

Anaphora and coreference: definite and non-definite antecedents

Roberto Musella

In this paper we will be concerned with the conditions under which two NPs can or cannot be anaphorically and/or coreferentially linked, and we will try to sketch out a theory of nominal anaphora. First of all, we will try to define correctly the *obiectum* of analysis (§ 3), by distinguishing between anaphoric and coreferential reading (§ 1-2), then we will formulate (§ 4) a 'relational' restriction on anaphora reading and we will verify (§ 5) its empirical adequacy. Then (§ 6), we will deal with instances of anaphora with non-definite antecedents and we will show that it is not necessary to formulate specific restrictions for these NPs types insofar as their anaphoric behaviour is analogous to that of definite NPs; rather, we will show (§ 6.4.) that, in the case of non-definite antecedents, it is only the availability of coreferential reading that is more restricted.*

0. Introduction.

Specifying the restrictions on anaphoric and/or coreferential interpretation has been one of the most investigated problems in the last three decades, mainly by the generative grammarians. All generativist theories¹ of anaphora state that there are restrictions on the coreferential interpretation of two nodes just in case one of them is in the "syntactic domain"² of the other.

Given, then, two NPs, the crucial question is whether either of them is in the syntactic domain of the other: if this is not the case, there are

* This article is based on my doctoral dissertation (Musella (1990)), carried out under the direction of Professor Raffaele Simone, to whom I wish to express my gratitude firstly for initiating me into linguistic studies and then for being constantly my landmark for over a decade now. I would also like to thank the two anonymous referees for their precious suggestions.

¹ In this work we will refer to the theories developed by the generativists (especially from Reinhart (1976) and (1981)) as the most consistent and complete attempt to elaborate a theory of nominal anaphora.

² The syntactic domain of a given node "is the portion of the tree consisting of those nodes which [this] node bears the structural relation to" (Reinhart 1983b: 14). This notion has had several definitions; the most widely accepted one is by Reinhart (1976). For a brief presentation and discussion of this notion, we refer to what will be said in §. 4.

no restrictions on the coreference options. On the contrary, if either of them is in the domain of the other, there are restrictions.³ This means that linear ordering has no relevance either in the definition of syntactic domain or, more generally, in the coreferential interpretation of two nodes in the same S(entence).

Though agreeing with generative grammarians about the possibility of formulating precise⁴ restrictions on anaphora interpretation and about the centrality of the notion of "domain of a NP", we will argue that:

- a) this notion has to be defined not in configurational, but in 'relational' terms: what is important in the anaphoric interpretation of two NPs is their Grammatical Function (Subject, Direct Object, ...);
- b) the linear order of the NPs is relevant when two NPs are not in the same domain.

We will see that these two assumptions make our theory empirically and explanatorily more adequate than configurational ones: among other things, our theory differs from configurational ones in that there is no need to formulate *ad hoc* restrictions on non-definite anaphora.

We will argue too that Reinhart's theory is even inadequate at an observational level, insofar as it does not distinguish between anaphoric and coreferential readings; therefore, first of all, in order to define correctly the *obiectum* of analysis, we will try to precisely characterize the notions of anaphora and coreference.

1. Anaphora.

Anaphora is an asymmetrical and intransitive relationship between two terms of the same text: the antecedent or "attachment point", which is the anaphorized term, and the pro-form, which is the anaphorizing one.

The essential feature of the anaphora relation is the unilateral dependence (hence the asymmetry) between the antecedent and the pro-form: the pro-form cannot be interpreted, unless it is linked to its attachment point. The anaphoric relationship, besides being asymmetrical, is, as already pointed out, also intransitive:⁵ if Y depends on X, it depends on no other term.

³ Prior to Lasnik (1976), this restriction was valid only when the antecedent was a pronoun. Lasnik claimed, on the other hand, that the restriction was always valid, even when NPs were full NPs. Lasnik (1986) went back to the question, adopting the weak version, inasmuch as in some languages – but not in English – an R-expression can be coreferential with another R-expression, even if one is in the domain of the other. We shall return to this point further on.

⁴ In fact there are those who, like Bolinger (1980) for instance, doubt even the very possibility of formulating restrictions on the anaphoric resumption of a NP. "After years of efforts as rule-making that have only led up one blind alley after another, a number of researchers have concluded that the key to 'pronominalization' is not to be found in syntax, perhaps even that the 'key' does not exist" (p. 289).

⁵ Wasow (1979) asserts the transitivity of anaphoric relationship, and he goes so far as to formulate a Transitivity Condition. Wasow's assertion is vitiated, in my opinion, by the non-

Obviously, this does not mean that X itself cannot be a pro-form and consequently depend on a term Z. However, if a pro-form cannot depend but on one antecedent, a single term can serve as an antecedent for more than one pro-form:

- (1) Bill said to himself that he would leave the next day.

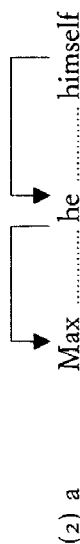
In this sentence *Bill* is the antecedent of both *himself* and *he*.⁶ Conversely, in a sentence like

- (2) Max confessed to Claire that he only trusts in himself.

himself is anaphorically dependent on *he*⁷ which in turn depends on a preceding third term, which may be *Max*.

The correct analysis of a sentence such as (2) demands, then, further refining of the notion of antecedent: we have to identify the "immediate antecedent", which can be a pro-form itself, and the "last antecedent", that is the first member of the anaphoric chain, which must be autonomously referential⁸ instead.

For ease of presentation, the last antecedent will be called "attachment point", and the term "antecedent" will be used as unmarked term, to which the attribute "immediate" will be added to mean, obviously, the "immediate antecedent". In (2) *Max*, beside being the immediate antecedent and the attachment point of *he*, is also the attachment point of *himself*; instead, *he* is only the immediate antecedent of *himself*.



2. Coreference.

Two NPs are linked by a coreference relationship when they both refer to the same *designatum*. Coreference is a symmetrical and transitive relationship: if X is coreferential with Y, the latter is coreferential with

distinction between the notions of coreference and anaphora: if A and B are anaphorically related and so are B and C, A and C must be coreferentially, but not anaphorically related.

⁶ Two pro-forms, however, cannot have the same antecedent if they are "anaphorically disjoint" (cf. §. 3). This principle allows the blocking of coreferential reading in sentences like: (a) *Bill said that he saw him.

⁷ In this article, we will not be examining the specificity of what generativists have called anaphors, i.e. reflexive and reciprocal pronouns, insofar as what we are interested in clarifying are the general principles which regulate the relationship between antecedents and pro-forms.

⁸ An NP is autonomously referential if it is possible to identify its referent without its being linked to any other NP. When this is not possible, the NP is not autonomously referential.

X (symmetry); if X is coreferential with Y and Y is coreferential with Z, then X is also coreferential with Z (transitivity).

- (3) Max confessed that he only cares for himself.

himself is coreferential with *he*, which is coreferential with *Max*, then *himself* and *Max* are coreferential.

2.1. 'Direct' and 'indirect' coreference.

In (3) the coreference relations, even that between *himself* and *Max*, are mediated by anaphora.⁹ When the coreference is mediated by anaphora, we call this 'direct' coreference. More generally, we will say that two terms are directly coreferential if one of them, NP₁, is the antecedent of the other, NP₂:

- (4) I told him₁ that he₂ is quite a bad tennis player.
 (5) Every man₁ loves himself₂.
 (6) Max ran a boy₁ over yesterday. The boy₂ broke his right leg.

Two terms are also directly coreferential if neither of them is a pro-form, that is to say they are both autonomously referential:

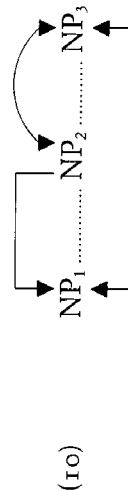
- (7) Oscar came in. Few minutes later, Oscar's wife called him and ...

Two terms — NP₂ and NP₃ — are indirectly coreferential, instead, if neither is the other's antecedent and if one of them, NP₃, is coreferential with a third term, NP₁, already occurred in the text and which is in turn (directly) coreferential with NP₂:

- (8) ... (Zelda₁) ... Those who know her₂ adore Zelda₃.

- (9) Max₁ said to himself₂ that he₃ should not do it.

In (8) *her* and *Zelda*, and in (9) *himself* and *he* may be indirectly coreferential, even if they are not anaphorically linked at all:



⁹ Corblin (1990), on the other hand, says that: "By coreference, is meant a relation between two (or more) expressions which do refer to the same particular, without being connected by means of the linguistically based relation of anaphora" (Corblin 1990: 83).

(the continuous line indicates anaphoric relationship,¹⁰ the dotted one indicates direct coreference, the curved one indicates indirect coreference).

3. Evidence.

Although the need for keeping the notions of coreference and anaphora separate has already been emphasized,¹¹ none of the current theories of anaphora makes any distinction between the (im)possibility or rather the (un)naturalness of coreferential reading, and the (im)possibility of anaphoric reading,¹² so that they look for only one generalization which can explain the (un)grammaticality of sentences as the following ones:

- (11) *He is an absolute dictator in Ben's office.

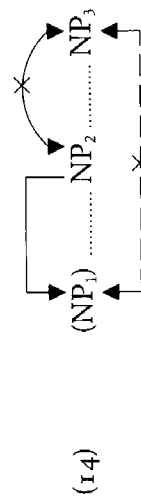
- (12) While she was eating, Liz felt a pain.

- (13) He had already shot himself before John quite knew what he was doing. (Bolinger 1980)

On the one hand, there are those who — like Kuno (1975) or Bolinger (1980) — deny the very existence of the backwards pronominalization, since they think that the backwards pronominalization is only, and always, 'apparent',¹³ and, on the other hand, there are those who — like Lasnik (1976), Reinhart (1976) and many others — state, instead, the real existence of the backwards pronominalization and don't make any distinction between real and apparent pronominalization.

In other words, in the sentences (11)-(13), according to the former there are the following relationships:

- i) sentence (11):



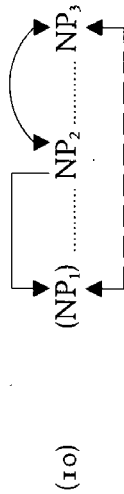
¹⁰ The term 'anaphora' is now used, and so will be later on, as an unmarked term to indicate both anaphoric relationships in a proper sense, and cataphoric ones.

¹¹ Cf. Postal (1971), Gross (1973), and above all Milner (1982).

¹² The one exception we know of is Evans (1980), who claims that it is necessary to distinguish between coreference and referential dependence. Reinhart (1983a, 1986), on the other hand, distinguishes between coreference and anaphora, but by anaphora she means just the relation between an antecedent and a pronoun which can be read as a bound variable. We will come back to this proposal of Reinhart later on; for a more detailed discussion of this proposal, see Musella (1990).

¹³ According to Kuno and Bolinger, in these cases the antecedent of the pronoun is not the full NP which is on its right: "Backwards Prenominalization requires that the referent of the pronoun be 'determinable or 'predictable' from the preceding context..." (Kuno 1975: 283).

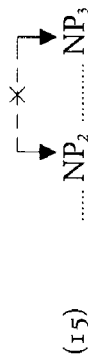
ii) sentences (12) and (13):



and what is important to make clear are the reasons which make it "necessary or desirable ... to reidentify the referent at a given point" (Bolinger 1980: 308).

According to the latter, however, the relationships are as follows:

iii) sentence (11):



(NP₂ and NP₃ are coreferentially disjoint and therefore cannot be coreferential, even 'accidentally' (= indirectly));

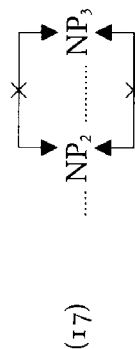
iv) sentences (12) and (13):



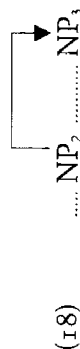
What must be clarified here are the syntactic conditions which allow a coreferential reading in one case, but not in the other.

In our opinion, things can be put in a different way:

a) in (11) the two NPs are anaphorically disjoint, that is to say that neither can *be* depend on *Ben* (for syntactic reasons that we'll see later on), nor can *Ben* depend on *be* (for obvious semantic reasons):



b) in (12) the pronoun can be read as anaphorically dependent on *Liz*, thus the correct analysis is the one formulated in (18):



The proof of the correctness of this analysis is given by a simple test which we'll call "indefiniteness test", which consists in substituting a definite NP with an indefinite NP. If even after this substitution, the coreferential reading is possible, then the two terms are in an anaphoric relationship, because, evidently, an indefinite NP may not have an antecedent:

(12) a. White she was eating, a girl felt a pain.

According to us, (12a) is acceptable.

c) the correct analysis of the sentence (13) is the one given in (10) as is clearly showed by the negative result of the test of the indefinite:

(13) a. *He had already shot himself before a boy quite knew what he was doing.

So, there are two facts which have to be explained:

- i) what makes the settling of the anaphoric relationship possible in (12) and impossible in (11) and (13)?
- ii) what makes the settling of the indirect coreferential relationship possible in (13) and impossible in (11)?

In this work, we will answer the first question only and not the second one which is a textual problem.

From now on, in order to avoid any confusion between an anaphoric and a coreferential reading, when discussing the (im)possibility of an anaphoric reading we shall write in capital letters the term which is (from the semantic point of view) the potential antecedent. So, in order to say that in a sentence such as (11) the anaphoric reading is impossible, the sentence will be rewritten like this:

(11') *He is an absolute dictator in BEN's office.

The asterisk indicates that *Ben* cannot be read as the antecedent for *he*.

4. Towards a new restriction.

In our opinion, while linear ordering is irrelevant in the definition of "domain", it nevertheless plays a very precise role in the restriction on anaphoric reading, also at intrasentential level. By using temporarily the notion of domain without defining it, we can thus formulate the general restriction on anaphora:

- (19) An NP, NP₂, can be read as anaphorically dependent on another NP, NP₁, iff:
- i) NP₁ has NP₂ in its domain; or
 - ii) NP₁ precedes, and is not in the domain of NP₂.

This restriction assumes that the linear order is relevant at a textual as well as an intrasentential level (continuity between text and sentence), unless one of the two terms is in the domain of the other (discontinuity between text and sentence). In the latter case only, the linear order is not relevant; where neither of the terms is in the domain of the other, on the other hand, linear order remains relevant:

(20) She came in. As soon as he saw Claire John left.

In (20) no anaphoric relation can occur between the two terms; obviously, this does not exclude the possibility of (indirect) coreference.

4.1. Grammatical domain.

In order to verify whether the restriction (19) is correct, we have to define the notion of "domain".

4.1.1. C-Command.

The most widely accepted definition of this notion is the one by Reinhart:

(21) The domain of the node A consists of all and only the nodes c-commanded by A. (Reinhart 1983b: 19).

There are several definitions of c(onsituent)-command. We will quote the simplified one formulated in Reinhart (1983b):

(22) Node A c-commands node B iff the branching node most immediately dominating A also dominates B. (p. 18)

Besides the empirical problems (acknowledged by Reinhart herself and to which we will return), a restriction based on c-command runs into thorny theoretical problems. Such a restriction is symmetrical: a given NP can simultaneously c-command and be c-commanded by another NP. The anaphoric relationship, instead, is asymmetric. I think this is enough to consider the c-command theoretically inadequate.¹⁴

4.2. The 'relational' domain.

Our hypothesis is that the notion of c-command works because, and solely to the extent to which, it reflects the relationships between the

¹⁴ Higginbotham (1983) and (1985), who also underlined the need to substitute the numerical indexing with a system of linking (like that adopted by us in this work), still considered the notion of c-command valid.

grammatical functions or grammatical relations (GRs). This hypothesis rests on the existence of a hierarchy of GRs or R(elational) H(ierarchy). The existence of a RH is not an *ad hoc* hypothesis of ours: it has been independently and, I believe, satisfactorily motivated.¹⁵

There is substantial agreement upon what should be the content of such a hierarchy:¹⁶

- (23) Subject (Sub),
- Direct Object (DO),
- Indirect Object (IO),
- Prepositional Phrase (PP),
- Genitive (Gen).

The RH is based on the assumption that the various arguments have a different structural degree of topicality, depending on what GR they have in the sentence.

The argument having the GR of Sub, therefore, is the one with the highest structural degree of topicality, that is to say the topic of an unmarked sentence; the DO has a lower structural degree of topicality than Sub, but higher than all the other GRs, and so on, up to the Gen, which is the GR with the lowest structural degree of topicality.

The restriction based on GRs assumes that:

(24) An argument A can serve as the antecedent of an argument B in the same sentence, iff B does not have a higher structural degree of topicality than A.

We can now re-formulate the restriction (22):

- (25) An NP₁, NP₂, can be read as anaphorically dependent on another NP, NP₁, iff:
 - i) NP₁ is relationally higher than NP₂; or
 - ii) NP₁ precedes and is not relationally lower than NP₂.

¹⁵ Keenan and Comrie (1977), for example, suppose the existence of what they call 'accessibility hierarchy' in order to explain the various possibilities that NPs have to be relativized.

¹⁶ In addition to the above quoted Keenan and Comrie (1977), see, among others, Antinucci (1977), Lappin (1982), and Berretta (1990). While there is a substantial agreement on the content of the hierarchy, the terminology adopted varies quite a lot: Antinucci (who does not distinguish between IO and PP) talks of 'Subject', 'Object', 'Dative', and 'Genitive'; Lappin (who does not include Gen.) talks of 'Subject', 'Direct Object', 'Indirect Object', and 'Oblique Argument'; Keenan and Comrie talk of 'Subject', 'Direct Object', 'Indirect Object', 'Oblique NP', 'Gen.', and 'Object of comparison' (which we will not consider); finally Berretta talks of 'Subject', 'Object and Dative', 'Genitive', and 'Oblique'.

5. Checking the restriction based on GRs.

5.1. One of the two arguments is the Sub.

Certainly this is the least controversial case: the NP-Sub is higher than any other node dominated by S(entence); for that reason, whatever its position in the surface linear order of the sentence is, it cannot serve as a pro-form but only as an antecedent. Some examples will help assessing the degree of correctness of this prediction.

5.1.1. ... the other term is a preposed or topicalised phrase.

The 'relational' hypothesis is based on the premise that topicalisations, PP preposing and, more generally, so-called *wh-movements* cause no change in the anaphoric options, if both terms belong to the same sentence and one is relationally higher than the other. One of the arguments used by Reinhart to reject a 'non-surface' approach is based just on the denial of this assumption: as a matter of fact, Reinhart maintains that in the case of topicalisations, as well as sentential PP preposing, "coreference [= anaphoric] options do change" (Reinhart 1983b: 103).

First of all, let us analyse a case of topicalisation:

- (26) a. JOHN esteems only himself.
b. Only himself, JOHN esteems.

According to Reinhart, a sentence such as (b) should not be acceptable, because "the pronoun is dominated directly by S. Hence it c-commands the antecedent" (Reinhart 1983b: 83). Actually, sentence (26 b) is acceptable. This means that the topicalization has not changed the coreference options at all.

Let us analyse a case of sentential PP preposing:

- (27) a. JOHN will end up in hospital because of his new job.
b. Because of his new job, JOHN will end up in hospital.
c. *He will end up in hospital because of JOHN's new job.
d. *Because of JOHN's new job, he will end up in hospital.

It is worth remembering that Reinhart's restriction should predict the acceptability of both sentences with PP preposing — (b) and (d) — since the two phrases are in two different syntactic domains.¹⁷ We believe, however, that only one of the two sentences above, namely (27 b), is acceptable, and it is acceptable because the pronoun is in the syntactic domain of the full NP. The correctness of our explanation is provided by

¹⁷ According to Reinhart, a sentential NP in final position is directly dominated by node S, whereas in the initial position it is directly dominated by node E.

a remark by Reinhart herself, which shows that when a pro-form is in the domain of its antecedent, the so called 'sloppy-identity interpretation'¹⁸ can occur. Indeed, in (27 b) this reading is possible:

- (27) e. Because of his new job, JOHN will end up in hospital, and so will FELIX.

The most obvious interpretation of the second conjunct of (27 e) is actually the one stating that:

- (27) f. ... FELIX will end up in hospital, because of Felix's new job.

This means that also preposed sentential PPs are in the domain of Sub. Therefore Reinhart's argument against the foundations of a non-surface restriction based on GRs is nullified: topicalisation does not change the possibilities of anaphoric reading, when either term is in the relational domain of the other.

5.2. One term is in the VP, the other one is a sentential PP.

- (28) a. Claire kissed MAX in front of his mother.
b. *Claire kissed him in front of MAX's mother.

Since neither DO c-commands the sentential PP, nor the latter c-commands the former, the two sequences are, in Reinhart's view, equally acceptable. In (28 b), instead, between the two terms there is no anaphora relationship, as the indefiniteness test shows:

- (28) c. *Claire kissed him in front of MY FRIEND's mother.

The same applies to all the other phrases dominated by the VP node:

- (29) a. Claire talked to HER DAUGHTER's husband in front of her.
b. *In front of her, Claire talked to HER DAUGHTER's husband.

Unlike Sub, therefore, the DO and the other verb-phrasal PPs do not have the sentential PPs in their domain.

5.3. The two terms are dominated by the VP node

One of the weak points of the Reinhart's restriction is, as Reinhart herself admits, the handling of "anaphora options of the NPs inside the VP" (Reinhart 1976: 155). It seems that a restriction such as (25) allows us to deal adequately also with the cases which would pose some problems for Reinhart.

¹⁸ The relation between a so-called 'pronoun of laziness' and its antecedent is called 'sloppy-identity'.

5-3.1. Direct Object and Prepositional Phrase.

- (30) a. Luis managed to bring MARY back to her senses.
 b. *Luis managed to bring her senses back to MARY.
 c. *Luis managed to bring back to MARY her senses.
 d. Luis managed to bring back to her senses MARY.

Since the verb-phrasal PP is in the domain of the DO, the linear order is not relevant: the antecedent must be the DO. Restriction (25) provides both for acceptability of sentences (30 a) and (30 d), and for the unacceptability of (30 b) and (30 c).

5-3.2. The terms are both PPs.

- (31) a. Claire talked a long time to JOHN about him.
 b. *Claire talked a long time about him to JOHN.
 c. *Claire talked a long time to him about JOHN.
 d. Claire talked a long time about JOHN to him.

The restriction formulated by Reinhart fails to explain the unacceptability of sentences (31 b) and (31 c): the two PPs are in different domains and therefore it should be possible to read them as anaphorically linked. Our restriction, however, succeeds in explaining the (un)acceptability of all these sentences: since neither PP is higher than the other, the identification of the antecedent and pro-form depends on the argument's relative position: the potential antecedent must precede the pro-form (condition (25 ii)). For this reason, in the sentences (31 b) and (31 c) the full NP cannot be read as the pronoun's antecedent.

5-4. One of the two terms is Gen

- (32) a. JOHN's father loves him.
 b. Him, JOHN's father loves.
 (33) His father loves JOHN.

Reinhart (1983b) and Chomsky (1985) claim that (32 a) and (33) are acceptable, while sentence (32 b) is unacceptable. If this were correct, it would prove that, at least in this case, topicalisation does change the anaphora options. However, if there can be no doubt as to the unacceptability of the anaphoric reading in (32 b), there are good reasons for believing that in (32 a) and (33) too the anaphoric reading is unacceptable: these sentences cannot be interpreted if we use them as an answer to a question such as the following: "What can you tell me about John's father?"

Indeed, in (32 a) the coreferential reading is possible only if the sentence serves as an answer to such a question as: "What can you tell me about John?", on account of the pronoun having as its antecedent the proper noun which is present in the question. Between the two terms, therefore, there is a coreference relation, but it is an indirect one.

According to (23), the DO is higher than the Gen, so that it would seem possible to say that (32 a) and (32 b) are unacceptable because the Gen cannot serve as an antecedent for a DO even when preceding it. But if sentences (32) are unacceptable for that reason, then a sentence like (33) should be considered as absolutely acceptable, which it is not. Sentence (33), then, is a counterexample to restriction (25). Moreover, its unacceptability also ruins the explanation for the unacceptability of sentences (32), for which a new one is now needed.

Our hypothesis holds that, in order to account for the unacceptability of (32) and (33), it is above all necessary to revise the hierarchy (23). More precisely, what is to be revised is the position occupied in it by the Gen. This GR differs from all the others as it is an adnominal complement. This implies that the Gen cannot be directly compared with the remaining GRs: a NP-Gen is in the domain of the arguments which are higher than its head, but it does not have in its domain the arguments which are in the domain of its head:

- (32) a. JOHN's father loves him.

- (34) JOHN loves his father.

in (34) *his* is in the subject's domain, in (32 a), *him* is in the domain of *John's father*, but not in that of *John*.

But, then, what makes (32) and (33) unacceptable? In (33) the pronoun precedes the full NP, so it cannot be the antecedent; in a sentence such as (32), though, *him* doesn't see *John's*, it can only see *John's father*. On the contrary, an adnominal complement can see the nodes which are not lower than its phrase:

- (35) I met JOHN at his parents-in-law last year.

In this sentence the pro-form (which is not in the domain of the DO) sees the DO and can therefore pronominalize it.

5-5. Problems.

Actually, there are some facts which our analysis fails to capture. In English, as Reinhart pointed out, "the object of an experiencing verb can in many cases control a pronoun in the subject" (Reinhart 1983b: 179):

- (36) Jokes about his wife upset MAX. (Reinhart 1983b)

- (37) His wife disturbs MAX.

Both in (36) and (37) *Max* can be read as the antecedent of *bis*, although it is relationally lower than *bis*.

There are counterexamples in Italian too:

(38) *Il proprio modo di parlare non piace a MAX.*
(*Max does not like his way of speaking.*)

(39) *Ad UMBERTO I successe suo figlio.*
(*His son succeeded Umberto I.*)

In these sentences, the subject anaphorizes a relationally lower argument. It is well known that in some cases the DO and, even more often, the IO have a higher structural degree of topicality than the Sub. Consequently, rather than being a counterexample to such a restriction as (25), these sentences are well-known exceptions to the underlying RH, and pose no specific problem for the anaphoric theory.

6. 'Non-definite' antecedents.

We have dealt so far with restrictions on 'definite anaphora', that is the anaphora of definite descriptions and proper nouns. However, such restrictions are not sufficient to predict the behaviour of 'non-definite' antecedents, that is of indefinite or quantified descriptions, and indefinite and interrogative pronouns. It is well-known, in fact, that in many cases in which it is possible to read a definite NP and a pro-form as coreferentially linked, the substitution of the definite NP with a non-definite one can render the coreferential reading unacceptable:

(40) a. *MAX* is loved by his parents.

b. *EVERY BOY* is loved by his parents.

(41) a. The secretary who works for him despises Stegfried.

b. *The secretary who works for him despises each of the managers.

(42) a. Claire did not answer him, when he said hallo to her.

b. *Claire did not answer anybody, when he said hallo to her.

(43) a. ?*MAX* said the theft had been committed by Max.

b. **A BOY* said the theft had been committed by a boy.

c. **A BOY* said the theft had been committed by the boy.

While in (40) replacing a proper noun with a quantified NP does not change the possibilities of anaphoric and coreferential reading, in sentences (41) and (42) the possibilities of anaphora and/or coreference do change: in (41 a) and (42 a) the two NPs can be read as coreferential, whereas in (41 b) and (42 b) this is not possible. Correspondingly, in (43 a) the

coreferential reading is awkward but possible,¹⁹ whereas in (43 b) and (43 c) it is absolutely unacceptable.²⁰

Let us begin our analysis with sentences as (43 b) e (43 c), i.e. the chain "non-definite NP full NP".

6.1. "Non-definite NP full NP".

6.1.1. "Non-definite NP non-definite NP".

(44) a. **A BOY* said the theft had been committed by a boy.

(45) a. **NO BOY* said the theft had been committed by any boy.

(46) a. **SOMEONE* said the theft had been committed by someone.

(47) a. **WHO* said the theft had been committed by whom?

All these sentences are unacceptable because two NPs, both non-definite, cannot be anaphorically linked: anaphora relationship may occur only if there is a pro-form, and a non-definite NP cannot serve as a pro-form. That the reason for its unacceptability lies here rather than where indicated by Reinhart ('a non-pronoun is in the domain of its potential antecedent') is shown by the fact that, even when the two non-definite NPs are in two different domains, anaphoric reading is unacceptable:

(44) b. **A BOY* confessed: the theft had been committed by a boy.

(45) b. **NO BOY* confessed: the theft had been committed by any boy.

(46) b. **SOMEONE* confessed: the theft had been committed by someone.

(47) b. **WHO* confessed? Whom the theft was committed by?

¹⁹ This contrasts with Principle C of the Binding Theory. This principle says that a R(efential)-Expression must always be free, i.e. must not be in the domain of its antecedent.

On this ground, a sentence like (43 a) is ungrammatical as well as a sentence like:

(a) He said the theft had been committed by Max.

The weakest version of Lasnik (1989), according to which an R-expression is pronoun free' might explain the acceptability of the coreferential reading in (43 a) and the unacceptability of the anaphoric reading in (a), but it would not, however, explain the unacceptability of the anaphoric and coreferential reading in (43 b) and (43 c) or even in (44 b)-(47 b).

²⁰ Reinhart (1976) hypothesizes that non-definite NPs (i.e. indefinite NPs, traces, and focus NPs), besides obeying the general conditions on anaphora, also obey the following stricter one: "Non-definite NPs ... can have anaphoric relationship only with NPs in their syntactic domain" (Reinhart 1976: 126).

What is not clear is what indefinite NPs, traces and focus NPs have in common and why exactly these should obey a further restriction.

6.1.1.2. "Non-definite NP definite NP".

Explaining why sentences (44)-(47) are unacceptable, however, does not account for the impossibility of interpreting the two NPs in the following sentences as being anaphorically or even only coreferentially linked:

- (44) c. *A BOY said the theft had been committed by the boy.
- (45) c. *NO BOY said the theft had been committed by the boy.
- (46) c. *EVERY BOY said the theft had been committed by the boy.
- (47) c. *WHICH BOY said the theft had been committed by the boy?

These sentences are unacceptable because the definite description anaphorizes a non-definite one before the latter 'has established a referent'. 'Establishing a referent' simply means introducing in the discourse a new referent, and assigning some predicate to it.²¹ The attachment point of a definite description, then, is not the mere indefinite description, but the indefinite description 'enriched' with the predicate assigned to it in the clause to which it belongs. Accordingly, in

- (48) Max ate AN APPLE. The apple was very good.

the *designatum* of the NP *the apple* is identified by the description "apple which Max ate". We shall call "identifying definite description" the description allowing the identification of the *designatum*.²²

The reason why an indefinite description cannot be anaphorized by a definite one belonging to the same sentence becomes obvious: the definite description would be applied when the indefinite description is yet to establish the referent. If in (44 a) we replaced the definite description with its identifying description, we should run into a *regressio ad infinitum*:

- (49) *A boy said the theft had been committed by the boy who said the theft had been committed by the boy who ...

²¹ Karttunen (1969) has shown that an indefinite NP does not always establish a referent. This does not happen, for instance, in a negative utterance such as: "Bill doesn't have a car", which cannot be followed by a sentence such as "It's black".

Conte (1988) has shown that 'negativity' is neither necessary nor sufficient for a referent not to be established, as exemplified in the following mini-passages:

(a) Strauss fingeva di possedere un aeroplano. *L'aeroplano è attrattato venti minuti fa. (Strauss pretended he owned an aeroplane. *The aeroplane landed twenty minutes ago).
 (b) Il terrorista non nascose una ferita al petto. La ferita era stata riportata...

[The terrorist did not conceal his chest wound. The wound had been caused...]
²² The identifying definite description can replace the corresponding 'pure' definite description, without adding any new information to the sentence.

6.2. "Non-definite NP pronoun".

In analysing the pronominalisation of non-definite NPs, we will distinguish between 'empty' NPs, that is negatively quantified NPs, and 'non-empty' NPs.²³

6.2.1. "Non-empty NP pronoun".

- (50) FEW CONGRESSMEN admire Kennedy, and they are very junior. (Evans 1980)
- (51) Just ONE MAN drank champagne and he was ill. (Evans 1980)
- (44) d. A BOY confessed the theft had been committed by him.
- (46) d. SOMEONE confessed the theft had been committed by him.
- (47) d. WHO said the theft had been committed by him?

Non-empty NPs pose no problem: they may be anaphorized both by pronouns which are not in their domain — (50) and (51)²⁴ — and by pronouns which are in their domain — (44 d), (46 d), and (47 d).

6.2.2. "Empty NP pronoun".

- (52) *No congressmen admire Kennedy, and they are very junior. (Evans 1980)
- (53) a. *No boy confessed the theft. The boy/he said that ...
- (54) a. *No injured soldier was awarded any medal. Since the soldier/they were unable to move, the minister went to the hospital.

These sentences are not interpretable: empty NPs cannot be anaphorized by NPs outside their domain. They can be anaphorized only by a pronoun within their domain:

- (55) NO BOY confessed the theft had been committed by him.
- (56) NOBODY forced Lucia to give him a kiss.
- (57) NOBODY was talking to his son's friend.

²³ "Empty NPs" are opposed to "non-empty NPs" not to "full NPs"; "Full NPs" are opposed to "pronouns".

²⁴ These sentences prove that Reinhart's hypothesis, according to which "Non-definite NPs ... can have anaphoric relationship only with NPs in their syntactic domain" (Reinhart 1976: 126), is wrong.

linked. In other words, no indirect coreference relation can occur between a non-definite NP and a pro-form. As a result the different relations existing between (a) and (b) in (40)-(42) can now be accounted for:

- (40) a. MAX is loved by his parents.
 b. EVERY BOY is loved by his parents.
- (41) a. The secretary who works for him despises Siegfried.
 b. *The secretary who works for him despises each of the managers.
- (42) a. Claire did not answer him, when he said hallo to her.
 b. *Claire did not answer anybody, when he said hallo to her.

In sentences (40 a-b), the possibility of coreferential reading is unchanged, as in both sentences there is an anaphora relation and in (b) the pronoun, being in the domain of the antecedent, can be read as a bound variable; in (42 a), on the other hand, the two NPs are anaphorically linked, but *he* is not in the domain of the antecedent, and therefore in (42 b) it cannot anaphorize an empty non-definite NP; finally in (41 b) coreferential reading is impossible, because the two NPs are not anaphorically linked: although they are coreferential in (41 a), they are only indirectly so.

7. Summary.

To summarize, we have seen that a theory of nominal anaphora must first distinguish between the (im)possibility of anaphora reading and (un)naturalness of coreferential reading. Then we have tried to show that the (im)possibility of interpreting two NPs as anaphorically linked depends basically on:

- i) their role in the Relational Hierarchy;
- ii) their position in the surface linear order.

According to us, this generalization is adequate not only from the descriptive point of view,²⁷ but also from the explanatory one (it can be easily translated into intuitive and informal terms): it is easy to accept the idea that an argument with GR of Sub, that is to say the most 'important' element of the sentence, could serve as the antecedent of a less 'important' argument, that is having a lower GR, for example a PP. Equally obvious is the generalization for which an argument 'coming first' can serve as antecedent of another one 'coming after'. However, it is not always as simple as that; the reason for many complications lies in the fact that the principles mentioned might come into conflict with each other.

²⁷ Of course, much evidence, specific to single languages, has not been examined, but this work aimed at outlining an anaphoric theory and, at that stage, not being exhaustive is certainly a limit, but not a defect.

As for anaphoric relations in complex sentences, we refer to what we have said in Musella (1990).

Therefore, while a non-definite and non-empty NP can be anaphorized by a pronoun anyway, a non-definite and empty NP can be anaphorized only by a pronoun within its domain.

6.3. Syntactic and referential use of pro-forms.

Pronouns can be used both referentially and non-referentially:

- (58) FRED said that he would leave the next day.
 (59) WHO said that he would leave the next day?²⁵

In sentence (58), *he* can be exchanged for *Fred*, that is a referential NP, and the two sentences will not differ in their truth conditions; on the contrary, in (59) *he* cannot be exchanged for any referential NP. In this case "there is a purely syntactic relation that links the pronoun to the syntactic position of its antecedent" (Bosch 1983: 41).

Our hypothesis assumes that the referential or purely syntactic interpretation of a pronoun depends on:

- i) its position,
- ii) the semantic nature of the antecedent.

A pronoun is to be read as a purely syntactic device, that is as a bound variable, when it is in the domain of its antecedent, and the latter is a non-definite NP.²⁶

6.4. Non-definite coreference.

What we have said leads us to conclude that a non-definite NP can be read as coreferential with a pro-form iff the two NPs are anaphorically

²⁵ The indistinctness between coreference and anaphora has weighed, maybe heavily, on the study of anaphoric and coreferential behaviour of interrogative pronouns. Actually, ignoring that in a sentence like:

(a) His mother loves John.

the two terms can certainly be coreferential but only indirectly, has led to the non realisation that a sentence such as:

(b) Who his mother loves?

cannot have a direct coreference reading for much the same reason which forbids anaphoric interpretation in (a). Hence, the search for specific restrictions on interrogative anaphora.

²⁶ Both Bosch (1983) and Reinhart (1983a, 1983b, 1986, and forthcoming), on the other hand, hold that the nature of the antecedent is irrelevant: "all pronouns can be interpreted as bound variables, regardless of whether the antecedent is quantified NP or not" (Reinhart 1983a: 62). It follows from this statement that the pronouns in the domain of their antecedent "are always ambiguous between the bound-variable interpretation and the referential interpretation" (Reinhart 1983b: 156). But, in our opinion, a sentence like "LUC trusts in himself" is not at all ambiguous.

A sentence like "Only LUC trusts in himself" is ambiguous; but, as is well known, a NP modified by *only* is a quantified NP.

Next, we have analysed the specificity of non-definite anaphora. We have divided non-definite NPs in two sub-categories: empty NPs (negatively quantified NPs) and non-empty NPs. We have shown that the latter have the same anaphoric possibilities of definite NPs, and that their more restricted coreferential possibilities depend on the fact that in a chain like

(60) (NP₁) NP₂ (= pronoun) NP₃ (= non-empty, non-definite NP)

NP₃ may not be indirectly coreferential with NP₂ because, obviously, NP₁ may not be its antecedent.

Finally, we have shown that the restricted anaphoric possibilities of empty NPs have a very intuitive cause: empty NPs don't establish a referent, so they may not be anaphorized by a referential pro-form.

Address of the Author:

Via F. Cileia, 26 — 80127 Napoli, Italia

Références

- Antinucci, F. (1977), *Fondamenti di una teoria tipologica del linguaggio*, Bologna, Il Mulino.
- Berretta, M. (1990), "Catene anaforiche in prospettiva funzionale: antecedenti difficili", *Rivista di Linguistica*, 2: 91-120.
- Bolinger, D. (1980), "Pronouns in Discourse", in T. Givón, (ed.), *Discourse and Syntax* (= *Syntax and Semantics*, 12), New York, Academic Press: 289-309.
- Bosch, P. (1983), *Agreement and Anaphora*, London, Academic Press.
- Chomsky, N. (1985), *Knowledge of Language. Its Nature, Origin, and Use*, New York, Praeger.
- Corblin, F. (1990), "Anaphoric and Referential Chains in Discourse", *Rivista di Linguistica*, 2: 67-89.
- Conte, M.-E. (1988), *Condizioni di coerenza. Ricerche di linguistica testuale*, Firenze, La Nuova Italia Editrice.
- Evans, G. (1980), "Pronouns", *Linguistic Inquiry*, 11: 337-61.
- Gross, M. (1973), "On Grammatical Reference", in F. Kiefer & N. Ruwet (eds.), *Generative Grammar in Europe*, Dordrecht, Reidel: 203-17.
- Higginbotham, J. (1983), "Logical Form, Binding, and Nominals", *Linguistic Inquiry*, 14: 395-420.
- Higginbotham, J. (1985), "On Semantics", *Linguistic Inquiry*, 16: 547-93.
- Karttunen, L. (1969), "Discourse Referents", Preprint No 70, International Conference on Computational Linguistics (Coling), *Sänga-Säby/Stockholm*; tr. it. in M.-E. Conte (ed.), *La linguistica testuale*, Milano, Feltrinelli, 1977: 121-147.
- Keenan, E. & Comrie, B. (1977), "Noun Phrase Accessibility and Universal Grammar", *Linguistic Inquiry*, 8: 63-99.
- Kuno, S. (1975), "Three Perspectives in the Functional Approach to Syntax", in J. Kreiman & A. Ojeda (eds.), *Papers from the Parasession on Functionalism of the Chicago Linguistic Society*, Chicago, University of Chicago: 276-336.
- Lappin, S. (1985), "Pronominal Binding and Coreference", *Theoretical Linguistics*, 12: 241-63.
- Lasnik, H. (1976), "Remarks on Coreference", *Linguistic Analysis*, 2: 1-22.

- Lasnik, H. (1989), "On the Necessity of Binding Conditions", in H. Lasnik (ed.), *Essays on Anaphora*, Dordrecht, Kluwer Academic Publishers: 149-167.
- Milner, J. C. (1982), *Ordres et raisons de langue*, Paris, Editions du Seuil.
- Musella, R. (1990), *Coreferenza e anafora*, Doctoral Dissertation, University 'La Sapienza', Rome.
- Postal, P. (1971), *Cross-over Phenomena*, New York, Holt, Rinehart & Winston.
- Reinhart, T. (1976), *The Syntactic Domain of Anaphora*, Doctoral Dissertation, MIT, (Cambridge (Mass.)).
- Reinhart, T. (1981), "Definite NP Anaphora and C-Command Domains", *Linguistics Inquiry*, 12: 605-635.
- Reinhart, T. (1983a), "Coreference and Bound Anaphora: A Restatement of the Anaphora Questions", *Linguistics and Philosophy*, 6: 47-88.
- Reinhart, T. (1983b), *Anaphora and Semantic Interpretation*, London, Croom Helm.
- Reinhart, T. (1986), "Center and Periphery in the Grammar of Anaphora", in B. Lust (ed.), *Studies in the Acquisition of Anaphora*, vol. 1: *Defining the Constraints*, Dordrecht, D. Reidel Publishing Company: 123-50.
- Reinhart, T. & Reuland, E. (forthcoming), "Anaphors and Logophors: an Argument Structure Perspective", in J. Koster & E. Reuland (eds.), *Long-Distance Anaphora*, Cambridge (Mass.), Cambridge University Press.
- Wasow, T. (1979), *Anaphora in Generative Grammar*, Ghent, E. Story Scientia.