On Binding Domains

Martin Everaert
Utrecht University

Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar
Department of Informatics, University of Lisbon
Stefan Müller (Editor)
2005
CSLI Publications
pages 503–518

http://csli-publications.stanford.edu/HPSG/2005

Abstract

In this paper I want to explore reasons for replacing Binding Theory based on the anaphor-pronoun dichotomy by a Binding Theory allowing more domains restricting/defining anaphoric dependencies. This will, thus, have consequences for the partitioning of anaphoric elements, presupposing more types of ‘anaphors/’pronouns’ than standard Binding Theory offers us.

1. Introduction

Mainstream generative accounts (Chomsky 1981; Pollard & Sag 1994; Manning & Sag 1999; Bresnan 2002, and Reinhart & Reuland 1993) sketch a very clear, uniform picture of anaphoric dependencies. Binding in the syntactic sense of the word is primarily limited to the predicational domain, formulated as in binding conditions A (cf. 1) and B (cf. 2): 1

(1)  
(a) An anaphor is bound in its Governing Category.
(b) A locally a-commanded short-distance reflexive must be locally a-bound.
(c) A nuclear (reflexive) pronoun must be bound in the minimal nucleus that contains it.

(2)  
(a) A pronominal is free in its Governing Category.
(b) A pronoun must be locally a-free.
(c) A nonnuclear pronoun must be free in the minimal nucleus that contains it

‘Reflexives’ are subject to condition (1), i.e. they are referentially dependent upon a hierarchically superior NP (cf. 3a), and the antecedent must be found within a certain domain (cf. 3b). 2

2 Anaphoric dependencies are indicated by italics.

I would like to thank the organizers of the workshop, António Branco and Manfred Sailer, and the editor of this volume, Stefan Müller, for giving me the opportunity to present my work, and their patience. Alexis Dimitriadis, Shakuntala Mahanta, Eric Reuland, Anca Sevcenco, Giorgos Spathas have contributed, in different ways, to this paper, for most without knowing it.
‘Pronominals’ obey condition (2), the reverse from (1): whatever the reference of the pronoun may be, it is not able to take a co-argument for an antecedent.

These standard generative binding conditions (cf. Everaert 2003 for a comparison of Binding Theories in several generative frameworks) describe recurrent patterns in the various languages of the world. Examples from Finnish (4a), Sakha (4b, personal communication Nadya Vinokurova), and Spanish (4c) illustrate that, in many languages, reflexives and pronominals are, indeed, in complimentary distribution:

(4)  a. Pekka näki itsensä/*hänet
    ‘Pekka saw himself/him’
  b. Misha bejetin/*kinini tapyyr
    Misha himself/him \loves
    ‘Misha loves himself/him’
  c. Juan se/*lo admira
    ‘Juan admires himself/him’

The examples in (5), from Italian, Dutch, Russian, and Icelandic, respectively, show that, in addition, reflexives must be locally bound, while pronominals allow non-local binding:

(5)  a. Gianni pensava che Maria *si/lo ammirasse
    ‘Gianni thought that Maria admired him’
  b. Jan vroeg mij voor *zich/hem te werken
    Jan asked me for himself/him to work
    ‘John asked me to work for him’
  c. Vanja dumaet čto Maša uvažaet *sebja/ego
    ‘Vanja thinks that Maša admires him’
  d. Jón veit að María elskar *sig/hann
    ‘John knows that Maria loves-IND himself/him
    ‘John knows that Maria loves him’

In all generative accounts (HPSG, LFG, P&P, etc.) there seems to be general agreement on the following properties being encoded in Binding Theory:

(6)  i. Reflexivization is local.
ii. A distinction must be drawn between two types of anaphoric element: anaphors (= reflexives and reciprocals) and pronouns.

iii. Any anaphoric dependency that is non-local is either exceptional, marked or does not fall under Binding Theory proper. In other words, anaphor resolution (as it is used in the literature on discourse) is outside the scope of Binding Theory.

In this paper I will focus on (6ii). However, it will become clear that this is only possible if we also address (6i). In other words, I will discuss:

(7) i. the notion ‘domain’/’locality’.
ii. the partitioning of elements that are sensitive to binding restrictions.

It is important to observe that I will be guided by the principle in (8), which is inspired by a view, formulated in (9), on what syntax might be:

(8) Binding Theory deals with those nominal expressions that encode their referential properties in the morpho-syntactic vocabulary (feature system) of a specific language.

(9) “One of the prerequisites for attaining the goals of the Minimalist Program (MP) developed in Chomsky 1995, 2000, to appear, is to draw the boundaries of syntax in a principled way. The MP proposes that the computational system of human language (C_{UL}) reflects the combinatorial properties of a purely morpho-syntactic vocabulary.” Reuland (2001: 440)

My starting point is that any grammatical feature that is morpho-syntactically encoded might be, in principle, be relevant for binding. Taking (8) as a fundamental principle will significantly widen the empirical scope of the Binding Theory. It defines it as an interface system, as discussed in Reuland (2001). Although what I propose is compatible with Reuland’s position, the focus is slightly different. Reuland (2001) is focused on the binding principles A and B, both part of syntax, replacing syntactic ‘identity derived by co-indexation’ from ‘identity derived by movement’. I am arguing that there might be reason to extend Binding Theory to discourse.

2. Partitioning of anaphoric elements

Nominals are generally partitioned as follows (Pollard & Sag 1994):
Since we generally accept that reflexives and reciprocals behave the same with respect to binding conditions, (10) is reduced to (10'), with the three binding conditions indicated:

(10')

Let us, for the moment, focus on binding condition A (cf. 1). It restricts elements classified as 'anaphors' to be bound locally. And local is defined in several ways:

3 'subcat-list', 'arg-structure', 'complete functional complex', 'predicate', etc.

Condition A, however, is not without exceptions. Quite early on it was noted that, cross-linguistically, there were many anaphors with antecedents essentially beyond the regular domain (Thráinsson 1976, Reis 1976, Inoue 1976, Yang 1983, Harbert 1983, and many others since). The examples in (12), Norwegian, Dutch, Japanese and Icelandic, respectively, illustrate this:

(12) a. Jon bad oss hjelpe seg
    Jon asked us help himself
    ‘John asked us to help him’

b. Jan laat mij voor zich werken
    Jan made me for himself work
    ‘John made me work for him’

A very different take on locality is the assumption that anaphora domains and NP-movement domains coincide (Reuland 2001, Hornstein 2001).
c. *Bill-wa John-ga zibun-o seme-ta to omot-ta*
   Bill John himself blamed that thought
   ‘Bill thought that John blamed him’

d. *Jón segir að Péturi raki sig á hverjum degi*
   Johnn says that Peter shave himself at every day
   ‘John says that Peter shaves him every day’

Following the terminology of Koster & Reuland (1991) we will classify the exceptions to binding condition A in (12a,b) as medium distance binding, and those in (12c,d) as long distance binding. Medium distance is reflexivization that is non-local, but the non-locality is restricted to a reanalysis/small clause domain. The phenomenon of long distance binding, a binding relation between an anaphor and a non co-argument antecedent, is tackled in different ways:

(13) **Long distance binding is:**
   c. accounted for by introduction of a fourth binding condition, principle Z: cf. (14) for a formulation of the principle

(14) **Principle Z** (Xue et al. 1994, and others; formulation from Branco 2005)
   An o-commanded long-distance reflexive must be o-bound.

It is this fourth binding condition, principle Z, that allows Branco & Marrafa (1997) and Branco (2005) to explore the possibility of deriving the binding conditions from a more general principle of quantification structure. Branco (2005) argues that the empirical generalizations captured in the definition of the four binding principles, conditions A,B,C and principle Z, are “just the effect of the specific quantificational force of the anaphors lexically encoded in their semantic values” (Branco 2005: 166). So, the question whether the four-way partitioning of binding conditions is motivated, and linked to well-motivated partitioning of lexical elements, becomes an important one.

In the way Principles A and Z are formulated a distinction is made between short-distance and long-distance binding. The question, of course, is whether such a distinction is motivated. And if so, could it be that this distinction is derived from other principles of grammar. Many have argued that it could be derived from the morphology of anaphoric elements. Pica (1985) argued that long distance anaphors are heads, short distance anaphors are ‘complex’.
Everaert (1986) argued that the fact that certain anaphors require strict local binding follows from their morpho-syntactic make-up. Alternatively, we could derive the distinction between short distance anaphors and long distance anaphors from a well-defined feature specification. Everaert (1991) argues that short distance anaphors could be seen as +A,-P specified, to be distinguished from +A,+P long distance anaphors. Defining the notions ‘governing category’/ ‘minimal governing category’ relative to the A(naphor)- and P(ronominal)-features, respectively, Everaert derives that <+A,+P> reflexives, bound in some governing category and in their minimal governing category, are necessarily locally bound, while <+A,-P> reflexives, bound in some governing category and not bound in their minimal governing category, are not.

I will assume that, indeed, something like principle Z exists, but that it is, perhaps, the only binding principle in the traditional sense of the word that exists. Following Everaert (1986) I would like to suggest that binding condition A is, a priori, non-local, but limited to the sentence-internal domain.

3. Domains

What would be a priori domains relevant for anaphoric dependencies? The first distinction seems to be the distinction between the domain in which syntax is relevant, sentence grammar (cf. 15a,b), and the domain where syntax is only indirectly relevant, discourse (cf. 15c,d). Within sentence grammar we might make a distinction between the domain in which predicate-based grammatical processes like passive apply (cf. 15a) versus the domain in which processes like wh-movement apply (cf. 15b). At the discourse level we distinguish discourse (15c) from deixis (cf. 15d), the latter being the more ‘local’ option in discourse.

(15) For \( y = \) reflexive, \( x = \) antecedent of \( y \):

a. (complex) predicate/clause

\[ \ldots \ldots \{CP, \ldots \ldots \} \ldots \ldots \]

b. sentence

\[ \{CP, \ldots \ldots \} \{CP, \ldots \ldots \} \ldots \ldots \]

c. deixis

\[ \{CP, \ldots \ldots \} \ldots \ldots \]

\[ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \]

d. discourse

\[ \ldots \ldots \] \[CP, \ldots \ldots \] \[CP, \ldots \ldots \]

---

4 Whether or not such generalizations hold true is not at issue here (cf. Everaert 1991).
In the Principles and Parameters theory, Lexical-Functional Grammar, Head-Phrase Structure Grammar, Binding Theory is focused on syntactic binding, limited to the predicational domain. Reflexives encode referential dependencies in the clausal domain, i.e. (15a). In all Binding Theories that I am acquainted with, with the exception of Reflexivity, there is room for debate whether (15b) could still be taken as a possible domain for regular ‘syntactic’ binding. But for all Binding Theories mentioned above, reference outside the sentence, i.e. (15c,d) is forbidden ground for anaphors (cf. Kang 1988 for discussion). For the domain of discourse, we exclusively have elements called pronouns, and the binding conditions have nothing to say about anaphoric dependencies in this domain.

Is there a reason to assume that anaphora are partitioned this way? In other words, is there reason to assume that we need more than the simple anaphor (for 15a) – pronoun (for 15b,c,d) distinction of BT? If we look at what defines an element as an ‘anaphor’ it is not straightforward that the anaphoric dependencies in (15a) and (15b) would be morpho-syntactically encoded differently from those in (15c) and (15d). It is not evident that a definition of anaphors rooted in Chomsky (1986) and Keenan (1988) according to which anaphors are referentially defective NPs predicts that reflexives could, for instance, never be taken as discourse anaphora (15d). Only if reflexive anaphors were necessarily interpreted as bound variables, subject to a c-command/o-command/ syntactic rank restriction, the predicted discourse restrictions on reflexive anaphors would follow naturally from whatever explains the (un)grammaticality of the examples in (16):

(16) a. *Every ex-husband feared that he would be neglected
   b. *Because she hated every ex-husband, Mary would certainly tell Zelda
      why she left him
   c. *Every ex-husband feared that I would be neglected. He …

In other words, we generally assume that the preferred domain for a ‘reflexive’ is (15a). There is no a priori reason that this should be the case, but most languages (like Dutch, Spanish, Russian, etc.) mentioned above offer us this as the primary distinction.

In a sense, English is rather atypical, because its reflexive anaphor can be used in all domains. That is, it is often used in more structural configurations than we might consider calling reflexive environments:

---

5 It has been observed that in various languages reflexives are used as honorifics. See Siewierska (2004: 224-228) for an overview on this particular, deictic, use of reflexives.
(17) a. Predicate:
Mary thinks that [John saw himself]
b. Sentence:
And that was exactly it, he thought. [He really didn’t care too much
[wat happened to himself]
c. Deixis:
There were five tourists in the room apart from myself
d. Discourse:
[Whom he [=Philip] was supposed to be fooling, he couldn’t imagine].
[Not the twins, surely, because Désirée, in the terrifying way of
progressive American parents, believed in treating children like adults]
and [had undoubtedly explained to them the precise nature of her
relationship with himself].

With the fourfold distinction given in (15), we could, in principle, expect a
language to make the following partitioning, giving every domain its unique
identifiable anaphoric element:

(18) a. anaphor\textsubscript{1} for (15a)
b. anaphor\textsubscript{2} for (15b)
c. pronoun\textsubscript{3} for (15c)
d. pronoun\textsubscript{4} for (15d)

As far as I can tell there is no language that straightforwardly offers us this
picture - four different forms - but there are many languages that offer a morpho-
syntactic partitioning of anaphoric elements that is clearly different from the
simple anaphor-pronoun distinction. In the following section I will give a very
limited sketch of some of the diversity one may find.

4. Anaphoric elements and their domains

The literature gives us overwhelming evidence that most/all languages seem to
have an anaphor\textsubscript{1}-type. To give an example, take the Norwegian reflexive seg
selv, which contrary to seg, can only be bound in its most immediate domain, as
is shown by the contrast between (12a), here repeated, and (19):

(12) a. Jon bad oss hjelpe seg
   ‘John asked us to help him’
(19) *Jon bad oss hjelpe seg selv
   ‘John asked us to help himself’
Likewise, reciprocals seem to be primarily clause-bound, as has been observed in Yang (1981). This is illustrated for Kannada in (20) (Amritavalli 2000: 67,89):

(20) a. [shyaama tannannu, priitissaane anta] raama, heeLidanu
   Shyama self,acc loves that Rama said
   ‘Rama said that Shyama loves him (=Rama)’

   b. *makkaLu, [naanu obbaranna obbaru, baide anta] heeLidaru
   children I one,acc one,nom scolded that said
   ‘The children said that I scolded one another’

But what about the other anaphor/pronoun types that could, potentially, exist? A language like Tamil gives a good illustration of the point I want to make.

4.1 Tamil
Tamil, as described in (Lehmann 1989, Annamalai 2000), has two pronouns referring to 3rd person antecedents: avan (that one, he; 3rd Person, Masculine, Accusative, -Proximate) and ivan (this one, he; 3rd Person, Masculine, Accusative, +Proximate). In addition Tamil has a pronominal form taan (3rd Person, -Plural, not specified for gender), which could be taken as the equivalent of English himself.

(21-22) illustrate the binding properties of taan; taan cannot be discourse bound (cf 21), but intra-sentential reference is not restricted to the local domain (cf. 22a,b)

(21) a. *kamalaa avan tann-ai veru-kkir-aan en-ru ninai-tt-aa
   Kamala he self-acc hate-pres-3sm say-vbp think-pst-3sf
   ‘Kamala thought that he hated him (=Kumaar)’

   b. *kumaar kaDekki poonan; ange tanaiKku oNNum piDikkale
   Kumar shop to go-pst-agr there self to anything like not
   ‘Kumar went to the shop; he did not like anything there.’

(22) a. kamalaa avan tann-ai veru-kkir-aan en-ru ninai-tt-aa
   Kamala he he-acc hate-pres-3sm say-vbp think-pst-3sf
   ‘Kamala thought that he hated himself’

   b. kamalaa avan tann-ai veru-kkir-aan en-ru ninai-tt-aa
   Kamala he she-acc hate-pres-3sm say-vbp think-pst-3sf
   ‘Kamala thought that he hated her’

---

6 Cf. Everaert 2005 for a discussion of this generalization.
7 A similar partitioning of anaphoric elements and similar distributional facts hold for Malayalam, Bangla, Telugu (cf. Jayaseelan & Haripasad 2001).
In Lehmann (1989) *taan* is described as a 4th person pronoun: “the occurrence of *taan* in a reflexive construction is only one of its occurrences and there is, therefore, no justification to call it a reflexive pronoun […] just because it can occur in a reflexive construction.” (p.97) In other words, because *taan* is not limited to the smallest domain (21a), but is regularly used in a wider domain (21b), like an anaphor, type, Lehmann does not want to call it a reflexive, contrary to Annamalai (2000).

In some cases, however, *taan* seems to behave like a true anaphor-type, necessarily clause bound, as is shown in (23):

(23) a. kumaar umaa tanne tiTTikiTTaaNNu sonnaan
   Kumar Uma self-acc scold-pst-VR-pst-agr-that say-pst-agr
   ‘Kumar said that Uma scolded himself’

b. *kumaar umaa tanne tiTTikiTTaaNNu sonnaan
   Kumar Uma self-acc scold-pst-VR-pst-agr-that say-pst-agr
   ‘Kumar said that Uma scolded himself’

Note, however, that it is the verbal auxiliary *kiDu* reflexive marking the embedded predicate, resulting in local binding (23a), blocking long-distance binding (cf. 23b).

The pronoun *avan* is the designated element for discourse binding (cf. 24a); local binding is excluded (24b), unless modified by an emphasis marker (24c):

(24) a. kumaar kaDekki poonan; ange avanukku oNNum piDikkale
   Kumar shop to go-pst-agr there he to anything like not
   ‘Kumar went to the shop; he did not like anything there.’

b. *kumaar avan-ai veru-kkir-aan
   Kumar he-Acc hate-pres-3sm
   Kumar hates himself

c. kumaar avaneyee verukaan
   Kumar he-acc-emph hate-prst-agr
   ‘Kumar, hates himself/him’

The differences/similarities between the proximate/obviative pronouns becomes clear in (25-26). (25) shows that both pronouns can be used deictically, but that for sentence internal reference *ivan*, the proximate element, is excluded:

(25) a. *ivan en tampi
Summarizing we can say that *taan* is an anaphor₂ element that is used for sentence internal reference (cf 15b); *ivan* is a pronoun₃ element, used for deictic contexts only (15d); *avan* can be used for deixis, discourse binding and sentence internal binding (15b,c,d). Strict local binding (cf 15a) is only realized when the anaphor₂ element *taan* is combined with a verbal reflexive marker, making it a reflexively marked predicate in the sense of Reinhart & Reuland (1993).

### 4.2 Roumenian and Mupun

There are other languages that, like, Tamil, seem to have a anaphor₂ element, whose distribution is defined as in (15b): the ‘reflexive’ *sine* in Roumenian (Sevcenco 2004) and the ‘logophoric pronoun’ *émi* in Fon (Kinyalolo 1993) and *dî* in Mupun (Frajzyngier 1997). I will limit my brief discussion here to Roumenian and Mupun.

The distribution of the Romanian anaphor *sine* (Sevcenco 2004) shows that it can be bound in both local and long distance contexts, as in (27), which involves the occurrence of *sine* in a clitic doubling structure, and (28), which is ambiguous between the reading in which Alex is the antecedent of *sine* and another reading in which George is the antecedent:

(27) Directorul se admiră pe sine.
    Director-the se REFCL ACC admires 3SG pe PREP ACC self.
    ‘The director admires himself’.

---

8 All languages seem to morpho-syntactically encode indexicals like *I, we, you* of the pronoun₄ type.
9 The fourth person pronouns in Mabaan as described in Andersen (1999) might offer another example.
10 What is interesting is that Romanian seems to have no ‘logophoricity’ constraints, in the semantic sense. But does have blocking effects.
George wants that Alex count on self.

‘George wants that Alex count on Alex/George’.

Logophoric systems are, generally, also defined by the domain given in (15b). The case of Mupun (Frajzyngier 1997) illustrates this:

(29) a. wu/wa/mo sat nò ta dì/dè/dù d'ee n-jos
   he/she/they say COMP stop he/she/they stay prep-Jos
   ‘He/she/they said that he/she/they stopped over in Jos’

b. wu/wa/mo sat nò ta wà/wà/wà d'ee n-jos
   he/she/they say COMP stop he/she/they stay prep-Jos
   ‘He/she/they said that he/she/they stopped over in Jos’

In (29a) the logophoric pronouns refer, necessarily, to the matrix subject. If one want to encode sentence external reference a regular pronoun is chose, as illustrated in (29b).

5. Conclusion

In the preceding section I have given some evidence for a richer classification of anaphoric elements that the anaphor-pronoun distinction. This is based on the assumption that we should distinguish four types of domains, as sketched in (15). Many languages indeed reflect these domains by morpho-syntactic encoding domain with dedicated anaphoric elements. The consequences for a proper formulation of the Binding theory are substantial. Given the postulation of four domains of anaphoric dependencies, and four anaphoric types, we might also need four binding conditions. However, not in the traditional sense of the word.

Anaphoric dependencies outside the scope of sentence grammar I leave undiscussed here. But, clearly, notions like Source, Self and Pivot, as introduced in Sells (1987) will play a crucial role.

For sentence grammar we, at least, need the equivalent of Principle Z, for instance:

(30) An anaphor is bound (=c-commanded by a co-indexed element)

This condition applies to any element that is standardly called a reflexive/reciprocal, but it also holds for logophors, or ‘4th person’ pronouns. This condition gives no domain restriction other than that the antecedent must be a sentence internal c-commanding NP. The fact that certain anaphors have a
restricted choice of antecedents, a co-argument, is the result of reflexive marking of the predicate of which the anaphors is an argument. Reflexive marking is either overtly visible through verbal morphology, or covertly through incorporation of a reflexive-marker (cf. Reinhart & Reuland 1991, Anagnostopoulou & Everaert 1999), generally morpho-syntactically encoded on the anaphoric element itself. One could take (31) as a binding condition,

(31) A reflexive marked predicate must be reflexive

but this condition is different from (30) in that it not directly refers to the anaphoric element itself.

References

Everaert, M. 2005. Long-Distance Reciprocals, ms Utrecht University


