Modeling syntactic constraints on anaphoric binding

Mary Dalrymple
SRI International
and
Stanford University

John Maxwell
Xerox-PARC

Annie Zaenen
Xerox-PARC
and
CSLI, Stanford University

Abstract

Syntactic constraints on antecedent-anaphor relations can be stated within the theory of Lexical Functional Grammar (henceforth LFG) through the use of functional uncertainty (Kaplan and Maxwell 1988; Halvorsen and Kaplan 1988; Kaplan and Zaenen 1989). In the following, we summarize the general characteristics of syntactic constraints on anaphoric binding. Next, we describe a variation of functional uncertainty called *inside-out* functional uncertainty and show how it can be used to model anaphoric binding. Finally, we discuss some binding constraints claimed to hold in natural language to exemplify the mechanism. We limit our attention throughout to coreference possibilities between definite antecedents and anaphoric elements and ignore interactions with quantifiers. We also limit our discussion to intrasentential relations.

1 General characteristics of syntactic constraints on anaphoric binding

The relation between an anaphor and its antecedent is semantic in nature. In the simple cases that we limit our attention to here, the two are coreferent. [1] This semantic relation is subject to syntactic constraints, however, and it is the statement of these constraints that we focus on.

In the LFG approach to these constraints proposed in Bresnan et al. (1985),[2] binding conditions are stated as conditions on f-structure configurations rather than conditions on c-structures. Two kinds of syntactic factors are shown to influence anaphoric binding possibilities: the grammatical function of the potential antecedent (in particular whether or not it is a subject) and the characteristics of the syntactic domain in which the potential antecedent and the anaphor are found (for example, whether that domain is tensed or whether it has a subject). In Bresnan et al. (1985), anaphors are consequently annotated for both domain and antecedent constraints.

Some constraints are stated in positive terms: the antecedent must be found within a particular domain or have a particular function. In other cases the constraints are negative: the antecedent and the anaphor cannot both be part of a particular domain, or the antecedent cannot bear a particular grammatical function. Under such negative conditions, the anaphor is *disjoint* in reference from its antecedent.

2 Modeling binding constraints with functional uncertainty

F-structure relations are in some cases not characterizable as a finite disjunction over paths: for example, dependencies between 'fillers' and 'gaps' in, for example, relative clauses and wh-questions. Functional uncertainty was developed for the analysis of such dependencies.

Kaplan and Maxwell (1988) and Kaplan and Zaenen (1989) develop a formal specification of relations involving disjunction over paths by allowing the argument position of functional equations to denote a set of strings. Suppose α is a (possibly infinite) set of symbol strings; then

(1) $(f\ \alpha) = v$ holds if and only if

a. $f = v$ and $\epsilon \in \alpha$, or

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

[1] This is of course not always the case. Reciprocals and binding by quantified NP's are two well-known cases in which the semantic relation is more complicated.

[2] For a summary of the views in Bresnan et al. (1985), see Sells (1985).

An equation with a string-set argument holds if and only if it holds for some string in that set. This kind of equation is trivially unsatisfiable if α denotes the empty set. If α is a finite set, this formula is equivalent to a finite disjunction of equations over the strings in α. Passing from finite disjunction to existential quantification enables us to capture the intuition of unbounded uncertainty as an underspecification of exactly which choice of strings in α will be compatible with the functional information carried by the surrounding surface environment.

Kaplan and Zaenen (1989) require that α 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.

Functional uncertainty can also be used in the case of negative constraining equations. In that situation, the requirement is that there be *no* path picked out by the regular expression that makes the equation true. That is, the negation of an expression involving functional uncertainty has the effect of negating an existentially quantified expression.

Kaplan and Zaenen (1989) consider only expressions of the form

$$(f\ \alpha)$$

where α is a regular expression. In expressions such as these, α represents a path through the f-structure f. We refer to paths of this type as *PathIn*, and to functional uncertainty of this type as *outside-in* functional uncertainty.

In Halvorsen and Kaplan (1988), expressions of the form

$$(\alpha f)$$

are introduced. We will refer to the path in expressions of this form as *PathOut*, and to functional uncertainty of this type as *inside-out* functional uncertainty. Expressions involving inside-out functional uncertainty are interpreted as denoting f-structures from which f is reachable over some path in α.

More formally:

(2) $(\alpha f) = g \in \{h \mid \exists s \in \alpha[(hs) =_c f]\}$

$(\alpha\ f)$ denotes some f-structure g through which there is a path in the set of strings α leading to f. The equation $=_c$ is a constraining equation checking for the existence of such an f-structure.

Relations between anaphors and their antecedents are also in some cases not characterizable as a finite disjunction of paths within f-structures; for this reason, the use of functional uncertainty in characterizing the anaphor-antecedent relation seems appropriate. In our view, modeling anaphoric binding constraints consists of specifying a set of f-structure paths relating anaphors with elements that are either possible or disallowed antecedents. We use inside-out functional uncertainty to characterize the relation between an anaphor and these elements.

To illustrate, the antecedent of the Norwegian anaphor *seg* must be a subject outside of the minimal complete clause nucleus[3] in which *seg* appears; this antecedent can be at an indefinite distance away from the anaphor, as long as only the highest nucleus in the domain contains a tense marker (Hellan 1988; p. 73):

(3) Jon bad oss forsøke å få deg til
Jon$_i$ asked us to try to get you to

å snakke pent om seg
talk nicely about him$_i$

Under an LFG analysis, the path between the antecedent and the anaphor in (3) contains three XCOMPs, as diagrammed in Figure 1. Assume that $\uparrow_A$ denotes the f-structure for *seg*, the structure labeled 9 in Figure 1. The set of nested f-structures containing 9 is characterized by the regular expression

(4) $(\text{XCOMP}^* \ \text{OBJ} \uparrow_A)$

In Figure 1, this set consists of the structures labeled 1, 2, 3, and 4. The expression in (5) designates the subjects of these four f-structures, those labeled 5, 6, 7 and 8:

(5) $((\text{XCOMP}^* \ \text{OBJ} \uparrow_A) \ \text{SUBJ})$

F-structures 5, 6, and 7 are the f-structures of the possible antecedents of *seg*: the subjects outside of the minimal clause nucleus in which *seg* appears. F-structure 8 is not a possible antecedent for *seg*, since it appears in the same minimal clause nucleus as *seg*; f-structure 8 will

[3] A *clause nucleus* is formed by any predicate (regardless of its syntactic category) and its dependents. A *complete clause nucleus* is a clause nucleus with a subject dependent.

$$1:\begin{bmatrix} \textsc{subj}\ 5{:}[\,] \\ \textsc{xcomp}\ 2{:}\begin{bmatrix} \textsc{subj}\ 6{:}[\,] \\ \textsc{xcomp}\ 3{:}\begin{bmatrix} \textsc{subj}\ 7{:}[\,] \\ \textsc{xcomp}\ 4{:}\begin{bmatrix} \textsc{subj}\ 8{:}[\,] \\ \textsc{obj}\ 9{:}[(\text{anaphor})] \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Figure 1: F-structure for sentence (3)

be excluded from the set of possible antecedents for *seg* by a negative constraint.

More schematically, the set of possible antecedents of an anaphoric phrase can be characterized by an expression of the form in (6):

(6) ((PathOut $\uparrow_A$) PathIn)

(PathOut $\uparrow_A$) picks out the set of f-structures which contain the anaphor and in which the antecedent must be located. PathIn characterizes the functional role of the antecedent. It is a general constraint on antecedent-anaphor relations that the the antecedent must *f-command*[4] the anaphor; for this reason, the PathIn is always of length one. The PathIn, then, consists of (and constrains) the grammatical function borne by the antecedent.

Conditions on the binding domain are formalizable as conditions on the PathOut, since the PathOut characterizes the domain in which both the anaphor and its antecedent are found. We will look in detail at one such constraint; before doing so, however, we make a simplifying assumption about the semantics of the anaphor-antecedent relation.

In the simple cases we are considering here, the relation is be represented as identity between the semantic content of the anaphor and its antecedent. Elaboration of this representation would require us to introduce the LFG mechanism of *projections* (Halvorsen and Kaplan 1988), which is beyond the scope of this paper.

Here we will use the informal notation in (7):

(7) $< \sigma >$ ((PathOut $\uparrow_A$) PathIn)$=< \sigma >\uparrow_A$

[4]Bresnan (1982) defines f-command as follows: for any functions GF1, GF2 in an f-structure, GF1 f-commands GF2 iff GF1 does not contain GF2 and every f-structure that contains GF1 contains GF2.

to indicate that the semantics of the anaphor, $< \sigma > \uparrow_A$, is to be identified with the semantics of its antecedent. The material in angle brackets stands for the mapping (not further specified) between the syntax and the semantics.

To prevent the anaphoric element from being contained in its antecedent, we formulate the constraint in (8), where $\uparrow_{ANT}$ stands for the f-structure of the antecedent:

(8) $\neg\ [(\uparrow_{ANT}\ \mathrm{GF}^+) = \uparrow_A]$

The effect of this constraint is very similar to the i-within-i condition in Government-Binding Theory (Chomsky 1981). It has been argued that this constraint should be relaxed (see e.g. Hellan (1988)) but the correct analysis of putative counterexamples is not clear. We will assume here that the constraint can be maintained.

We now describe how to model a domain constraint that holds of some anaphors: some anaphors must be bound within the minimal complete nucleus — the minimal nucleus containing a subject.

Let F_1 designate an f-structure containing the anaphor. We can characterize F_1 in the following way:

(9) $F_1 = (\mathrm{GF}^+ \uparrow_A)$

where GF denotes the set of grammatical function labels.

For F_1 to be a valid binding domain for anaphors subject to this constraint, it must not contain any smaller f-structure that properly contains the anaphor and a subject. That is, F_1 must be the *smallest* complete nucleus. We will define DPF ('domain path f-structure') as any of the f-structures that contain the anaphor and are properly contained in F_1:

(10) $(\mathrm{DPF_1\ GF^+}) =_c \uparrow_A$
$\mathrm{DPF_1} =_c (\mathrm{F_1\ GF^+})$

It is these intermediate f-structures that must not contain a subject:

(11) $\neg(\mathrm{DPF_1\ SUBJ})$

The constraint that an anaphor must be bound within the minimal complete nucleus can, then, be stated as follows:

(12) a. $<\sigma> (\mathrm{F_1\ GF}) =<\sigma>\uparrow_A$

b. $\neg(\mathrm{DPF_1\ SUBJ})$

These two equations ensure identity between the semantic content of the anaphor and its antecedent, where the antecedent is the value of some GF of an f-structure $\mathrm{F_1}$ that contains the anaphor. There may not be a f-structure $\mathrm{DPF_1}$ that is properly contained in $\mathrm{F_1}$ which has a subject.

3 Examples of anaphoric binding

We now illustrate the use of these binding constraints with some of the conditions that have been proposed for English, Marathi, and Scandinavian pronouns and reflexives.[5]

The English reflexive pronoun was described in Bresnan et al. (1985) as having to be bound in the minimal complete nucleus, as illustrated by the following contrast:

(13) a. He_i told us about $\mathrm{himself}_i$.

b. We told him_i about $\mathrm{himself}_i$.

c.*He_i asked us to tell Mary about $\mathrm{himself}_i$.

As discussed in Section 2, this pattern of grammaticality judgments can be modeled by the constraints given in (9) through (12).

The antecedent of the Marathi reflexive *swataah* must be a subject, but may be at an indefinite distance from the anaphor, so long as the antecedent and the anaphor appear in the same minimal tensed domain. This requirement can be translated into the following path specification.

(14) a. $<\sigma>(\mathrm{F_1\ SUBJ}) = <\sigma>\uparrow_A$

b. $\neg(\mathrm{DPF_1\ TENSE}) = +$

where $\mathrm{F_1}$ and $\mathrm{DPF_1}$ are as defined above

According to these equations, the antecedent of the anaphor must be contained in an f-structure $\mathrm{F_1}$; further, there must not be an f-structure $\mathrm{DPF_1}$ properly contained in $\mathrm{F_1}$ that has a TENSE attribute with value +.

A more interesting case arises when a binding relation is subject to both a negative and a positive constraint. An example is the Swedish anaphor *honom själv*. Its antecedent must appear in its minimal complete clause nucleus, but it must be disjoint from subjects. This anaphor occurs felicitously within the following sentence:

(15) Martin bad oss berätta för honom
Martin_i asked us to talk to him_i

om honom själv
about $\mathrm{himself}_i$

Conditions on *honom själv* do not prohibit *Martin* and *honom själv* from being interpreted as coreferent, though *Martin* bears the grammatical function SUBJ. This is because *Martin* appears outside the binding domain of *honom själv* and is thus not considered when either positive or negative binding constraints are applied.

In our framework, two constraints are required for *honom själv*. One, (16)a, states the positive constraint: the domain in which the antecedent of *honom själv* must be found. The other, (16)b, states the negative constraint: *honom själv* must be disjoint from the subject in that domain.

(16) a. $[\mathrm{F_1} = (\mathrm{GF^+}\ \uparrow_A) \wedge$
$<\sigma>(\mathrm{F_1\ GF}) = <\sigma>\uparrow_A \wedge$
$\neg(\mathrm{DPF_1\ SUBJ})]$

b. $\neg\ [F_2 = (GF^+\ \uparrow_A) \wedge$
$<\sigma> (F_2\ \mathrm{SUBJ}) =<\sigma>\uparrow_A\ \wedge$
$\neg(DPF_2\ \mathrm{SUBJ})]$

The negative constraint rules out coreference only between the anaphor and the subject of the minimal complete clause nucleus; it does not prevent coreference between the anaphor *honom själv* and a subject *Martin* outside the binding domain. In general, negative binding constraints do not hold in a larger domain than is specified by the positive equation.

[5] Data are from Bresnan et al. (1985), Hellan (1988), and Dalrymple (in prep.).

For the Norwegian anaphoric form *hans*, the only specifications are negative (Hellan(1988), Bresnan et al. (1985)); it must be disjoint from the immediately higher subject. We can encode this requirement as:

$$(17)\ \neg\ [F_1 = (GF^{+}\ \uparrow_A) \wedge \\ <\sigma> (F_1\text{SUBJ}) = <\sigma> \uparrow_A \wedge \\ \neg(DPF_1\ \text{SUBJ})]$$

This is the same negative requirement as was illustrated above, in example (16). As no positive requirement is given, no antecedent relation is imposed. It is assumed that another module, presumably the discourse component, will supply a referent for the pronoun.

4 Conclusion

We have sketched a way to use inside-out functional uncertainty to constrain the relation between an anaphor and an antecedent. A formal theory of anaphoric binding will involve a specification of a universal inventory of anaphoric binding possibilities and possible dependencies between them.

A general discussion of such a theory is beyond the scope of this paper, but we conclude by indicating how our approach captures a few of the cross-linguistic properties of anaphoric binding.

If the domain and the antecedent binding requirements for an anaphor are both positive or both negative, the requirements must be satisfied by the same element. This is enforced by requiring that only one positive and one negative equation can be associated with each anaphor.

Additionally, only elements that are *superior* to the element should be considered in applying the constraints. GF1 is *superior* to GF2 if (1) GF1 asymmetrically f-commands GF2, or (2) GF1 and GF2 f-command each other, and GF1 is higher on the hierarchy of grammatical functions given in (18):

(18) SUBJ > OBJ > OBJ2 > OBL > ADJ

As noted above, the f-command requirement is enforced by the requirement that the PathOut be non-null and the PathIn be of length one. The modelling of the functional hierarchy given in (18) within our framework is, however, a task that remains to be done.

A final observation is that inside-out functional uncertainty can interact with outside-in functional uncertainty as used in the analysis of dependencies between 'fillers' and 'gaps', as in the following:

(19) a.*Bill said that Sue likes himself.

b. Himself, Bill said that Sue likes.

Preliminary research indicates that no special machinery is needed to model the right interactions in these cases.

References

Bresnan, J. 1982. Control and complementation. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 282–390. Cambridge, Mass.: MIT Press.

Bresnan, J., P.-K. Halvorsen, and J. Maling. 1985. Logophoricity and bound anaphors. MS, Stanford University.

Chomsky, N. 1981. *Lectures on Government and Binding.* Dordrecht: Foris Publications.

Dalrymple, M. in prep. *Syntactic Constraints on Anaphoric Binding.* PhD thesis, Stanford University.

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

Hellan, L. 1988. *Anaphora in Norwegian and the Theory of Grammar.* Dordrecht: Foris Publications.

Kaplan, R. M., and J. Maxwell. 1988. An algorithm for functional uncertainty. In *Proceedings of COLING 88.*

Kaplan, R. M., and A. Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In M. Baltin and A. Kroch (Eds.), *Alternative Conceptions of Phrase Structure.* Chicago University Press.

Sells, P. 1985. *Lectures on Contemporary Syntactic Theories.* Stanford University: CSLI/University of Chicago Press. CSLI Lecture Notes, Number 3.