An alternative to the Binding Theory

Joke Dorrepaal *
Onderzoeksinstituut voor Taal en Spraak
3512 JK Utrecht
Joke.Dorrepaal@let.ruu.nl

Abstract

This paper presents an alternative approach to the Binding Theory (cf. Chomsky (1981), Pollard and Sag (1992), a.o). It will be argued that the core issue of such a theory should be locality constraints rather than descriptive content. We demonstrate that a reformulation of the Binding Theory in this spirit accounts more adequately for phenomena like kataphora and weak cross-over. We will furthermore place the approach in the perspective of a theory of discourse, and argue that the notion descriptive content belongs at the level of discourse rather than in sentence grammar.

1 Standard Binding Theory

Binding Theory refers to a set of principles which governs the syntactic anaphoric behaviour of NPs. The term 'Standard Binding Theory' shall be used to refer to the current approach taken to Binding in the literature (cf. in GB (Chomsky, 1981), in HPSG (Pollard and Sag, 1992), etc.). The differences which exist among these theories do not affect the argument presented in this paper.

Data The theory serves to account for grammaticality judgments in cases like (1), where equal subscripts denote coreferentiality. The examples show that the type of NP determines whether an anaphoric relation is possible or not. In the minimal pair (1)b-c, for example, the pronoun can be coreferential with the subject *Otje*, whereas the reflexive cannot. Similarly, wheras the pronoun can be coindexed with the proper noun in (1)c, a full definite NP in that same position cannot be coreferential (cf. (1)d).

*This research has been carried out in project LRE-61-062, 'Towards a Declarative Theory of Discourse', funded by the Commission of the European Communities.

(1) a $\quad$ Otje$_i$ likes herself$_i$

b $\quad$ *Otje$_i$ thinks Bob likes herself$_i$

c $\quad$ Otje$_i$ thinks Bob likes her$_i$

d $\quad$ *Otje$_i$ thinks Bob likes the woman$_i$

In principle, coindexing between NPs is random as long as agreement is respected. The aim of the theory is to restrict - or filter - the possible coindexation relations between two NPs. The (im)possible relations are dependent on the types of NPs involved in the relation and the principles governing these types.

Types Standard Binding Theory classifies NPs into three groups. Each group is subject to a principle which governs the possible anaphora relations for the particular type of NP. These principles are known as Principle A, B, and C respectively. The NPs are classified in the following way:

- *ana*: reflexives and reciprocals
- *pro*: pronouns
- *np*: non-pronominal full NPs

Principles Each type of NP is subject to a principle which restricts the anaphoric possibilities. These principles are formulated below:

- Principle A: If an anaphor is of type *ana*, it must be locally bound to an antecedent
- Principle B: If an anaphor is of type *pro*, it must be locally free
- Principle C: If an anaphor is of type *np*, it must be free

The auxiliary notions (locally) bound/ free can de defined as follows:

- An NP is locally bound if it is coindexed with a locally commanding NP
- An NP is locally free if it is not coindexed with a locally commanding NP
- An NP is free if it is not coindexed with a commanding NP

What remains to be explained is the notion *command*. This notion cannot be explained in a theory-neutral way. In GB, the notion *command*, called c-command, is defined configurationally, i.e. on the surface syntactic tree representation. In HPSG, *command* is called o-command and is defined in terms of an obliqueness hierarchy, which is reflected by the order of elements on the SUBCAT-list. We shall present a simplified definition of both notions here (cf. (Reinhart, 1976), (Pollard and Sag, 1992)).

In GB:

- A c-commands B iff the first branching node dominating A dominates B
- A locally c-commands B iff A c-commands B, and A and B are contained in the same minimal S or NP

In HPSG:

- A locally o-commands B iff there is a SUBCAT-list on which A precedes B
- A o-commands B iff A locally o-commands some C dominating B

2 Standard Problems

There is one important claim in the Binding Theory presented above which raises some questions. The claim is that all full NPs behave the same as far as binding relations are concerned. Proper nouns, non-pronominal definite NPs, indefinite and quantified NPs are all subject to Principle C, and therefore all treated equally. Now, bearing in mind that the distinction between definite NPs and indefinite/quantified NPs has often been characterized as one between 'familiarity' and 'novelty' respectively (Christophersen (1939), Karttunen (1976), Heim (1982)), we wonder how this relates to these binding principles.

Consider the following data, where definite NPs are contrasted with nondefinite NPs:

(2) a When he$_i$ arrived home, John$_i$ kissed his wife
b *When he$_i$ arrived home, every man$_i$ kissed his wife

(3) a John$_i$ is a fool but John$_i$ doesn't mean any harm
b *A man$_i$ is a fool but a man$_i$ doesn't mean any harm

(4) a John$_i$'s mother loves John$_i$'s father
b *A boy$_i$'s mother loves a boy$_i$'s father

(5) a His$_i$ friends say John$_i$ is very intelligent
b *His$_i$ friends say [every boy in my class]$_i$ is very intelligent

The examples above all respect principle C of the Binding Theory. The full NPs are free - they are coindexed with some NP but not with a *commanding* NP. And yet, the a) examples are much better than the b) examples.

The problem has been noted before (cf. the appendix to Lasnik (1989)) and solutions have been proposed, but they were not adopted by the standard version of the Binding Theory. One such solution proposes to add constraints on the antecedent in the relation. Postal (Postal, 1971), for example, claims that when a definite pronoun is to the left of an NP, this NP must be definite for it to serve as antecedent. Wasow

(1972) suggests that the relevant distinction is between referring expressions, which he calls 'determinate', and other NPs. The former class includes specific NPs and generic NPs. Nonspecific nongenerics are indeterminate, in Wasow's terms.

Solutions like these seem to describe the data more adequately, but are otherwise not really 'natural' extensions of the Binding Theory. The Binding Theory is about anaphors, in that it is concerned with:

- the type of anaphor and
- whether the ANTECEDENT *commands* the anaphor

The alternatives proposed by Wasow and Postal (a.o) add constraints concerning

- the type of the antecedent
- whether the ANAPHOR *precedes* the antecedent

These constraints would have to be added explicitly to the Binding Theory, as a special instance of Principle C.

Another type of solution is to say that the difference between examples like (5)a and (5)b should be accounted for in the semantic representation. It is defendable that a property like 'definite' or 'determinate' belongs to semantics rather than syntax. However, this does not really solve the problem. In Reinhart (1983), it is argued that examples like (5), known as cross-over cases, remain a problem. Reinhart discusses the following example:

(6) a The secretary who works for him$_i$ despises Siegfried$_i$
   b *The secretary who works for him$_i$ despises each/a manager$_i$
   c *The secretary who works for each manager$_i$ despises him$_i$
   d Each x: x a manager [the secretary who works for x despises x]

> [..] while free coreference is possible in a), the pronoun cannot be interpreted as bound by the operator in b). The puzzling problem about these facts, known as the '(weak) cross over' cases, is that they do not follow from semantic considerations: The unavailability of anaphora interpretation in c), is easily explained on semantic grounds: since in the semantic interpretation of the sentence the pronoun is not in the scope of the operator corresponding to *each manager*, it obviously cannot be interpreted as a variable bound by this operator. This however is not the case in b). A logical formula which can, reasonably, correspond to these sentences is the one given in d), in which the pronoun is in the scope of the operator and its translation to a bound variable is perfectly acceptable" (Reinhart (1983), pp 55-56).

The quotation above supports the idea that the unavailability of anaphora in the case of (6)b should be ascribed to properties of surface constituent order rather than to scope. After all, the logical formula for (b) could be (d), and there is no reason why the quantified NP cannot bind the variable in this formula. All this favours an analysis in terms of the syntactic binding theory. We shall propose an alternative to the Binding Theory, which appears to account more adequately for the examples in (2)-(6).

3 The notion Antecedent

The first step towards the alternative binding theory is to redefine the notion 'antecedent'.

The antecedent is usually the NP which

- precedes the anaphor
- has more descriptive content than the anaphor

The question is what happens if there is a clash between these two intuitive characteristics of the anaphora relationship. Consider an example where the more descriptive NP *follows* the underdetermined NP:

(7) a In one of his$_i$ papers, Lasnik$_i$ discusses bound variables
b In one of his$_i$ papers, the man$_i$ discusses bound variables

It is common use to call *Lasnik* the antecedent in a), and to refer to the pronoun as the anaphor. Apparently, descriptive content is more important than precedence. Also in cases where the descriptive content is minimal, as in b), the NP which has relatively more descriptive content is regarded as the antecedent.
Our claim is that the antecedence should prevail over descriptive content. The anaphoric relation is *not* a relation between two NPs Ant and Ana, where Ant has more descriptive content than Ana. Anaphora is a relation between two NPs L and R, where L precedes R.

It was pointed out to us that maybe, in an example like (7)a, there is no direct anaphoric relation between the pronoun and the full NP. It could be the case that both *he* and *Lasnik/the man* are coreferential with an NP in the preceding sentence. The reply to this observation is twofold. First, (van Deemter, 1991) gives strong evidence that kataphora really exists:

(8) a *Mary, Dorit and Bill are a strange lot. She is weird. The others ..
b Mary, Dorit and Bill are a strange lot. Ever since her childhood, Dorit has been extremely lazy. The others ..

Example (8)a is unacceptable and (b) is felicitous. The pronoun cannot be used anaphorically in these examples. The fact that (8)b is correct stems from the fact that the pronoun refers to the second occurrence of 'Dorit'. The pronoun is a kataphor.

And second, in a set $\{NP_i,\ NP_j,\ NP_k\}$, it doesn't make much difference whether we say that NP$_j$ and NP$_k$ are coindexed directly or NP$_j$ and NP$_k$ are both coindexed with NP$_i$. In either case, we have to account for (im)possible coreferentiality of NP$_j$ and NP$_k$.

4 An alternative Theory

In this section, we will define our alternative approach in a similar way in which we presented the standard theory. At the end of this section, we shall repeat the essential differences and discuss how the alternative theory relates to the data.

Types We distinguish the following types of NPs:

- *ana*: reflexives and reciprocals
- *ppro*: pronouns
- *defnpro*: full **definite** NPs, i.e. proper nouns and NPs with a definite article

Note that the third type of anaphoric NP excludes nondefinite NPs.

Principles Let R, and L be NPs, then:

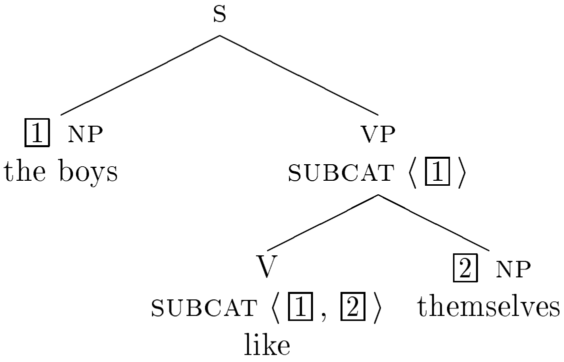
- Principle A: R should be coindexed with an L iff R is of type *ana* and L locally commands R
- Principle B: R may be coindexed with L iff R is of type *ppro* and L globally commands or non-commands R
- Principle C: R may be coindexed with L iff R is of type *defnpro* and L non-commands R

Note that *iff* - if and only if - in the definition above excludes additional random coindexing. All coindexing is guided by the principles A-C. The notion *command* in our theory corresponds to o-command in HPSG. As in HPSG, we define *command* on the subcatlist:

- L locally commands R if L precedes R on a SUBCAT-list
- L globally commands R if L locally commands Y and Y dominates R
- L non-commands R if
  - X dominates L, Y dominates R and X locally commands Y, or,
  - X dominates L and X locally commands R

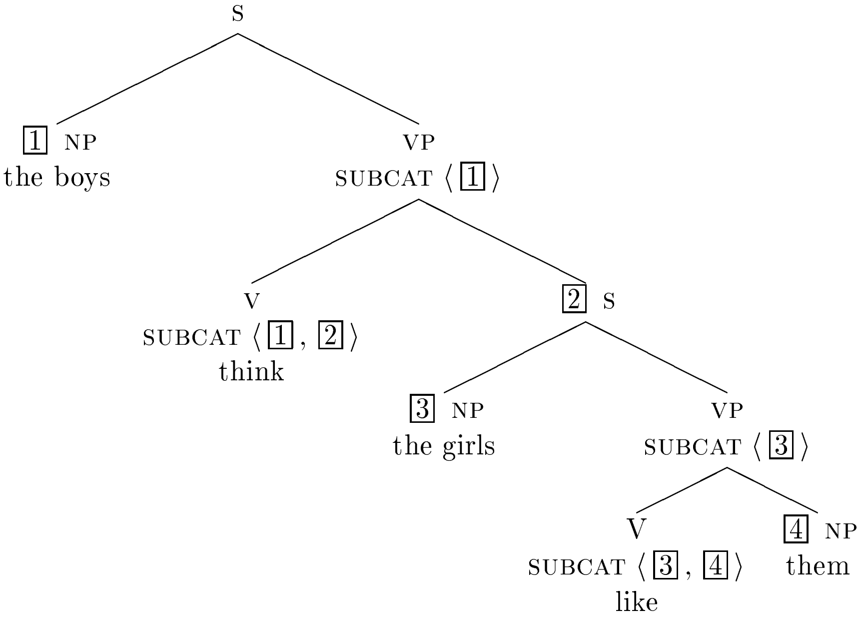
Examples We can illustrate this by the following examples:

(9) The boys like themselves



The NP *the boys* locally commands the anaphor *themselves* because there is a SUBCAT-list on which the first precedes the latter. The NP 'themselves' is of type *ana* and according to Principle A, *the boys* and *themselves* must be coindexed.

(10) The boys think the girls like them



The NP *them*, a pronoun, is locally commanded by the NP *the girls* because the latter precedes the pronoun on a SUBCAT-list. However, *them* and *the girls* may not be coindexed because pronouns should be *globally* commanded, not locally. It *is* possible to coindex the pronoun with the NP *the boys*. The relation between them is that of global command: NP$_{\boxed{1}}$ locally commands S, which dominates NP$_{\boxed{4}}$.

Data revisited

The main difference with standard binding theory is that we propose a mechanism of controlled coindexing instead of random (co)indexing. The coindexing is controlled by the Binding Principles. Given that those principles require the righthand element in the relation, R, to be definite, this approach accounts for the data which pose problems in standard theory:

(11) a When he$_i$ arrived home, John$_i$ kissed his wife
b *When he$_i$ arrived home, every man$_i$ kissed his wife

(12) a John$_i$ is a fool but John$_i$ doesn't mean any harm
b *A man$_i$ is a fool but a man$_i$ doesn't mean any harm

(13) a John$_i$'s mother loves John$_i$'s father
b *A boy$_i$'s mother loves a boy$_i$'s father

(14) a His$_i$ friends say John$_i$ is very intelligent
b *His$_i$ friends say [every boy in my class]$_i$ is very intelligent

In the a) examples, the proper nouns are all subject to Principle C, which results in (optional) coindexing with a non-commander. For the quantified/indefinite NPs in the b) example, no coindexing with L is allowed.

5 Binding Theory in a Broader Perspective

In our more general approach to anaphora, the syntactic binding constraints are merely a subset of all constraints governing anaphora. Additionally, we distinguish semantic constraints - to account for the difference between bound and discourse anaphora - and discourse contraints. The idea is that for an anaphora relation to hold, all these constraints have to be respected simultaneously.

The motivation for adding this section to our paper is that indeed, syntactic constraints may interact with other constraints. Consider for example the following data:

(15) a *When he$_i$ came in, a man$_i$ switched on the light.
b When he$_i$ came in, John$_i$ switched on the light.
c ?He$_i$ came in. John$_i$ switched on the light.

We explained above that unlike standard Binding Theory, our approach accounts for the difference between (15)a and (15)b. The anaphoric relation can only be established if the second NP is definite. However, if we take a discourse approach to anaphora, the question is how we account for (15)c, or even, how we *wish* to account for (15)c. The following are the options:

(16) a John$_i$ arrived. He$_i$ came in. John$_i$ switched on the light.
b *Bill$_j$ arrived. He$_i$ came in. John$_i$ switched on the light.
c Bill$_j$ arrived. He$_j$ came in. John$_i$ switched on the light.

The question is how to rule out kataphora across sentences, as in 16. Our hypothesis is that anaphors should be *less specific than or equally specific as* their antecedents. This hypothesis shall be referred to as the Specificity Condition. Note that this condition is a discourse constraint: it works only extrasententially.

Assume that a sort hierarchy like the following represents specificity of sorts, where A $\ll$ B encodes that A is more specific than B:

Maaike, Brigitte	$\ll$ GIRL				
		$\ll$ FEMALE			
Marlies, Mary	$\ll$ WOMAN				
			$\ll$ HUMAN	$\ll$ ANIMATE	
John, Fred	$\ll$ BOY				
		$\ll$ MALE			
Harry, Bill	$\ll$ MAN				
					$\ll$ ENTITY
Soccer	$\ll$ SPORT				
Coffee,Tea	$\ll$ DRINK	$\ll$ LIQUID		$\ll$ NON-ANIMATE	
Brazil, Spain	$\ll$ COUNTRY				

This sort hierarchy encodes the sorts of the objects we talk about, and their relative specificity. The Specificity Condition claims that the anaphor should always be less (or equally) specific than the antecedent. Consider the following examples (assuming that the pronouns *she* and *he* are of sort FEMALE and MALE respectively):

(17) a Maaike$_i$ was in Spain$_j$ last year. The girl$_i$/She$_i$ liked the country$_j$
b *Maaike went to a sunny country$_j$ last year. She liked Spain$_j$
c *A girl$_i$ went to Spain last year. Maaike$_i$ liked the country

In (17)a, the anaphors are less specific than the antecedents according to the sort hierarchy defined above. In (17)b-c, there is an anaphor which is more specific than the antecedent, which results in bad, or at least unnatural, discourses.
Now let's return to the original example, (16)b, repeated below:

(18) *Bill$_j$ arrived. He$_i$ came in. John$_i$ switched on the light.

This example would obviously be ruled out: the anaphor *John* is more specific than the antecedent *he*. Our hypothesis as it stands is not entirely adequate yet. Consider (16)a again, repeated below:

(19) John$_i$ arrived. He$_i$ came in. John$_i$ switched on the light.

This discourse may be a counterexample to the Specificity Condition if you think of the second occurrence of *John* as being anaphoric to *he*. After all, the proper name is more specific than the pronoun. There is one obvious solution that we shall not adopt.

One might say that the the second occurrence of *John* is not coreferential with the pronoun but with the first occurrence of *John*. This solution is undesirable from the perspective of recency. We wish to be able to recognize the most recently mentioned object, no matter whether it is referred to with a pronoun or otherwise. Another solution is to say that the anaphor may not be more specific than the descriptive content of the *chain of antecedents*. This latter solution we will adopt. It seems indeed possible to use an anaphoric phrase which addresses information from different antecedents in the chain:

(20)	Speaker 1:	Did I tell you about Gloria?
	Speaker 2:	You mean the secretary?
	Speaker 1:	No, I'm talking about the dentist
	Speaker 2:	Oh of course, it's just that I haven't heard from Gloria the dentist for ages

The following definition provides the final version of the Specificity Condition:

Specificity Condition
An anaphor may be coindexed with an antecedent iff the antecedent is part of an A(ntecedent)-chain which is at least as specific as the anaphor.

At least a few critical remarks with respect to this condition are in order. Consider the following examples [1]:

(21) a A man_i entered the restaurant. He first started to shout insults at the women. Next, the chauvinist pig_i even molested one of the waitresses.

b Maaike went to a sunny country_i last year. She wanted to go to Spain, but eventually went to Portugal_i

c Once upon a time there was a king_i. King Rednose_i was a wealthy man.

In all these cases, the anaphor is more specific than the antecedent. Example (21a) demonstrates that in fact, the specificity condition should not be imposed on the antecedent itself but on the whole preceding context, or at least 'on the antecedent and what is being said about it'. It reminds us of the infamous example of a man ordering a hamburger in a restaurant and leaving without paying. The comment of one of the waitresses is "Look, the hamburger left without paying his bill". We realise that our approach fails here. As to the discourse (21)b, we are not sure whether this really is a case of coreference. Consider the following variant, where the supposed antecedent and anaphor are reversed:

(22) Maaike went to Portugal last year. She wanted to go to Scotland, but eventually went to a sunny country

[1] We thank the anonymous reviewers for the data.

This example is equally acceptable, despite the fact that indefinite NPs never have an anaphoric identity relation with a previously mentioned object.

The example in c) is frequent in language, esp. in children's stories and in news bulletins. The only defense against this counterexample we have is that it is not entirely neutral language use. I remember once having read a sports article and though I am quite a soccer-fan, I had no idea whether they talked about one, two or three footballplayers. It ran more or less as follows:

(23) The keeper caught the ball.
The clown of the Mexican team made a signal to the trainer.
Campos was injured.

Now, after the World Cup, I understand that they were talking about Campos, the clownish keeper of the Mexican national team ..

6 Concluding Remarks

In most linguistic theories, the anaphora problem is tackled from only one perspective. Syntactic theories have primarily been concerned with syntactic sentence-internal constraints. However, Standard Binding Theory also has an 'inbuilt' notion of specificity, viz. by requiring that the antecedent be more specific - has more descriptive content - than the anaphor. In this paper we have tried to demonstrate that the binding theory is much more adequate if it ignores the issue of descriptive content. We have also tried to show that this issue should be treated at the level of discourse rather than at sentence-level. We therefore proposed a binding theory which deals with locality constraints only, and a discourse principle which accounts for descriptive content, or specificity as we call it.

References

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

P. Christophersen. *The Articles - A study of Their Theory and Use in English.* Copenhagen: Munksgaard, 1939.

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

L. Karttunen. Discourse referents. In J.McCawley, editor, *Syntax and Semantics 7.* New York, 1976.

H. Lasnik, editor. *Essays on Anaphora.* Kluwer Academic Publisher, 1989.

C. Pollard and I. Sag. *Information Based Syntax and Semantics, Volume 2.* Center for the Study of Language and Information. Stanford, 1992.

P. Postal. *Cross-Over Phenomena.* Holt, Rinehart and Winston, 1971.

T. Reinhart. *The Syntactic Domain of Anaphora.* PhD thesis, MIT, 1976.

T. Reinhart. Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics & Philosophy*, 6:47–88, 1983.

K. van Deemter. *On the Composition of Meaning.* PhD thesis, Universiteit van Amsterdam, 1991.

T. Wasow. *Anaphoric Relations in English.* PhD thesis, MIT, 1972.