{ANA}) = (\text{SUBJ,OBJ,OBJ2, OBL, POSS}) \)

(c) \( (\text{ANA GENDER}) = (\text{ANA GENDER}) \)

(35a) prohibits the anaphoric controller from being SUBJ, OBJ, or OBJ2 in the governing f-structure or any of its sub-f-structures. (35b) assigns any nominal f-structure in the rest of the sentence as a possible anaphoric controller, subject, of course, to the constraint in (35a). (35c) makes sure that there is no gender clash between the pronoun and its controller.

(36a) * He was well aware that Walter had maligned Ron.

(b) Geri was well aware that her husband supported Ron.

(c) At home with Jimmy one evening, Amy finally decided to set him straight.

(36a) shows that neither a subject nor object in an f-structure subordinate to the governing f-structure of the pronoun is allowed to be coreferential with it. (36b) shows that other functions in such embedded f-structures are allowed to be coreferential. (36c) shows that there are no constraints that either the controller or pronoun command the other.

Parasitic Gaps

Using the / and \( \sim \) notations in long-distance schemata also allows the beginning of a treatment of parasitic gaps in LFG. Unlike the sentences of (26) above in which a full noun phrase could appear in the position of either gap, and the functional paths from the filler to each gap could be described by the long-distance schema for wh-binding (25), a full NP cannot grammatically appear in the position of a parasitic gap, and the path from the filler to the parasitic gap cannot be described by (25). This suggests that parasitic gaps are a different kind of gap than usual gaps associated with wh-phrases. Chomsky suggests that like PRO, a parasitic gap is a base-generated pronominal (Chomsky 1992, p. 41,51). In LFG terms, this means that the equation introducing the PRED 'PRO' is attached to a lexical item rather than to a c-structure rule that introduces a gap. These gaps are bound by phrases occurring in non-argument positions, namely the complementizer position, or the position of extraposed constituents (Chomsky 1982, p. 41). In LFG terms, these are f-structures labelled by TOPIC or FOCUS. Another condition is that the real gap not command the parasitic gap (Chomsky 1982, p. 40).

Following are some examples of parasitic gaps, with the parasitic gap symbolized by \( \_p \):
Who did Jackie marry ___ in order to eventually rob ___ blind?

These are the articles John filed ___ without showing Bill ___.

This is a book you ought to read ___ before going to see the movie of ___.

This is the theoretician whose murder the article about ___ failed to mention ___ was probably caused by his students' feelings about his theory.

From these examples, we see that a parasitic gap can be an OBJ, and OBJ2, or an OBL. Chomsky also comments that SUBJ parasitic gaps are only allowed in "an Exceptional case marking construction" such as

someone who John expected ___ to be successful though believing ___ to be incompetent

(Chomsky 1982, p. 54)

where LFG would attribute both the SUBJ and OBJ functions to the gap.

Therefore, a parasitic gap is (1) a lexically introduced PRO (2) in an OBJ, OBJ2, or OBL position (3) whose controller is a FOCUS or TOPIC (4) and which is not commanded by the real gap. These constraints can be summarized in the following group of equations, which can be freely added to an item in the lexicon:

\[(39a) \quad (\alpha \alpha \text{PRED}) = '\text{PRO}'\]
\[(39b) \quad (\alpha \alpha \text{ANA}) = (\alpha / \ldots \neg \{\text{TOPIC},\text{FOCUS}\})\]
\[(39c) \quad \neg((\alpha \alpha \text{ANA}) \equiv (\alpha / \ldots \neg \{\text{SUBJ},\text{OBJ},\text{OBJ2},\text{OBL}\})\)

where \(\alpha\) is the same in all three equations, and \(\alpha \in \{\text{OBJ,OBJ2,OBL}\}\). (39a) introduces the PRO in an OBJ, OBJ2, or OBL position. (39b) binds it to an f-commanding TOPIC or FOCUS. (39c) says that the controller cannot f-command the parasitic gap; since the equations in (26) unify the f-structures of the controller and the independent gap, the f-structure of the independent gap is the f-structure of the controller, so saying that the f-structure of the controller cannot f-command the parasitic gap (as (39c) does) via functions other than TOPIC and FOCUS, is the same as saying the independent gap cannot f-command the parasitic gap. The proper interpretations of all of the sentences in (37) and (38) are determined by the equations in (39).

Notice that these equations are merely descriptions of the phenomena, rather than explanations of them. They are specific to one language (in the case of the introduction of PRO forms in the lexicon) in some cases, specific to a small set of lexical items in others (in the cases of reflexives and personal pronouns), and specific to particular c-structure rules in others (in the case of the introduction of
This is both problematic and fortunate. It is problematic in that the formalism itself gives no explanation of the phenomena. It is fortunate in that it makes no claims for universality of phenomena found only in English or related languages.

The Japanese Reflexive 'Jibun'

Consider the Japanese reflexive 'jibun', for example. In the sentence

(40) Sato wa Tanaka ga Nakamura ni Hara ga jibun no
    TOPIC SUBJ OBJ2 SUBJ self 's
    ie de korosareta koto o hanashite
    house in kill-passive-past thing OBJ tell
    shimatta koto o satotta.
    perfective thing OBJ realize-past

Sato realized that Tanaka had already told Nakamura that Hara was killed in self's house.

(McCawley 1976, p. 53)

In this sentence, the reflexive 'jibun' is three ways ambiguous; it can refer to Sato, Tanaka, or Hara. In GB terms, the governing category of 'jibun' in (70) dominates the phrase

(41) Hara ga jibun no ie de korosareta
    SUBJ self 's house in kill-passive-past

Hara was killed in self's house.

Since 'jibun' is a reflexive, it ought to be an anaphor in GB. According to Principle A of the GB binding theory, an anaphor is bound in its governing category; but 'jibun' can be coreferential with Sato or Tanaka, which are not in its governing category. Therefore, 'jibun' cannot be an anaphor. According to Principle B of the GB binding theory, a pronominal is free in its governing category; but 'jibun' can be coreferential with Hara, which is in its governing category. Therefore, 'jibun' cannot be a pronominal. The only other possibility is that 'jibun' is an R-expression. According to Principle C of GB binding theory, an R-expression is free, that is, it is not bound by SUBJ, OBJ, or OBJ2; but 'jibun' can be bound by 'Hara', 'Tanaka' and 'Sato', which are all subjects ('Sato' is marked with the TOPIC marker 'wa', but it also functions as subject of 'satotta', meaning 'realized'). Therefore, Government and Binding Theory has nothing to say about the Japanese reflexive, at least nothing correct.
The conditions governing 'jibun' are as follows:

(42a) Only higher animate nouns can be reflexivized.
(b) 'Jibun' is used invariably, regardless of person, gender, and number.
(c) The antecedent must be the subject of a sentence (the subject-antecedent condition).
(d) The antecedent must command the reflexive (the antecedent-command condition).

(Inoue 1976, p. 118)

These can be captured in LFG simply by adding the following equations to the lexical entry of 'jibun':

(43a) ( \[I\text{ANA}\] = (\[I/\ldots\cdot\text{\textasciitilde}\text{SUBJ}\])
(b) (\[I\text{ANA ANIM}\] )

The first equation says that 'jibun' is anaphorically bound by a SUBJ which f-commands it. The second says that its controller must be animate. Since no equations mention person, gender, or number, no agreement is necessary between 'jibun' and its controller.

CONCLUSION

This paper proposes a slight addition to the new formalism proposed in recent work on long-distance dependencies by Ron Kaplan and others. An extension to this new formalism, using the upward-looking mechanisms symbolized by / and \textasciitilde, enables Lexical-Functional Grammar to state analogues to Chomsky's Binding Principles. The conditions on the occurrence of parasitic gaps in English can be stated simply. Principles of anaphoric reference in even nonconfigurational languages, which have been resistant to description within Chomsky's Government and Binding Theory, are also easy to state within the new formalism.
REFERENCES


