

# The Proper Treatment of Coreference Relations

Louis-H. Desouvrey

(Revised July 2006)

A novel approach to coreference relations is proposed. It is shown that there are no coreference principles per se in the grammar. Rather three constraints independently needed account for this phenomenon: the Obligatory Contour Principle (OCP), the Avoid Ambiguity Constraint (AAC), and the Freedom Constraint. The OCP and the AAC deal with features lexical elements are specified for. Referring elements are indeed distinguished from each other by a relational feature, which represents what the element stands for in the real world. Given nonlinear type representations whereby each feature belongs to its own plane, R-features spread from a referential expression on an adjacent referring element, either a pronoun or an anaphor. The ban on line crossing in the representation, which constrains the spreading, accounts for the adjacency between anaphors and their antecedents. The complementarity in the distribution of anaphors and pronouns follows from the feature specification of these elements as well as the interaction of the OCP and the Ambiguity Constraint.

**Keywords:** coreference, constraints, domains, pronouns, anaphors, syntactic features.

## 1. Introduction

This article explores a novel approach to possible coreference relations between referring elements in syntactic structures. It deals with the set of facts that traditionally rely on the Binding Theory, as developed by Chomsky (1981, etc.) and many others (for instance see Harbert 1995, from which most of the data come from, for a still good survey of the literature). As is well-known, current Binding Theory settles the problem of possible coreference between the three types of referring elements, namely referential expressions, pronouns and anaphors, with three conditions, one for each referring element, and other assumptions peculiar to the Principles and Parameters Theory, for instance the c-command relation, A and A-bar positions, category labels, etc.

On conceptual grounds, the standard approach on coreference has too many conditions (but see Safir 2004); since there are three referring elements, one would like to have at most two conditions, the third referring element would be left unspecified for whatever condition in a way similar to underspecification theory (e.g. Archangeli, 1984).

Furthermore, a general problem with such an approach is that it presents a number of complications due in part to its use of several functional heads, its ignoring intrinsic properties of lexical items and the inviolable rigidity of the grammatical principles it makes use of.<sup>1</sup> Theories that dispense with such apparatus can in principle capture coreference relations in a simplest way. It appears that the theory of features and constraint interaction developed in Desouvrey (2000) possesses the power and the flexibility for providing a fine-grained analysis of coreference relations. Assuming this theory is a model of language acquisition by infants, I will show that possible coreference relations between referring elements, as well as the form that the latter may take in the syntactic structure (i.e. pronominal or anaphoric), mostly follow from the nonlinear representation which is being used conjointly with a few constraints, independently needed elsewhere in the grammar of English and other languages.

Specifically, I wish to show that the distribution of referring elements is governed by the interaction of the constraints stated in (1), (2), and (3), of which (1) and (2) are violable, as in Optimality Theory (e.g. Prince and Smolensky 2002, etc.). These constraints are not extrinsically ranked; they apply whenever possible and certain operations may take place to rescue a constraint violation. Constraint (1) deals actually with anaphors and pronouns; referential expressions need not be mentioned in the determination of coreference, as I will show. In fact, the distribution of referential expressions falls under constraint (2), which has considerable interest in nonlinear phonology and morphology (cf. Goldsmith 1976, McCarthy 1986) and in syntax as well (cf. Desouvrey 2000, 2001a, 2001b, 2005, etc.). As such it comes at no cost to this analysis, that is, it need not be reckoned in the evaluation of the complexity of the proposal. Similarly, the constraint against ambiguity is not formulated for the purpose of the present article; it plays, in effect, a fundamental role in the syntax of operators in both English and French (Desouvrey 2001a). For these constraints to work, I shall take a fresh look at syntactic structures, as well as of the nature of pronouns and anaphors.

---

<sup>1</sup> Standard Binding Theory is often challenged by a number of scholars working in current generative theory. The last attempt I know of is Safir (2004) who comes up with the Form-to-Interpretation Principle (FTIP) in order to derive complementarity in the distribution of anaphors and pronouns. FTIP is intended to replace both Binding Theory principles A and B in a general framework where the notion competition between forms plays a crucial role.

- (1) Freedom Constraint (FC)  
Referring elements must be free in their minimal domain.
- (2) Obligatory Contour Principle (OCP)  
Two elements bearing identical R-features are banned in the same syntactic domain.
- (3) Avoid Ambiguity Constraint (AAC)  
Morphological ambiguity must be avoided whenever possible.

The paper is organized as follows. The next section introduces basic assumptions on syntactic representations and an account of some core facts. Section 3 discusses the case of *wh*-elements and other quantified expressions, and section 4 presents additional facts regarding complementation structures and relative clauses. Section 5 briefly discusses the results and concludes the article.

## **2. The basics of coreference**

I will show that the basic relation between referring elements is that of 'antecedent of,' similar in some respect to the use of this term in Linking Theory (e.g. Higginbotham 1983; also see Safir 2004). This relation is asymmetric: if A is the antecedent of B, B is not the antecedent of A. The 'antecedent of' relation will be shown to follow from the definition of each referring element and it is made obvious by the use of tiered representations, as in nonlinear phonology. In the next subsection I introduce the definitions of referring elements and the relevant part of their feature tree structure. In section 2.2 I discuss basic facts of coreference in the light of the definition of syntactic domain. Finally in section 2.3 I consider the interaction of the constraints and the nonlinear representation.

### **2.1 Referring elements and their representations**

As mentioned above, referring elements include syntactic objects like referential expressions (nouns / NPs), pronouns and anaphors. Every theory of coreference should include a definition of these objects prior any attempt to account for their distribution. I argue for the following definitions of the referring elements relevant in this article:

- (4) Definitions of the referring elements
- a. A pronoun (is a lexical item that) refers only to a referential expression.
  - b. An anaphor refers to a referential expression or a pronoun.
  - c. A referential expression refers to an entity in the real world.

The definition of anaphors is obvious, since we know that their antecedent may be a noun or a pronoun (cf. *Mary/she likes herself*). The definition of referential expressions may not be controversial either, since it is consistent with the current view in Binding Theory. However, the present definition of pronouns clashes with traditional conceptions of these elements. Current theories adopt the view that pronouns refer to nouns (or NPs). They also assume that pronouns refer to other pronouns. This latter view is incorrect, however. I argue that pronouns refer only to NPs, but not to other pronouns.

In Binding Theory, it is assumed that in a sentence like *he<sub>i</sub> saw Mary yesterday. She kissed him<sub>i</sub>*, both masculine pronouns may refer to the same individual, and thus they are bound to each other. That is, *him* refers to *he* and vice versa. In fact, once one carefully considers the sentences, it appears that these pronouns refer to the same referential expression, which is known in the context of the discourse, but this does not entail that one is the antecedent of the other. To see this, one may try the following reasoning. First, let us assume that use of a pronoun is felicitous in a structure only if its referent is known; otherwise the sentence is not interpretable. Regarding the above sentences, if a referential expression is not previously mentioned in the discourse, are they interpretable? Obviously no, for none of the pronouns may referentially license the other. Therefore, it is clear that the subject pronoun do not refer to the object one. On this view, we can conclude that a language like English includes among its lexical items the element 'pro-noun' or more accurately 'pro-NP', but not 'pro-pronouns', or 'pro-pro-NPs', which would be the case if a pronoun could refer to another pronoun.

Nothing in these definitions states how a referring element comes to refer to another one. This situation must be dealt with by adopting some appropriate mechanism. In current Binding Theory, as is well-known, indexes are assigned to every referring element. And relations of coreference or disjoint reference between two given elements depend on whether they have the same index or not. In Safir (2004), relying on work by

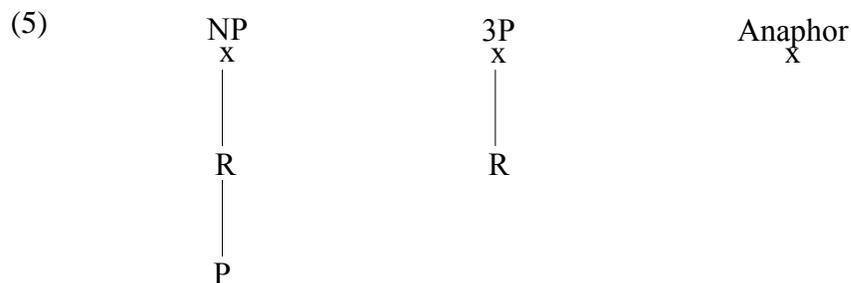
Higginbotham, indexes are abandoned for an arrow which links an element to its antecedent. However, under the assumption that the behavior of grammatical entities is a function of the features they are specified for (in my view syntactic features being similar to genes in biological systems), the difference between referring elements, as defined in (4), must be due to their 'genetic' composition. I wish to suggest that each referring element in a syntactic structure either bears or comes to acquire a referential feature, or R-feature, which is intended to translate what the element stands for in the real world. R-features will be named after capital letters P, Q, and so on.<sup>2</sup>

Now we need a structure to integrate R-features and it must allow us to capture their interaction. It happens that nonlinear phonological representations perfectly respond to this requirement. Suppose that features and morphemes bearing them appear on different tiers and are linked to one another by an association line, making up a plane in the sense of Euclid. So every identical feature will be in the same plane and may not interact with a feature in another plane (see below for further clarification). With this in mind, consider now the above definitions. An NP need not refer to anything in a syntactic structure and we may posit that it has an R-feature of its own, that is, its referent, in the normal case, is in the real world. Since both anaphors and pronouns refer to a syntactic entity, as opposed to a real world entity, they may not have an R-feature of their own. Furthermore, since anaphors may refer either to an NP or a pronoun, they must be distinguished from pronouns in some way. Suppose indeed that R-features are dependent of an abstract referential node, or R-node, which is a class node as currently used in nonlinear phonology. The R-feature and the R-node make it possible to structurally represent the definition of the three referring elements: a referential expression has an R-feature and an R-node, a (third person) pronoun has an R-node, but no R-feature, while an anaphor has neither an R-feature nor an R-node. This is

---

<sup>2</sup> It is not without interest to compare syntactic features and parameters, as used in the Principles and Parameters Theory. Features, in my view, are minimal units in syntax, as in phonology. Some of them may depend on inventory. In any event, they are not constraints of the grammar, but the material over which constraints apply. A parameter, on the other hand, is the minimal explanation for some grammatical phenomenon, hence a principle. To take a concrete example, the fact that wh-elements in certain languages appear in a derived position is explained by the wh-movement parameter, to which English says 'yes' and Chinese 'no', for instance. In this feature-based system, wh-movement depends on a constraint, the OCP, which applies whenever some feature is present, in this instance Case. Thus, the prediction is: if wh-elements (or whatever elements) in a given language are specified for Case, they are likely to move from their original position under OCP, otherwise they remain in situ (see Desouvrey 2001a for details).

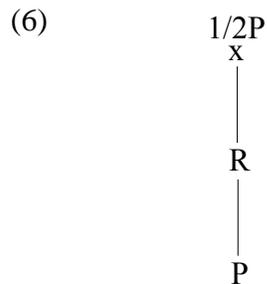
illustrated in (5) with a set of third person elements. It should be noted that the R-node is itself dependent to some node, which might well be the one that holds the phi-features of the morphemes. Such a node is not crucial in English; hence I ignore it in this paper, leaving its identification for further research (see Desouvrey 2006). Thus in (5), x is the root node or whatever dominating node.



To repeat, the representations in (5) translate the fact that only NPs and pronouns have referential contents; they are R-elements. An anaphor has no referential content in itself, and therefore it must have an appropriate antecedent, an R-element. Since a pronoun does not have an R-feature, it must acquire one from an appropriate antecedent, in fact only a referential expression. On the other hand, a referential expression is autonomous; it need not have an antecedent. Also, in the representation of an NP, there are two planes, the R-plane to which belong the node R and its association line, and the P-plane, made of the feature P and the association line linking it to the R-node; the third person pronoun, on the other hand, has a single plane, the R-plane, while the representation of the anaphor shows no plane.

Unlike third person pronouns, the referents of first and second person pronouns are always known directly from the real world, independently of any context. Therefore, it must be the case that such pronouns are fully specified for an R-feature, just like referential expressions, as shown in (6). Moreover, the feature structures in (5) and (6) are not complete. In effect the three types of referring elements possess other features, for instance Case, person, gender, animacy. To this respect, one may distinguish two types of syntactic features: relational features and stative features. Relational features are R-features, which can be acquired structurally under conditions to be discussed shortly. Case is also a

relational feature, but it will not be discussed here, see Desouvrey (2000, 2005, etc.). Stative features are such like animacy, number, gender, person, etc. Indeed, we know that referring elements morphologically vary according to these features. I take the animacy (or human) to be a property of (or perhaps is) the R-node itself. The other stative features, also known as phi-features, might be the mother node of R-features and will be ignored throughout. The reason for ignoring them is that their effect is always predictable. To illustrate this point with an example, *\*she likes himself* is ill-formed, obviously because there is a gender mismatch between the anaphor and its intended antecedent, not because any of the constraints mentioned above is violated. Besides, we know that anaphors are divided into reflexive, reciprocal, relative, etc., and each of them bears particular features that distinguish it from the others. For instance, reciprocal and reflexive anaphors have the same behavior with respect to coreference, as is standardly recognized, but while a reflexive anaphor exists in two versions, singular and plural, reciprocal anaphors are inherently plural, and to perform their semantic functions, they require a plural antecedent. Also, a relative anaphor substantially differs from reflexive and reciprocal anaphors in that it must have an antecedent in a different clause (see below).<sup>3</sup>



As mentioned above, the relation of coreference is in fact a relation of 'antecedent of,' which is asymmetric. Indeed, the feature structures of referring elements given above contain a three-way asymmetry: each referring element has a different feature structure. The nonlinear representation crucially has a further interest. It allows feature spreading. Feature spreading makes it possible to genuinely track down the antecedent of pronouns and

---

<sup>3</sup> It seems that there are more anaphors in natural languages than traditionally thought. In my view the so-called complementizer is a relative anaphor (see below) and according to preliminary observations, the English possessive marker 's and the coordinate morpheme *and* can be best analyzed as anaphors.

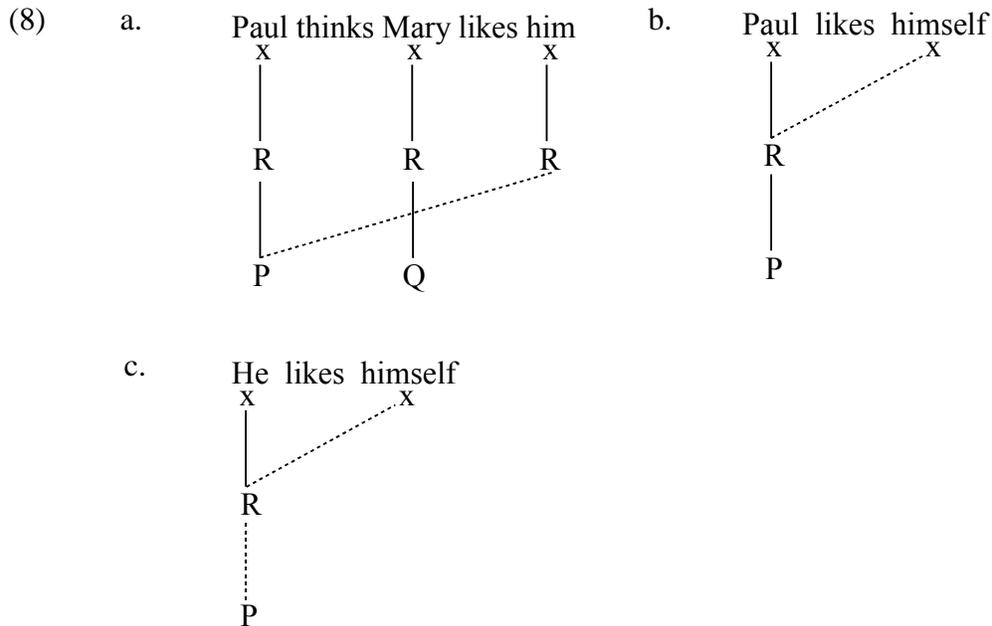
anaphors. It turns out that the spreading of referential features is constrained by the following well-formedness conditions, which are consistent with basic tenets of nonlinear phonology.

- (7) Conditions on feature spreading
- a. An R-feature (terminal) can spread only to a bare R-node.
  - b. An R-node (class node) can spread only to a root node.
  - c. Lines must not cross in the representation.

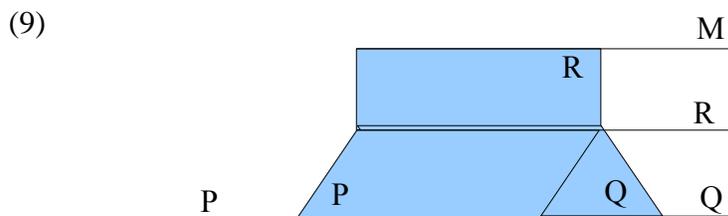
The feature structure for pronouns makes it clear that such elements cannot be the antecedent of a referential expression; and similarly an anaphor cannot be the antecedent of a pronoun or a referential expression. Indeed, by virtue of (7a), a pronoun can acquire an antecedent only by the spreading of a R-feature from a referential expression. The antecedent of an anaphor is the referential expression (or the pronoun) that spreads an R-node on it.

Before turning to concrete examples, we may note that a pronoun and its antecedent need not be in the same sentence, provided the antecedent is sufficiently identifiable in the discourse. To capture this fact, the R-feature of a pronoun whose antecedent is absent from the structure will be linked with a dash line, which means it is a floating feature, along the lines of a common practice in nonlinear phonology.

The sentences in (8) illustrate the process of acquiring an antecedent. First of all, recall that each feature is in the same plane with its association line and only identical features are in the same plane. So (8a) contains three different planes: the R-plane, the P-plane and the Q-plane; the feature [P] spreads on the bare R-node of the pronoun, consistent with (7a); there is no line-crossing effect, since P and Q are not in the same plane. In both (8b) and (8c) the R-nodes spread on the anaphor, which therefore acquires the feature associated to them respectively. (I will return on complementation structures like (8a) in section 4.) Why Q does not spread on the pronoun in (8a), which discards it as a possible antecedent, is discussed shortly.



To clarify further, the three-dimensional representation of (8a) is shown in (9), where M is the morpheme tier. The planes P, Q and R are contiguous to each other, as if they were pivoting around the R-node tier. So the spreading of P to R in (8a) does not take place in the same plane as Q. Similarly, the spreading of the R-node to the anaphor in (8b) and (8c) involve only the R-plane, the space between the M and and the R tiers.



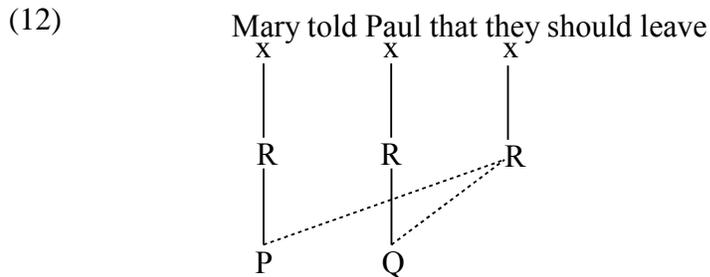
Since every animate R-node belongs to the same plane, R-node spreading may not be always possible. Thus in (10a) the subject pronoun can spread its R-node to the anaphor, yielding a well-formed sentence. In (10b) however, the subject cannot spread its R-node through the intervening object pronoun, given (7c), hence the ill-formedness of the structure. This is a remarkable result; the fact that anaphors must be adjacent to their antecedents follows from the representation which is being used and not from an independent stipulation.

- (10) a. They expected each other to win
- 
- b. \*They expected me to help each other
- 

In the same vein, the fact that anaphors do not allow split antecedents, as is well-known, follows straightforwardly. As seen for instance in (11), by virtue of the no-crossing-lines condition (7c), the R-node of the subject cannot spread through the R-node of the direct object, hence the ill-formedness of the structure.

- (11) \*Mary asks Paul about themselves
- 

On the other hand, a pronoun allows split antecedents, since the latter come from terminal features which are not in the same plane. As seen in (12), each R-feature belongs to a different tier, and therefore separately targets the R-node of the pronoun, the tier of which is adjacent to planes P and Q. A well-formed sentence is hence obtained. (Notice that the plural specification of this pronoun presumably allows it to receive many R-features; perhaps it has two R-nodes.)



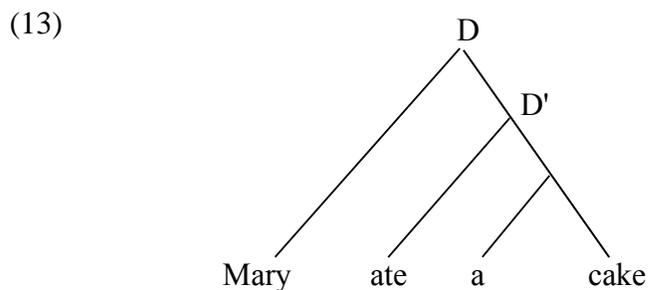
Before concluding this section, we may note that it will never be the case that there are two different referential expressions (or morphemes) with the same R-feature. However, the pairing of R-features with real world entities is totally subjective in that it depends on the speaker's knowledge (and use) of the world reality, which may not be perfect. Therefore the grammar has nothing to say about disagreement among speakers, ignorance, etc, though it provides the speaker the tool to deal with them. For instance, in the famous sentence *the evening star is the morning star* there is no OCP effect since both NPs bear different R-features, say P and Q. They refer to the same entity (P=Q) in the real world only if the speakers agree with each other and that the sentence is conform to the real world, in which case the sentence is grammatically true. Similarly, in the statement  $5^2 = 25$  each element has its own R-feature and it is up to the speaker to accept their identity or reject it. The very fact that it is intrinsically true (in our world) does not involve any linguistic knowledge.

Summing up, the properties of anaphors and pronouns follow from basic tenets of nonlinear phonological representations. An element is the antecedent of another only if it can spread its feature on that element. In the spreading process, lines must not cross, a condition which makes it possible to capture the fact that anaphors must be adjacent to their antecedents. In the next subsections, it will be shown that the complementarity in the distribution of anaphors and pronouns follows from the OCP and the constraint against ambiguity.

## 2.2 Syntactic domains

Given this nonlinear representation, which is based on feature hierarchy, syntactic features are to be found in lexical elements and not in abstract functional heads as in the Principles and Parameters Theory. Following Desouvrey (2000, 2001a, 2001b, etc.), I assume a very

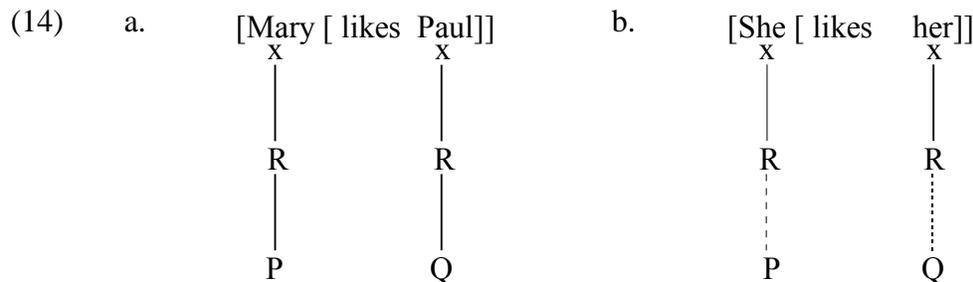
bare syntactic structure, which is only made up of actual lexical elements. Such a syntactic structure is free from functional projections, and therefore inflectional and derivational morphology need not take place in the syntax. On this view, a simple sentence like (13) consists of the shown constituents, namely 'a cake', 'ate a cake' and 'Mary ate a cake', all of them being resulted from successive application of the operation Merge, as discussed in Desouvrey (2000), expanding from Chomsky (1995).



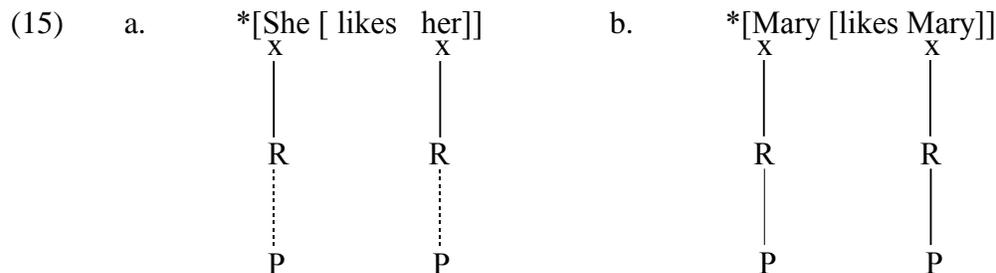
Labels of the constituents are irrelevant and therefore will be omitted throughout. Instead, the letters D and D' are used to indicate specific domains. D' is the domain of the complement of the verb, or complement domain, while D is the domain of the subject, or subject domain. The crucial point here is that the complement domain is embedded within the subject domain, but the contrary is not true. That is, the subject is in the same minimal domain as the object, but the object is not in the same minimal domain as the subject. It turns out that this simple view of syntactic domains, which is derived from a basic constituency analysis, independently needed, is a major key to understanding the phenomena of coreference.

In order to see the interaction of constraints and syntactic domains, I first discuss the effect of the OCP and the Ambiguity Constraint on the appearance of pronouns and anaphors and then I turn to see how the syntactic domains, as defined above, restrict the distribution of the latter (Freedom Constraint). Consider first the case of disjoint reference. Recall that all animate R-nodes are in the same plane, but each terminal R-feature is in its own plane. Thus disjoint reference obtains whenever each argument bears its own R-feature, either inherently as in the case of referential expressions, or by association of floating features that exist in the discourse, as seen in (14a) and (14b) respectively. Notice

that (14b) is not ambiguous to the extent that speakers of English will never assume that both pronouns may bear the same referential feature. Why this is the case will become clear in due course.



Things would be different, however, if both elements were bearing the same R-feature, as seen in (15). I suggest that configurations like (15a) in which two identical R-features stand in the sentence domain are precisely ruled out by the OCP, as stated in (2).<sup>4</sup> In (15b), on the other hand, there is obviously no OCP effect, since both R-nodes get the R-feature by spreading, or more accurately by linking with a floating feature. Nevertheless, (15b) is ruled out by the Ambiguity Constraint, stated in (3). Indeed if (15b) were possible, it would be ambiguous with (14b), the case of disjoint reference.



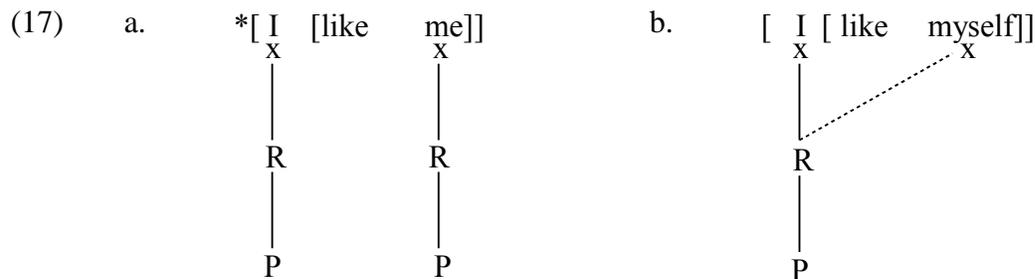
I suggest that such structures would have to be repaired by a process of dissimilation if they were used as inputs.<sup>5</sup> Since referring elements differ from each other by their feature

<sup>4</sup> It turns out that with R-features, the OCP takes effect in the sentence domain and even beyond, and it is not avoided by movement of one of the element bearing the R-feature. With Case features, however, OCP effects occur always in smaller domains, generally the complement domain of the verb, and in most cases trigger the movement of the verb argument; thus clitic movement, wh-movement, adverb movement are all caused by a Case OCP. The difference between Case and reference with respect to the OCP is due to the fact that Case is a clause-bound grammatical feature, while reference is a discourse feature.

<sup>5</sup> The literature often mentions the case of Thai in which (15a) is well-formed. This does not rule out the analysis presented above, since in my view the OCP is a structural constraint, hence violable. What



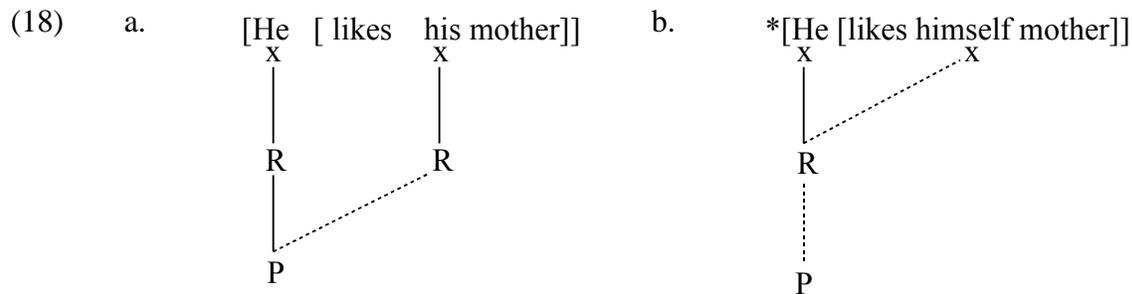
pronouns in complementary distribution and speakers know that the antecedent of an anaphor must be the closest R-node.<sup>6</sup> It is important to further distinguish the OCP from the AAC. To this effect, structure (15b) can be compared with (17a). Structure (17a) carries no ambiguity at all, since first and second person pronouns always bear their own referential feature. Nevertheless, it is ill-formed. The ill-formedness of this structure is predicted by the OCP, since both pronouns bear the same R-feature, [P]. The delinking of the R-node of the object pronoun would readily yield the desired result, as seen in (17b).<sup>7</sup> In contrast (15b) (as well as (16c)) is structurally perfect, only the intended coreference relation between the pronouns makes it deviant. Thus the OCP and the AAC are not to be confused, although they have the same effect on the syntactic structure, namely feature manipulation.



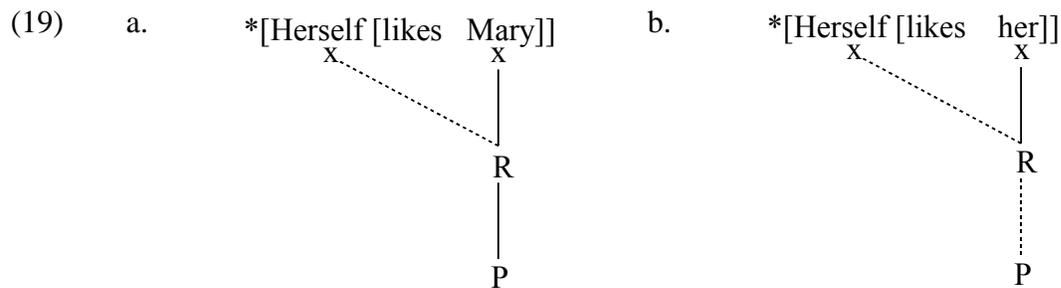
Notice that ambiguity cannot be eliminated in every context. Structure (18a) has two readings, one of coreference and one of disjoint reference. Hence its ambiguity. But the structure cannot be improved by an alternative input or a delinking process, as seen in (18b). The impossibility of (18b) is related to the nature of the pronouns. A personal pronoun is an argument of the verb, but not a possessive determiner, and there is no determiner anaphor in the inventory of English. A reflexive anaphor may not be used as a determiner anymore than a determiner can be used as an argument on its own; cf. *\*She likes the/my*.

<sup>6</sup> An anaphor may be used emphatically without an antecedent in the sentence, as in "The paper was written by Ann and myself" (from Safir 2004). This case, which is quite marginal, is possible when the referent is easily traceable and in addition no R-element in the sentence may be wrongfully taken to be an antecedent. At this point, use of an anaphor as a pronoun is not crystal-clear.

<sup>7</sup> Notice that in major Romance languages, there are no first and second person anaphors; so the resolution of the OCP is never possible in such contexts.

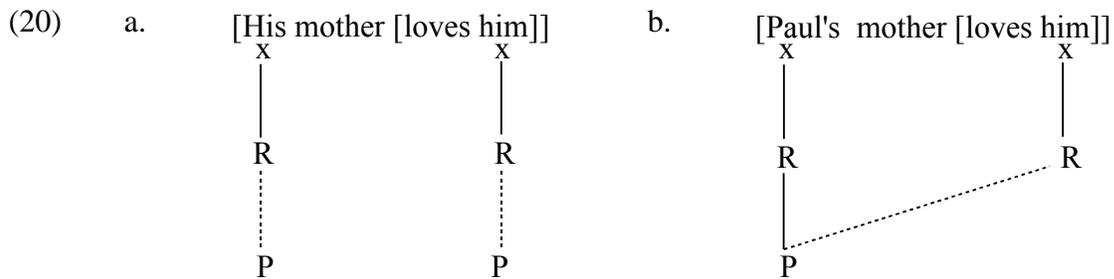


Now one can logically ask why the dissimilation process seen above does not act upon the subject, such as to derive the structure in (19). Or what amounts to the same thing, why an anaphor cannot be used as subject. It is at this point that the notion of domain enters the stage. Indeed, these structures are straightforwardly ruled out by the Freedom Constraint (1), which requires that anaphors and pronouns be free in their domain. Since the subject domain includes the object domain, the subject anaphor cannot be free whenever its antecedent is the object. Therefore, (19a) and (19b) are ruled out. Notice that in current Binding Theory, this result is ensured by the c-command requirement, by virtue of which the anaphor must be c-commanded by its antecedent. The view of domain adopted in the present proposal makes it possible to dispense with such an unnatural structural relation.<sup>8</sup>

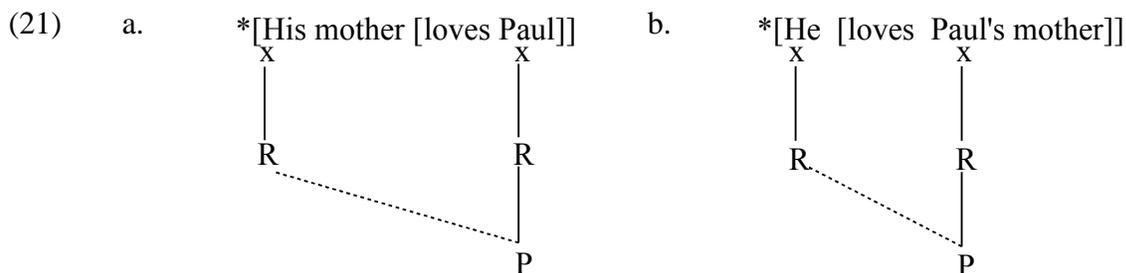


In contrast, the sentences in (20) are well-formed, since the pronouns have no antecedent in their domain. In fact, the pronouns in (20a) have no antecedent at all in the sentence, since a pronoun may not be the antecedent of another pronoun. In (20b) the antecedent is in the subject domain, not the object domain that includes the pronoun, and therefore the structure is perfect.

<sup>8</sup> The present analysis is superior to the c-command-based analysis, since the definition of domain is invariant and follows from the constituency analysis, in contrast to the c-command which is variously defined in the literature, according to the desired result.



The sentences in (21) further illustrate the domain effect. In both (21a) and (21b) the antecedent of the pronouns is included in the subject domain, yielding ill-formedness, as expected.



Given the facts just discussed, one might be tempted to assume that feature spreading must proceed from left to right. Such an assumption would be consistent with Chomsky's (1976) Leftness Principle, according to which a variable cannot be the antecedent of a pronoun to its left. This principle would explain the ill-formedness of (19), without invoking any domain-related analysis (and the c-commanding requirement as well in Chomsky's system). This would be incorrect, however. The following sentences (cf. Reinhart 1983) are well-formed, but they display right-to-left spreading of the referential expression to the possessive. Notice that the latter belongs to a phrase which is not in the subject domain of the sentence (also see below). (Indexes are conveniently used as a descriptive device.)

- (22) a. Near his<sub>i</sub> child's crib nobody<sub>j</sub> would keep matches.  
 b. For his<sub>i</sub> birthday, each of the employees<sub>j</sub> got a Mercedes.

To sum up, we have seen that anaphors arise in the structure from the delinking of an

R-node belonging to a pronoun or a full referential expression. The delinking process is triggered by either the OCP or the Ambiguity Constraint. It is also shown that anaphors can never be in the subject position (the larger domain), given that the subject domain necessarily includes the antecedent of the anaphor. Similarly pronouns cannot be used as subject whenever their antecedent is in the object domain, consistent with (1).

At this point, the analysis is complete and the role of each constraint is already made clear. The rest of the paper will mostly deal with complexities.

### 2.3 Constraint interaction

Up to this point, we have seen the individual play of the OCP and the AAC. Both trigger a delinking process that yields an anaphor in simple sentences. Given the feature structure of referring elements, the OCP and the AAC may not conflict with each other, since the former prevents two instances of the same terminal feature from appearing on a tier, while the latter opposes bare R-nodes, which can be marked by any R-feature from any plane, causing an unwanted ambiguity. However, the OCP and the AAC interact with the no-line-crossing constraint (NLC) (7c), as I will show. The effect of such an interaction has consequences for the domain the antecedent of the anaphor can appear in.

To see the interaction of constraints, I will consider two types of structures in this section: structures with complex arguments and ditransitive verbs; bi-clausal structures will be considered at a later section. Consider for instance a sentence with a complex object like (23a). There are three referring elements in this structure, the subject pronoun (*they*), the complement of the preposition (*them*), and the object of the verb (*books*). Since the latter is not animate, its R-node may not be on the same tier as those of the subject and the indirect object (the R-feature of *books* is not shown for simplicity, and even its R-node will be omitted in further examples). The bare R-node of the second pronoun can take any R-feature, which makes it ambiguous, hence the deviance of the sentence. Fortunately, this ambiguity can be rooted out by delinking the bare R-node, a process that yields (23b), which is perfectly grammatical.<sup>9</sup>

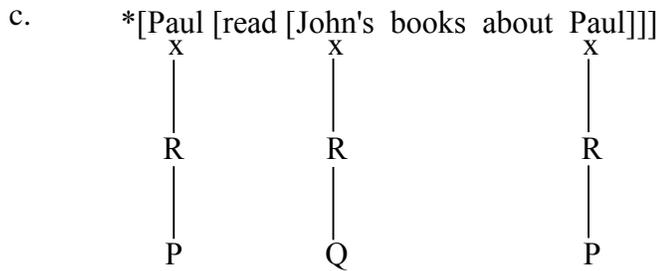
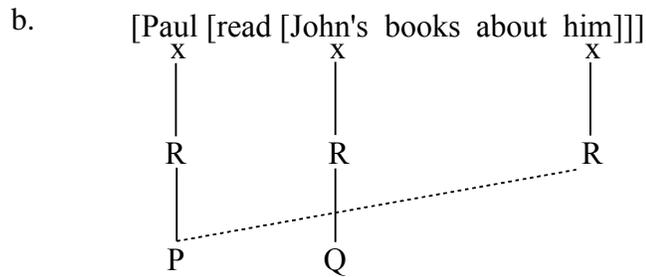
---

<sup>9</sup> The hesitant judgement on (23a) may be due to the long distance between the R-nodes, as induced by the complexity of the object. Speakers may just be puzzled over its nature, a clause or a simple argument.

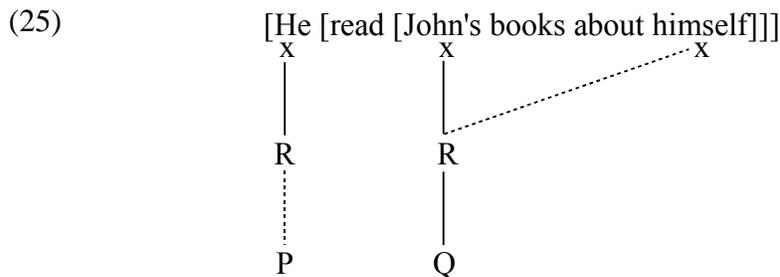
- (23) a. ??[They [read [books about them]]]
- 
- b. [They [read [books about themselves]]]
- 

It is important to note that in (23) the R-node of the head noun does not interfere with the animate R-nodes of the pronouns. However, if the complexity of the structure increases by adding a further animate argument, the no-crossing-lines requirement will block any resolution of the AAC. Let us illustrate this with (24). (24c) is not ambiguous, but it violates the OCP in the P-tier. As far as the OCP is concerned, either the terminal feature or the whole R-node may be suppressed in order to obtain a well-formed structure. However, the intervening R-node of *Mary* eliminates the choice, for only the delinking of the P-feature is possible, as seen in (24b). It appears that (24b) is ambiguous, in violation of the AAC, for the pronoun may be coreferent or not with the subject. But there is no way to avoid this problem. In effect, if the R-node is deleted as in (24a), the intended coreference is no longer possible by virtue of the ban of lines crossing.

- (24) a. \*[Paul [read [John's books about himself]]]
-

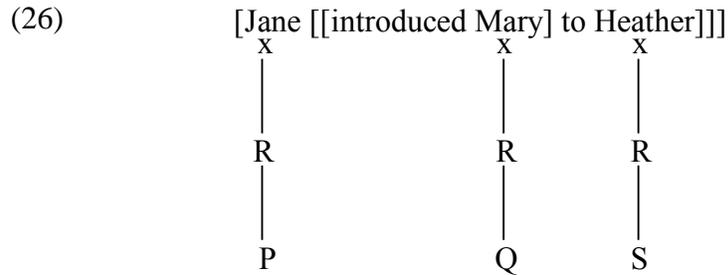


Notice that although (24b) is ambiguous, the interpretation that the second pronoun is coreferent with the intervening NP is not available. If *John* were the intended coreferent, the R-node of the pronoun would have been deleted under AAC, and this would yield a structure with no ambiguity, where the R-node of *John* easily spreads on the anaphor, as seen in (25).

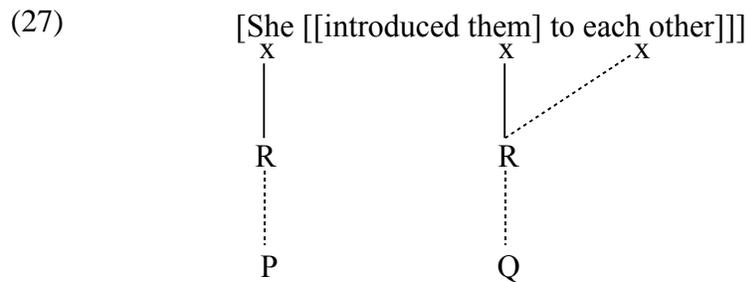


Let us turn now to ditransitive verbs. Since domains are related to arguments, a simple transitive verb contains two domains, the subject and the object domain, as seen above. But what if the verb, in addition, has an indirect object? If the indirect object is in a PP, it goes without saying that a such constituent makes up a further domain. Since the V' (so to speak) and the PP are merged together, it must be the case that these two independent domains come down to a single one. Now, under the natural assumption that in the merging of the two constituents the V', but not the PP projects, the domain of the direct object is

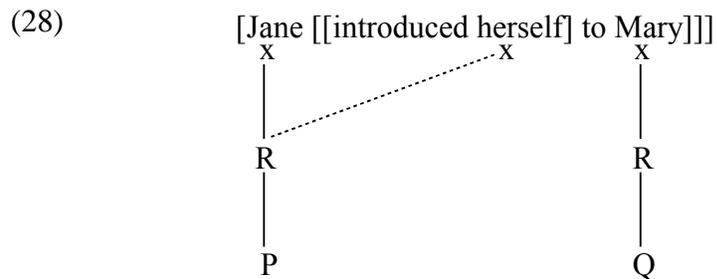
expanded so that it includes the domain of the indirect object. Thus, in a structure like (26) the domain of *Mary* includes the domain of *Heather*, but the domain of *Heather* does not include that of *Mary*.



Under this view, when an anaphor is the indirect object it is in its own domain and therefore it can pick out the direct object as its antecedent, as seen in (27). Needless to say that its antecedent cannot be the subject, given the no-lines-crossing condition.



However, if the anaphor is the direct object, as in (28), its domain includes that of the indirect object, and therefore the latter is ruled as a possible antecedent by the FC. In a such a case, the antecedent must come from the largest domain, that of the subject.



### 3. Operators

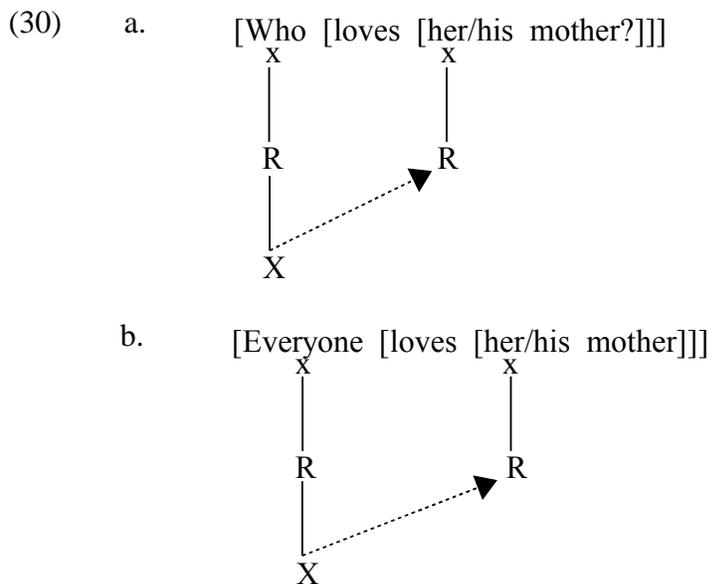
The analysis sketched above can account for the distribution of quantifiers, wh-elements and crossover effects with the same accuracy. What we need to know adequately is the referential properties of such elements, in other words their definitions. I argue that wh-operators and quantifiers are fully specified referring elements, that is, they are specified for R-features, just like ordinary referential expressions.

Under current assumptions in the literature, wh-operators are taken to be pronouns, although their traces are treated as referential expressions. In the present theory, traces are not used, but wh-operators do move from a base-generated position to the clause initial position, as discussed in Desouvrey (2001a). Thus it is possible that operators are simply what their supposed traces are, namely referential expressions. We know that a personal pronoun must be linked to an R-feature provided by a referential expression (its antecedent). If a pronoun has no known antecedent, the structure it appears in is semantically deviant. Now if wh-operators were pronouns, they would be used only in context where they can be linked to an antecedent. However, wh-operators do not need an antecedent in the structure or the discourse, just like quantifiers. For instance, a speaker could begin a long discourse with such sentences like (29); for (29b) the speaker or someone else may implicitly or explicitly provide a name; but the key fact is that the operator (or the quantifier) remains autonomous, that is, they may not have the same relationship with any element as that that exists between an anaphor and a noun or a pronoun.

- (29) a. Everyone loves his mother.  
 b. Who is ruling the world?

In fact, there are a few differences between wh-operators and NPs. I suggest in Desouvrey (2000, 2001a) that wh-operators are specified for a special feature, referred to as  $[\sigma]$ . This feature is intended to handle the fact that operators have a scope in structures they appear in. Let us say that a  $\sigma$ -specified element is a vector, as opposed to a scalar, which has no scope capabilities. On this view, operators must assign an R-feature in one direction

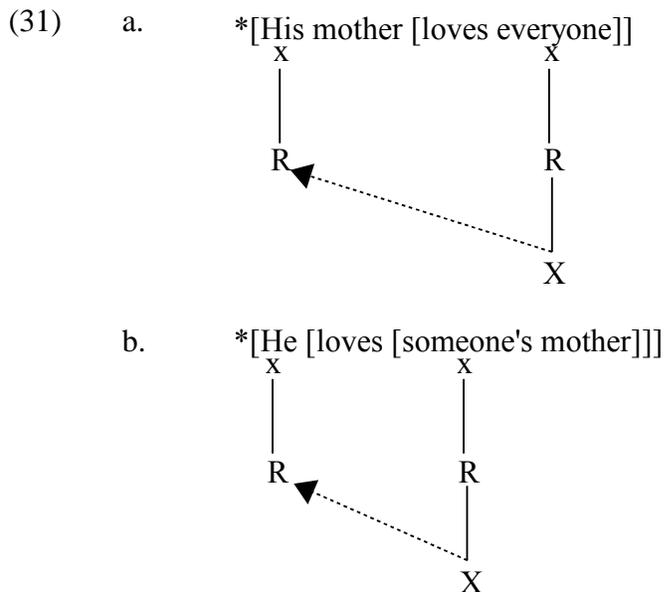
only, either from left to right or from right to left, i.e., their R-feature is oriented. On the other hand, an NP (scalar) can assign a R-feature to another element to its left or to its right. Suppose further that quantifiers are specified for scope as well, and therefore are vectors. Since operators and quantifiers are not specific, their R-feature will be noted X, and the spreading line will be arrowed in order to highlight their vectorial properties. In most cases the vectorial nature of operators is irrelevant, and therefore they behave just like NPs. The examples in (30) show that is indeed the case. The operator in (30a) may be the antecedent of the pronoun and, if so, it spreads it a feature; so is the quantifier in (30b). As in cases discussed above, these structures are well-formed because the pronoun is free in its domain.



If the vector element is not the subject of the sentence, however, ungrammaticality obtains for two reasons. First FC is violated, since the pronoun (in subject position) is no longer free in its domain. Second, as discussed in Desouvrey (2003), English (tensed) verbs and modals are operators, and therefore they interact with quantifiers and wh-operators. Their interaction is in the realm of what is currently known in the literature as superiority effects. In the present theory, superiority effect is simply a constraint that forbids vectors to exchange their relative positions in the course of the derivation. With respect to reference, I suggest that a vector cannot spread an R-feature over another vector. That is to say, spreading of an R-feature past another vector amounts to out-scope that vector, as if the

element were moved past the other.

On this view, consider the sentences in (31). They are ruled by FC, since the quantifier is included in the domain of the pronoun it is the antecedent of. As well, since the tensed verb is a vector, it blocks the spreading of feature X to the subject pronoun. Now one may ask how the non-linear representation handles this? Unfortunately, I have to leave this for future research, because at this time the location of feature  $[\sigma]$  in the tree structure is not crystal-clear.



Furthermore, if an object operator, which is in a clause initial position in English, is the antecedent of a pronoun in the subject domain, ill-formedness obtains as well, as illustrated in (32). Such cases are well-studied in the literature and are known as crossover effects. (32a) is usually referred to as a weak crossover, while (32b) is a strong crossover.

- (32) a.  $*\text{Who}_i \text{ does his}_j \text{ mother love?}$   
 b.  $*\text{Whose}_j \text{ mother does he}_i \text{ love?}$

The initial inputs for structures in (32) are shown in (33), where the operators are put back in their initial positions. I set aside the cause of the movement as well as the reason why *do* must be there, but see Desouvrey (2001a). These inputs may not be correct

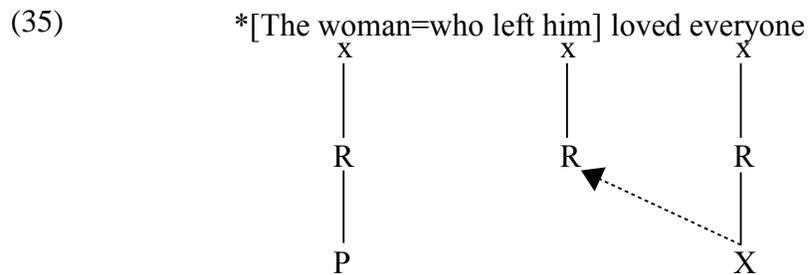
however, since (a) the pronoun is not free in its domain (b) the wh-operator over-cross (or out-scope) the modal verb, just as in (31). However, the ill-formedness persists under extraction of the operator, which suggests that there is an additional problem in these structures. As a matter of fact, after extraction the operator comes to R-mark the pronoun from left to right, a different direction from the input, and therefore the output is ruled out.

- (33) a. \* $[\text{His mother } [ \text{does } [ \text{love } \text{who?} ] ] ]$
- 
- b. \* $[\text{He } [ \text{does } [ \text{love } \text{whose mother?} ] ] ]$
- 

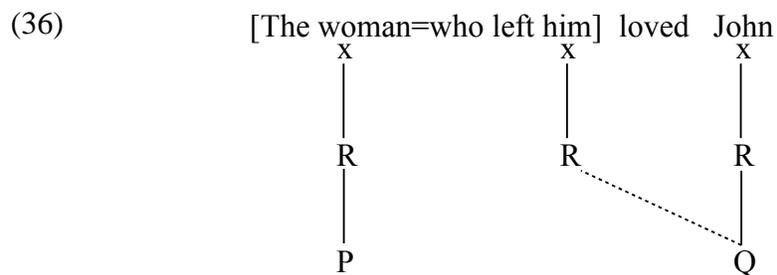
The generalization that comes out of these structures is that coreference relations between an operator and a pronoun are checked at the starting point of the derivation (the input structure) and the direction of the R-spreading must be preserved throughout.

Let us consider now an example with a quantifier. In structure (34), seen in Huang (1995) (his (42b)), all referring elements are conveniently co-indexed with their antecedents: the antecedent of relative *who* is *the woman* and the antecedent of pronoun *him* is the quantifier. The input structure will make it possible to track down the cause of the ill-formedness. Following Desouvrey (1995, 1997) the input for (34) must consist of two independent clauses: the main clause and the relative clause, the latter being adjoined to the head noun for reason of coreference, as shown in (35). In this linearized structure, the quantifier over-crosses the tensed verb, and therefore ill-formedness obtains.

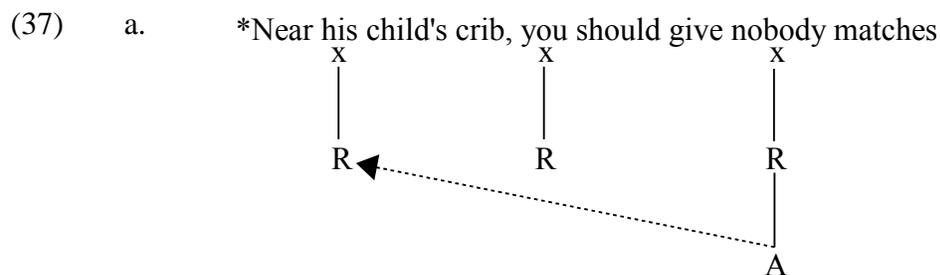
- (34) \*The woman<sub>j</sub> who<sub>j</sub> left him<sub>i</sub> loved everyone<sub>i</sub>

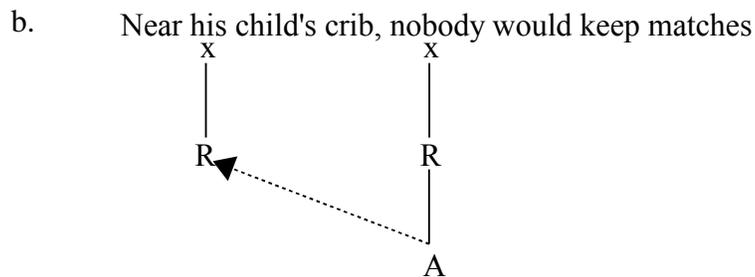


This analysis is supported by the fact that if the antecedent of the pronoun is a scalar, a well-formed structure apparently obtains after embedding of the relative clause, as shown in (36) (Huang 1995, (41b)).



The contrast in (37) (cf. Reinhart 1983) is a further evidence for this analysis. In (37b), the quantifier can be the antecedent of a pronoun to its left. This sentence is well-formed, since none of the problems seen above, namely over-scope and spreading redirection are present. In (37a) however, the sentence is ill-formed as expected, since the quantifier assigns its R-feature to the pronoun by over-crossing the vector modal.





To conclude, I have suggested in this section that operators, including quantifiers, bear their own referential feature, and therefore behave mostly like ordinary referential expressions. In addition, their R-feature cannot be reoriented by further syntactic processes and they enter in a superiority effect relation with other non-referential vectors, like modals and tensed verbs.

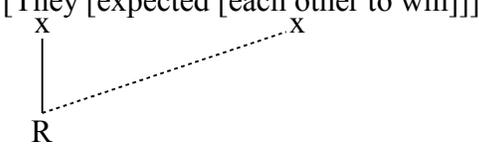
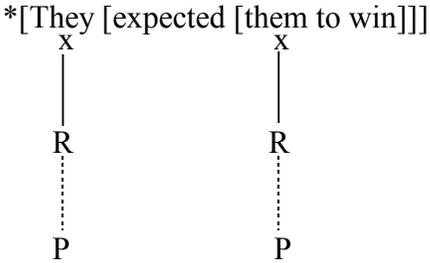
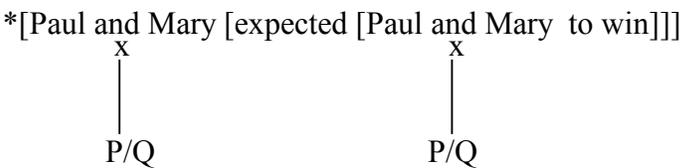
#### 4. The relative anaphor

In this section I turn to coreference between elements belonging to different clauses. Tensed subordinate clauses (in English) are introduced by a complementizer, which may or may not be overt, but not non-tensed ones. Both types of clauses are involved in a striking contrast, illustrated in (38). The problem is that the appearance of an anaphor in a tensed embedded verb yields ill-formedness (38b), which disappears if the embedded verb is not tensed (38a). I show that this state of affairs is due to a Case conflict that arises in subordinate tensed clauses and is avoided by the use of the so-called complementizer.

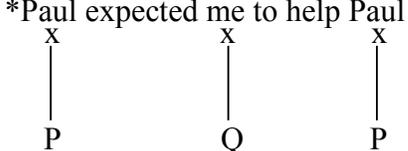
- (38) a.      They expected each other to win.  
       b.      \*They expected each other would win.

Consider (38a). It displays no particular interest; as seen in (39a), the R-node of the pronoun spreads normally on the anaphor, yielding a perfectly grammatical sentence (from now on, R-feature will be shown only when relevant). Notice here that the anaphor cannot be replaced by a pronoun, as the ill-formedness of (39b) shows. This must be the case, since the R-node of the subject can spread without being intercepted by an intervening association line. Thus (39b) is ruled out because it contains an avoidable ambiguity. Notice further that referential expressions are not possible here either, consistent with the OCP, as

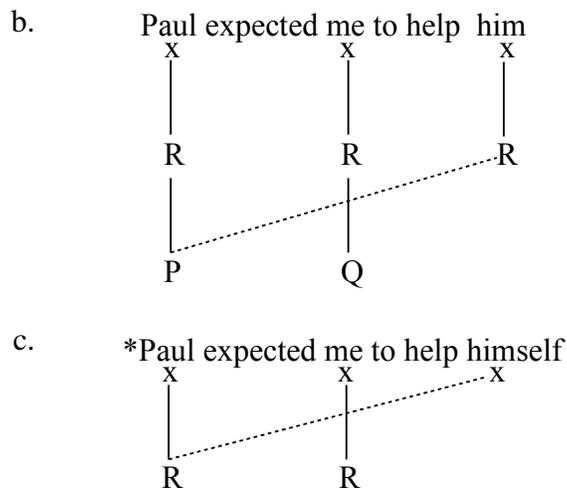
seen in (39c) (P and Q are linked individually to R and are not in the same plane, and in fact there is a double OCP effect here, one in the P-plane the other in Q-plane.<sup>10</sup>)

- (39) a. [They [expected [each other to win]]]  

- b. \*[They [expected [them to win]]]  

- c. \*[Paul and Mary [expected [Paul and Mary to win]]]  


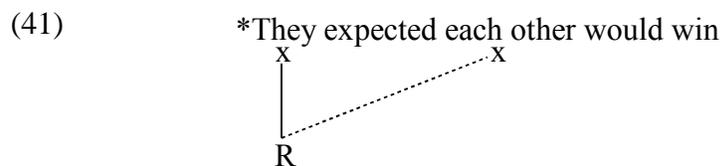
This analysis is independent of whether the embedded verb is transitive or not. What matters is the interaction of constraints and the representation. As can be seen in (40a), OCP triggers the delinking of the P-feature on the second instance of *Paul*, yielding (40b), which is ambiguous. The ambiguity of (40b) is not avoidable, since the intervening R-node is an obstacle to the spreading of the subject R-node on the anaphor, hence the ill-formedness of (40c). In other words, the no-lines-crossing condition blocks the elimination of the ambiguity, which therefore remains in the structure.

- (40) a. \*Paul expected me to help Paul  


<sup>10</sup> I am simply assuming that coordinate phrases have a single R-node in order to avoid a long digression on the nature of conjunction *and* (see footnote 3). However, we may note briefly that the internal structure of this coordinate phrase is most likely the following: [Mary [and Paul]], but not a ternary structure like [Mary and Paul].



Let us turn now to (38b), the interesting problematic case. A structure along the lines of (41), which is parallel to (39a), may not work for this sentence. The problem is that the anaphor is in a position where it is assigned two different Cases, nominative by the embedded auxiliary and accusative by the matrix verb. Hence there is a mismatch between Case and grammatical relations, a matter that is discussed in Desouvrey (2000). Notice that in (38a) the infinitive verb does not assign nominative Case, and therefore this problem does not arise. The strategy for avoiding the Case-mismatch in depending clauses with tensed verbs will help to understand the impossibility of the anaphor in (41).



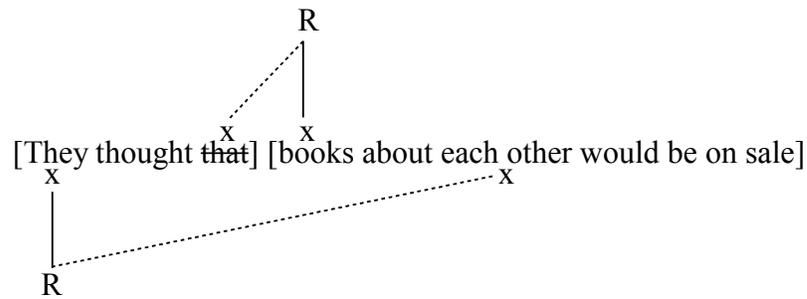
The Case mismatch arises because both verbs share a common argument, the anaphor. Indeed, the anaphor is at the same time the direct object of one verb and the subject of the other, since no null argument is possible in this context, unlike in (38a), where the infinitive verb has a null subject, as standardly assumed. Intuitively, each verb in (41) must have its own argument (not shared) that absorbs its Case in order to avoid the Case mismatch. In Desouvrey (1996, 1997), I suggest that in natural languages, complementizers are intended to replace the missing argument in such contexts. In fact, the complementizer is nothing but a relative anaphor and it is the complement of the matrix verb with which it forms an

independent structure from the second clause. The antecedent of the relative anaphor is intended to be the whole complement clause, but actually it is the subject of that clause. On this view, matrix clauses in English-like languages are basically head final relative clauses. From this new perspective, the input structure for (38b) must be (42a), where ~~that~~ indicates that the complementizer may or may not be overt in English, as we know. It is thus obvious that the double-Case problem disappears, since the relative anaphor, as an argument of the matrix verb, absorbs the accusative Case of the latter. Now the reason why the structure shown in (42a) is ill-formed should be clear. By its very nature, a relative anaphor must be linked to an R-antecedent, which must not be a clausemate. Clearly there is no such element in (42a)--the subject of the matrix verb cannot be the antecedent of the relative--and therefore ill-formedness obtains. However, if the subject of the complement clause is an R-element, a well-formed sentence readily obtains, as seen in (42b).

- (42) a. \* $\left[ \begin{array}{c} \text{They} \\ \text{X} \end{array} \right] \text{ expected } \left[ \begin{array}{c} \text{that} \\ \text{X} \end{array} \right] \left[ \begin{array}{c} \text{each other} \\ \text{X} \end{array} \right] \text{ would win}$
- $\left. \begin{array}{c} \text{X} \\ | \\ \text{R} \end{array} \right\} \text{---} \left. \begin{array}{c} \text{X} \\ | \\ \text{R} \end{array} \right\}$
- b.  $\left[ \begin{array}{c} \text{They} \\ \text{X} \end{array} \right] \text{ expected } \left[ \begin{array}{c} \text{that} \\ \text{X} \end{array} \right] \left[ \begin{array}{c} \text{they} \\ \text{X} \end{array} \right] \text{ would win}$
- $\left. \begin{array}{c} \text{X} \\ | \\ \text{R} \\ \vdots \\ \text{P} \end{array} \right\} \text{---} \left. \begin{array}{c} \text{X} \\ | \\ \text{R} \\ \vdots \\ \text{P} \end{array} \right\}$

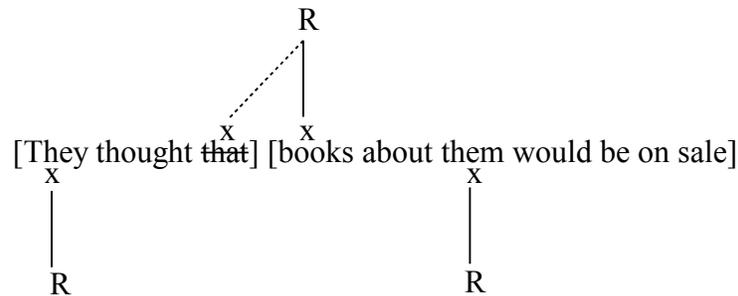
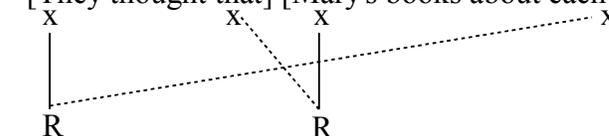
Clearly (42a) is ill-formed because an anaphor must have an antecedent which can provide it with an R-node. As a further evidence for this analysis, one may note that the subject of a complement clause may contain an anaphor, as far as the latter does not count as the antecedent for the relative. Thus in (43), although the whole bracketed phrase is intended to be the antecedent of the 'complementizer,' only its head *books* actually plays this role. Since *books* is an R-element, the complementizer is correctly linked to it, yielding a perfect sentence.

(43)



Incidentally, notice that a pronoun can successfully appear in place of the anaphor, as seen in (44). Let us see how and why this non-complementarity is possible. There are two ways to see these complex structures: either on a local level or on a sentential level. On a local level, the complement clause is perfectly autonomous from the matrix clause, although both are connected with each other via a relative anaphor. Thus the appearance of the pronoun is warranted as in any root sentence, since the intended antecedent is not in the clause. At the sentential level, which possibly arises from a reanalysis that integrates the complement clause and the relative anaphor into a single structure, as linguists do traditionally, the ambiguity of the pronoun is eliminated by delinking its R-node, which yields the anaphor. Of course, if an intervening R-node is present, the delinking process may not take place, as seen in (44b).

(44) a.

b. \*~~[They thought that]~~ [Mary's books about each other would be on sale]

Before concluding this section, consider an interesting problem related to the OCP, namely the contrast in (48). According to Lasnik and Uriagereka (1988), (48a) is worse than

(48b). There is no ambiguity problem there, but the OCP is definitely violated. Thus the problem is to explain why the violation of the OCP is mild in (48b), but strong in (48a). The answer may be that (48a) is a step in a doomed derivation, since a better sentence can never be obtained from it. Indeed, in the first place, feature manipulation on *John's* is not possible in both sentences, since there is no corresponding element to derive; cf. *\*he's teacher*, *\*him's teacher*. However, *John* is manipulable, as we know. A *John* without an R-feature yields *him*, but without the R-node it yields *himself*. So to rescue (48a), one would have to manipulate the subject *John* so as to obtain either a pronoun or an anaphor. However, this would be incorrect, since we know that neither a pronoun nor an anaphor can be the subject if the antecedent is in the same domain (the Freedom Constraint). Indeed, replacing the subject by a pronoun yields ill-formedness, as shown in (49a). In (48b), on the other hand, a better sentence can be obtained by manipulation of the direct object *John*, since the resulting pronoun remains free in its domain, as evidenced by the grammaticality of (49b). To put it differently, (48a) is not a possible input, while (48b) is and only needs a further derivational step (feature delinking) to become fully acceptable.

- (48) a. \*John can't stand John's teacher.  
       b. John's teacher can't stand John.
- (49) a. \*He can't stand John's teacher.  
       b. John's teacher can't stand him.

To conclude, the main point of this section is that the complementizer is a relative anaphor, and therefore it must have an R-element as its antecedent. An anaphor cannot be the antecedent of another anaphor because as we know anaphors lack an R-node of their own. Anything else follows from the interaction of the constraints and the representation, as discussed throughout.

## 5. Concluding remarks

My objective were to account for relations of coreference between referring elements. Although the discussion was limited to English, for which an important body of literature is

available, I strongly believe that such an approach can easily account for similar phenomena in other languages. The key features of the above proposal are the following. An OCP effect prevents the appearance of two R-feature elements in the syntactic structure. One of the feature involved in the OCP effect must be delinked up to structure preservation, a process known as dissimilation in the phonological literature. A constraint against ambiguity favors the delinking of the whole R-node instead of the terminal R-feature so as to obtain an anaphor. However, in many contexts, the ban on line crossing in the representation opposes the appearance of an anaphor, since the latter must receive an R-node by spreading. The other important constraint is that pronouns and anaphors must not be in the same domain as their antecedent. Domains are defined recursively; the larger domain in the structure contains the smaller ones. Thus an anaphor in the larger domain cannot be free, since its antecedent must be in a smaller domain. The nonlinear representation makes it clear that the antecedent of a pronoun or an anaphor may not be a pronoun or an anaphor respectively.

Compared to the complexity of the facts, and their account under current Binding Theory, this proposal is amazingly simple and explanatory. Besides the fact that the three conditions of Binding Theory are done away (the FC is not related to any principle of Binding Theory), numerous apparatus and related claims are eliminated. The c-command relation, any reference to tensed clause or to nominative Case as well as notions such as governing categories and other barriers are eliminated. The assumptions needed are not ad hoc to the extent they are part of a well-articulated theory of features and constraint interaction. Furthermore, the constraints are not peculiar to the phenomena of coreference. The OCP and the constraint against ambiguity are independently needed elsewhere in the grammar. For instance, the OCP deals with other features, such as Case, in a way that forces the feature bearer (an adverb, a clitic, etc.) to move to another position, as discussed in Desouvrey (2000), (2001a), (2001b), (2004), etc.. As for domains, they basically correspond to the constituents and therefore they do not add further complexity to the system.

It should be noted as well that none of the constraints looks unnatural. The OCP is surely based on an economy preference, since basically it rules out any repetition of a feature in the structure. The constraint against ambiguity is self-explanatory and it is

virtually impossible to think of a natural language, or a logical system in general, that favors any form of ambiguity. The FC, which requires the antecedent to be in the larger domain than the element it is the antecedent of, is after all related to the notion of prominence. Traditionally it is recognized that the subject is more prominent than the object. In the present proposal, a noun/NP, which bears an R-feature, is more prominent than a pronoun that only has an R-node, and the latter is more prominent than an anaphor that does not have any R-node. It is thus highly plausible that a referring element is required to be in a more prominent position than the element it references. That is, in a coreference relation involving a noun and a pronoun or an anaphor, the former must be the subject. Similarly in a coreference relation between an anaphor and a pronoun, the latter must be the subject. In current theories of coreference, if the principles or any structural relations are replaced by their contrary, the plausibility of the whole system may not be affected (see a discussion in Desouvrey 2000).

## References

- Archangeli, Diana (1984). Underspecification in Yawelmani Phonology. Doctoral dissertation, MIT.
- Chomsky, Noam (1976). Condition on Rules of Grammar. *Linguistic Analysis* 2, 303-351.
- Chomsky, Noam (1981). Lecture on Government and Binding. Dordrecht: Foris.
- Chomsky, Noam (1995). The Minimalist Program. Cambridge, MA: MIT Press.
- Desouvrey, Louis-H. (1996). Case Tiers, Clause Structure and the Nature of the Complementizer-trace Effect. Ms.
- Desouvrey, Louis-H. (1997). Relativization in French Without Complementizer. In Blair, L., C. Burns and L. Rowsell (eds.), Proceedings of the 1997 ACL, Calgary Working Papers in Linguistics, University of Calgary, pp. 73-84, 1997b.
- Desouvrey, Louis-H. (2000). Romance Clitics and Feature Asymmetry: An Autosegmental-Based Approach. Doctoral dissertation, UQAM.
- Desouvrey, Louis-H. (2001a). Effects of ambiguity in syntax: A study of wh-operators in English and French. Ms. (under review)
- Desouvrey, Louis-H. (2001b). Adverbs, Negation and OCP Effects. Ms.
- Desouvrey, Louis-H. (2005). Romance clitic clusters: The Case connection, In Heggie

Lorie, and Francisco Ordonez (eds), Clitic and Affix Combinations: Theoretical Perspectives. John Benjamins.

Desouvrey, Louis-H. (2006). Underspecification and Long-Distance Antecedents: The Case of Chinese Ziji, <http://semanticsarchive.net>.

Goldsmith, John (1976). Autosegmental Phonology. Doctoral dissertation, MIT.

Harbert, Wayne (1995). Binding Theory, Control and Pro. In Gert Webelhuth (ed.), Government and Binding Theory and the Minimalist Program, 179-240. Cambridge, MA: Blackwell.

Higginbotham, James (1983). Logical Form, Binding and Nominals. Linguistic Inquiry 14, 395-420.

Huang, C.-T. James (1995). Logical Form. In Gert Webelhuth (ed.), Government and Binding Theory and the Minimalist Program, 127-175. Cambridge, MA: Blackwell.

Lasnik, Howard and Juan Uriagereka (1988). A Course in GB Syntax. Cambridge, MA: MIT Press.

McCarthy, John (1986). OCP Effects: Gemination and Antigemination. Linguistic Inquiry 17, 207-263.

Prince, Alan and Paul Smolensky (2002). Optimality Theory: Constraint Interaction in Generative Grammar, <http://roa.rutgers.edu>.

Reinhart, Tanya (1983). Anaphora and Semantic Interpretation. London: Croom Helm.

Safir, Ken (2004). The Syntax of Anaphora. Oxford University Press.

-----  
Author's email addresses:

lhd33@yahoo.com

louish8@gmail.com