CORNELL WORKING PAPERS IN LINGUISTICS

Volume 4 • Spring 1983
CORNELL UNIVERSITY
WORKING PAPERS IN LINGUISTICS
Number 4, Spring 1983
A Special Issue: Papers from the Cornell Conference on Government and Binding Theory, July 8-9, 1982.
Edited by Wayne Harbert
Editorial Committee:
Leonard Babby Wayne Harbert Carol Rosen
John Bowers Sally McConnell-Ginet Margarita Suñer
Wayles Browne Carlos Piera
The Cornell Working Papers in Linguistics is an informal publication of the Department of Modern Languages and Linguistics at Cornell University. It is intended as a forum for presentation and rapid dissemination of current research by the faculty and students of the DMLL. Since the papers represent work in progress, comments and criticism are invited. All correspondence may be addressed to Working Papers in Linguistics, DMLL, 203 Morrill Hall, Cornell University, Ithaca, New York 14853.

Due to the irregular schedule of publication, we cannot offer subscriptions. Brochures describing current issues as well as back issues will be sent to those on our mailing list.
PREFACE

The first Cornell Conference on Government and Binding Theory, organized by Robert Freidin, Wayne Harbert and Margarita Suñer, was held on July 8-9, 1982 at Cornell University. The program consisted of eighteen thirty-minute papers plus two talks by invited speakers:

Howard Lasnik  
Binding or Pragmatics?: The Nature of Condition C.

James Higginbotham  
Some Remarks on Binding Theory and Logical Form.

A list of the participants follows the table of contents. The decision to publish the proceedings was not made until after the conference, and as a result many of the participants were unable to prepare written versions of their contributions.

The organizers wish to extend their thanks to the participants for their stimulating papers, to Margaret Milliken, Stuart Milliken, Hilary Sachs and Laurie Zaring for assisting with numerous details of organization during the conference, and to Sheila Haddad for her help in preparing the manuscript.
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>John S. Bowers</td>
<td>Parasitic Gaps</td>
</tr>
<tr>
<td>Robin Clark</td>
<td>Clitics as $\text{~A}$-Binders</td>
</tr>
<tr>
<td>Steven Franks</td>
<td>Case and Control in Polish</td>
</tr>
<tr>
<td>Isabelle Haïk</td>
<td>Indirect Binding and Referential Circularity</td>
</tr>
<tr>
<td>Randall Hendrick</td>
<td>Corroboration for the ECP</td>
</tr>
<tr>
<td>James Higginbotham</td>
<td>Some Remarks on Binding Theory and Logical Form</td>
</tr>
<tr>
<td>C.T. James Huang</td>
<td>LF, ECP and Non-vacuous Quantification</td>
</tr>
<tr>
<td>John Jensen</td>
<td>Latin Passive Without NP Movement</td>
</tr>
<tr>
<td>Juliette Levin</td>
<td>Free Relatives and the Null Head Parameter</td>
</tr>
<tr>
<td>Judith McA’Nulty</td>
<td>$\text{~A}$-Binding</td>
</tr>
<tr>
<td>Carlos Piera</td>
<td>Spanish Comparatives, Deletion Operations and the ECP</td>
</tr>
<tr>
<td>Gilbert Rappaport</td>
<td>On Anaphora and Control in Russian</td>
</tr>
<tr>
<td>Margarita Suñer</td>
<td>Free Relatives and the pro-Head Hypothesis</td>
</tr>
<tr>
<td>Name</td>
<td>Title</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>John Bowers</td>
<td>Parasitic Gaps</td>
</tr>
<tr>
<td>Robin Clark</td>
<td>Clitics as $\bar{A}$-Binders</td>
</tr>
<tr>
<td>Lori Davis</td>
<td>Distribution of Pleonastic NPs</td>
</tr>
<tr>
<td>Steven Franks</td>
<td>Case and Control in Polish</td>
</tr>
<tr>
<td>Jacqueline Guéron</td>
<td>On so-called Articles in French</td>
</tr>
<tr>
<td>Isabelle Haïk</td>
<td>Indirect Binding and Referential Circularity</td>
</tr>
<tr>
<td>Wayne Harbert</td>
<td>Inverse Attraction, the Overt Head Requirement and Matching</td>
</tr>
<tr>
<td>Randall Hendrick</td>
<td>Corroboration for the ECP (Paper not presented at conference)</td>
</tr>
<tr>
<td>James Higginbotham</td>
<td>Some Remarks on Binding Theory and Logical Form</td>
</tr>
<tr>
<td>C.T. James Huang</td>
<td>LF, ECP and Non-vacuous Quantification</td>
</tr>
<tr>
<td>John Jensen</td>
<td>Latin Passive without NP Movement</td>
</tr>
<tr>
<td>Howard Lasnik</td>
<td>Binding or Pragmatics?: The Nature of Condition C.</td>
</tr>
<tr>
<td>Juliette Levin</td>
<td>Free Relatives and the Null Head Parameter</td>
</tr>
<tr>
<td>Maria Rita Manzini</td>
<td>A Theory of Control and its Consequences</td>
</tr>
<tr>
<td>Robert May</td>
<td>On Some Applications of the ECP to LF and the Notion of c-command</td>
</tr>
<tr>
<td>Judith McA'Nulty</td>
<td>$\bar{A}$-Binding</td>
</tr>
<tr>
<td>Carlos Piera</td>
<td>Spanish Comparatives, Deletion Operations and the ECP</td>
</tr>
<tr>
<td>Gilbert Rappaport</td>
<td>On Anaphora and Control in Russian</td>
</tr>
<tr>
<td>Claudia Ross</td>
<td>AGR and the Properties of Empty Categories</td>
</tr>
<tr>
<td>Timothy Stowell</td>
<td>The Thematic Grid</td>
</tr>
<tr>
<td>Margarita Suñer</td>
<td>Free Relatives and the pro-Head Hypothesis</td>
</tr>
</tbody>
</table>
Parasitic Gaps

John S. Bowers

Cornell University

One of the most exciting consequences of the recent shift in emphasis from the elaboration of rule systems to the formulation of general principles is that it has become possible for the first time to propose and defend explicit hypotheses concerning the fundamental empirical questions of linguistic theory. One of the most fundamental problems of theoretical linguistics is determining which properties of grammatical systems are learned on the basis of experience and which are determined by the genetically transmitted structure of the mind/brain. To the extent that the principles proposed by linguists succeed in characterizing in an illuminating fashion the range of structures that actually occur in the languages of the world, to that extent one is justified in assuming as a tentative hypothesis that these principles embody, in an abstract form, real properties of the mind. Those properties of grammatical systems that cannot be reduced to general principles, on the other hand, we can tentatively assume to be learned on the basis of experience.

Suppose that we succeed in coming up with a set of principles that appear to characterize in an interesting way the fundamental structure of human languages. We can then ask a further question: Are these principles purely linguistic, in the sense that they characterize a special linguistic ability, distinct in kind from other cognitive abilities that human beings are born with, or are they manifestations of more general cognitive
abilities, ones that are perhaps used in other tasks besides the acquisition of grammar? A priori there is no way of knowing the answer to this question. It is conceivable, for instance, that there are no specifically linguistic principles involved in the acquisition of language. On the other hand, it might turn out that the principles that permit human beings to acquire language constitute a separate "organ" of the mind that is for all practical purposes autonomous and distinct from other aspects of cognition.

These two questions - determining what is innate and what is learned, on the one hand, and determining the domain of the innate principles, on the other - are in principle independent of one another. From the fact that a certain principle appears to be part of the human organism's genetic endowment, one cannot necessarily conclude that it is part of universal grammar (meaning by this a theory of the innate faculté de langage), since what appears to be a purely linguistic principle might, on further investigation, turn out to be nothing more than a special application of some more general cognitive ability to the specific data of linguistic experience. To provide a conclusive answer to the domain question it is not sufficient to consider only a narrow range of data that is, in some sense, "purely linguistic." Rather, we must consider how principles discovered on the basis of linguistic data are to be integrated into a more comprehensive theory of mind. To take a relatively straightforward example, consider the recently discovered fact that very young infants exhibit categorical discrimination of speech sounds. This fact could, in principle, be due to a genetically transmitted perceptual ability that is specifically linguistic, i.e. limited to the perception of speech sounds, and therefore unrelated to other perceptual abilities of human beings. On the other hand, it could equally well be just a special application of innate perceptual abilities that permit infants to discriminate all sorts of acoustic information, or even perceptual information from different sensory modes. In this case, as in every case, it is an empirical question what the domain of applicability of the hypothesized innate ability is.

In the case just mentioned, it is not difficult to imagine ways of going about deciding which of the two alternatives is correct. However, when we consider more abstract properties of language, the task of devising ways of answering the domain question becomes considerably more problematical. Even here, though, it may be possible to adduce evidence in favor of one or another of the theoretically possible alternatives. Thus Jackendoff (to appear) suggests, for example, that the principles of grouping that underlie phrase-structure analyses of syntax are not specifically linguistic, but reflect much more general properties of cognition. As support for his view, Jackendoff points
out that similar principles are necessary to explain how people group phrases in musical compositions and also to explain how they group items in visual displays. It would argue for the narrowly linguistic character of some hypothesized principle, on the other hand, if it could be shown that it was highly specific to the kinds of structures that are necessary in linguistic description and that similar principles did not seem to be necessary in other areas of cognition. Some of the principles proposed in recent versions of transformational grammar would seem, on the face of it, to have this character. A principle such as subjayency, for instance, or the binding conditions proposed in Chomsky's recent work, are stated in terms of the concepts and structures of syntax and appear to have no obvious analogue in other areas of cognition.

On the other hand, even principles that look at first glance as if they are highly specific to language might turn out, on closer inspection, to be particular applications of principles that are needed to explain other cognitive abilities. An obvious place, one would think, to look for such principles is in the area of thought. There is surely no doubt that thought occurs independently of language and it is reasonable to assume that there must exist severe constraints on what Fodor (1975) has termed the "language of thought." The possibility is therefore open that what appear to be principles of UG are actually just reflexes of deeper cognitive constraints governing the language of thought. This possibility is especially likely in the case of putative constraints on semantic representation, as Jackendoff again has pointed out. In general, it is safe to assume that the more 'syntactic' the principles governing linguistic form are, the less likely it is that they can be reduced to deeper principles governing the structure of thought. The more 'semantic' the constraints on linguistic form are, the more likely it is that they are just a projection of such principles.

Chomsky (1981b) has suggested recently that the phenomenon of 'parasitic gaps' may provide crucial evidence for the existence of innate principles of universal grammar. The following quotation makes clear the general line of argument he has in mind: "Surely one would not expect that the parasitic gap phenomenon is sanctioned by special principles of UG or that there is a component of UG dealing specifically with the properties of these constructions. Furthermore, it is highly implausible to suppose that the basic properties are learned from exposure to examples identified as acceptable or unacceptable...What we expect, then, is that the properties of the parasitic gap constructions, and the very existence of the phenomenon, will reduce to independently established principles of UG." Chomsky's claim, then, is that all the properties of parasitic gaps can be derived from the interaction of independently motivated principles of UG. It
follows that speakers of English do not learn any specific rules
governing constructions with parasitic gaps, nor are there even
any principles of UG that deal specifically with these construc-
tions. Rather, the existence and the special properties of para-
sitic gaps are just a side-effect, as it were, of principles
that are needed for other reasons entirely.

The form of argument that Chomsky is trying to develop here
is of potential importance for a number of reasons. First of all,
it would demonstrate unequivocally that there exist abstract pro-
erties of grammars that cannot be derived inductively from the
data of linguistic experience, but must be the result of innate
principles that guide the process of grammar-construction.
Secondly, if there are indeed no principles of UG that deal
specifically with parasitic gap constructions, then the existence
of the phenomenon provides a particularly dramatic confirmation
for the existence of abstract principles whose interaction is
capable of producing such unexpected results. Finally, an argu-
ment of this form would of course lend strong support to the
particular principles of Government-Binding theory, from which
it is claimed the properties of parasitic gaps can be derived.

In this paper I shall try to accomplish two things. First,
I shall argue that Chomsky's analysis of parasitic gaps is incor-
rect, while agreeing with the general claim that there are no
specific rules or principles governing their distribution. In
particular, I shall show that the parasitic gap is not a variable
that is 'licensed' under certain conditions to appear in positions
where variables are normally excluded. Rather, it is just the
null form of a theoretically possible type of pronoun whose exis-
tence has apparently not been noticed before. Second, I shall
discuss the domain problem with regard to the typology of pronouns
and anaphors, attempting to show that the fundamental properties
that distinguish pronominal types are semantic, rather than syn-
tactic. This result will lead me to suggest that the principles
governing the distribution of pronouns are not principles of UG,
but rather are principles that must be assumed in any case if we
want to characterize the language of thought.

I start out with some interesting observations in a recent
paper by Bosch (1981). Bosch points out that there are two
different types of pronoun in English that differ quite systema-
tically. The first type he calls 'referential pronouns' (RP)
and the second he calls 'syntactic agreement pronouns' (SP). RPs
are distinguished from SPs by two fundamental properties: (1) pro-
noun-like NPs such as the fool, the idiot, the nerd, etc. can
substitute for RPs, but not for SPs; (2) SPs, but not RPs, can
have as their antecedents quantified NPs such as no one, everyone,
etc. Bosch illustrates these contrasts with data of the following
sort:
(1) a. John₁ said that he would go.  
    *the idiot₁
b. John₁ left, and then he came back again.  
    *the idiot₁
(2) a. No one₁ admitted that he was the thief.  
    *the fool₁
b. No one₁ left, and then *he came back again.  
    *the fool₁

Consider now the positions in which parasitic gaps typically appear. A representative set of examples is contained in (3):

(3) a. Which article did John file t without reading e?  
b. This is the kind of food you must cook t before you eat e.  
c. Here is the influential professor John sent his book to t in order to impress e.  
d. Who is John as pleased with t as he is angry at e?  
e. This is the type of book that no one who has read e would give t to his mother.  
f. He is a man whom everyone who meets e admires t.  
g. Who did your interest in e surprise t?  
h. John offended t by not recognizing e immediately, his favorite uncle from Cleveland.

It is a remarkable fact that the overt pronouns that can appear in parasitic gap positions behave uniformly like RPs, rather than SPs, as is shown by data of the following sort:

(4) a. John filed the article₁ without reading it.  
    the sucker₁
b. You must cook this kind of food₁ before you eat it.  
    the junk₁
c. John sent his book to the influential professor₁ in order to impress him.  
    the idiot₁
d. Bill is as pleased with Mary₁ as he is angry at her.  
    the bitch₁
e. No one who has read it would give this book₁ the thing₁ to his mother.  
    f. Everyone who meets him₁ admires John₁.  
    the s.o.b._₁
g. Mary's interest in him₁ surprised John₁.  
    the poor boy₁
h. John offended his favorite uncle from Cleveland₁ by not recognizing him₁ the old' coot₁.
(5) a. *John filed every article without reading it.

b. *John warned everyone twice before he reported him.

c. *Bill sent his book to no one in order to impress him.

d. *Bill is as pleased with everyone as he is angry at him.

e. *A man who has read it wouldn't give anything that crap to his mother.


g. *Mary's interest in him surprised no one.

h. *John offended no one by not recognizing everyone.

This fact can hardly be an accident, suggesting that parasitic gaps are simply RPs - whatever they might be - in their null form. The problem of determining the distribution of parasitic gaps thus reduces to the problem of determining the distribution of RPs.

What, then, is an RP? The fact that RPs (unlike SPs) can be full NPs as well as pronouns suggests, naturally enough, that they are simply R-expressions. We could then regard the parasitic gap as the null pronominal form of the class of R-expressions. Consider next the distribution of RPs and SPs. The basic condition governing RPs (and hence, by hypothesis, parasitic gaps) is that the antecedent not c-command the RP, or conversely. SPs, in contrast, must be coreferential with a c-commanding category. In other words, SPs must be bound, whereas RPs are free. The second condition is of course just Chomsky's Condition C, the basic binding condition for R-expressions. However, in current GB theory pronouns are governed by Condition B, which requires only that they be free in their governing category. In other words, the current theory groups together all occurrences of pronouns, whether anaphoric or deictic, under Condition B and puts lexical nouns into the separate class of R-expressions. I am proposing, in contrast, to group the deictic use of pronouns together with lexical nouns into the class of R-expressions, leaving those pronouns that are obligatorily bound to form a separate class of SPs (in Bosch's terminology). Notice that under this proposal the distribution of RPs, and hence of parasitic gaps, follows automatically from Condition C. The reason is that a pronoun can be free in one of two ways. It can either fail to be coindexed with
anything, which is simply the deictic use of pronouns, or it can be coindexed with a non-c-commanding category, which is just the basic property of RPs and parasitic gaps. However, in order to fully motivate this proposal we must consider the semantic interpretation of SPs and RPs, respectively.

It was observed some time ago by Partee (1973), following a suggestion by Geach (1968), that pronouns are of two fundamentally different types. Certain pronouns, namely, those that can be bound by a quantifier, act like logical variables, while others, which Geach termed "pronouns of laziness", simply act as substitutes for nominal expressions. This second type of relation is what Lasnik (1976) has termed 'accidental coreference' and Evans (1980) 'coreference with no dependency.' It is immediately evident that what Bosch calls SPs are simply bound variables, since they can be bound by quantifiers, while RPs are just pronouns of laziness, since they can't. Furthermore, as Partee (1973) showed, both controlled PRO and anaphors must be interpreted as bound variables, since they too can be bound by quantifiers. Looked at from this point of view, the fundamental distinction is between bound variables, on the one hand, which cannot have independent reference, and which are interpretable only by being bound to a quantified expression, and lexical nouns and deictic pronouns, on the other, which are independent in reference, but which may, under the appropriate conditions, exhibit accidental or pragmatic coreference. Notice that the only difference between anaphors and the bound variable interpretation of pronouns, according to this theory, is that the former must be bound in the lowest S or NP domain containing them, whereas the latter must be bound outside this lowest category. It is therefore not surprising to find, as Bouchard (1982) notes, that in some languages both anaphors and pronouns have the same lexical realization.

We have, then, two classes of expressions which differ in both syntactic distribution and meaning. SPs and anaphors are interpreted as bound variables and must be bound by a c-commanding category. RPs and R-expressions, on the other hand, can either refer outside the sentence containing them or be coindexed with a non-c-commanding category and are never interpreted as bound variables. Notice that each of these fundamental types has a corresponding null form. The null form for bound variables is simply controlled PRO. In English, the null form for RPs occurs only in parasitic gap constructions, i.e. it is restricted to positions where it is coindexed with a non-c-commanding variable. However, in many languages (e.g. Japanese, Chinese, Sinhala, Walbiri, etc.) the null form of referential pronouns can occur quite freely in any syntactic position, its interpretation being pragmatically determined by the context in which it is uttered (cf. Huang (1982), for extensive discussion of this phenomenon in Chinese). Hence a null pronoun of this type is needed for such
languages in any case, quite independently of the facts regarding parasitic gaps in English.

The preceding observations, if correct, suggest rather strongly that the fundamental criteria for distinguishing pronominal types are semantic, rather than syntactic. Specifically, let us suppose that nominal expressions are of two types: referential and non-referential (±R). [-R] expressions are subject to the general condition that they must be bound in LF, while [+R] expressions are free in LF. RPs and lexical nouns are [+R], while SPs and anaphors, are [-R]. The properties discussed earlier follow at once from these simple conditions. Referential pronouns, being free, can't be coindexed with a c-commanding category. This condition can be satisfied in one of two ways: (1) a pronoun can fail to be coindexed at all, in which case its reference is determined pragmatically, or (2) it can be 'accidentally' coindexed with a non-c-commanding category. The latter is the case of RPs, including parasitic gaps. Non-referential pronouns, on the other hand, must be bound. If they are bound in their governing category, they are anaphors. If they are bound outside their governing category, then they are pronouns (including PRO).4

Returning now to the domain question, suppose that the distinction between referential and non-referential expressions, as well as notions such as 'free', 'bound', etc. are not part of UG, but belong instead to that part of cognitive theory that is concerned with the language of thought. Not only is this a reasonable assumption, but it is in fact quite difficult to imagine a representation of the language of thought that failed to make use of such notions. It follows that the task of a human being who is trying to acquire a language (say, English) is to figure out how the actually occurring phonetic forms of English to which he is exposed are related to the fundamental cognitive structures permitted by the principles governing the language of thought. Let us call the level of representation that mediates between surface phonetic forms, on the one hand, and the structures of thought, on the other, the level of 'logical form' (LF). It follows that representations at the level of LF are simply a projection onto the data of English of cognitive principles governing the structure of thought.5 If this view can be maintained, then it may well turn out that the fundamental structures (as opposed to the accidental and idiosyncratic features) of human languages are determined by principles that are not exclusively, or even primarily, linguistic. To some this may seem like a negative conclusion. The positive side of it is, however, that in attempting to discover and formulate general principles governing the structure of natural language, we can be assured that we are approaching an understanding of the fundamental principles of cognition that underlie human knowledge and behavior.
FOOTNOTES

1 It follows that variables, in the sense of GB theory, i.e. traces bound by a category in a Α-position, cannot be treated as R-expressions. For a detailed justification of this view, see Bowers (1983a, 1983b).

2 This view requires that definite as well as indefinite NPs be treated as operators. That such an analysis is not only possible, but indeed required by certain facts concerning sloppy vs. non-sloppy identity, is demonstrated in Reinhart (1983). Reinhart's treatment of bound anaphora and coreference, which became available to me after revising this paper, parallels my own in most respects. I disagree with her view, however, that the cases where non-coreference is required can be explained by means of pragmatic principles. The judgments of native speakers are much too firm to make a pragmatic approach at all plausible.

3 It follows that PRO is not unguarded, as in the standard GB typology of null elements. See Bouchard (1982), Bowers (1983), for independent arguments in support of this view.

4 Bouchard (1982) has suggested that Chomsky's (1982) functional definition of empty categories should be extended to the overt pronominal categories. Such an approach is implicit in the analysis suggested here, as well.

5 In order to justify an independent theory of UG, it would have to be shown that there are one or more levels of syntactic representation, governed by principles that are purely syntactic. In GB theory, the levels of D-structure and S-structure are levels of just this sort. Likewise, in such a framework principles, such as the binding conditions, subadjacency, and so forth, are conceived of as syntactic principles. For arguments that all of these principles can be better stated at the level of LF, see Bowers (1983).

REFERENCES


Partee, B. (1973) Deletion and variable binding. Linguistic Agency University of Trier.

Clitics as $\bar{A}$-Binders

Robin Clark

U.C.L.A.

For the purpose of this paper, I will assume an informal definition of "clitic" which, while clearly too broad, has been implicit in much recent work on clitics. Specifically, a clitic is an element which, when present, allows an apparent violation of the complementation requirements of a head. Recall that the Projection Principle (Chomsky, 1981) necessitates the satisfaction of lexical requirements at every level of representation. It follows that a clitic is an element which allows an empty category to satisfy the lexical requirements of a head element. For example, although the French verb "voir" (see) is transitive, "Jean l'a vu" (Jean saw it) is grammatical because the clitic "l" allows the NP complement of "voir" to be empty. Minimally, an adequate account of clitics must specify the nature of this empty category and the relationship between the clitic and the empty category.

In terms of recent proposals, perhaps the most thorough and articulately argued is that of Borer (1981). Borer takes clitics to be a particular morphological operation on the head element of a phrase. The Case feature of the head is spelled out as a group of features made up of Case, gender, person and number (features typical of pronominals):

(1) Clitic Spell-Out (Borer, 1981)

\[
[X, [\alpha \text{Case}]] \rightarrow [X, [\alpha \text{Case}, \alpha \text{Gender}, \alpha \text{Number}, \alpha \text{Person}]]
\]

The clitic itself is part of the feature matrix of the head element.
The spell-out of the Case feature of the head renders the head incapable of assigning Case to a complement NP.

The assumption that the complement NP does not receive Case when the clitic is present has two consequences. First, since the empty category lacks Case (and, hence, is invisible at LF), it cannot be a variable.\footnote{Further, since the empty category is, plausibly, governed by the head, it cannot be PRO. Given a three-way partition of the set of empty categories into PRO, variable and NP-trace, the empty category must be NP-trace. The second consequence is that, if the clitic is present, the complement NP position cannot contain lexical material unless the NP can receive Case from some element other than the head.} Further, since the empty category is, plausibly, governed by the head, it cannot be PRO. Given a three-way partition of the set of empty categories into PRO, variable and NP-trace, the empty category must be NP-trace. The second consequence is that, if the clitic is present, the complement NP position cannot contain lexical material unless the NP can receive Case from some element other than the head.

The latter consequence accounts for the complementary distribution between clitics and lexical material in the complement position found in some languages (e.g., Standard French). It also derives "Kayne's Generalization" (Jaeggli, 1980) that when a language allows the clitic to occur with a lexically-filled NP (the phenomenon of "clitic doubling"), the complement NP must be governed by a preposition which will assign it Case. Suppose that, in some structure in a clitic doubling language, the clitic and a lexical complement NP were both present, but that the lexical NP is not governed by a preposition. Since, by assumption, the clitic has absorbed the Case feature of the head, the complement NP has no means of receiving Case. Such a structure would be ruled out by the Case Filter (or, alternatively, by the criterion under the Visibility Hypothesis).

A number of languages which allow clitic doubling support the above analysis. Hebrew, River Plate Spanish and Rumanian, for example, all require that a preposition govern the complement NP in clitic doubling structures. We can account for these languages by claiming that, although the clitic has absorbed the Case normally assigned to the complement NP, these languages have a rule of Dummy Case Marker Insertion (DCMI) which inserts a preposition that will assign Case to the complement NP:

\[
\begin{align*}
\text{(2) Dummy Case Marker Insertion (DCMI)} \\
\emptyset \rightarrow \text{Dummy Case Marker}/ \left[ \_{\text{X}} \text{ NP} \right] \\
\end{align*}
\]

\[
\begin{array}{c}
\text{XP} \\
\text{X} \quad \text{N}
\end{array}
\begin{array}{c}
\text{XP} \\
\text{X} \quad \text{N} \\
\text{DCM} \quad \text{N}
\end{array}
\]

It follows that clitic doubling is possible only if the grammar in
question contains the appropriate rule of DCMI.

The presence of clitics and the presence of DCMI are independent parameters. Thus, there are four logical possibilities:

(3) Clitics DCMI
    I   +   +
    II  +  -
    III -  +
    IV  -  -

The predictions made by Borer's analysis are explicit and testable. Languages of type I should be able to allow clitic doubling while type II languages should be incapable of clitic doubling under any circumstances (type III and type IV languages are, of course, not relevant for our purposes).

With the above predictions in mind, consider the following sentences in Palauan: 2

(4a) Ak me-nglelebed a bilis.
    I RM hit CM dog
    "I hit a dog."

b) Ak cholebed-i\textsubscript{ij} [NP\textsubscript{j} e].
    I hit-it
    "I hit it."

c) A sensei a cholebed-au\textsubscript{ij} [NP\textsubscript{j} e].
    CM teacher CM hit-you
    "The teacher hit you."

d) A sensei a cholebed-ak\textsubscript{ij} [NP\textsubscript{j} e].
    CM teacher CM hit-me
    "The teacher hit me."

The verb meaning "hit" requires a following NP, as can be seen in (4a). In (4b), the element "-ii" allows the complement NP to be empty. Since the element "-ii" licenses an apparent violation of the Projection Principle, it meets our informal definition of a clitic (above). The theory of clitics would therefore treat "-ii" as a spell-out of the verb's Case feature, rendering the verb incapable of assigning Case.

Consider, now the sentences in (5):

(5a) Ak cholebed-ii a bilis\textsubscript{j}.
    I hit-it CM dog
"I hit the dog."

b) *Ak cholebed-ii er a bilis.
   I hit-it Prep CM dog
   "I hit the dog."

c) Ak me-ngelbed er a bilis.
   I RM hit Prep CM dog
   "I hit the dog."

Sentence (5a) should come as something of a surprise since both the clitic and a lexical NP are present, but there is no Case assigning element (other than the verb) present to assign Case to the doubled NP. Sentence (5b) shows that, in structures where a preposition is present to assign Case to the doubled NP, the sentence is ungrammatical. This situation is exactly what the theory outlined above predicts to be impossible. In general, however, the preposition "er" may be freely inserted before direct objects when a clitic is not present, as in (5c).

Insertion of the preposition "er" does have an interesting effect, as illustrated in (6):

(6a) Ak ulemes a bsibs.
    I saw CM termite
    "I saw a termite."

b) Ak ulemes er a bsibs.
    I saw Prep CM termite
    "I saw the termite."

In (6a), the direct object is interpreted as indefinite, bound by an existential quantifier at LF. The broadest statement that can be made is that, in the mapping from s-structure to LF, Quantifier Raising (QR) applies to NPs associating each A-position with an existential quantifier where possible. The subject and the object of a preposition normally receive definite interpretations. This follows from the Empty Category Principle (ECP, see Chomsky, 1981 for discussion), since Palauan prepositions are not proper governors, and the subject position can only be properly governed, under certain circumstances, by INFL. The empty category left by QR from these positions will not be properly governed and, hence, such structures will be eliminated by the ECP. The definite singular interpretation of the direct object in (5b) follows from this analysis. The particle "er" is not a proper governor, hence, if QR were to apply, the empty category would not be properly governed. QR cannot associate this position with an existential quantifier.
CLITICS AS $\bar{A}$-BINDERS

Significantly, the direct object in a sentence with a clitic has a definite interpretation, just as when the direct object is governed by "er" (as in (5)), despite the complementary distribution of clitics and the particle "er." This is suggestive of a solution to the problem of clitic doubling in Palauan.

As a first step, let us suppose that clitics in Palauan do not result from a sepp-out of the Case feature of the verb, but are, rather, independent of the verb in some sense. In this case, there should be no reason to assume that the verb's ability to assign Case is in any way altered by the presence of a clitic. This is the desired result for Palauan where the direct object may be doubled without the support of a dummy Case marker.

Given the above, the direct object is permitted to be lexically filled in the presence of a clitic, but, as we have seen, the direct object may be empty when (and only when) the clitic is present. Since we are assuming that the position is Case marked, we can take the empty category to be a variable. 4 Adapting an idea from Aoun (1981), I will assume that a well-formedness condition on variables is that they be $\bar{A}$-bound. Notice that, given our assumption on QR, the position will be bound by an existential quantifier at LF, in any event. This ignores the problem of how the definite interpretation of these NPs is derived. I will return to this problem below.

Notice that the variable will be $\bar{A}$-bound at LF, but we have not yet guaranteed that it will be $\bar{A}$-bound at s-structure. In order to give an adequate account of clitics, we must have some means of expressing the relation between the clitic and the complement NP position. An obvious means of accomplishing this task is provided by coindexing. The complement NP position in Palauan may be either a variable or lexical; hence, it counts as an R-expression with respect to the Binding Theory which requires R-expressions to be A-free. If the relationship between the clitic and the complement NP is expressed by coindexing, it follows that the clitic must be in an $\bar{A}$-position and $\bar{A}$-bind the R-expression. 5 If the clitic occupied an A-position, the result would violate the Binding Theory. Since the clitic is present at s-structure (and a d-structure), the variable will be A-bound at s-structure as required.

We have so far noted three outstanding properties of Palauan clitic constructions. First, Palauan clitics do not absorb Case. Second, the presence of the clitic forces the definite interpretation of the NP that the clitic binds. Finally, the clitic must occupy an $\bar{A}$-position. Following the idea that the clitic $\bar{A}$-binds the complement NP, we can treat the clitic as, in some sense,
an operator (perhaps as a "specificity" operator).

I believe that there is a more general approach to clitics in Palauan that includes the definite interpretation of the direct object while capturing the complementary distribution between clitics and the particle "er." Recall that QR applies quite generally to NPs in Palauan, although the direct object is normally the only position from which QR is allowed by the ECP. Suppose that the clitic is a special type of operator which converts the existential quantifier into a specificity (or, perhaps, a cardinality) operator. In this treatment, the clitic requires an existential quantifier in order to be well-formed. The clitic and the existential quantifier must, of course, λ-bind the same variable.

By way of exemplification, consider the derivation of (7) ("I hit the dog") from s-structure to LF:

\[(7) \, [S\, [s\,ak \, INFL \, [VP\,cholebed-\,ii_j \, [NP_j \, a \, bilis]]]]\]

First, QR will apply to NP_j, creating an existential quantifier:

\[(7') \, [\exists x_j \, e \, dog \, [s\,ak \, INFL \, [VP\,cholebed-\,ii_j \, x_j]]]\]

Next, QR applies to the clitic:

\[(7'' ) \, [S\,cl_j \, \exists x_j \, e \, dog \, [s\,ak \, INFL \, [VP\,cholebed \, x_j]]]\]

Finally, the clitic and the existential operator combine to create the specificity operator:

\[(7''' ) \, [S\,for \, unique \, x_j \, e \, dog \, [s\,ak \, INFL \, [VP\,choleged \, x_j]]]\]

In contrast, consider the analogous mapping for the ungrammatical sentence (8):

\[(8) \ast \, [S\, [s\,ak \, INFL \, [VP\,cholebed-\,ii_j \, [PP\,er \, [NP_j \, a \, bilis]]]]]\]

QR cannot apply to NP_j since the empty category left behind would not be properly governed. QR can, however, apply to the clitic:

\[(8') \, [S\,cl_j \, [s\,ak \, INFL \, [VP\,cholebed \, [PP\,er \, [NP_j \, a \, bilis]]]]]\]
Since we are treating the clitic as an operator, we can assume that it must bind at least one variable at LF. Since the direct object is in situ, the clitic will not have a variable to bind. Furthermore, the clitic can only be interpreted in tandem with an existential quantifier. Since QR could not apply to the direct object, there is no existential quantifier for the clitic to work with. Sentence (8) is ruled out in two ways.  

The ungrammaticality of (8) was accounted for, in part, by appeal to the Bijection Principle. In its strongest form the Bijection Principle requires a one-to-one correspondence between variables and operators. Although this formulation of the principle is clearly too restrictive (see, e.g., Chomsky, 1982, for some discussion), we must have some constraints on the variable-operator relation. One plausible requirement is that there be no vacuous quantification. Each operator must bind at least one variable. Equally plausible is the requirement that no variable be unbound. This latter requirement could be captured by some form of the Binding Theory, perhaps along the lines of Aoun (1981). As stated above, I assume that the requirement that variables be locally A-bound is independently necessary. We might, furthermore, require that when two A-binders bind the same variable, they must recombine to form a single complex operator. One consequence of this requirement might be that the new complex operator will bind the union of the sets of variables bound by the original operators. It is unclear to me what consequences this requirement might have.

The requirement that variables be locally A-bound conjoined with the ECP has the result that, in Palauan, the only position available to be bound by the clitic is the direct object position governed by the verb that the clitic is attached to. A variable in any other position will either not be locally A-bound by the clitic at s-structure, or will not be properly governed.

We have achieved the desired result for Palauan by treating the clitic as an element that is independent of the head in that it does not absorb the Case feature of the head. Certain facts about QR and the Binding Theory led to an interpretation of the clitic as an A-binder which forces, in conjunction with an existential quantifier, a definite interpretation of the position that it binds.

A natural question is whether clitics can, in general, be treated as A-binders. Notice that we shall have to preserve the insight that clitics absorb Case in some languages in order to capture the complementary distribution that generally holds between clitics and lexical complements (not to mention DCMI in many clitic
doubling languages). One parameter of Universal Grammar must be involved with the distinction between Case absorbing clitics and non-Case absorbing clitics.

Furthermore, if the notion of $\overline{A}$-binder" is taken to be co-extensive with the notion "operator," then we will probably be unable to claim that all clitics are $A$-binders. If we accept the equation "$A$-binder=operator," then we would expect that clitics would quite generally license parasitic gaps, which is contrary to fact. There are two possible solutions to this. One could claim that operators are a subset of the set of $A$-binders and that only operators license parasitic gaps. Thus a clitic could be an $A$-binder, but not an operator and would not be eligible to license parasitic gaps.

A second possibility lies in the fact that parasitic gaps typically occur in adjunct positions and that the operator that allows the parasitic gap must c-command the parasitic gap as s-structure. Even in Palauan, the clitic will not appear in COMP until LF, at s-structure the clitic will be attached to the verb (I am ignoring genitive clitics for ease of exposition). One could claim that the verb, and therefore the clitic, does not c-command the adjunct at s-structure, so it could not possibly license a parasitic gap inside the adjunct. This solution, of course, presupposes a detailed analysis of adjuncts of certain form. In any event, it seems clear that the parasitic gap issue does not necessarily prevent us from treating clitics as $A$-binders in general.

We have seen that clitics cannot be uniformly treated as a spell-out of the Case feature of a head element, and that a clitic may behave like an operator in some languages. The basic distinction might be captured in a parameter that specifies the feature ($+$ Operator), where ($-$ Operator) implies that the clitic does not absorb Case. I will leave the full ramifications and testing of this prediction for future research.

**FOOTNOTES**

*I would like to acknowledge Hagit Borer, Sandy Chung, Mike Hammond, Maryellen MacDonald, Mario Montalbetti, Paul Schachter, Tim Stowell and Tracy Thomas-Flinders for all their help. Any slip-ups or stupid statements are, of course, my complete responsibility.

1The assumption that variables are distinguished from NP-traces by virtue of a Case feature is no longer standard. I adopt
here solely for expository convenience.

2Palauan is a language spoken in Micronesia. Most of the examples in the article are from Josephs (1975). In the glosses, the CM (category marker) "a" is an element which appears before non-pronoun NPs and before the VP when the subject is non-pronominal. The RM (Realis marker) "me-" marks realis mood on verbs. The full paradigm of direct object clitics is:

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Prsn</td>
<td>-ak</td>
<td>-id (inclusive)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-emam (exclusive)</td>
</tr>
<tr>
<td>2nd Prsn</td>
<td>-au</td>
<td>-emiu</td>
</tr>
<tr>
<td>3rd Prsn</td>
<td>-ii</td>
<td>-terir</td>
</tr>
</tbody>
</table>

3Palauan has a set of subject clitics associated with hypothetical aspect. When the subject clitic is present, the subject itself is optionally empty. See Josephs (1975) for details.

4Recently (e.g., Chomsky, 1982), it has been argued that clitic-trace is 'pro.' As will be seen below, the empty category in question is treated as a variable at LF. I will leave open the question of whether empty categories may change their status with respect to the Binding Theory from one level of representation to another and treat the clitic-trace in Palauan as a variable both at s-structure and LF.

5I will leave the question of how to represent this configurationally open since it goes beyond the scope of this paper.

6One could, equally, claim that the clitic triggers reconstruction at LF. The NP would be put back in place and the definite interpretation would result from its being in situ at LF. I will not explore the consequences of this analysis.

7This process is, perhaps, related to "absorption" as defined in Higginbotham & May (1981).

8Strictly speaking, this "er" is not the same as the preposition "er." For finer details of this analysis, I can only refer the reader to Clark (1982).

9It may be the case that the clitic leaves behind a trace which it could bind. This empty category, if it exists, would be that of an operator, by assumption. We can suppose that operators, in order to be intelligible, must bind a non-operator—i.e., a pure nominal. Otherwise, we could conceivably admit the existence
of quantifiers which quantify only themselves.

10This relation is, quite plausibly, highly constrained since the language acquisition device has no direct evidence of the nature of LF representations.

11See Chomsky (1982) and the references cited therein for an extensive discussion on the nature of parasitic gaps.
CLITICS AS $\Lambda$-BINDERS

REFERENCES

Aoun, J. (1981); The Formal Nature of Anaphoric Relations; MIT dissertation

Borer, H. (1981); Parametric Variation in Clitic Constructions; MIT dissertation

Chomsky, N. (1981); Lectures on Government and Binding; Foris Publications, Dordrecht, the Netherlands

Chomsky, N. (1982); Some Concepts and Consequences of the Theory of Government and Binding; The MIT Press; Cambridge

Clark, R. (1982); "Palauan Clitics"; UCLA Master's Thesis

Higginbotham, J. & R. May (1981); "Questions, Quantifiers and Crossing"; The Linguistic Review 1:41-80

Jaeggli, O. (1980); On Some Phonologically Null Elements in Syntax; MIT dissertation

Josephs, L. (1975); Paulauan Reference Grammar; The University Press of Hawaii; Honolulu

Kayne, R.S. (1975); French Syntax; MIT Press; Cambridge

May, R. (1977); The Grammar of Quantification; MIT dissertation

Montalbetti, M. (1981); Origins of Phrase Structure; MIT dissertation
Case and Control in Polish

Steven Franks
Cornell University

This paper is concerned with providing syntactic evidence that can be brought to bear in evaluating a theory of control. Any such theory should have mechanisms powerful enough to account for the case agreement I am about to discuss. It should have different enough ways for establishing control to be able to draw the line between different kinds of control in the right places. Here I suggest one of those places.

1.1. The theory of empty categories is an essential part of Chomsky's Government and Binding model of grammar. Let me sketch some of its primary aspects. PRO is a phonologically-null pronominal anaphor. PRO contrasts with "\( \tau \)", the trace left behind by movement rules, in the following respects: (i) it is ungoverned, (ii) it may have a remote antecedent, or even lack one altogether, and (iii) its antecedent bears an independent thematic \( (\theta-) \) role in its own clause.

The second and third properties are basically those of pronouns. They are hardly surprising, since PRO is, after all, a pronominal. The first property derives from theory-internal considerations: the first and second principles of the Binding Theory conspire to exclude the possibility of PRO having a governing category. This is, of course, precisely the result Chomsky desires, since restricting PRO to ungoverned positions accurately captures its distribution.
PRO, then, occurs in ungoverned positions, such as subject of infinitive. Since the case assignment principles discussed by Chomsky (1981a) depend on government, PRO will never get to be case-marked. This is also a desirable result if variables have case, since PRO is an anaphor and so should lack case. Therefore, since in tenseless clauses the subject position is ungoverned, the most typical position for PRO is subject of a tenseless clause. Chomsky argues that clauses invariably have subjects. This means that the subject of a non-finite clause, which is ordinarily ungoverned, will ordinarily be PRO.

1.2. A variety of complex factors (cf. Chomsky, 1981a, 74-79) conspire to determine the reference of PRO. It seems at least that grammatical control must be distinguished from pragmatically construal. In the former instance, there is a strictly delimited grammatical domain within which PRO must find an antecedent. Thus, PRO in (1a) must be controlled within the matrix S, whereas this is not true of (1b).

(1) a. Mary promised John [S, [S PRO to buy herself a scarf]]
   b. John believes [S, [S it is immoral [S, [S PRO to eat eggs]]]]

In this paper I tentatively adopt the theory of Manzini (1982 and this volume). It appears that the controlling NP must command the domain of PRO, where "domain" is defined as in Reinhart (1976). In terms of Manzini's theory of control, the antecedent must be in the "domain-governing" category for PRO. Manzini's definition follows:

(2) a. \( \gamma \) is a domain-governing category for \( \alpha \) iff
   \( \gamma \) is a governing category for the domain of \( \alpha \)
   \( \gamma \) contains a subject accessible to \( \alpha \)
   b. A PRO is bound in its domain-governing category

Ignoring certain issues that will arise later, this means that a PRO is controlled within the minimal S or NP containing a governor for the S that PRO is in - for all practical purposes, the relevant domain is the next S or NP up. This is an attempt to capture Chomsky's (1981a, 78) intuition that "very loosely, it appears that PRO searches for a possible antecedent within its clause, and if it can't find one there, looks outside." Manzini has little to say either about precisely how the particular antecedent for PRO is selected, or about how arbitrary PRO, as in (1b), is construed. The evidence to be discussed here, however, confirms the need to separate construal processes from grammatical control.
1.3. Since this paper depends on evidence provided by morphological case, a few preliminary comments on how case arises are in order. I assume NP's that are arguments bearing thematic roles must have case in order to be phonological. Case is either inherently associated with a particular thematic role, or else is dependent on a particular syntactic governor. The point is that the minimal morphological case system will mark just overt thematic NP's, these being the only ones that receive "abstract" case. Also crucial is the assumption that AP's do not bear thematic roles, hence need not have case. Now, this system is extended in various ways in various languages. In Slavic, all terminal [+N] elements, whether nominal or adjectival, must bear morphological case. However, whereas nouns are typically assigned case, adjectives get case through agreement with nouns. This enables the adjective to be predicated of that noun at the appropriate logical level.

1.4. There is much, if not complete, overlap between case-marked and governed positions. Thus, PRO is typically regarded as lacking case features. What features, then, does PRO have? Chomsky (1981a, 20) explicitly claims that PRO "has no properties beyond those of a minimal pronominal element: namely, the features person, number, and gender, which must of course match those of the coindexed element." Indeed, it seems obvious that PRO must be allowed to have at least these "inherent" features, although whether they simply "match" those of the controller, or are acquired from that controller can be debated. After all, anaphors bound by PRO must bear the appropriate features. Consider, for example, the sentence in (1). The anaphor herself must be bound within its governing category, which is the embedded clause. For this reason, herself has PRO as its antecedent. PRO, in turn, is controlled by Mary. Now, according to Chomsky, PRO must come supplied with the correct pronominal features. Then, sometime in the derivation, PRO will be coindexed with both Mary and herself. PRO participates in a sort of "equation of identity," which somehow enables it to "transmit" morphological features.

Let us reserve the term "transmission" for this phenomenon, bearing in mind that in adopting the term we are not a priori assuming a feature-assignment model over a feature-checking one (but cf. section 3.1.). Either, for example, PRO in (1) is unspecified for number and [α plural] becomes [-plural] when control is established, or else PRO is inserted with the feature-specification [-plural] and this is simply matched against the controller (again, only when control has been established). The point is that PRO has slots for pronominal features; it can bear them in a way it cannot bear, say, [+subjunctive]. In fact, PRO must bear pronominal features. If, for some reason, something in the predicate needs to agree with a subject PRO,
and that PRO has not had its value fixed by dint of being controlled, then PRO gets pronominal features by some independent "default" mechanism. Indeed, it has been noted that non-control PRO also carries pronominal features, and that the features of this "obviative" PRO are subject to parametric variation; obviative PRO is human and singular in Spanish, French, and English, but it is masculine and plural in Italian. It appears, then, that when PRO is proximate it assumes the features of its controller, but even when it has arbitrary reference it has features, insofar as these are indicated by the form of coindexed nouns and adjectives.

Recall that one of the properties of PRO making it distinct from all other empty categories is that it is ungoverned and, consequently, caseless. If it is true that PRO simply cannot bear case, then we should not expect the kind of matching and transmission that takes place with inherent features to exist for case. So, although nothing stops PRO from having values fixed for person, number, and gender features, it cannot be case-marked. Thus, PRO should not be able to serve as a "bridge" for transmitting case features. The problem is that most of the time there is no way to tell whether or not PRO is acting as a case-marked entity. The hallmark of an empty category is, after all, that its properties can only be observed indirectly. The features of PRO are witnessed in the agreement of lexical items with overt realization. This is perfectly fine for inherent features, since anaphoric material will be able to agree in them. Case, however, is dependent on government - in general, any nominal (except for another PRO) bound by PRO will be governed itself and as such will receive its own case. In (1), for example, herself is governed by buy; it is objective because of its syntactic configuration, and not nominative, even though its ultimate antecedent is Mary.

2.1. What I want to discuss is an exception to this general situation. There are two constructions in Polish in which the feature composition of PRO is crucial in determining case agreement on a coindexed predicate adjective. If, as I show, adjectives can agree in case with PRO, then an approach is needed which might allow PRO to be ungoverned, yet still bear morphological case features. Although the data are exceedingly complicated and not always consistent, it does appear that there are some definite limitations on which PRO's are able to transmit case. After presenting these facts, I discuss the two issues the problem of case-transmission entails. Firstly, what factors determine the accessibility of PRO for case-transmission and, secondly, what is the actual mechanism subsumed in the notion of "case transmission" that allows PRO to "transmit" case at all.
Although the two constructions have traditionally been regarded as distinct, and there is some slight variation in their relative proclivity for transmission, I will present them in a parallel fashion, since I believe they illustrate the same general process.

The relevant phenomenon is that of predication: adjectives in Polish agree in case with the NP they are predicated of. Typical examples are as follows.

(3) a. zostawiłem go samego
    leave (1sg, past) him (acc) alone (acc)
    'I left him alone'

b. zastałem go chorego
    find (1sg, past) him (acc) sick (acc)
    'I found him sick'

These are the only forms possible. I thus assume an obligatory rule of agreement. Presumably, as mentioned in section 1.3., the obligatoriness of this rule results from a logical form requirement of case identity for predication.

The same principle is in effect for agreement with subjects: predicate adjectives in Polish are nominative when there is a nominative subject to agree with. However, they appear in the instrumental when there is no case-marked subject. Compare (4a), where there is case-transmission, with (5a).

(4) a. Jan/pro jest przystojny/*przystojnym
    John(nom) is handsome(nom)/(instr)
    'John/he is handsome'

b. Jan idzie sam/*samemu
    (nom) goes alone(nom)
    'John is going alone'

(5) a. dobrze jest [PRO być] *młody/młodym
    good (adv) is be(inf) young(nom)/(instr)
    'it is good to be young'

b. dobrze jest [PRO iść] *sam/samemu
    good (adv) is go(inf) alone(nom)/(dat)

In (4a) przystojny 'handsome' can agree in case with its subject Jan. In (5a), however, the PRO subject is not nominative. It is not even controlled by a case-marked NP - instead, it has arbitrary reference. Hence, the predicate adjective młodym 'young' comes out in the instrumental, this being the case of
predication in Polish. Verbs like być 'to be' take instrumental complements. Hence, one cannot claim that mroidym in (5a) is agreeing with an instrumental PRO.

The other construction that seems to obey the same set of conditions is the so-called "second dative." This was first discussed in a transformational framework by Bernard Comrie (1974), and more recently by Barry Schein (1980), while non-transformational analyses have been proposed by Carol Neidle (forthcoming) and Gerald Greenberg (1982). However, all of this "second dative" literature, except Comrie's article, deals with the phenomenon exclusively in Russian.

First let me run through the paradigm of facts. Basically, the lexical items sam 'alone' and jeden 'one', like other predicative adjectives, agree in case with a coindexed subject. This is illustrated in (4b). In (4b) sam is nominative like the NP it predicates. If there is no case-marked subject to agree with - that is, if the subject is PROarb, as in (5b) - no agreement is possible and dative must appear. We therefore want to state a rough rule of agreement: a predicate adjective agrees in case with a coindexed NP.

2.2. What happens if the subject coindexed with the predicate adjective is controlled PRO? This is exactly where transmission comes into play. Consider (6a) and (6b), where PRO is controlled by a nominative NP and the predicate adjective also appears in the nominative.

(6) a. Jan chce [PRO być przystojny/*przystojnym] (nom) wants be(inf) (nom) / (instr) 'John wants to be handsome'

b. Jan chce [PRO isć sam/*samemu] (nom) wants go(inf) (nom)/(dat) 'John wants to go alone'

Transmission allows PRO to be nominative, at least for the purpose of satisfying the agreement rule. So, just as PRO must be masculine and singular in order to be controlled by Jan, it is also nominative, like Jan. Longer chains, as in (7), are also possible.

(7) a. Jan chce [PRO umieć [PRO być bogaty/*bogatym]] (nom) wants be able(inf) be(inf) rich(nom)/(instr) 'John wants to be able to be rich'
(7) b. Jan chce [PRO umieć [PRO iść (nom) wants able(inf) go(inf) sam/*samemu]]
    alone(nom)/(dat)
'John wants to be able to go alone'

In (7), the first PRO appears to "transmit" nominative from the
matrix subject Jan to the second PRO, with which in turn the
predicate adjective agrees. Let us then also state an informal
rule of case transmission: a controlled PRO may agree in case
with its controller. Thus, there are two independent processes
involved here, at least under the assumption that predication-
agreement is clause-bounded and case-transmission depends on
control.

2.3. When one begins to scrutinize transmission a little more
carefully, one finds that not all instances of what might other-
wise be called "control" satisfy the conditions for transmission.
Let us now try to figure out just what those conditions are.
Perhaps, only nominative controllers are able to transmit case.
Consider the oblique controllers that cannot transmit case in
(8)-(11).

(8) a. pro poprosiśem Jana [PRO być asked(1sg) (acc/gen)
    *szczery/szczerym/*szczerego]
    sincere(nom)/(instr)/(acc/gen)
    'I asked John to be sincere'

    b. pro poprosiśem Jana [PRO iść (acc/gen)
    *sam/samemu/*samego]
    alone(nom)/(dat)/(acc/gen)
    'I asked John to go alone'

(9) a. pro kazaśem Janowi [PRO być told(1sg) (dat) be(inf)
    *szczęśliwy/szczęśliwym/*szczęśliwemu]
    happy(nom)/(instr)/(dat)
    'I told John to be happy'

    b. pro kazaśem Janowi [PRO iść *sam/samemu]
    'I told John to go alone'

(10) a. dobrze mi jest [PRO być good(adv) me(dat) is be(inf)
    *młody/młodym/*młodemu]
    young(nom)/(instr)/(dat)
    'it is good for me to be young'
(10) b. dobrze mi jest [PRO iść *sam/samemu]
good(adv) (dat) is go(inf) alone(nom)/(dat)
'it is good for me to go alone'

(11) a. ważne jest dla mnie [PRO być]*
important(adv) is for me(gen) be(inf)
*dobry/dobrym/*dobrego]
good(nom)/(instr)/(gen)
'it is important for me to be good'

b. ważne jest dla mnie [PRO iść *sam/samemu/*samego]
(gen) (inf) (nom)/(dat)/(gen)
'it is important for me to go alone'

None of these oblique controllers transmit their case to PRO. One might therefore try to restrict case-transmission to nominative as a way of prohibiting it from taking place between Jana, Janowi, mi, or mnie and PRO in (8) through (11). Keep in mind, however, that these still control PRO, and PRO must match in inherent gender, number, and person features, even though it is unable to agree in case. Consider the following example (from Comrie, 1974, 138), where, to the extent that an infinitive is acceptable here, person, number, and gender agree independently of case.

(12) zmusiłem Barbarę(acc) sprzedać wóz samej(dat)/
*samą(acc)
'I forced Barbara to sell the car alone'

It thus seems that only a restricted subtype of control is relevant to the transmission of case features, whereas pronominal features must match in any version.

One can easily show that the presence of surface nominative is not the correct criterion for case-transmission. On the one hand, as illustrated by the examples in (13) and (14), there are nominative controllers that do not transmit their case to the predicate adjective.

(13) a. [PRO być *szybery/szyberyem] to twój
being(noun) (nom)/(instr) that your(nom)
obowiązek
obligation(nom)
'being sincere is your obligation'

b. [PRO wykonanie tego przedsięwzięcia
completion this(gen) enterprise(gen)
*sam/samemu] to twój obowiązek
(nom)/(dat) (nom)
'completion of this enterprise on your own is your
obligation'
(14) a. Jan sadzi że ważne jest [PRO być
  (nom) thinks that important is be(inf)
*przystojny/przystojnym]
  handsome(nom)/(instr)
  'Jan thinks that it is important to be handsome'

b. Jan sadzi że ważne jest [PRO iść
  *sam/samemu]
  go(inf) alone(nom)/(dat)
  'John thinks that it is important to go alone'

Since PRO in (14) is free (it is arbitrary in that it can refer
to Jan, or to someone outside the sentence), we are not surprised
to find transmission impossible. The indices on Jan and PRO in
(14) are independent. Control, then, rather than simple con-
strual, is the relevant variable. Regarding (13), it is not a
priori clear whether it is agreement or transmission (or both)
that fails to operate.

I argue in work in progress that Slavic differs from Eng-
lish in not assigning case to the specifier (subject) position
of NP's. Hence, only items that do not bear case may appear in
this position and it follows that there can be no agreement of
predicate adjectives with subjects of NP's. The non-agreement
instrumental of (13a) is expected as a lexical property of być,
and so should be independent of the availability of a case-
marked subject. The dative on samemu in (13b) remains somewhat
of a mystery. It seems, however, that there are two sources
of dative: the "assigned" syntactic one as discussed in the
literature, and the "default" adverbial one of footnote 4. Now,
the dative in (13b) must be the default one since there is no
syntactic source for dative within NP's in Polish, as will be-
come clear when the dative is explicated in section 3.2. One
test for this hypothesis would be whether parallel constructions
are grammatical in Russian, since Russian lacks the phenomenon
of adverbial dative. Indeed, it turns out that examples like
(13b) are impossible in Russian, supporting my claim about
Polish.6 The conclusion is thus that there is no way a predicate
adjective in a NP can get case through agreement with its sub-
ject, since this will never be cased. In order to restrict
transmission in these examples, we optimally want to say that
these are not structures of obligatory control. In this way,
control will be a necessary condition for PRO to get case by
transmission. Once again, the theory of Manzini (1982) seems
adequate to account for the lack of control, since (cf. (2a)),
the sentences in which these PRO's find themselves lack a domain-
governing category.7

Control is clearly different in (13) than it is in (6) and
(7); twój in (13) does not c-command PRO.
The other side of the coin is that agreement does take place with a non-nominative subject. This proves that syntactic position is a relevant factor for case-transmission. There exists in Polish a peculiar phenomenon that comes in handy here in showing the relevance of configuration. Heads of quantified NP's exhibit quirky genitive in Polish. Relevant examples, where the entire quantified subject NP is in the genitive, are given in (15) and (16).\(^8\)

(15) a. wielu studentów jest młodych/*młodymi
   many(gen) students(GEN) is young(GEN)/(INSTR)
   'many students are young'

   b. wielu studentów idzie samych/*samym
   (gen) (GEN) goes alone(GEN)/(DAT)
   'many students go alone'

(16) a. wielu studentów chce [PRO być młodych/*młodymi]
   (gen) (GEN) wants be(inf) (GEN)/(INSTR)
   'many students want to be young'

   b. wielu studentów chce [PRO idzić samych/*samym]
   (gen) (GEN) wants go(inf) (GEN)/(DAT)
   'many students want to go alone'

In (15) we have simple predicate-agreement, and in (16) transmission of genitive to an embedded PRO subject. Although some of my informants showed vacillation in judging the (a) sentences (15) and (16), transmission in the (b) sentences proves that any problem in the (a) sentences of (15) and (16) lies in the agreement rule itself. There seems to be a restriction, at least for some speakers, against a predicate adjective agreeing in case with a genitive subject, regardless of whether that subject is lexical (as in (15)) or PRO (as in (16)). So, on the basis of the (b) examples, we can still conclude that transmission is possible from subject controllers, even when they are non-nominative.

2.4. The fact that case-transmission always takes place from subject controllers can be stated in terms of c-command relations between PRO and its controller, as in Manzini's theory. A PRO is coindexed with an NP in its domain-governing category, and then is able to share case features with that NP. "Control" thus formally defined must be distinct from "pragmatic control" or "construal," even though the two may sometimes share important properties, like apparent obligatoriness. The claim here is that grammatical control is a necessary prerequisite for case-transmission.
It is important to keep in mind, of course, that the conditions under which pragmatic control is possible are, although still poorly understood, certainly far broader than those for either grammatical control or case-transmission. Chomsky (1981a, 77) points out that "c-command is not necessary for control" and cites such examples as in (17) and (18).

(17) [PRO to clear myself of the charges] is important to me
(18) [PRO finishing his work on time] is important to John's development

PRO in (17) and (18) must match in intrinsic features with its pragmatic controller, even though it is not c-commanded by that controller and so will not be able also to acquire case features. I thus reiterate my earlier point that intrinsic features must always match for construal to be possible, and so are distinct from case features.

If the claim is that control is sufficient for case-transmission to take place, it is not clear how to handle the object-control situations in (8) and (9), where no transmission occurs. Perhaps the impossibility of transmission in (8) and (9) (also in (10) and, trivially, (11)), is due to the controller not c-commanding PRO. Then one ought to posit an appropriate structure. The structure in (19a) accommodates this assumption.

\[
\begin{array}{cc}
(19) a. & (19) b. \\
\begin{array}{c}
\text{VP} \\
\text{V'} \\
\text{V} & \text{NP}
\end{array} & \begin{array}{c}
\text{V'} \\
\text{V} & \text{NP}
\end{array} \\
\ldots \text{PRO...} & \ldots \text{PRO...}
\end{array}
\]

NP cannot c-command anything in S', since the first branching node is V'.

Unfortunately, there are several problems with structure (19a). Firstly, it is ad hoc in that there is no evidence to prefer it over (19b), where NP and S' are sisters. Secondly, NP does c-command PRO under more recent definitions of government, specifically, that of Aoun and Sportiche, since S' is dominated by the maximal projection of V. A different idea one might adopt as offering the correct generalization is developed in Kayne (1981, forthcoming). The claim would be that there must be an "unambiguous path" from PRO to its controller, where is traversing an unambiguous path "one is never forced to make a choice between two (or more) unused branches, both pointing in the same direction" (Kayne, 1981, 146). By Kayne's definition, then, government of S' by NP is ruled out in both (19a) and (19b).
There are two serious problems, however, with explaining case-transmission in terms of unambiguous paths. The first is theory-internal: Kayne wishes to reformulate every c-command requirement with the notion of an unambiguous path, but I am exploiting it here only to explain a single phenomenon. If relations between nodes are better conceived of as involving paths rather than configurations, then in a consistent theory this will be true for all such relations. In particular, the same unambiguous path requirement that prevents case-transmission from non-subject controllers to PRO should also prohibit control itself. The other problem with using the lack of an unambiguous path between PRO and an object controller to rule out case-transmission is simply that a less pervasive type of explanation is called for. What we are seeking is some extra codicil that is not part of the theory of control, but will operate to restrict case-transmission. This auxiliary constraint will have the status of a parameter, since it does not seem to be in effect in some languages. Classical Greek, according to Andrews (1971), was one. But parameters are best demonstrated by variation in closely related languages, so let me cite Czech and Slovak as others (cf. Comrie, 1974). Consider the following example of case-transmission from an object controller in Czech, from Comrie (1974, 141).

(20) donutil jsem ho(acc) přijít samotného(acc)/∗samotnému
      (dat)
      'I forced him to come alone'

So the theory of control is the aspect of core grammar that provides the fundamental property of case-transmission; it establishes that only antecedents that are proper controllers are eligible for transmission. Now, what we want in a theory of control is that it correctly predict when case-transmission is potentially possible. Transmission in actuality for any given language may only constitute a subset of those circumstances that the theory of control will permit.

If an existing theory of control separates the transmission constructions from the non-transmission ones, then case-transmission can be regarded as providing support for that theory. Although, to my knowledge, the data don't completely support any treatment I know of, one that I originally deemed hopeful was Williams' predication theory of control. Williams does distinguish two kinds of control: obligatory and non-obligatory. However, Williams' requirements for obligatory control are so rigid as to exclude almost all matrix verbs. Even want is a verb of non-obligatory control, where control is established by a so-called "arb-rewriting" rule. Yet want is a paradigm example of a case-transmission verb, so Williams' distinction between
obligatory and non-obligatory control does not give the desired results. Want does, incidentally, exhibit similar properties of non-obligatory control in Polish as in English, as shown in (21).

(21) a. pro chciałem [PRO spotkać się o piątej]
   wanted(isg) meet(inf) (refl) at five
   'I wanted to meet at five'

   b. *starałem się spotkać o piątej
      *'I tried to meet at five'

In (21a), with chciałem 'I wanted', the antecedent for PRO need not be complete, but in (21b), with starałem się 'I tried', it must be. This is the exact same test Williams employs to differentiate these verbs in English. Moreover, one cannot even argue that transmission takes place with both obligatory control and control established by arb rewriting, since most instances of arb rewriting do not correspond to case-transmission situations. Instead, I adopt Manzini's theory as providing a closer fit to the facts of Polish.

An alternative account for the subject/non-subject asymmetry involves the assumption that the case feature of PRO must be bound by the case feature of its controller (J. Guéron, pers. comm.). Since Polish differs from English in that objects ordinarily cannot bind reflexives, we would expect that whatever prohibits coreference between Jerzemu and sobie in (22) also prevents the case feature of PRO from being bound by an object controller.9

(22) Jan opowiada Jerzemu o sobie,*j
    'John told George about himself'

Unfortunately, such an explanation suggests that the Czech equivalent of (22) will be ambiguous, like its English counterpart. This does not seem to be the case (C. Townsend, pers. comm.), although I have not checked the data with a native speaker. Nevertheless, the idea that case (and, indeed, all inflectional features) must be bound is independently plausible, and I am trying to develop it in other work. For the nonce, I leave the solution to the problem of parameterizing case-transmission open.

2.5. A natural question to wonder about is how transmission works with more deeply embedded PRO's. It seems, for one thing, that the greater the syntactic distance from the controller, the less likely speakers are to admit transmission. In Comrie's original paper, he called this the "cohesion principle," but left the term ill-defined. Here I want to attempt to fill in some of the details of what constitutes "cohesion."
One enticing possibility involves the assumption that case-transmission literally implies movement of features. In this way, for example, transmission can be ruled out in (14), there being two intervening clausal boundaries between Jan and PRO. This is because Subjacency constrains movement to over at most one bounding node. A feature movement analysis, buttressed by some kind of locality principles, should also have ramifications for clauses embedded in Noun Phrases. And indeed there does appear to be variation in the ability to transmit case into such clauses, as a comparison of (23) and (24) shows.

(23) a. Jan zJożył obietnice [PRO być *szczery/szczerym] 'John made a promise to be sincere'

b. Jan zJożył obietnice [PRO iść *sam/samemu] 'John made a promise to go alone'

(24) a. Jan ma zamiar [PRO być szczery/*szczerym] 'John has an intent to be sincere'

b. Jan ma zamiar [PRO iść sam/*samemu] 'John has an intent to go alone'

Case-transmission does not take place in (23), where it freely does in (24). A Subjacency account is feasible if one argues that the embedded clause is outside the object NP in (24), but inside it in (23). The appropriate trees are diagrammed in (23a)' and (24a)'.

(23) a.'
(24) a.

```
(24) a.'
    S
     /\  \
    /   \
   /    \
  NP   VP
   /\   /\ \
  /   /   \
 /    /    \
 Jan  ma  zamiar PRO być szczery
```

Movement of the feature [+nominative] from Jan to PRO would violate Subjacency in (23a), but not in (24a). There are, however, several difficulties, both mechanical and conceptual, with the claim that locality principles similar to those on constituent movement are relevant for case-transmission. Firstly, it is not clear what it would mean to "move" a feature (and, in fact, "copy" would be a better term anyway). It is certainly not like more familiar kinds of movement, since Θ-theory restricts movement to non-thematic positions, but the empty category target of movement here is PRO in a Θ-marked position. In addition, movement of case features entails a distinct device for all other features; since no Subjacency-like requirements appear to hold for intrinsic morphological features, an entirely different matching mechanism must be in operation. This would mean that intrinsic features are truly "matched," whereas case features are copied or "transmitted."

If locality principles are at the crux of the necessary "cohesion" between PRO and its controller, then we might expect syntactic movement to obey the same restrictions as case-transmission. And, in fact, it seems our analysis of the structures of (23) and (23) is correct. Consider the evidence in (25).

(25) a. *co Jan zJożył obietnicę czytać
   'what did John make a promise to read'

b. co Jan ma zamiar czytać
   'what did John have an intent to read'

c. *co Jan ma
   'what did John have'

d. co Jan zJożył
   'what did John make'

A comparison of (25a) and (25b) indicates that the S' is inside the object NP in (23); movement would then constitute a violation of Ross's Complex Noun Phrase Constraint. Likewise, (25c) cannot be answered with a phrase zamiar + S', since these do not form a constituent.
Note, however, that this argues for the structures posited in (23a) and (24a), and not for the feature-movement theory per se. Assuming these structures, however, can case-transmission be ruled out in the appropriate contexts given Manzini's theory of control?

Comparing the structures in (23a)' and (24a)', we see that the domain of PRO - S' in both instances - is governed by N in (23a)' but V (or V') in (24a)'. Therefore, the domain-governing category for PRO in (23a) would be the object NP, except that since there is no accessible SUBJECT (cf. Chomsky, 1981a) there is no domain-governing category. In (24a), on the other hand, S is the domain-governing category and PRO is bound within it. This, of course, means that a different mechanism would be needed to establish control in (23), but that is an entirely independent issue. In any event, it seems reasonable to claim that S' in (23) is an argument of obietnice 'promise' and that the semantics of obietnice itself determine the controller of PRO.

It appears the choice between the theoretically-implausible feature-movement analysis and the otherwise-motivated control theory here boils down to whether we want to claim the nature of control in (23) is distinct from that of (24). It is not certain that the feature-binding idea mentioned in 2.4. cannot rescue the situation. Binding of a reflexive inside NP is perfectly possible in (26).

(26) Jan_1 zjóżył [NP obietnice [S_1, PRO_1 kupić sobie_1 samochód]]
  'John made a promise to buy himself a car'

However, one may maintain sobie is bound by PRO and PRO is of course coreferential with Jan. So, the feature-binding theory could rule out transmission in (23), regardless of whether PRO is simply construed as Jan or, if Manzini's definition of domain-governing category is appropriately revised, it is obligatorily controlled by Jan. In conclusion, although the movement explanation of feature-transmission seems empirically adequate, I can find no reason to prefer it to control. It need never be specially invoked, and since control is independently necessary, there being no case-transmission from non-controllers, I assume control, rather than Subjacency, rules out long-distance transmission.

2.6. When a filled COMP intervenes between PRO and its controller, the facts of case-transmission become exceedingly complex and speakers show much vacillation. A filled COMP does consistently make case-transmission harder, which is in perfect accord with Manzini's theory. As Manzini explains for English control into a tensed clause, if S' branches and S is the domain of PRO,
then PRO won't have a domain-governing category (COMP not being a governor and $S$ a maximal projection of INFL). The problem is that (pragmatic or semantic) control is still relevant. Compare, for example, (27) and (28).

(27) a. Jan nie wie [$S$, czy [$S$ PRO być](nom) not knows* whether be(inf)?szczery/szczyrewym][] sincere(nom)/(instr)
   'John doesn't know whether to be sincere'

   b. Jan nie wie [$S$, czy [$S$ PRO iść sam/samemu]]
      (nom) go(inf) alone(nom/(dat))
   'John doesn't know whether to go alone'

(28) a. Jan nie wie [$S$, jak [$S$ PRO być *szczery/(nom) not knows how be(inf) sincere(nom)/szczyrewym][]
   (instr)
   'John doesn't know how to be sincere'

   b. Jan nie wie [$S$, jak [$S$ PRO iść *sam/samemu]]
      (nom) go(inf) alone(nom)/(dat)
   'John doesn't know how to go alone'

Although these sentences appear to be structurally identical, control in (27) is obligatory, whereas in (28) it is optional. Correspondingly, case-transmission can take place in (27), but not in (28). Why this should be so I do not know (but cf. fn. 11). It is true, though, that whether is often syntactically distinct from other wh-words. Maybe the appropriate variable should be what gets to count as the domain of PRO: with how (and all other lexical wh-words) it is $S$, but with whether (and other grammatical items in COMP) it is $S'$. The claim is then that whether is more of a grammatical formative than a lexical morpheme, and that this somehow allows it not to count for "L-containment." An $S$' that branches, but has a COMP filled with something devoid of lexical content, doesn't count as branching (quite as much). Purpose clauses provide similar examples, as in (29).

(29) a. Jan pracuje [$S$, żeby [$S$ PRO być](nom) works in order be(inf)? bogaty/bogatym][]
   rich(nom)/(instr)
   'John works to be rich'

   b. Jan pracuje [$S$, żeby [$S$ PRO mieszkać sam/samemu]]
      (nom) reside(inf)(nom)/(dat)
   'John works to live alone'

It therefore looks like a COMP with grammatical material in it doesn't really count as a barrier to control or case-transmission.
It is significant that these two phenomena once again go hand-in-hand. Whatever dictates "remote" control in (27)-(29) also enables case-transmission to take place. Granted, the "remote" control strategy augments the general theory of control, but this is to be expected given the marked nature of case-transmission in these examples. One might now ask whether a COMP filled with a grammatical formative renders S' equally transparent for the purposes of Subjacency. The impossibility of syntactic movement in (30) shows that it does not.

(30) *co Jan nie wie czy robić
'what does John not know whether to do'

Movement over a filled COMP, regardless of its contents, is ungrammatical. Thus the feature-movement approach of the previous section breaks down when more unusual instances of control, albeit inexplicable as such in Manzini's theory, are examined.

There are even stranger and more idiosyncratic factors at work here. Case-transmission in sentences like (31) can take place from a neuter, feminine, or plural subject, but is for some reason (for many speakers) impossible from a masculine subject, as in (32).

(31) a. dziecko jest za mały [S, żeby child(neut) is too little(neut) in order  
[ S PRO być intelligent(ne/masc)]
    be(inf) intelligent(nom)/(instr)
'the child is too little to be intelligent'

b. dziecko jest za mały [S, żeby [S PRO iść samo/  
'sam/alone')]
'the child is too little to go alone'

(32) a. chłopiec jest za mały [S, żeby [S PRO być  
boy(masc) is too little(masc) in order be(inf)  
?intelligent(nom/masc)]
    intelligent(nom)/(instr)
'the boy is too little to be intelligent'

b. chłopiec jest za mały [S, żeby [S PRO iść  
'sam/alone')]
'the boy is too little to go alone'

I have no explanation for this puzzling phenomenon, but would like to attribute it to some special (and extraneous) property of masculine singular nouns. Note, in addition, that the embedded clauses differ from the purpose clauses in (29) - where the gender of the controller has no effect on transmission - in that they are even more deeply embedded. So again we have this
awkward notion of "syntactic distance" popping up, suggesting that locality principles still might be appropriate.

3.1. Let me now turn to some speculations about the nature of the actual mechanism of "transmission." The two possible feature models one might envision are assignment and checking. In an assignment model, such as the Extended Word-and-Paradigm theory being developed by Stephen Anderson (1982) and his students (Thomas-Flinders, 1981), there is a very simple and straightforward heuristic for telling derivational features from inflectional ones: derivational features are inherent and inserted at D-structure, inflectional features are syntactically conditioned and acquired by S-structure. In a checking model, as assumed in Chomsky (1981a) for somewhat elusive theory-internal reasons and as required by Kiparsky's "lexical phonology" theory, words are inserted as complete sets of features.

Under either approach the first rule of the Phonology is going to be one replacing these features with a phonological matrix. Since nominals must have case in order to be phonologically "visible," (cf. Chomsky, 1981a) only nominals with case will get to have their "words" actually replaced by "word-forms" with phonetic content. Note, also, that if the rules on the Logical Form side of the grammar can only operate on these feature complexes (not on phonological entities!), then this ensures that LF is interpreted from S-structures in either approach.

However, my treatment of predicate-adjective agreement in Polish can only be accomodated within a feature-checking model. The reason is that the adjective agrees with PRO, but PRO cannot have case at S-structure, since it is unable to appear as a lexical item. Now if PRO lacks feature values until control is established, which fixes the value of, in particular, case features, the predicate adjective cannot agree with it until after that point. This means that the predicate adjective must be inserted at D-structure with case features, but that these are checked at LF. Case on PRO is thus available for checking only in the "right-hand" LF side of the grammar and PRO predictably lacks a phonological matrix (since NP's must have case in order to be mapable onto one). PRO nevertheless is able to function as though it has morphological features in LF matching processes (which is what I believe agreement is), so that case on predicate adjectives can be properly checked. A checking model can thus show how the LF and PF components appear to interact, although they are actually independent, in a way a feature-assigning model cannot. PRO satisfies the Binding Theory by being ungoverned, but as a pronominal requires morphological features, not surprisingly, including case.
What then is the difference, so neatly characterized by an assignment model, between inherent and inflectional features? Somehow, inflectional features must be checked and approved; they are dependent on their syntactic environment. So [+feminine] on an adjective must be checked, whereas [+feminine] on a noun need not be? Is this any more than notationally different from saying a noun is inserted with gender and number, but an adjective acquires them through agreement? In the long run, I think not. However, it is still unclear how something like John tried [[Bill to come]], with Bill simply and accidentally inserted with case, should be ruled out in a checking model. The notion that a feature must be "approved" needs to be better articulated; it certainly is unlike other filters.

What I want to suggest here is that the "inflectional" features, those that are assigned in Anderson's theory and need to be approved in a checking model, are precisely those that must be bound. Since their value derives exclusively from coindexing, the insights afforded by both notions - assignment and approval - are captured. With particular reference to case, a case feature can be bound either by a feature of its governor ("assignment") or by the case feature of a coindexed nominal ("agreement"). This is similar to the "thematic grid" idea presented by Stowell in this volume. Thirdly, there are NP's that must simply come complete with case and θ-role (such as instrumental time and place "adverbials" in Slavic). In claiming that inflectional features are "bound" I mean that they have indices independent of the feature matrix of the whole. In effect, they have an independent existence. Now, we might simply stipulate that inflectional features have specified values that need to be bound, or this could be effected by maintaining α values for inflectional features at insertion, which are fixed through coindexation. These two proposals may be equivalent.

Recall now example (13), where transmission into NP's is discussed. It turns out that agreement of sam with a genitive complement to N does take place.

(33) [przybycie mojego przyjaciela samego/*samemu] arrival(nom) my(gen) friend(gen) (gen)/(dat)
    zaskocio nas
    'my friend's arrival alone worried us'

Interestingly enough, such complements can also bind reflexives, as in (32).

(34) [powieść mojego przyjaciela o sobie] zaskocio nas
    'my friend's tale about himself worried us'
The tentative proposal is that whatever allows feature-binding in (31) also permits anaphor-binding in (34). The examples in (33) and (34), as well as (8/9) and the impossibility of object control (perhaps only when there is a Θ-marked subject), suggest feature-binding within nodes is subject to the same conditions as Chomsky's anaphor-binding, which involves binding the entire feature matrix. One potential problem for this theory is the apparent contrast between (3) and (22). If this turns out to be a real inconsistency, we might need to say the binding of a case feature of an adjective ("agreement") differs from the binding of the case feature of PRO ("transmission") in some substantive way.

Unfortunately, we are still left with the problem of why, if PRO can acquire case through control-binding, it cannot appear as an overt pronoun. Three possibilities, among which I leave the reader to choose, follow: (i) PRO differs from other NP's in that it has features, but lexical items come with inflectional features unbound yet specified; so, whereas [+nom] on a predicate adjective and [+case] on PRO both get bound in LF, only the former is subject to lexicalization, (ii) PRO is not in the same case-chain as Θ-chain, (iii) PRO is a pronominal anaphor and as such lacks a governing category; hence, it differs from all other empty categories in that it has no overt counterpart.

3.2. One possible objection to the claim that PRO actually gets to bear case features is that the Agreement rule really ought to be just extended to permit agreement with any antecedent. Then the restrictions on transmission become restrictions on agreement. In other words, a predicate adjective looks to its antecedent for case, and if that antecedent lacks case, it just keeps on looking back to coindexed controllers until it finds a case-marked one. The "ultimate antecedent" approach cannot be summarily rejected. It would allow PRO not to have case actually itself. If predicate adjectives are coindexed with PRO at LF by predication, then again case must be checked ("bound") (whether control is established at S-structure or in LF). The problem remains, however, that not all instances of control allow the adjective to agree with its ultimate antecedent. This is surely strange if object control is indistinguishable form subject control, but makes sense if only certain controlled PRO's have their case features bound.

A further source of confusion has to do with explaining the "default" cases. Instrumental is not especially problematic, since it is the typical case of predication and is governed for NP complements of być. The second dative, however, is somewhat more troublesome. One might assume, along with Comrie (1974), that the dative arises by dint of agreement with a dative PRO
subject. This seems reasonable, since verbs lacking agreement typically take dative overt subjects, as in (35).

(35)  a. nie ma komu pracować
not has who(dat) work(inf)
'there is no one to work'

b. mnie tu ani żyć, ani umierać
me(dat) here neither live(inf) neither die(inf)
'I can neither live nor die here'

Dative subjects of infinitives in main clauses, as in (35b), are much rarer in Polish than in Russian, but they are common in embedded clauses, as in (35a). However, Schein (1980) has argued that such dative subjects in Russian are not really subjects, but are assigned dative because they are actually in copular constructions and the copula as a modal governs the dative.

If the second datives are not agreeing with a dative subject, how do they get to be dative (and not, for example, instrumental by the predication rule)? Moreover, why is it only these two special words sam and jeden that are able to appear in the dative? A hint to the solution to this mystery is provided by the observation that these two words are morphologically unique: they differ from other adjectives in that they resemble nouns in the non-oblique cases. I propose, therefore, that the dative is assigned directly to these words in the verb phrase, just as it would be to a true NP. This quirky state of affairs results from them being non-distinct from nouns at the appropriate level of abstraction. They are thus able both to get case by agreement like other adjectives, or by assignment as would a noun in the same position (presumably, sister to V'). It is interesting to note that the Czech samotný of (20) is morphologically an adjective and it seems only to be able to get case by agreement.

On the other hand, there is some evidence from other languages supporting the argument that the second dative is agreeing with the would-be case of an ungoverned PRO. Andrews (1976) has noted that predicate einn 'one' in Icelandic agrees with PRO subjects of infinitives. But certain Icelandic verbs select quirky oblique (that is, accusative, genitive, and dative) subjects and einn agrees in case with them. Moreover, even when such clauses appear as embedded infinitivals with PRO subjects, einn still agrees and bears the correct oblique case. An example from Andrews is given in (36).

(36) hun vonast til [S, ad [S PRO reka á land
she(nom) hopes to that S drift(inf) to land
eina]]
alone(acc)
'she hopes to drift ashore alone'
The verb reka 'drift' takes accusative subjects. PRO must somehow get accusative from the infinitive in order that eina be able to agree with it. So, even if my analysis of Polish does not show PRO can be cases, Icelandic indicates that un gover PRO must indeed bear case so that the predicate adjective can agree with it.

3.3. In summary, I have argued that Polish provides evidence that PRO has case features accessible for the purpose of predicate agreement. Since PRO cannot be governed, it must acquire case by some independent mechanism called "transmission." I have demonstrated that transmission is mediated by control. However, more than just control is involved. Hence, I proposed that inflectional features, including case, must be bound to account for various situations where control should permit transmission, but doesn't. I must admit that attempts to weed out the irrelevant facts have left me somewhat confused, but I think the central point - that grammatical processes can be dependent on control coindexing - is incontrovertible. I conclude with the suggestion that further research into the nature of control be aimed at explicating these processes.

*I wish to thank S. Pazdziora, W. Stasiak, and S. Baranczak for their help with Polish. This work has also benefited from discussion with W. Browne, G. Greenberg, J. Guéron, W. Harbert, J. Herschensohn, M.-R. Manzini, R. Smits, and E. Willim, among others.

FOOTNOTES

1R. Smits has pointed out to me that the English facts are far from clear. Consider examples like (i), where non-control PRO appears to be first person plural.

(i) it is unclear [how PRO to save ourselves]

2Predicate adjectives also agree with deleted or "pro" subjects, Polish being a pro-drop language. Notice, incidentally, that this shows that "pro" is distinct from "PRO" in that it has case. Agreement of predicate adjectives can even be used as a test for the status of an empty category. For example, the impossibility of agreement in (ii) indicates that "e" lacks case.

(ii) e chce się Janowi być szczery/*szczery
     wants refl (dat) be(inf) sincere(instr)/(nom)
     'John wants to be sincere'

See Franks (1982) for further discussion.
Predicate NP's differ from adjectives in that they must be instrumental. Therefore, Poles often accept or even prefer instrumental predicate adjectives, construing them to be periphrastic for an entire NP (or else extending the instrumental rule to adjectives). Because of this, my focus is on whether or not nominative is possible at all, regardless of its felicity status.

Once again, what is crucial is that the possibility of nominative exists, since in the colloquial language dative is always a viable alternative with a nominative controller. In fact, according to Comrie (1974), who cites Brodowska-Honowska (1964), for many speakers the dative form samemu has been completely adverbialized and is used regardless of number and gender. Sam thus differs from the other predicate adjectives discussed: rather than being subcategorized (as an obligatory complement), it functions as an adverb.

Note that in (9b) and (10b) the dative is the default case, not the transmitted one. This is indicated by the obligatory nature of the instrumental in the (a) examples. Also, some speakers reject (8) altogether.

Sentences like (13a) should be acceptable in Russian as well as in Polish, but I have been unable to construct any testable examples.

Another extremely interesting situation where predicate adjectives appear not to need to agree with nominative subjects (or PRO subjects controlled by overt subjects) is that of gerunds. Consider the structures in (iii).

(iii) a. [(Jan) bedąc *młody/młodym] (on) being young(nom)/(instr) he(nom) zwiedził wiele krajów visited many countries 'being young, he has visited many countries'

b. [(Jan) idąc (*)sam/samemu] (on) going alone(nom)/(dat) he(nom) przybył na czas arrived on time 'going alone, he arrived on time'

According to Doroszewski (1973), agreement is impossible in (iib). However, some speakers prefer the nominative sam, while nevertheless rejecting agreement in (iii). This asymmetry remains a mystery. Although subjects of gerunds do not seem to get case by the same mechanism subjects of tensed clauses do, they are yet marginally able to feed agreement. Unfortunately, the question of how case-assignment in gerund clauses works is far beyond the scope of this paper.
Quantifiers themselves in Polish, whether in subject or object position, appear to be genitive with masculine human nouns and nominative elsewhere. If the generalization that they are invariably accusative is made (W. Browne, pers. comm.), this strange situation becomes clear: the accusative form looks like the genitive for masculine human nouns, and like the nominative otherwise (for nouns of the declensional class to which the numerals belong). Predicate adjectives appear in the genitive regardless of the form of the quantifier. Hence, agreement is with the head noun, which will always be genitive after a quantifier.

The problem is not one of the object controller being outside the governing category for the reflexive, since sentences like "John told Bill to pour himself some wine" are ambiguous in Polish and Russian. The crucial point is that "John promised Bill to pour himself some wine" is not ambiguous. In the former sentence there are two distinct subjects accessible to bind the reflexive (John and PRO = Bill), but in the latter there is only one, since PRO = John. An analysis involving feature movement has also been entertained in Stowell (1980).

Manzini leaves open the question of how pragmatic and semantic types of control are to be effected. Notice that whether it forces obligatory control, and that this fact does not follow from Manzini's theory.

The fact that some verbs in some languages take quirky subjects is not easily incorporated into the Government and Binding framework, since verbs do not subcategorize their subjects (cf. also Mohanan (1982)). I have three proposals about how to remedy this situation: (i) oblique subjects are really D-structure VP constituents, (ii) case-marking is tied into selection rather than subcategorization; Ø-role assignment is dependent on case, and (iii) S is the maximal projection of V (and not INFL), so that V does indeed govern its subject (parallel to N in NP's).

REFERENCES


Indirect Binding and Referential Circularity*

Isabelle Haïk

MIT

There exists a phenomenon that can be described as coreference between NPs interpreted as variables, best illustrated in "donkey sentences", as in (1):

(1) Many people who owned a donkey$_i$ liked it$_i$

Sentences like (1), in which a pronoun and its antecedent (here, an indefinite) are interpreted as variables bound by a wide scope NP, show that the grammar should include a principle of Indirect Binding, subject to the same conditions as direct binding, except that it is a relation between non-coindexed NPs. (We will come back to this in 1.3.) The purpose of this paper is to show that binding -- direct or indirect -- intervenes in a crucial manner in a domain of facts involving circularity of referential dependence. We are going to see that certain constructions are ruled out because they are referentially circular, and that certain others should be ruled out in the same fashion, but nevertheless are interpretable and grammatical.

1.1. The problem. Consider (2):

(2) *[Her$_j$ husband]$_i$ likes [his$_i$ wife]$_j$

It seems natural to explain the ill-formedness of (2) by virtue of referential circularity (see, for example, Jacobson, 1979, citing Langendoen; Vergnaud, 1974; Higginbotham and May, 1979, and Brody, 1981): we can tell the reference of his wife only if we know who his is, hence we have to look at the antecedent of his. The antecedent is her husband, whose reference depends on that of her, which has as its antecedent his wife. By transitivity, his wife is dependent on itself. This phenomenon also occurs with one, or
three, or more NPs, as in (3) and (4):

(3) *[A picture of it]_i

(4) *[His wife]_i told [her daughter]_j [her father]_k was angry

All these examples are violations of Higginbotham's (1982) filter, which states that no NP may be referentially dependent on itself:

(5) Not D*(X,X)

where D*(X,Y) means that X is dependent on Y, possibly by transitivity.

Leaving aside the interesting generalization treated in Brody (1981) that circular reading holds of constituents other than NPs (see fn.1), we will adopt Higginbotham's filter to rule out circularity of referential dependence.

The problem which is to be solved is that certain constructions seem to violate this filter, like Bach-Peters sentences, and other types of crossing coreference sentences (see Jacobson, 1979, for a study of these constructions). For example, (6) and (7) should be ruled out by filter (5):

(6) [The man who loved her]_i kissed [his wife]_j

(7) [The man who loved her]_i embarrassed [the woman he kissed]_j

We are going to see that an ill-formed structure, like (2), is ruled out because circularity is due to the necessity of looking at the antecedent of the pronoun in order to interpret the NP that contains the pronoun. In fact, suppose that the following principle holds:

(8) An NP which contains a pronoun is referentially dependent on the antecedent of the pronoun.

The notational convention we adopt in order to express referential dependence of NP_i on NP_j is to append the index of NP_j on that of NP_i: NP_i/j is to be interpreted as "NP_i is dependent on NP_j."

Looking back at (2), we see that, by virtue of (8), both NPs are dependent on themselves:

(9) [Her husband]_i/j/i likes [his wife]_j/i/j

The proposal I will present for the difference between the ill-formed cases like (6) and (7) is that, owing to Indirect Binding, one of the NPs of (6) and (7) will be considered to be interpretable without having to check what the reference of the antecedent of its internal pronoun is. Before giving an account of (6) and (7), I have to justify the principle stated in (8). Remember that the construing of referential dependence as in (9) follows from principle (8).

1.2. The Extended Name Constraint.

I want to make a correlation between referential dependence as seen in (9), and referential dependence obtained when an NP is interpreted as in the scope of another NP. Consider (10):
(10) Two men saw a woman

When a woman is in the scope of two men, we can say that it is referentially dependent on two men, because the value assigned to a woman is a function of that assigned to two men: for each of the men, there is, associated with him, a woman.

In the general case, any scope interpretation is possible, and for example in (10), we can also talk of one woman, whose reference is not dependent on that of the men. However, some scope interpretations are forced, disallowing any choice. For example, consider (11):

(11) Everyone\textsubscript{j} likes [some film he\textsubscript{j} saw]\textsubscript{j}

In (11), NP\textsubscript{j} does not have the choice of having wide scope: it has to be interpreted as referentially dependent on everyone. In his dissertation, May gives an explanation of this in terms of QR and c-command at LF:

(12) \[ S\textsubscript{y} [\text{some film he saw}] [S\textsubscript{x} everyone] [S x likes y]]

In (12), the logical form of (11) which corresponds to wide scope of some film he saw, the pronoun he is not c-commanded by everyone, meaning that there is no binding of the pronoun by everyone, rendering coindexation between the two senseless. In order for the pronoun to be bound at LF by everyone, it has to occur in a c-commanded position, forcing NP\textsubscript{j} to be interpreted as in the scope of everyone.

Notice that another explanation is available for the imposibility of the reading represented in (12), in terms of the Name Constraint. (See May, 1977, for a first formulation, and Guéron, 1981, for an extension to all constituents. See also Fiengo and Higginbotham, 1981, for a similar proposal.) The Name Constraint (NC), formulated in (13), accounts for the imposibility of wh-extraction from a definite NP, as in (14):

(13) A name may not contain a free variable.

(14) *Who did you like that picture of t?

Notice that the facts of (14) and of (11) should be expressed by the same generalization, namely: when an NP contains an element interpreted as a variable -- like the pronoun he in (11), and the wh-trace in (14) -- and when this element is not bound inside the NP, the NP is said to be open with respect to the index of the element, and must be interpreted as in the scope of the quantifier that binds the element. In other words, (14) is ruled out because that picture of t, containing a free wh-trace, should be interpreted as in the scope of the wh-quantifier, which is incompatible with its status of name. The NC may thus be restated so as to account for (11) as well as for (14):
(15) Extended Name Constraint (ENC): If NP\textsubscript{i} contains a free variable\textsubscript{j}, then NP\textsubscript{i} is within the scope of NP\textsubscript{j}.

This formulation entails referential dependence consecutive to scope interpretation, this is why it mentions the term variable. However, considering an NP that contains an ordinary pronoun, not interpreted as a variable, referential dependence still holds, in case the pronoun draws its reference from an antecedent. In other words, we want to put in parallel pronouns interpreted as variables and pronouns which do not come with a specified reference but draw it from an antecedent. Notice then that the principle postulated in (8) is nothing else than the ENC, with this view that pronouns which draw their reference from a linguistic antecedent behave like variables with respect to referential dependence. Then the ENC, independently motivated to govern scope interpretations, has just to be revised as (15'), in order to include all cases of referential dependence:

(15') ENC (revised): If NP\textsubscript{i} contains a free variable\textsubscript{j} or pro\textsubscript{j}, where pro\textsubscript{j} draws its reference from an antecedent, then NP\textsubscript{i} is within the scope of NP\textsubscript{j}, i.e. NP\textsubscript{i}//NP\textsubscript{j}.

We may return to the problem posed by (6), repeated below, having in mind the ENC:

(6) [The man who t\textsubscript{i} loved her\textsubscript{j}]\textsubscript{i} kissed [his\textsubscript{i} wife\textsubscript{j}]

Inside both NPs, the pronouns are free, forcing referential dependence to be drawn from each of the NPs to the other, leading to circularity, and (6) should be ruled out. In fact, we are going to see that one of the pronouns, here her, is bound inside the NP that contains it, and, more precisely, that it is indirectly bound by the wh-trace. Since this accounts relies on Indirect Binding, let us turn to the definition and properties of this principle.

1.3. Indirect Binding. Indirect Binding is a process of binding that occurs when an NP has scope over another one: the NP in scope is referentially dependent on the wide scope NP. The idea behind it is that the narrow scope NP is treated like a variable bound by the wide scope NP. Moreover, following Reinhart (1977), I assume that binding, direct or indirect, holds if and only if c-command at s-structure holds. As an illustration, consider (16)a-b, in which some student is in the scope of many people:

(16)a. [Some student\textsubscript{1/2} spoke to [many people\textsubscript{2}]

b. [Many people\textsubscript{2} spoke to [some student\textsubscript{1/2}]

We see that, even if NP\textsubscript{2} has scope over NP\textsubscript{1} in both (16)a and b, NP\textsubscript{2} is an indirect binder of NP\textsubscript{1} only in (16)b, because only then does it c-command the NP it has scope over.
Given these assumptions, suppose now that a pronoun coindexed with a narrow scope NP also occurs in the sentence in which the wide scope and the narrow scope NPs appear, and suppose also that the two NPs are in a configuration in which Indirect Binding holds. Then the pronoun, if it occurs in a position in which it is indirectly bound, may be interpreted, like its antecedent, as a variable, bound by the wide scope NP. The most striking illustration of this process of binding is provided by donkey sentences, in which the indefinite NP, a donkey, is in a relation of coreference with the pronoun, both being bound by the wide scope NP:

(1) [Many people who \(t_i\) owned [a donkey] \(j/i\)] \(j/i\) liked it \(j/i\)

(1) is a case of Indirect Binding of index \(j\) by index \(i\). It means that all occurrences of \(j\) are interpreted as functions of NP. This principle is crucial to explain why the pronoun \(it\) is licit in (1), but not in, e.g., (17):

(17) *Some people who were selling a donkey arrived, and I bought it
Like any variable, it has to be in the scope of the quantifier that makes it a variable. Obviously, its antecedent a donkey is not this quantifier. The one in question is the wide scope NP.

We now have the tools to explain our previous problem, namely, why (6) is not ruled out by circularity, along with (2), repeated here:

(6) The man who loved her kissed his wife
(2) *Her husband kissed his wife

1.4. Referential Circularity Circumvented. The idea is to make it possible for the subject NP, in (6), to be spared from looking at the antecedent of the pronoun that it contains. Indirect Binding allows such an escape: suppose we assume that NP, the man who \(t_i\) loved her, is prior in reference to NP, his wife. If we do so, then, according to the ENC, NP has to be closed with respect to \(j\), i.e. the pronoun her has to be bound inside NP.

Now, the pronoun, inside NP, may be indirectly bound by the wh-trace. This means that her, and its antecedent, are construed as functions of NP, with NP, prior in reference. Notice that if NP is assumed to 'indirectly' bind NP, then all occurrences of NP have to be indirectly bound by some occurrence of NP: like any variable, they must occur in the domain of scope of their binder.

In (6), NP, c-commands NP, and thus is allowed to indirectly bind it. As a consequence, the pronoun her, inside NP, may in fact very well be indirectly bound by the wh-trace: this binding relation is internal to NP, hence NP, does not have to worry about the antecedent of her. In other words, NP, obeys the ENC by indirect binding of her inside it, and her and its antecedent are both
interpreted as dependent on NP\textsubscript{j}. This construction is saved from circularity because only one NP is dependent on the other.

To better understand how (6) is saved, and in order to illustrate the necessity of the assumptions, let us consider ill-formed structures. (Similar examples are considered in Jacobson, 1979, and Higginbotham, 1982.)

1.5. Referential Circularity. Consider (2) again:

(2) *[Her\textsubscript{j} husband\textsubscript{i}] \textsubscript{j} likes [his\textsubscript{i} wife\textsubscript{j}]

We easily see that the pronouns can be internally bound in none of the NPs. Hence the construction is circular. Consider (18):

(18) *[The man she\textsubscript{j} loved t\textsubscript{i}] \textsubscript{i} kissed [his\textsubscript{i} wife\textsubscript{j}]

Jacobson's (1979) analysis of (18) is that her is obtained by replacement of its antecedent his wife. When his wife occurs in the position of her, weak crossover is violated inside NP\textsubscript{i}, as seen in (19):

(19) [The man [his\textsubscript{i} wife\textsubscript{j}] loved t\textsubscript{i}] \textsubscript{j}...

Hence, whatever rules (19) out should also rule (18) out. Under the present analysis, the pronoun her is not internally bound to NP\textsubscript{i} (the wh-trace, which could indirectly bind it, may not do so because it does not c-command it). Hence NP\textsubscript{j} which contains a pronoun which draws its reference from an antecedent, is referentially dependent on this antecedent, NP\textsubscript{i}. The same reasoning applies for NP\textsubscript{j}, and circularity cannot be avoided.

Consider also (20):

(20) *[His\textsubscript{j} wife\textsubscript{i}] \textsubscript{j} kissed [the man who t\textsubscript{j} loved her\textsubscript{i}]\textsubscript{j}

The pronoun her is bound inside NP\textsubscript{i} by the wh-trace, so (20) should not be ruled out. Remember, however, that Indirect Binding means that an NP is treated as a variable, and that any occurrence of a variable has to be bound. If her is properly c-commanded by \textsubscript{j}, the same is not true of its antecedent, his wife. This NP may thus not be interpreted as a variable bound by NP\textsubscript{j}. However, this implies that her may not either be treated as a variable indirectly bound by \textsubscript{j}, and the construction may not escape circularity.

Lasnik's (1976) examples, mentioned in fn.2, can also be accounted for by binding. Nevertheless, it is interesting to notice that we give different accounts of it:

(21) *[The story about [the pilot that t\textsubscript{i} shot at it\textsubscript{j}]]\textsubscript{j}

amused [the Mig that t\textsubscript{j} chased him\textsubscript{i}]\textsubscript{j}

Lasnik suggests that crossing coreference sentences are saved from circularity if we analyze the second pronoun as a variable bound by its antecedent, the first NP. In (21), the pronoun him is not in the scope of its antecedent, prohibiting a variable analysis of the
pronoun, and hence leading to circularity. Our analysis also makes crucial use of the structural relation between NP\textsubscript{i} and NP\textsubscript{j}, but the reason for that is a little different: we want NP\textsubscript{i} to c-command NP\textsubscript{j} in order for it to be able to be an indirect binder of NP\textsubscript{j}, and hence be able to be closed with respect to \textsubscript{j}. There is no requirement of direct binding on him, though. The difference of points of view is obliterated by the fact that the pronoun him occurs inside NP\textsubscript{j}, and then there is no way to tell which of him or of NP\textsubscript{j} has to be c-commanded by NP\textsubscript{i}.

As a conclusion for this part, we have seen that Indirect Binding, justified to account for donkey sentences, provides an escape from circularity of referential dependence, in sentences like (6) and (7), because it allows us to disregard the antecedent of the pronoun. The second part of the paper is devoted to embedded constructions inside NPs.

2.1. Circularity and subcategorized complements. Consider again (3), or (22):

(3) *[A picture of it\textsubscript{i}]

(22) *[The writer of his\textsubscript{i} book]\textsubscript{i}

Chomsky (1981) rules them out by the i-within-i condition, a filter that forbids constructions in which the index of a constituent occurs inside that constituent. This filter misses the generalization made by Brody (1981), Jacobson (1979) and Higginbotham and May (1979) that (3) and (22) are cases of circular readings, which should be excluded similarly to, for example, (2).

It is easy to see that (3) and (22) are cases of referential circularity: by convention (8), which is restated in the ENC (15'), NP\textsubscript{i} is referentially dependent on the antecedent of the pronoun that it contains, namely itself. This is prohibited by filter (5). Circularity could be avoided if the pronoun was bound somehow, inside NP\textsubscript{j}, but it is not: we assume that head nouns are not coindexed with the NPs they are the heads of. (Moreover, even if they are in fact coindexed, it seems reasonable to postulate that only \textsubscript{Nmax} can bind another NP, if binding has anything to do with reference.) We are going to see that circularity may be circumvented by direct binding, in certain constructions.

2.2. Non-subcategorized Complements. Consider the following NP:

(23) [The man next to [his\textsubscript{i} dog]]\textsubscript{i}

Following suggestions by Andrew Spencer and David Pesetsky (personal communication), we observe that, in the ill-formed cases of the previous section, like (3), the NP is subcategorized by the head noun and receives a \(\theta\) role from it, whereas it is not subcategorized, and henceforth does not receive a \(\theta\) role in the well-
formed constructions, like (23) (cf. Baltin, 1981), for a similar
distinction deriving from other considerations). Moreover, in
(23), the relation between the head noun (rather, the head NP, as
we shall see) and the PP is one of predication: there is man such
that he is next to his dog.

Let us give a structural account of (23), which will explain
why it is not circular. Following Chomsky (1981), subcategorized
positions must be governed by a head in order to be assigned a
θ-role by it. Now, if we say that base-generated positions of PPs
inside NPs which are governed by the head noun are subcategorized
positions and thus must receive a θ-role from the head, (like NP
complements), then we do not want the PP of (23) to occur in a
position governed by the head noun, since there is no θ-role to
assign to it. Hence a maximal projection must intervene between N
and PP, the only reasonable one being NP. This allows a struc-
tural difference to be drawn between (3) and (23). Their respec-
tive structures are shown in (24)a-b:

(24)a. \[
\begin{array}{c}
\text{Spec, } \overline{\text{N}}_i \\
\text{a } \overline{\text{N}} \\
\text{picture of it}_i \\
\end{array}
\]

b. \[
\begin{array}{c}
\text{N}_i \\
\text{PP} \\
\text{the man} \\
\text{next to his}_i \text{dog} \\
\end{array}
\]

In fact, (24)b is structurally identical to a relativized NP, with
PP instead of S predicated of the head NP. Now, under the assump-
tion that any NP may receive a referential index, the predicated
NP of (24)b may actually bear an index, as in (25):

(25) \[
\begin{array}{c}
\text{the man} \\
\text{next to his}_i \text{dog} \\
\end{array}
\]

The internal NP_i c-commands the coindexed pronoun, and thus binds
it, internally to the large NP_i. Thus, the large NP_i is closed
with respect to i, and then is not construed as referentially
dependent on itself. The conclusion is that direct binding by the
head NP in a PP-predicative construction like (23) allows the NP
not to be referentially circular.

Now that we have defined the contrast between NPs with sub-
categorized complements and those with non-subcategorized ones in
terms of internal binding of the pronoun pro_i, coindexed with the
large NP, we may think that we can predict the well-formedness of
the constructions in accordance with the kind of complements pro_i
appears in. These expectations are deceived in some, very
interesting, cases, which we study in the next section.
2.3. **Complex NPs.** Consider the following paradigm:

(26)a. I met [the author of [a book that \( t_j \) made him_{ij} famous]]_{ij}

b. *I met [the author of [a book he_{ij} wrote \( t_j \) in two days]]_{ij}

The structure of NP_{ij}, in (26)a-b, is: \( N + \) subcategorized complement restricted by a relative clause containing pro_{ij}, schematically, \([\text{NP}_{ij} \ldots N \text{of} \ldots \text{NP}_{ij} \ldots \text{pro}_{ij} \ldots \]). NP_{ij} contains a pronoun pro_{ij}, hence it should be construed as referentially dependent on the antecedent of the pronoun, i.e. itself, which is circular. To avoid circularity, we have to consider that pro_{ij} is bound internally to NP_{ij}, hence allowing us to disregard its antecedent. Direct binding is not available, because no NP_{ij} c-commands pro_{ij}. Indirect Binding should thus be considered: if pro_{ij} is interpreted as a function of another NP, this NP could be its indirect binder, if it c-commands the pronoun.

Considering the pronoun in (26)a-b, and its configurational relation with the wh-trace, it seems that the contrast between the two sentences can only be due to the fact that the pronoun is c-commanded by \( t_j \) in the well-formed sentence, and not in the ill-formed one. What else than Indirect Binding of the pronoun by the wh-trace could explain such a minimal contrast as that between (26)a and (26)b? We propose that nothing else explains it; however, we are going to see that pro_{ij} is not an ordinary indirectly bound pronoun.

Indirect Binding of him_{ij} by \( t_j \) would mean that him_{ij}, or NP_{ij}, i.e., the author, is interpreted as a function of \( t_j \), or NP_{ij}, i.e., the book. That is to say, from the reference of the book, one could find the reference of the author. The problem posed by (26)a is that this is semantically and structurally impossible: the book may not have a reference prior to that of the author. For example, as Higginbotham pointed out to me, it is impossible to give the title of the book:

(27) *I know the author of a book that made him famous, namely 'David Copperfield'

We can assume that if we cannot name the book, a fortiori, we cannot construe the reference of the author on it. This semantic observation is reflected by the syntactic fact that Indirect Binding is impossible structurally, because NP_{ij}, the would-be indirect binder of NP_{ij}, does not c-command NP_{ij}, the NP which contains it:

![Diagram](image-url)
60

Still, if there is a difference between (26)a and b, it must be due to the relation between the pronoun and the wh-trace, suggesting that Indirect Binding saves (26)a and not (26)b.

2.3. Sloppy Identity Pronouns. We propose that the solution of the problem posed by the above paradigm is that pro, is a pronoun of laziness, in the sense of that given to the pronoun it in Karttunen's example, (28):

(28) The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress

More precisely, we propose that the pronoun of (26)a is interpreted like pronouns in sloppy identity readings, as in Karttunen's example, and in VP-deletion sentences, as in (29):

(29) John forgot his appointment, but so did Tom

We give the following interpretive rules for these pronouns:

(i) They are interpreted as functions of other NPs.
(ii) The name of the function is given by the head noun of some previous NP.

We assume that (i) holds only if the pronoun is indirectly bound by the NP it is interpreted as a function of.

For Karttunen's example, it names the value of "is a paycheck of" applied to the indirect binder of it, namely the less wise man: the wh-trace indirectly binds \_\_\_. By (ii), the name of the function is given by the head noun of the "antecedent" of it.

Notice that an analysis of these pronouns based on Indirect Binding leads us to expect the sloppy identity phenomenon to be impossible when the pronoun is not c-commanded by some indirect binder. This is verified:

(30) *The man who gave his paycheck to his wife was poorer than the man whose pocket it fell from

We see that, whether or not the NP whose pocket is reconstructed, the pronoun it is neither c-commanded at that level (after reconstruction), nor at s-structure, confirming our analysis.

Ross (1967), also cited in Sag (1976, p.129), and Lasnik (1976) point out that, in VP-deletion sentences, the sloppy reading is possible only when the antecedent of the non-overt pronoun commands, or precedes and kommands -- we now say binds -- the pronoun:

(31) John scratched his arm, and so did the boy who knows Bill

(31) may not mean that the boy who knows Bill scratched Bill's arm. We do not intend to give here a complete analysis of VP-
deletion sentences. For an interpretive analysis, see Sag (op. cit.), and Williams (1977). In their analysis, the non-overt pronoun corresponds to a variable bound by the subject of the deleted VP. Under our analysis, the pronoun corresponds to a function of the subject, and is indirectly bound by it. In other words, the logical form of a sentence like (31) looks like (32)a, rather than (32)b:

(32)a. \( \lambda x \), John (x scratched x's arm) and \( \lambda y \), NP\(_i\) (y scratched y's arm)

For reasons having to do with the fact that a VP-deleted conjunct contains only a subject as an NP, it seems that the two analyses make the same empirical predictions. We leave Gapping, which should distinguish between the two, for further research (see Pesetsky, 1982, for an interesting analysis of Gapping).

Returning to the original paradigm, let us see how (26)a, repeated below, is interpreted:

(26)a. I met [the author of [a book that t\(_j\) made him\(_i\) famous]],

We saw that him may not be interpreted as drawing its reference from the coindexed NP, NP\(_i\), because this leads to referential circularity, and we also saw that it cannot be interpreted as coreferential with NP\(_i\), with both it and NP\(_i\) interpreted as functions of the book, since the book may not be prior in reference.

Suppose now that it is interpreted as a sloppy identity pronoun: it is thus interpreted as a function of an indirect binder, here the wh-trace coindexed with NP\(_i\), the book. The name of the function is given by the head noun of the NP with which pro\(_i\) is coindexed, namely, 'author of'. The semantic information conveyed by the pronoun is thus that it refers to the author of the book, whoever he is. That is to say, pro\(_i\) does not draw its reference from NP\(_i\). However, since NP\(_i\) is a description for the author of the book NP\(_i\), and since pro\(_j\) refers to the author of the same book, pro\(_j\) and NP\(_j\) refer to the same author. This fact is due to the construction, and not to a direct link of coreference between NP\(_i\) and pro\(_i\). Hence, the only requirement for the pronoun him\(_i\) to be interpretable is that it be bound by some indirect binder. Moreover, since him\(_i\) is not coreferential with NP\(_j\), NP\(_i\) itself is not required to be indirectly bound by NP\(_i\) and hence to be c-commanded by it. This is why the construction is well-formed.
Hence, contrary to intuition, him is not construed as coreferential with the large NP -- even if, eventually, they refer to the same individual -- but merely as picking from it the name of the function it stands for, here, 'author of'. Comparing (26)a with (26)b, we see that the sloppy identity reading of the pronoun is not possible because no indirect binder c-commands it.6

Again, the nice fact about (26)a-b is that him, and NPi may not be construed as coreferential variables indirectly bound by NPi. This was suggested by the impossibility of giving the name of the book.7 Another piece of evidence is provided by the following facts. In French, (33) may have interpretation (34):

(33) [L'auteur de [trois livres]j]i
    'the author of three books'

(34) For each of the books, the author of the book
That is to say, NPi may have scope over NPi. This reading, however, is impossible in (35):

(35) [L'auteur de [trois livres qui t]j lont rendu célèbre]i
    'the author of three books that made him famous'
This is explained in a parallel fashion to why we cannot name the book: if NPi has scope over NPi, then it must be closed with respect to 'it', and hence must internally bind le 'him', implying that it should indirectly bind it and its antecedent. However, we saw that indirect binding of NPi by NPi is impossible, because NPi, the containing NP, is not c-commanded by Nj.8

In conclusion, let us summarize the points made in this paper. We suggest that certain constructions are circular when they violate Higginbotham's filter (5). Certain constructions, like crossing coreference sentences, seem to violate this filter, but in fact, they are made possible by Indirect Binding, which allows NPs to obey the Extended Name Constraint (15), and thus spares us from drawing the reference of an NP from the antecedent of a pronoun it contains.

Constructions with embeddings escape from circularity in case the pronoun occurs inside a non-subcategorized complement, owing to direct binding, with the structure as proposed in (25).

Complex NPs inside subcategorized complements, like (26)a, escape from referential circularity when the pronoun they contain is interpreted as a sloppy identity pronoun. The interpretive rules applied to such pronouns give them as functions of
an indirect binder, with the function given in some previous material.

FOOTNOTES

*I wish to thank J. Higginbotham, D. Pesetsky, and participants in the Cornell Conference on Government and Binding Theory, July 1982, for their useful comments.

1. Brody (1981) gives a different account of circular readings, and treats NP constituents as well as other kinds of constituents, as in (i):

(i) (=1c) *Tom [wanted to appear to Ø]x

His treatment of this phenomenon is based on the idea that circular readings involve a contradiction between two principles. The first one is that a-c dependency is transitive, where a-c dependency means dependency on the antecedent (a), or dependency by compositionality (c). The second principle is that, if A is an anaphor which a-c depends on an antecedent B, then B may not a-c depend on A, i.e. the relation between an anaphor and an antecedent is asymmetric. Brody's claim is that circular readings are not ruled out by logical necessity, as they seem to be, but rather because they violate a principle of grammar, namely, the unavailability of infinite processes of interpretation. For a justification of these principles and a discussion of Higginbotham and May's (1979) approach, see Brody, op.cit.

2. That binding requires some structural relation between the binder and the bindee is stated in Lasnik (1976), in terms of precede and kommand. (Kommand uses cyclic nodes as relevant domains, instead of branching nodes, as in Reinhart, 1976, or, more recently, maximal projections, as in Chomsky, 1981: a node A kommands a node B iff the minimal cyclic node that dominates A dominates B. A cyclic node is S or NP.) Lasnik also suggested that the Bach-Peters paradox constituted by crossing coreference sentences (infinite pronounization process), as in (i), could be solved by analyzing the kommanded pronoun as a bound variable. Lasnik points out that when the pronoun is not kommanded by its antecedent, the sentence becomes bad, as in (ii) (see section 1.5 for further discussion):

(i) The pilot that shot it hit the Mig that chased him
(ii) *The story about the pilot that shot it amused the Mig that chased him

3. That NP₂ may not indirectly bind NP₁, when it does not c-command NP₁, at s-structure may be analyzed as a case of weak crossover. For example, under an analysis with QR, as in May (1977), and the Bijection Principle, proposed by Koopman and Sportiche (1981), (16)a is a violation of the Bijection Principle.
(The B.P. states that there must be a one-to-one correspondence between quantifiers and variables.) The ill-formed LF of (16)a would be (i):

\[(i) \ast [S[many \text{ people}]_x [S[some \text{ student}]_y \text{ spoke to } x]]\]

Applying QR to the wide scope quantifier makes it bind the two variables, some student (indirectly bound), and x (directly bound). According to Koopman and Sportiche (personal communication), QR could apply to some NPs only, and not all, to represent scope. Under my system, in which narrow scope NPs are treated as variables, they are allowed not to move. (In Haïk, 1982, I claim that QR is not necessary to obtain scope interpretation, and that QR is inadequate to represent the distinction between having wide scope and 'not being in the scope of'. If I adopted QR, it seems that moving only wide scope quantifiers could yield the desired results. However, some complication would arise when more than two NPs interact, because the middle ones are to be interpreted as variables bound by the NPs which have scope over them, and still have to have scope over the NPs with narrowest scope.) See Heim (1982), who also treats indefinites as variables bound by wide scope quantifiers and other elements that take scope, like adverbs of quantification, following Lewis's (1975) analysis.

4. According to Vergnaud (1974 and 1982), the genitive preposition of may be analyzed as a Case-marker, devoid of semantic content. With genitive constructions, we can assume that the of NP is a complement of the head noun. Some other prepositions seem not to be analyzable as Case-markers, but the preposition is subcategorized for, and circularity rules the NP out, as in (i):

\[(i) \ast [A \text{ letter to its}_i \text{ addressee}]_i\]

5. Note that, contrary to constructions with PPs, we do not want head NPs of relative clause constructions to be indexed, because then there would be no difference between, say, (i) and (ii):

\[(i) \ast [The \text{ man whom she}_j \text{ likes } t_i]_i \text{ kissed [his}_i \text{ wife}]_j\]

\[(ii) [The \text{ man who } t_i \text{ loves her}_j]_i \text{ kissed [his}_i \text{ wife}]_j\]

If the man is allowed to bear an index, then, since it c-commands the pronoun pro, both in (i) and (ii), it should be able to indirectly bind it, and both sentences would be ruled in. In fact, the difference between (i) and (ii) is due to the presence of the wh-trace, the relevant binder of pro\_j.

It may appear a stipulation that the head NP is indexed when predicated by a PP, and not when predicated by an S. There may exist an explanation of this, if we make the assumption that the
head NP in both constructions is in an A-position. Then the impossibility of indexing the head NP in a relativized construction would follow from principle C of the Binding Theory: the wh-trace, a variable, would be A-bound by the head NP, as shown in (iii):

(iii)

\[ \begin{array}{c}
\text{NP}_i \\
\text{the man} \\
\text{who...t_i...} \\
\end{array} \]

This problem does not arise with a PP, since there is no variable in danger of getting A-bound.

6. It may be misleading to coindex him and NP, when him is construed as a sloppy identity pronoun, since they are not coreferential, as well as between his paycheck and it in (28). I hope it is clear that coindexing these kinds of pronouns with other NPs simply means that the head noun gives the name of a function, nothing else.

7. Notice that the name of the author can be given:

(i) I like the author of a book that made him famous, namely, Charles Dickens (from Higginbotham, 1982)

This is perfectly compatible with our analysis of these constructions.

8. One could wonder why the pronoun le could not be interpreted as a sloppy identity pronoun, thus preventing NP, the large NP, from having to be indirectly bound. This might be due to the fact that pronouns may not be construed as in the scope of other NPs, which would turn out to be the case if le was interpreted as a sloppy pronoun bound by trois livres, since this is a plurality. I will leave the matter here (for the statement that pronouns may not be construed as in the scope of quantifiers, see Haïk, 1982 and Heim, 1982).

REFERENCES

Guéron, J (1981): "Logical operators, complete constituents, extraction transformations", in J. Koster and R. May (eds.) Levels of Syntactic Representation.

Haïk, I (1982): "Indirect Binding", ms MIT.


Higginbotham, J(1982):"Remarks on binding and LF", talk: Cornell U.


Reinhart, T (1976): The Syntactic Domain of Anaphora, Ph.D dissertation, MIT.

Ross, J (1967): Constraints on Variables in Syntax, Ph.D dissertation, MIT.


Corroboration for the ECP

Randall Hendrick
University of North Carolina
at Chapel Hill

0. Introduction

The fundamental task before linguistics is the characterization of Universal Grammar, a set of cognitive principles that define the conceptual space in which grammars fall. Specifying Universal Grammar in this sense will contribute to an explanation for how a child is able to acquire knowledge of a grammar on the basis of data limited both qualitatively and quantitatively. Evidently the language learner acquires a grammar that has some properties that are not present in the corpus of data that he experiences. Those properties that are not present in the experience of the language learner owe their existence to Universal Grammar.

Within this context the study of 'gaps' takes on considerable significance. 'Gaps' are structural positions in a sentence which intuitively have an element missing. The sentences in (1) all exhibit gaps in this sense. In these examples the underscoring represents the gap.

(1)a. what did they put ____ on the front page
    b. they promised Bill ____ to leave early
    c. Jane seems ____ to be nervous

Determining whether these gaps have any properties bears on the question of Universal Grammar outlined above. The fact that gaps lack any phonetic realization makes it unlikely that their properties can be viewed as present in the data that the language learner experiences. Consequently those properties can reasonably be regarded as originating in Universal Grammar. In this way the study of gaps gives a comparatively unimpaired view of Universal Grammar.

The Government Binding Theory is a relatively sophisticated
set of principles which interact to determine the distribution of
gaps. Oversimplifying somewhat, the presence of gaps in (1) is
accounted for by the Projection Principle of Chomsky (1981) which
requires that if a predicate has a subject or an object at deep
structure, it will have a subject or an object at every other
linguistic level, including surface structure and the level of
Logical Form. In this way, the verb put of (1a) must continue
to have an object even after movement of what and the infinitive
of (1b) must have a subject at surface structure if it is to have
one at Logical Form. Consequently Government Binding Theory
recognizes abstract, phonetically empty categories in (1a) and (b)
as represented in (2).

(2)a. what\textsubscript{i} did they put [\text{e}]\textsubscript{i} on the front page
b. they\textsubscript{i} promised Bill [\text{e}]\textsubscript{i} to leave early

These phonetically null categories are not homogeneous however and
must be distinguished further. Let us call the empty category in
(2a) 'trace,' and that in (2b) 'PRO.' Government Binding Theory
claims that trace and PRO differ in the ways outlined in (3). These
differences are of special interest because they reflect
properties of Universal Grammar.

(3)a. Trace but not PRO is subject to the Subjacency Condition of Chomsky (1977).
b. Trace must be governed but PRO must not be.
c. Trace and its antecedent occupy only one argument position but PRO and its antecedent occupy two
distinct argument positions.

Property (3a) follows from bounding theory, (3b) follows from
binding theory and the Empty Category Principle, and (3c) follows
from the \( \emptyset \) Criterion. These sub theories of Universal Grammar
interact to virtually insure that PRO and trace are in complementary
distribution. For an exposition of these sub theories of Universal
Grammar see Chomsky (1981), (1982) and the references cited there.
I will not pursue the distinguishing properties of trace and PRO
any further here. At this point I would like to focus my attention
on the requirement that traces must be governed in some sense. This
requirement in known as the Empty Category Principle (ECP) and is
reproduced as (4).

(4) An empty category \([\text{e}]\) must be 'properly governed,' where
\( \alpha \) properly governs \( \beta \) if and only if \( \alpha \) governs \( \beta \) and:

\begin{enumerate}
  \item a. \( \alpha = +N, +V \), or
  \item b. \( \alpha \) is \textit{coindex}ed with \( \beta \)
\end{enumerate}
The empirical consequences of the ECP are that movement operations will exhibit subject-object asymmetries. Traces in object position will always be properly governed by the verb, as in condition (4a), whereas a trace in subject position must be properly governed by a coindexed antecedent, as in condition (4b). Since government in essence a local relation holding between adjacent constituents, the ECP will permit objects to move freely since their traces are inevitably governed by the verb. Subjects on the other hand are restricted to moving to an adjacent, unfilled COMP since otherwise it will be impossible to provide a governor for the trace in subject position.

My remarks are organized in the following way. The first section outlines some corroborating evidence for the ECP culled from the interaction of Subject Auxiliary Inversion and WH-movement in English. The second section considers the claim advanced in Chomsky (1981) that the ECP holds at the level of Logical Form. This claim is naively falsified by the rule of Quantifier Raising (QR). In the second section I suggest a revision of QR, following Williams (1977), which makes the operation of QR consistent with Chomsky's claim that the ECP holds at Logical Form.

I. Corroborating the ECP

In this section I want to show that the general thrust of the ECP finds independent verification from phenomena involving Subject Auxiliary Inversion (SAI) in English. Of course SAI functions to move an AUX to the left of a subject as in (5).

(5)a. can you ___ see Jane today?
   b. have you ___ seen Jane today?
   c. are you ___ seeing Jane today?

(6)a. who can you see ___ today?
   b. who have you seen ___ today?
   c. who are you seeing ___ today?

(7)a. who can see Jane today?
   b. who has seen Jane today?
   c. who is seeing Jane today?

This operation applies not just in yes-no questions like those in (5) but in WH-questions like (6) as well. It turns out that there is a well known subject-object asymmetry in the interaction of SAI and WH-movement: although SAI apparently applies when an object is moved by WH-movement, there is no apparent application of SAI when a subject is so questioned. Thus we can contrast the sentences in (6) which exhibit SAI with those in (7) which superficially do not.
Classically, generative grammar has explained this asymmetry by claiming that SAI is ordered prior to WH-movement. Such an ordering statement within a Standard Theory model will give the desired results: the sentences in (7), it would be argued, have undergone SAI just as the sentences in (6) have, the only difference being that when WH-movement is subsequently applied to (6) and (7) the effects of SAI are 'undone' or reversed in (7).

The ordering of SAI before WH-movement can be accomplished in one of several ways, either by some extrinsic ordering statement, by some intrinsic ordering solution relying on an elaboration of the structural description of SAI, or by the principle of the strict cycle is S and S are cyclic and SAI operates within S. None of these potential solutions is compatible with the Revised Extended Standard Theory (REST) however. Any extrinsic ordering statement would require access to negative evidence and is subject to the critique presented in Baker (1979) which shows that such solutions are unlearnable. Any attempt at an intrinsic ordering solution which requires an elaboration of the structural description of SAI is a step away from core grammar and the leading idea that surface structure is simply the product of a general rule, 'move constituent.' Finally any attempt to invoke the cycle encounters the problem that, as Freidin (1979) shows, the cycle is virtually a theorem of the binding conditions that hold at Logical Form and has no independent existence. Yet nothing in the binding conditions will account for the asymmetry apparent in (6)-(7).

The ECP however is capable of providing a principled explanation for the contrast in (6)-(7) consistent with the assumptions of the REST. The answer is that SAI is free to move AUX between a WH-phrase and a subject in (6) without violating the ECP since the trace of the WH-phrase is governed lexically by the verb.

(8)a. who can you see [e] today?  
    b. who can see Jane today?  
(9)a. who can see Jane today?

This can be seen in (8a). On the other hand, SAI has superficially not applied in (7) because it in fact has not applied: both SAI and WH-movement cannot operate in (7) without violating the ECP. This result can be observed in (8b). There who is not in a position to properly govern its trace because it fails to meet the locality requirement crucial to the notion of government. Instead (7) must have the structure in (9) where only WH-movement has applied. I return in section 2 to consider why it should be the case that SAI and not WH-movement should fail to apply in sentences like (7). I will also return shortly to some questions of execution that insure
the inverted auxiliary suffices to block proper government of the trace in (8b). For the moment I will explore a bit further the leading idea that the ECP prevents SAI when a subject is moved by WH-movement. I want to outline two pieces of independent evidence for this position.

The first piece of evidence comes from the distribution of do in questions. The distribution of do raises a problem for trace theory within the REST that is avoided under the general line of analysis just advanced. In the analysis of the English auxiliary system outlined in Akmajian, Steele and Wasow (1979), it is assumed that do is present in deep structure, provided by a rule of the categorial component something like (10), and that it is deleted by a rule like (11) when adjacent to a verb.

\[(10)\quad \text{AUX} \rightarrow \text{TENSE} \left\{ \text{Modal} \right\} \]

\[(11)\quad \text{Do Deletion} \]

\[X - \text{do} - V - Y \Rightarrow 1 - \emptyset - 3 - 4 \]

If (11) does not apply, do adjacent to a verb is stressed and is emphatic. This analysis explains why do may or may not appear when the AUX is adjacent to a verb but must appear when the AUX is separated from the verb. In (12a) do has deleted because it was adjacent to the verb but in (12b) it is not free to so delete because the negative prevents Do Deletion from applying.

\[(12)\quad \text{a. they lied} \]
\[\quad \text{b. they did not lie} \]

\[(13)\quad \text{a. who did Jane see} \]
\[\quad \text{b. *who Jane see} \]

\[(14)\quad \text{a. who saw Jane} \]
\[\quad \text{b. who did Jane see} \]

This is the traditional analysis of the distribution of do. It extends quite naturally to questions. In (13)-(14) it appears that do is required when an object is fronted by WH-movement but not when a subject is so fronted. The traditional analysis of these facts is that if Do Deletion is ordered to follow WH-movement these facts follow immediately. The idea is that in (13a) Do Deletion cannot apply because do is not adjacent to the verb. In (14a) however Do Deletion can apply because after SAI and WH-movement have operated the AUX is once again adjacent to the verb.

This explanation is endangered however by trace theory. If movement rules leave traces, then (14a) would actually have the structure in (15) after SAI and WH-movement have applied.
(15) who_i did [e_i] see Jane

In (15) did is no longer adjacent to see and it should behave exactly as (13) does. That is, if (15) is the correct structure underlying (14), (14a) should not be possible. At this point if we want to keep our analysis of the auxiliary constant and not abandon trace theory, we are forced to claim that (15) is not the structure underlying (14). Instead (14) must have the structure in (16) where SAI has not applied.

(16) who_i [e_i] [TENSE do] see Jane

In (16) it is possible to perform Do Deletion because do is adjacent to the verb. We can explain why the structure in (15) does not underlie the sentences in (14) by appeal to the ECP. If SAI had applied to produce (15) it would not be possible for who to properly govern its trace. As a result SAI cannot apply when a subject is fronted by WH-movement. Of course no similar problem arises if an object is fronted, as in (13). In these cases SAI applies without producing a violation of the ECP since the trace of the WH-phrase is still properly governed by the verb.

A second piece of evidence supporting the claim that SAI can not apply when a subject is fronted without violating the ECP comes from reduced questions. Reduced questions are yes-no and WH-questions in which the auxiliary is deleted. The sentences in (17) are examples of reduced questions.

(17)a. are you happy today?  ➔ you happy today?
    b. how happy have you been lately?  ➔ how happy you been lately?

These structures, common to informal English, are studied in some detail in Hendrick (1982). The characteristic of this construction relevant for our purposes is that the deletion of the auxiliary only occurs if the auxiliary is fronted. So there are no declaratives in informal Standard English which exhibit this reduction; only questions exhibit it. I want to use (18) as a diagnostic for SAI.

(18) An auxiliary is deleted in informal Standard English only if it has been fronted by SAI.

There is a curious asymmetry between subjects and objects with respect to reduced questions. It turns out that reduced questions are possible only when an object is fronted by WH-movement and not if a subject has been so moved. This fact is exemplified in (19).
CORROBORATION FOR THE ECP

(19)a. who have you been seeing today?
   b. who you been seeing today?
   c. who has been seeing Jane today?
   d. *who been seeing Jane today?

Sentence (19c) does not permit the deletion of AUX, as (19d) shows, whereas (19a) does undergo the operation to produce the well formed (19b). These facts would follow immediately from (18) in conjunction with the leading idea of this section. In (19c) the auxiliary could not have been fronted by SAI without violating the ECP. Consequently it is impossible to delete the auxiliary without violating (18). In (19a) however no similar problem presents itself. There SAI can operate without contravening the ECP and as a result the auxiliary is free to delete in accordance with (18).

I conclude from the foregoing that the ECP prevents SAI from applying when a subject has been fronted by WH-movement. SAI can only apply when a non subject has been fronted by WH-movement.

Before going on to the next section of these remarks, I want to devote some attention to a 'question of execution.' In particular I want to consider briefly just how the auxiliary between a WH-phrase and its trace prevents the WH-phrase from properly governing that empty category. There are two ways in which this result can be achieved. The first is to adopt the definition of government in (20), a definition adopted in Chomsky (1980).

(20) $\alpha$ governs $\beta$ if $\alpha$ c-commands $\beta$ and no major category or major category boundary appears between $\alpha$ and $\beta$.

This definition requires that $\alpha$ c-command and be adjacent to $\beta$ before $\alpha$ can be said to govern $\beta$. Under this conceptualization of government a WH-phrase will be unable to properly govern a trace in the structure (21) because AUX intervenes.

(21) who$_i$ can see John

In order for this explanation to go through notice that we must assume that AUX (or INFL) counts as a major category for the purposes of (20). This assumption is perhaps natural given the frequent claim that AUX (or INFL) is the head of S.

This solution is only as strong as the requirement in (20) that $\alpha$ and $\beta$ be adjacent. It has been suggested that the existence of double object constructions, if they have a structure like (22), warrant a revision of (20) that does not require adjacency for government.
It seems that the verb must govern NP₂ as well as NP₁ in (22) since it is possible to front NP₂ by WH-movement, a possibility that would violate the ECP unless the trace in the NP₂ position was governed by the verb. Yet only NP₁ satisfies the definition of government in (20). I will not enter the debate whether V should govern, in some unmarked fashion, NP₂. I will simply note that unless the adjacency requirement of (20) can be sustained we need some other explanation for why the WH-phrase fails to govern its trace in (21) and that the double object construction bears on this issue.

A second potential solution, not based on requiring adjacency for government, is to follow Kayne (1980). In that work both S and 5 are argued to be barriers to government and that only one such barrier can be overcome or 'erased' in instances of exceptional government. Let us assume that SAI is an adjunction to S. Such an assumption would mean that (21) has the structure in (23).

(23) who₁ ₑₛ can ₑₛ ₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑₑ euler see Jane]

If, as Kayne suggests, S is a barrier to government and only one such barrier can be deleted, there is no way that who can come to govern its trace in (23) and the structure will be ill-formed. If SAI had not applied who would have been able to govern its trace because in that case only one barrier to government would have intervened. Of course this solution would be subject to the problems with Kayne's proposal observed in Chomsky (1981: 296) but those difficulties may be obviated by further research.

One of the two approaches outlined above must be correct however if the general idea advanced in this section is to be sustained. I will not pursue the issue further here.

I conclude from the foregoing discussion that the ECP offers an attractive explanation for the interaction of SAI and WH-movement in English. SAI can be said to exhibit the ECP effect. I will now go on to consider at what grammatical level the ECP should hold.

II. The ECP at LF

Given a conception of the organization of the grammar like that in (24) it is worthwhile to ask at what level the ECP holds.
(24)  
\[
\begin{array}{c}
\text{deep structure} \\
\downarrow \\
\text{'move } \alpha \text{'} \\
\downarrow \\
\text{surface structure} \\
\downarrow \\
\text{phonic representation} \quad \text{logical form}
\end{array}
\]

In particular we might consider whether the ECP holds at the level of surface structure, or phonetic representation, of Logical form, or of all three.

It has been claimed that the ECP holds at least at LF. The argument is attractive and consists of showing that the contrast between (25a) and (26a), sometimes called the 'superiority facts,' follow from the ECP.

(25)a. I wonder who saw what  
b. I wonder (for which y, y a person (for which x, x a person (x saw y)))

(26)a. *I wonder what who saw  
b. I wonder (for which x, x a person (for which y, y a person (x saw y)))

There is a rule, called WH-Raising, that applies to surface structure to help derive LF. WH-Raising functions to adjoin a WH-phrase in situ to some higher COMP. This rule will thus construct the LF's in (25b) and (26b). Only the ungrammatical (26b) fails to have the variable x in subject position adjacent to its WH-quantifier and we could explain the ungrammaticality of (26a) by extending the ECP so that it held at LF. (26b) would be an ill formed LF because the variable in subject position is not properly governed. There appears to be some reason for claiming that the ECP holds at LF.

Attractive though this conclusion is, it has certain problems. In particular it would lead us to expect that all rules which help to derive LF should exhibit the superiority facts. This expectation is frustrated however by Quantifier Raising (QR). May (1977) suggests that QR functions to raise a quantifier to sentence initial position at LF, adjoining it to S. QR will apply twice in a sentence like (27a) and, depending on which quantifier is extracted first, result either in (27b) where everyone is given wide scope, or in (27c) where everyone is given narrow scope.

(27)a. everyone loves someone  
b. (\forall x, x a person (\exists y, y a person (x loves y)))  
c. (\exists y, y a person (\forall x, x a person (x loves y)))
This is a relatively straightforward account for the ambiguity of (27a). Yet despite its attractiveness, notice that (26b) in which everyone is given wide scope apparently violates the ECP since the variable in subject position is not properly governed by its quantifier. Evidently (27a) poses significant problems in the attempt to extend the ECP to LF.

Now, as it turns out, it is important for the argument outlined in the first section of this essay that the ECP hold at LF. Recall that our discussion of the interaction of SAI and WH-movement revealed that the ECP prevented both from applying when the WH-phrase originated in subject position. We might ask however why it is that English exhibits (29) instead of (28) given that both are consistent with the claim that SAI and WH-movement from subject position cannot both apply. Unless such an explanation is forthcoming we have failed to explain the interaction of WH-movement and SAI in English.

(28) *can\[e\]i see Jane?
(29) whoi [e]i can see Jane?

(28) is ill-formed, except as an echo question. This result can be obtained by assuming that the ECP holds at LF. WH-Raising would have to apply in (28) to raise the WH-phrase in situ to COMP so that it could be interpreted at LF. Such a raising would result in an LF like (30).

(30) (for which x, x a person ( can x see Jane))

In (30) the variable x is not properly governed by the WH-quantifier because can intervenes. (30) is excluded for the same reason that (8b) is excluded: the empty category in subject position is not properly governed, in violation of the ECP. Of course this explanation for the ungrammaticality of (28) only follows if the ECP holds at LF after WH-Raising.

The claim that the ECP holds at LF is doubly important: on the one hand, it provides an explanation for the superiority facts in (25)-(26), and on the other hand it also accounts for the ungrammaticality of (28). In this light the anomalous behavior of QR with respect to the ECP is embarassing and we would like to develop an explanation for why QR is apparently insensitive to the ECP. In the remainder of this section I want to suggest that Williams (1977) provides a way of avoiding this problem.

Williams (1977) provides a broad survey of a number of rules of sentence grammar responsible for the derivation of LF. For our purposes it is worth observing that Williams formulates QR as an
adjunction either to S or to VP. Adding VP as a landing site for QR opens an explanation for the ambiguity of (27a) that is consistent with the claim that the ECP holds at LF. QR would, under Williams' formulation, give any of the structures in (31).

(31)a. (∀x, x a person (∃y, y a person ( x loves y )))
  b. (∃y, y a person (∀x, x a person ( x loves y )))
  c. (∀x, x a person ( x [VP ∀y, y a person (loves y)])

Of course as we have observed before, (31b) which corresponds to the wide scope reading of someone is permitted by the ECP. (31a) which represents the wide scope reading of everyone violates the ECP. Yet everyone can still be interpreted as having wide scope because QR also provides the LF in (31c) where someone is within the scope of everyone but does not interfere with the proper governing of the empty category in subject position. In this way Williams' formulation of QR allows us to exclude (31a) as a violation of the ECP without giving up an explanation for the ambiguity of sentences like (27a).

Williams (1977) outlines independent evidence for believing that QR has VP as well as S as a landing site. Here I will present another argument to the same effect based on the interaction of QR and facts concerning null anaphora.

Besides providing evidence for his formulation of QR, Williams also presents evidence in favor of a second rule, the Derived VP Rule that has its origins in work reported in Partee (1973). The Derived VP Rule (DVPR) 'will affix a lambda and a variable to a VP and it will place a variable bound by the lambda in the position of the logical subject of the verb of the VP' (Williams 1977:115). In this way the DVPR will derive (33) from (32).

(32) Jane talks.
(33) Jane [ vp \lambda x (x talks)]

In (33) the variable x occupies the position of the subject of talk. The schema in (33) can be read as 'Jane has the property of talking.'

QR and the DVPR interact in such a way that they make an interesting empirical claim. As Williams points out, the sentence in (34) may have the LF of either (35) or (36), depending on whether one selects the VP or S as the landing site of QR.

(34) Jane saw everything
(35) Jane [ vp \lambda x ( \forall y (x saw y))]
(36) $\forall y \text{ (Jane } \lambda x \text{ (x saw y))})$

In (35) the VP variant of QR has been performed while in (36) the S variant of that rule has operated. Now these structures lead to an interesting consequence when (37) is given some consideration.

(37) Bill did $[\lambda a a]$ too

Williams proposes to treat the null anaphora in (37), studied in some detail in Wasow (1972), by the 'VP Rule'. Williams proposes that 'the VP rule copy the antecedent VP into the position of the missing VP, replacing the deltas of the missing VP by the corresponding terminals of the antecedent VP' (Williams 1977:127). If the VP Rule interprets (37) by copying the VP in (35), the result will be (38), while if it copies the VP in (36), the result will be (39).

(38) Bill did $[\lambda a (\forall y \text{ (x saw y))}]$
(39) Bill did $[\lambda a (\text{ x saw y})]$

Of these resulting structures, only (38) is well formed: (39) is uninterpretable because it contains an unbound variable, $y$. Williams observes that if his analysis is correct, 'only VP's that contain quantifiers that have been interpreted as subordinate to the lambda operator can be used as antecedents by the VP rule' (Williams 1977:127). In other words, in instances of VP anaphora like (37), only the VP variant of QR leads to well formed results. I now propose to present some independent evidence suggesting that this prediction is correct.

Consider the sentence in (40).

(40) A girl got 100% on every exam.

Sentence (40), like sentence (27a), is ambiguous, depending on which quantifier receives wide scope. That is, (40) will have either (i) the reading in which there is some girl such that she received 100% on every exam, or (ii) the reading in which every exam was such that a girl received 100% on it. QR will capture these readings by assigning the LF in (41) to the first reading and the LF in (42) to the second reading.

(41) $\exists y, y \text{ a girl (y [VP } \forall x, x \text{ an exam (got 100% on x)})}$
(42) $\forall x, x \text{ an exam } y \text{ a girl (y got 100% on x)}$

Application of the DVPR to the structures in (41)-(42) will produce (43) and (44) respectively.
(43) \( \exists y, y \ a \ girl \ (y \mapsto [\forall x, x \ an \ exam \ (y \ got \ 100\% \ on \ x)]) \)

(44) \( \forall x, x \ an \ exam \ (\exists y, y \ a \ girl \ (y \mapsto [\forall y \ y \ got \ 100\% \ on \ x]) \))

Williams' analysis predicts that, confronted with the elliptical discourse in (45), only (43) can provide an antecedent for the VP anaphora in (45b).

(45)a. I heard a girl got 100% on every exam.
   b. That's odd; I heard a boy did [\forall y \ \Delta \Delta \Delta \Delta \Delta]

Copying the VP of (44) into (45b) will copy an unbound variable, \( x \), and result in an uninterpretable structure. Only (43) presents a VP that has all of its variables appropriately bound. In this way while (40) in isolation is ambiguous, depending on whether a girl or every exam is given wide scope, (45) can only have the reading in which every exam is given narrow scope and a girl is given wide scope. Williams' prediction with respect to (45) seems to be correct and every exam does in fact seem to require narrow scope in (45).

In this way we are able to provide a relatively natural explanation for the contrast between (40) and (45) if we assume that QR has VP as well as S as a landing site. Moreover it allows an explanation for the ambiguity of sentences like (40) and (27a) without letting QR violate the ECP. We thus are in a position to maintain the claim that the ECP holds at LF and preserve our explanation of the superiority facts in (25)-(26) and the ungrammaticality of (28).

III. Footnotes

*This work was supported in part by the University Research Council of the University of North Carolina at Chapel Hill. I want to thank Carlos Piera for discussion of the material presented here.

1In this discussion I am abstracting away from examples of 'exceptional government' that occur in subject raising constructions. The ECP will account for the contrasts between the (a) and (b) sentences below.

(i)a. *the only person who I don't know when I can get to see ___ is John
   b. *the only person who I don't know when ___ can get to see me is John
(ii)a. which boy did she say ___ visited you
    b. which boy did she way that ___ visited you


2 I intend to show that the 'general thrust' of the ECP is correct. I say 'the general thrust' because it may well be that further research can improve the explanatory power of the ECP by showing that the restrictions in (4a) and (4b) are a false generalization, or by some other modification. No doubt some such advance is imminent. The value of these remarks, in my view, is to show that SAI exhibits what we might call 'ECP effects'. Any attempt to replace the ECP in (4) would thus need to provide some explanation for why SAI exhibits such ECP effects.

3 I am abstracting away here from sentences in which both a subject and an auxiliary are missing in informal Standard English.

4 Of course other possibilities may be conceivable though they escape me.

IV. References


CORROBoration FOR THE ECP


Some Remarks On Binding Theory And Logical Form*

James Higginbotham

MIT

In this discussion, I will outline a point of view on the nature of representations at the linguistic level LF of Logical Form, and I will consider two concrete issues to which the properties of LF-structures seem to be pertinent. My general view is that, over a wide domain, LF-structures are restricted to a few varieties that I shall describe, and that principles of interpretation apply to each of these varieties in a uniform manner, as explained more fully below. The concrete issues are, first, the elaboration of the theory of binding in terms of an asymmetric relation of linking, applying between anaphoric elements and their potential antecedents; and, second, the explanation of the curious absence of "raising" nouns, that is, of derived nominals such as (2), corresponding to sentences such as (1):

(1) John is likely [t to leave].

(2) *John's likelihood [t to leave]

I will discuss both of these questions in terms of the general view on the nature and interpretation of LF-structure.

The linguistic level LF, following Chomsky (1976) and subsequent writings, is taken to be the image of S-Structure under some rule or rules to be determined, and is assumed also to be
the sole linguistic level relevant to interpretation. Within a
programme of research that aims at comprehending the basis for
human linguistic knowledge, the properties of LF may be of parti-
cular interest; for if LF-structures may differ significantly
from S-Structures, then the remoteness of logical forms from lin-
guistic experience will tend to make the relation between the
evidence available to the child, on the one hand, and the pro-
properties of the system grasped, on the other, so marginal that we
should expect the evidential gap to be closed largely on the
basis of unlearned principles of grammar. A good working hypo-
thesis would be that the shape of LF is in fact completely fixed
(apart from the meanings of lexical items), and that is the
hypothesis that I shall adopt here.

LF-structures look toward (the rest of) syntax, and toward
semantics. The past several years have seen many proposals on
the nature of rules mapping from S-Structure to LF (the syntactic
side), and on the interpretive principles that might be supposed
to apply at LF (the semantic side): I have in mind work on the
scope of quantifiers and WH-expressions, rules of predication
in the sense of Williams (1980), principles of assignment of
anaphor-antecedent relations, and the like. I will take this
work as a point of departure.

Suppose, following May (1977), that in configurational
languages quantificational NP's are assigned scope at LF through
the transformational rule QR, a rule of adjunction. Then, for
instance, from the S-structure underlying (3) we will obtain at
LF the structure (4), where the trace \( t \) of the application of QR
marks the place of a variable:

(3) John saw everybody.

(4) \([\text{Everybody}] \text{ John saw } t\].

In general, QR Chomsky-adopts a quantificational NP to some
admissible position, or "landing site," its scope then extending
over that site; i.e., over its c-command domain. In the case of
(4) in particular, the variable \( t \) is within the scope of the
quantifier everyboby.

In May's theory, the relative scopes of quantifiers are
entirely fixed at LF. Hence, we can explain ambiguities of
scope in terms of the availability of multiple representations
at LF; inversely, we can relate the occasionally surprising lack
of ambiguities to conditions on QR, or to independent filters
that rule out certain structures at LF, and so forth. Of course,
in all such explanations, semantic principles are presupposed:
one of my purposes in this discussion will be to make these pre-
suppositions more explicit.
Alternatives to May's account of the scope of quantifiers may and have been envisaged; see for instance Vanlehn (1978), Hintikka (1982), and references cited therein. More recently, May (1982) argues for a revision of his own, earlier, account. Rather than explore these alternatives, however, I would like to press the question whether May's (1977) treatment of quantifical scope can be extended to all cases of scope-assignment (for instance, of WH-expressions and adverbs); and further whether any other rule than scope-assignment mediates between S-Structure and LF.

Consider possible transformational rules mapping between S-Structure and LF. If we assume the projection principle of Chomsky (1981), then these rules are forbidden from assigning to a given S-structure an LF-representation to which it fails to conform in categorial selection (see Chomsky (1981), p. 38 ff.). QR is consistent with the projection principle; specifically, the selection observed in (3) is not disturbed in (4), because the NP everybody appears there as an operator, to which selection does not apply. Few other hypothetical rules could be consistent with the projection principle, which if assumed acts as a narrow gate through which any proposed rule must pass.

Assuming QR, the following structures will be the major types available at LF:

(5) a. \[X^n ... X^{n-1} ...]  
b. \[X^{\text{SPEC}} X [X ...]]  
c. \[X^O [X ... O(t) ...]]

(5a) is the familiar structure of head and arguments; (5b) creates maximal projections by adding specifiers; (5c) is the form of structures consisting of an operator O on X with an argument t(O) within its scope. The types (5a) - (5c) are all available at S-Structures as well as LF, the first two by base rules, and the third by any sort of peripheral movement with Chomsky-adjunction, as discussed for instance in Baltin (1981).

Let us concentrate for a moment on the structures (5b). For X = S, we assume that SPEC X = COMP, so that SPEC X may be filled by a complementizer, by words such as if or whether, by WH-expressions, or by certain frontable "negative" quantifiers, as in (6):

(6) [Not a single book] did John read t.

From the point of view of interpretation, we might take (6) as a case where the semantic equivalent of QR has been carried out in
the syntax. The WH-expressions, whether related to argument positions in S, or, like if and whether, given in place, might be supposed sentential operators. Finally, complementizers don't seem to contribute to interpretation at all. For the case X = S, then, we could regard (5b) as semantically equivalent to (5c), in the sense that SPEC X is an operator O, provided that it is semantically significant.

What of the case (5b) with X = NP? Bypassing complexities of the sort discussed, e.g., in Jackendoff (1977), we will have SPEC X = Det (the, every, which, ...), or SPEC X = NP. We suppose that determiners bind variables, so that phrases like every man are taken in the way suggested by (7):

(7) [Every x: man (x)]

These phrases become restricted quantifiers once QR has applied. This sort of account can be elaborated for the cases of determiners that may build non-quantificational singular terms, such as the, and for WH-expressions; we are left with a type of variable-binding different from that shown in (5c), but, it appears, readily understandable. Finally, we have the structures (8):

(8) [NP NP₁'s N]

(John's cat, my beliefs, ...); our assumption here will be that they are interpreted as shown in (9):

(9) [the x: N (x) & R (x, NP₁)]

where R expresses some contextually understood relation. In the case of John's cat, for instance, with the understood relation of ownership, we would have (10):

(10) [the x: cat (x) & John owns x]

We elaborate below on some consequences of this view of the interpretation of structures of the form (8).

If (5) gives the basic types of structures available at LF, with subtypes and interpretive rules as sketched above, then we have a very tight theory of both the structure and semantics of LF for the most fundamental and productive phrase-types. We are left with a picture of LF-structures that might be expressed succinctly by saying that they instantiate the schemata of generalized quantification theory. May's QR as originally formulated is not only the sort of rule one should expect if this picture is approximately correct: it is virtually the only sort of rule available, or needed. We can envisage extensions of QR
that assign scopes to arbitrary elements that are neither heads nor arguments of heads. The elements of LF built by an extended scope-assignment rule applying to S-Structures would bear more than a casual relation to the structures proposed some years ago by James McCawley and others in Generative Semantics.

Pursuing our picture in detail, complications and exceptions are bound to arise. For example, I've argued elsewhere (Higginbotham (1981), forthcoming) that the LF-structure for a perception sentence like (11):

(11) John saw Mary leave.

is in fact (12):

(12) [[Mary leave] John saw t]

with t in effect an NP-trace (see Pesetsky (1981) for a discussion of categorial selection of traces, with interesting applications to certain quantifier-phrases in Russian). By associating (11) with the logical form (12), I've tried to account for some of the syntactic and semantic peculiarities of perception-verb sentences (and of certain causatives, to which the account naturally extends). In the present connection, the point is that (12) does not directly instantiate a quantificational schema; so, following in part a suggestion of Jon Barwise, one recognizes a hidden existential quantification in the adjoined small clause, giving a representation that comes out as (13):

(13) [E x: leave (Mary, x)] John saw x.

with the interpretation, roughly, "There was a departure by Mary that John saw." Small-clause complements to perception verbs, then, are for me one example where the guiding idea expressed above has to be spelled out further. The phenomena of referential opacity and idioms are other areas, raising issues that I won't consider here.

With this much background, I wish now to consider some of the central questions of binding theory, the first of the concrete issues to be discussed here. The structure of binding has for some years been expressed by means of coindexing, the assignment of indices being subject to conditions to be stated in binding theory. The numerical indices, of course, are meaningless in themselves; what they do is serve to indicate the relation of anaphor to antecedent. In the account of Chomsky (1981), which I take as point of departure, each element is assigned some index or other (with perhaps special conventions in the case of applications of movement rules), and the theory of binding governs the output of such indexing.
Now, it is obvious that coindexing, as against the direct assignment of anaphoric relations, loses information. Because coindexing does not show the direction in which antecedence was assigned, there are in general many assignments of anaphoric relations that yield identical indexed structures. We may, therefore, ask whether the fundamental conditions on binding are expressed in terms of the assignment of anaphoric relations, or else over the indexed structures that result therefrom (or, in principle, something intermediate between the two). I think that there is good evidence that the indexed structures should not be taken as fundamental; evidence, in other words, that the extra information given by the direct assignment of anaphoric relations is needed for adequate linguistic descriptions.

Let us represent the assignment of the anaphor-antecedent relation to two positions by linking those positions with a headed arrow, whose head is the antecedent, as in (14):

(14) John said he thought Mary liked him.

The linking shown is one among several that would reduce to the indexed structure (15):

(15) John₁ said he₁ thought Mary liked him₁.

We suppose that linking is given by a rule that applies freely between arguments at the level of S-Structure, the same rule applying automatically in the case of an application of movement rules. Our replacement for coindexing is thus (16):

(16) Link X to Y.

where Y is the antecedent. We do not undertake here a full articulation of (16); but we will give some considerations that argue in its favor.

One problem with coindexing, as noted in Chapter 5 of Chomsky (1981), is that there is no direct way to represent anaphoric interactions between plurals, or between singulars and plurals. Thus, to take a well-worn example, the interpretation of (17) in which both John and Mary are to be included in the reference of the pronoun they is not immediately given at LF:

(17) John told Mary they should leave.

Howard Lasnik in (1981) raises further problems of this sort. The limitations imposed by coindexing ill-suit the general structure of the theory of Chomsky (1981), or any theory that otherwise allows the explicit representation of optional coreference.
Now, it is easy to see that the expressive problems posed by (17) can be overcome if we but allow elements to be anaphoric to more than one antecedent. The linking rule (16) does allow this, and so in particular permits (18):

(18) John told Mary they should leave.

An indexed reduction of (18) (abstracting from the directions of the arrows) might assign multiple indices to plurals, as in (19):

(19) John\(_i\) told Mary\(_j\) they\(_{i,j}\) should leave.

with appropriate changes in the clauses of the binding theory.\(^2\)

Replacement of coindexing with linking; i.e., replacement of the rule of free indexing with the rule (16), solves the problem of representing split antecedents. It does so, however, at the cost of introducing into binding theory a new complexity, one that arises precisely because linking, by representing the direction of antecedence, gives more information about a structure than coindexing does. Consider, in this respect, the simplest case for binding theory, namely the datum that in sentences like (20) we must have "disjoint reference":

(20) He saw John.

In Chomsky (1981) this datum is to follow from condition (C) of the binding theory, stated here as (21):

(21) An R-expression is not coindexed with a c-commanding argument.

(C) rules out the indexing shown in (22):

(22) *He\(_i\) saw John\(_i\).

where it is assumed that the representation (22) of (20) would have to be available if "non-disjointness" of reference could, in some appropriate sense, be assigned to the two NP-positions in (20). Now, on the present construal, with linking in place of coindexing, we require to rule out both (23) and (24), because either of these would give an interpretation in which he, John were coreferential:

(23) He saw John.

(24) He saw John.

This complication is a result of our assumption that antecedence is represented by the asymmetric relation of linking, rather than the symmetric one of coindexing.
Suppose we split the condition (C), or (21), of Chomsky (1981) into two conditions, as follows:

(25) If X c-commands Y, and Y is an R-expression, then Y is not linked to X.

(26) If X c-commands Y, then X is not linked to Y.

(25) rules out (23), and (26) rules out (24). We thus reproduce the empirical effect of (21) for the case under consideration.

Our treatment of (20) in terms of linking, and the conditions (25)-(26), is, other things equal, more complex than that of Chomsky (1981). However, we can gain some support for the more complex notion if we find that languages may obey one of (25)-(26) but not the other; for then we will have located a linguistic parameter that is not immediately expressible in terms of conditions on coindexing. Howard Lasnik seems to have found such a language in Thai. In that language, unlike English, epithets exhibit a full range of pronominal behavior. In particular, epithets may be c-commanded by their antecedents, so that the Thai analogue of (27) is grammatical:

(27) John thinks the nut is smart.

As in English, epithets may not c-command their antecedents; (28) is ungrammatical in both English and Thai:

(28) The nut thinks John is smart.

On the other hand, (C) of the binding theory of Chomsky (1981) is freely violated in the case of names, since the following is perfectly grammatical in Thai:

(29) John thinks John is smart.

Given linking, we can characterize Thai as a language for which (25) is subject to exceptions, although (26) holds in full generality. If extensions of Lasnik's data continue to support this view of one difference between Thai and English, then we have some evidence for linking over coindexing.

Another advantage of linking over coindexing is that it enables us to account straightforwardly for cases of circularity, discussed from a different point of view in Higginbotham and May (1981), and considered also in Brody (1981). Standard cases of this phenomenon are sentences like (30):

(30) His wife saw her husband.
where intuition suggests that we must have at least three persons; i.e., that we cannot have her anaphoric to his wife, and simultaneously have his anaphoric to her husband.

First, consider that from linking we can derive a generalized notion of antecedence. Roughly, the antecedents of \( X \) include the elements to which \( X \) is linked, the elements to which they are linked, ..., and so on. Thus in (31):

(31) John thought he would shoot himself.

with the links shown, we will say the John is in the extended sense an antecedent of himself; appropriately, since it is the name that ultimately fixes the interpretation of the reflexive pronoun. More precisely, we define:

(32) \( Y \) is an antecedent of \( X \) if \( X \) is linked to \( Y \) or, for some \( Z \), \( X \) is linked to \( Z \) and \( Y \) is an antecedent of \( Z \).

Antecedence of \( Y \) to \( X \) is a special case of dependence of \( X \) on \( Y \), in a sense that I believe approximates that of Evans (1980). The intuitive idea is that an item is dependent upon those elements from which it receives its interpretation, as both pronouns are dependent upon John in (31). Besides these simple cases, there are also more complex cases of multiple dependence, illustrated by (33):

(33) John told [his wife] that she was beautiful.

To interpret she, we must interpret his wife, which requires in turn interpreting the pronoun that it contains. The pronoun she is thus dependent upon John, because the latter is used in fixing the interpretation of the former. In view of this type of case, we define a general notion of dependence by (34):

(34) \( X \) is dependent on \( Y \) if (i) \( Y \) is an antecedent of \( X \); or (ii) \( Y \) is contained in an antecedent \( Z \) of \( X \); or (iii) for some \( W \), \( X \) is dependent on \( W \) and \( W \) is dependent on \( Y \).

We denote the relation of dependence by \( D^* \), and we propose the following condition on LF-representations:

(35) Not: \( D^* (X, X) \).

(35) is in a sense an interpretive condition, reflecting the fact that the interpretation of an item cannot be given in terms of that item itself.5

It now follows that sentences like (36) are "circular," with the linking shown, because (35) is violated:
Again, we have the (37) is not circular:

(37) His wife saw John, her husband.

(37) escapes the condition (35), because linking can be assigned as shown in (38):

(38) [His wife] saw John, [her husband].

These consequences, and others of a like nature, argue the advantages of the linking rule (16) as against coindexing.

I turn now to a consideration of some of the "crossover" phenomena in connection with the present account of the assignment of anaphor-antecedent relations. The phenomena have been most extensively discussed in connection with examples like (39)-(42):

(39) Who did he see t
(40) He saw everybody.
(41) Who did his father see t
(42) His father saw everybody.

In each of these examples, the interpretation of the pronoun as a bound variable is excluded, "strongly" in the case of (39) and (40), "weakly" in the case of (41) and (42). For the "strong" cases, the analogy between (39) and (40), on the one hand, and the simple (20), repeated here, on the other, leaps to the eye:

(20) He saw John.

That is, just as we observe disjoint reference in (20), triggered by the structural relation of c-command that obtains between the pronoun and the name, so we observe the impossibility of binding in (39), where the pronoun c-commands the site of WH-movement, and in (40), where it c-commands the quantificational NP to which it would be bound. Hence, if traces of WH-movement or QR are assimilated to names (R-expressions) in binding theory, the status of (39) and (40) follows at once.

The above considerations, due to Chomsky (1976) and elaborated in subsequent writings, dispose of the simplest cases of "strong crossover." Several more recent discussions (Higginbotham (1980a), (1980b), Reinhart (1980), Koopman and Sportiche (1981), Huang (1982)) have concentrated chiefly on the "weak" cases,
(41)-(42). I will not consider the "weak" cases here; instead, I will point out some issues for the "strong" cases that are not immediately solved by Chomsky's suggestion that operator-traces be assimilated to R-expressions for the purposes of the binding theory, and I will analyze these cases in terms of the notion of linking presented here, extending some ideas in Higginbotham (1980b).

From one point of view, taking variables as R-expressions is rather unnatural, at least if proper names are the paradigm of R-expressions; for, unlike names, variables are always cross-referencing devices, and have no inherent semantic content. More significantly, there are several cases for which taking variables as R-expressions fails to block binding that is, intuitively, "strongly" unavailable. One such case, of a sort discussed in Higginbotham (1980a) and (1980b), is exemplified by examples like (43):

(43) *[[Which biography of which artist] do you think he wants to read t].

The account suggested in (1980a) is demonstrably inadequate, for reasons given in (1980b). Besides the suggestion of (1980b), to be proposed below in a somewhat different form, I am aware of two other types of approaches to examples like (43). The first, which may be found at least as early as Wasow (1972), is that some form of reconstruction, or "layered traces," will assimilate (43) to cases like (39), or (44):

(44) *[With whom] did he speak t.

The second, due to van Riemsdijk and Williams (1981), proposes that binding theory applies at a level of "NP-Structure," at which the complex phrase in COMP in (43) appears in the position of its S-Structure trace. Since, for binding theory, the pronoun c-commands the phrase which artist, it follows that binding is blocked in (43) in the same way as in (40), where the pronoun c-commands a phrase subject to scope-assignment only after the conditions on binding have applied.

Reconstruction in the case of (44) is unexceptionable from the point of view toward LF taken in this paper. The logical form of (44) is (45):

(45) [Whom] he spoke with t'.

In the case of (43), however, since the entire phrase in COMP and the phrase which artist that it contains are both operators, subject to scope assignment, the view that logical forms instantiate quantificational schemata would mandate the logical form (46):
(46) [Which artist] [which biography of \( t \)] do you think he wants to read \( t \).

Worse than this, however, appeals to reconstruction or layered traces in the case of (43) and the like must be accompanied by special stipulations of two sorts. First, it must be said how scope is to be assigned to elements, such as the phrase which artist in (43), contained in a reconstructed phrase. Second, there must be some provision for the fact that WH-movement affects possibilities for optional coreference, although not for pronominal binding; thus, (47) is better than (43):

(47) [[Which biography of Picasso] do you think he wants to read \( t \)].

The latter point raises questions also for the analysis of van Riemsdijk and Williams, as pointed out in Zenon de Fourier (1980).

Reconstruction, conceived as an approach to (43) and similar examples, and the more precise suggestions of van Riemsdijk and Williams have in common the aim of assimilating the more recalcitrant cases of "strong" crossover to the more basic cases, such as (39)-(40). The latter, in turn, are taken to reflect an extension of the principles of grammar that prohibit a pronoun from having its antecedent within its own c-command domain. These principles, encapsulated in condition (C) of the binding theory of Chomsky (1981) (our (21) above), accord the status of R-expressions to operator-bound traces. Both the hypothesis of reconstruction and the proposal that binding theory applies at a level of NP-Structure prior to WH-movement have the effect of "putting back" the offending operator, or a trace of it, within the c-command domain of the pronoun that fails to admit binding. Higginbotham (1980b) suggested a somewhat different approach, one which assigned to (43), for instance, the logical form (46), and at the same time prohibited the binding of the pronoun on the grounds that it was not accessible to its operator, where accessibility was defined inductively on phrase-markers. Rather than rehearse (1980b), I will define accessibility in terms suitable to the present context, where linking has replaced coindexing, and then show, first, how this notion provides an analysis of (43) and similar examples as cases of "strong" crossover, and, second, show how the accessibility principle solves two further problems that seem quite beyond the scope of the methods suggested so far, stemming from the original proposal of Chomsky (1976).

It was said above that we might envisage extensions of QR to other elements than quantifiers, and in particular to all elements that were neither heads nor the arguments of heads. Elements subject to scope-assignment will be called operators. We may leave the extension of this concept partially open, but
operators certainly include quantificational and WH-elements, and adverbials. Not only may operators be assigned a scope, but, we may suppose, they must be assigned a scope to create a well-formed LF-representation. We assume a condition, really an extension of May's (1977) Condition on Quantifier Binding, stated as (48):

(48) All operators are assigned scope.

(48) makes scope-assignment to the embedded WH-phrase which artist obligatory in (43); since WH-phrases are operators, they cannot appear in the position of arguments at LF.

Recall that we assumed that linking was automatic under applications of movement rules. In consequence, application of scope-assignment in, say, (49):

(49) Everybody saw his father.

gives the logical form (50):

(50) [[Everybody] [[t saw his father]].

Furthermore, linking apart from movement was to be between argument positions at S-Structure. We further stipulate that such linking is to a position, and is preserved under movement. Thus, to derive the reading of (49) with his interpreted as a bound variable, we begin with the S-Structure (51):

(51) [[Everybody] saw his father.

and then apply QR, obtaining (52):

(52) [[Everybody] [[t saw his father]].

With these assumptions in place, our question is: what prohibits the example (43)? We will assume that it is the underrivability of the appropriate logical form. To make our suggestion precise, we require a few definitions.

By a formal variable I shall mean an empty category in an argument position that is linked to a non-argument; the operator to which a formal variable is linked will be called its binder. A sequence (v₁, ..., vₙ) of formal variables such that each vᵢ, 1 ≤ i ≤ n-1, is contained in the binder of vᵢ+1 will be called a V-chain. Some examples: in (50) above, (t) alone is a V-chain; in (46), (t', t), since the formal variable t' is contained in the binder of t. In (53):

(54) Every turn from every exit on some freeway is dangerous.
whose logical form is (54):

\[
(54) \quad [\text{Some freeway}] \ [\text{every exit on } t'' \text{]} \ [\text{every turn from } t']
\]

\[
\text{t} \ is \ dangerous.
\]

the sequence \((t'', t', t)\) is a V-chain, as are each of its consecutive subsequences. Sentences whose logical forms show still longer V-chains are easily constructed, and arise whenever we have "inversely linked" quantification, in the sense of May (1977).

Our fundamental idea is that for a pronoun to have for its antecedent a formal variable \(v\), it must be accessible to \(v\), through a certain V-chain. Suppose, then, that a pronoun \(P\) is dependent, in the sense of definition (34), upon a formal variable \(v\). Let \(C\) be the longest V-chain \((v_1, \ldots, v_n)\) such that \(v_1\) is \(v\), and the binder of \(v_n\) does not contain \(P\). Under these circumstances we say:

\[
(55) \quad P \ is \ accessible \ to \ v \ if \ v_n \ c \-commands \ P; \ and
\]

\[
P \ is \ not \ accessible \ to \ v \ if \ P \ c \-commands \ v_n.
\]

The criterion (55) partially fixes the extension of the notion "accessible to." The condition (56) is assumed:

\[
(56) \quad \text{If a pronoun } P \text{ is dependent upon a formal variable } v, \text{ then } P \text{ is accessible to } v.
\]

Let us see how (55)-(56) rule out binding in (39), repeated here:

\[
(39) \quad *\text{Who did he see t.}
\]

Given the linked S-Structure (57):

\[
(57) \quad [[[\text{Who} \ did \ he \ see \ t]]]
\]

(recall that linking is automatic when a movement rule applies), we wish to rule out any possible dependency of the pronoun \(he\) on the operator \(\text{who}\). Such dependency could be expressed directly by (58):

\[
(58) \quad [[[\text{Who} \ did \ he \ see \ t]]]
\]

or indirectly, by (59):

\[
(59) \quad [[[\text{Who} \ did \ he \ see \ t]]]
\]

But we have assumed that, apart from movement, linking applies only between argument positions, so that (58) is ruled out straightaway (that the linking rule applies only between argument positions is our analogue to Chomsky's (1981) view that
binding theory is fundamentally concerned with "argument binding," in his sense. In (59), he has the trace $t$ of WH-movement for its antecedent, and is therefore dependent upon a formal variable, whose V-chain is the one-element sequence $\{t\}$ consisting of that variable itself. Hence, in terms of our definitions above, $v_v = v = t$, and since $P = he$ c-commands $t$, it is not accessible to it. So (59) violates (56). 6

Reasoning similar to the above shows that the standard cases of strong crossover are derivable from our assumptions. More significantly, our analysis extends automatically to cases like (43) above, although here the reasoning is somewhat more involved. Given the S-Structure (60) for (43):

(60) [[Which biography of which artist] do you think he wants to read $t$.]

nothing prevents the linking of he to which artist, because at this level that phrase occupies an argument position. For this reason also, (47), with linking as shown in (61), is admitted at S-Structure:

(61) [[Which biography of Picasso do you think he wants to read $t$.]]

The crucial difference between the cases is that the phrase which artist, being an operator, must undergo the rule of scope-assignment. This rule applies without destroying the link established at S-Structure between the pronoun and the argument position, so that the result of its application must be (62):

(62) [[Which artist] [[which biography of $t'$] do you think he wants to read $t$.]]

But (62) violates (56). For in (62) the pronoun he has for its antecedent the formal variable $t'$, which heads up the V-chain (63):

(63) ($t'$, $t$)

Since he c-commands $t$, it is not accessible to $t'$. Our reasoning shows that scope-assignment cannot apply, because that would violate (56); but scope-assignment must apply, because WH-phrases are operators, and this is a contradiction.

Our analysis of (43) has the advantage over those discussed above that it explains why WH-movement, although it affects possibilities for optional coreference, does not affect possibilities for pronominal binding. We proceed now to the discussion, promised above, of two further advantages of the system proposed here.
The following generalization, drawn in part from the seminal work of Jacobson (1977), and expressed here in terms drawn from our previous discussion, appears to hold without exception.

(64) A pronoun P can be dependent upon an operator O only if the rules of grammar would permit O to be its antecedent.

In other words, a pronoun that is considered as merely dependent upon a variable obeys the same principles of crossover that it would if considered as having that variable for its antecedent. The following contrast is, I believe, typical:

(65) [[Which man [[who] t admires his wife]] t' tries to please her]

(66) [[Which man [[who] t admires his wife]] does she try to please t']

(65) is well-formed, admitting the interpretation shown in (67):

(67) [Which x: man (x) & x admires x's wife] x tries to please x's wife.

the LF-representation with this interpretation being secured through the linking given in (68):

(68) [[Which man [[who] t admires [his wife]]] t' tries to please her]

(66), however, does not admit the interpretation corresponding to (67) of (65); that is, the one shown in (69):

(69) [Which x: man (x) & x admires x's wife] x's wife tries to please x.

To see how this contrast illustrates the generalization (64), observe that the interpretation (69) could only be secured by the linking shown in (70):

(70) [[Which man [[who] t admires [his wife]]] does she try to please t']

in which the pronoun she is dependent, in the sense of (34), upon the formal variable t, the trace of WH-movement. This formal variable is in an obvious sense the same as the formal variable t'; thus, we may regard the pronoun as dependent upon it. But the pronoun cannot have t' for its antecedent, because of the rules governing strong crossover; hence, as per (64), it cannot be dependent upon it either, and the interpretation (69) is appropriately blocked.
Now, the generalization (64) is strictly deducible from (56), a condition that applies to dependence generally, and not just to the special case of dependence that arises when a variable is the antecedent of a pronoun. In particular, (70) violates (56), because the pronoun she is dependent upon t', but not accessible to it. A similar account can be given of the examples in Jacobson (1977), and others.

We might attempt an analysis of (65)-(66) and like contrasts in terms of indexing, foregoing the further elaborations that the substitution of linking for coindexing requires. The problem, in this case, would be to rule out structures such as (71):

(71) [[Which man [[who], t, admires [his, wife],]] does she try to please t']

Notice, however, that no matter how this is done it appears that the fundamental idea of Chomsky (1976) must be radically modified, and perhaps abandoned; for the assimilation of variables to R-expressions for purposes of the binding theory does nothing to rule out (71).

But just here, it appears, the analysis of crossover in terms of conditions on accessibility makes a virtue of necessity; for, to the degree that the crossover phenomena are governed by these conditions, there is no reason to regard variables as R-expressions, a move that may have consequences for the general theory of empty categories. The analogy between the simplest cases of "strong" crossover and the disjoint reference condition, as exhibited in "He saw John" and the like may prove to have been misleading.

A second advantage of our account is that it applies immediately to the central cases of what Haïk (1981) calls "indirect binding." Haïk observed that crossover conditions apply even in the case of "donkey-sentence" anaphora; i.e., in contrasts like (72)-(73):

(72) Which man who owns a donkey hates it?

(73) ??? Which man who owns a donkey does it hate?

We will assume, as in Kamp (1980), Heim (1980) and (1982), that anaphoric relation shown in (72) is not a case of binding of the pronoun by the indefinite description a donkey, or at least not so to be construed at the level LF. In this case, we shall take the LF-representation of (72) to be (74):

(74) [[Which man [[who] [a donkey] t owns t']] t'' beats it.
in which the indefinite description does not c-command the pronoun. Similarly, we assume (75) for (73):

(75) [[Which man [[who] [a donkey] t owns t']] does it hate t'']

In each of (74)-(75), we are interested in the possibility of making the trace t' of QR applied to the phrase a donkey the antecedent of the pronoun; and in each case we are required by (56) to consider its accessibility to t', as determined by the V-chain (t', t''). The results are just as they are in the standard crossover cases, yielding the data shown in (72)-(73).

I hope on another occasion to present further details and ramifications of the system of binding theory here proposed. To return to our earlier theme, we may observe that it is a system that, besides possessing some intrinsic advantages, is consistent with our general hypothesis that LF-representations instantiate quantificational schemata; and we have seen that specific departures from that hypothesis, involving reconstruction and layered traces, are neither necessary nor sufficient to account for the crossover phenomena.

In the last part of this discussion, I will take up the second issue mentioned in the beginning, namely the explanation of the absence of raising nouns, as illustrated in (76):

(76) *John's certainty/likelihood/necessity [t to leave]

The phenomenon is surely related to the absence of movement passives from embedded clauses in nominals, as in (77):

(77) *John's belief/knowledge/expectation [t to be a nice fellow]

Our explanation will make essential use of the system sketched above for the interpretation of nominals whose specifiers are themselves NP's.

Williams (1980) and Kayne (1981) have both offered accounts of (76)-(77) and the like, Williams in terms of conditions on indexing, and Kayne in terms of the Empty Category Principle (ECP) of Chomsky (1981). Rappaport (1981) poses some problems for Kayne, and has independently developed an analysis similar to that to be proposed here, to be presented in (1982). I find the analyses of Kayne and Williams both unsatisfactory as general explanations; but rather than first give criticisms, I will proceed directly to sketch my own suggestions.
We may begin by reviewing some properties of simple and derived nominals, excluding both gerunds and nominals of the "mixed" type (as in John's proving of the theorem). Inside $\bar{N}$ we have structures of the type of (5a) above, namely (78):

$$\left[\bar{N} \, N \, A_1 \, A_2 \, \ldots \, A_p \right]$$

where the $A_i$ are arguments to $N$. Now, in contradistinction to $V$, all arguments to $N$ are optional, so that, for example, we can have (79) but not (80):

$$\text{(79) John's/the } \left[\bar{N} \text{ purchase}\right]$$

$$\text{(80) } \ast \text{John purchased}$$

Subject of $\bar{N}$ is optional too: only a few isolated nouns, such as sake and behalf, require an NP specifier (furthermore, these nouns are not modifiable). But the optionality of Subject is a special case of the more general fact.

Although the arguments to a nominal are all optional, those that occur inside $\bar{N}$ must be compatible with the thematic selection imposed by the head noun. Subject, however, need not be thematically selected; this is most obvious in the case of simple nominals like (81):

$$\text{(81) John's cat}$$

where there is nothing to select.

To the above familiar points, Rappaport (1980) makes an addition: thematic roles inside $\bar{N}$ are in English assigned through prepositions, where, as she shows, the "proper" preposition for the role must be chosen.

Suppose now in view of the above that an $\bar{N}$ of the form shown in (78) is always interpreted as (82):

$$\text{(82) } \left[\bar{N}(x) & R_1(A_1, x) & R_2(A_2, x) & \ldots & R_p(A_p, x)\right]$$

where the $R_i$ spell out the thematic roles of the respective arguments $A_i$. This hypothesis is not only compatible with, but might even be said to explain, the fact that arguments to $N$ are optional - it is not implausible to suppose that conjuncts can simply be omitted.

The thematic roles of the arguments $A_i$ to a head noun $N$ are not, on this view assigned through $N$ itself; rather, they are assigned through prepositions, whose significance may vary with the choice of $N$, and also contextually; our treatment in
this respect concurs with Rappaport (1982). To put the point another way, it is suggested that the representation of, say, (83) is not (84), but (85):

\[
(83) \ [N \ [N \ \text{gift}] \ [_{pp} \ \text{to Mary}]]
\]

\[
(84) \ \text{gift(-to) (Mary, x)}
\]

\[
(85) \ \text{gift (x) & to (Mary, x)}
\]

where the preposition to has the semantic status of a demonstrative - its content can vary with its context, both linguistic and extra-linguistic.

Expressions of the category \( \bar{N} \) have the status of open sentences, of the form of (82). NP then may be completed (i) by a binder (a SPEC not itself an NP), such as the, a, which, every, etc., giving rise to the familiar quantifier-variable structure as in, say, (86):

\[
(86) \ [\text{every x: gift (x) & to (Mary, x)}]
\]

or (ii) by a subject NP, assigned some role or other, with a default binder that, in English, presumably has the features of the definite article. Thus, (81) might come out as (87):

\[
(87) \ [\text{the x: cat (x) & R (John, x)}]
\]

where R expresses some relation, determined in context.

Consider now nominals like (88) and (89):

\[
(88) \ \text{John's gift}
\]

\[
(89) \ \text{John's gift to Mary}
\]

In (88), which by hypothesis comes out as (90):

\[
(90) \ [\text{the x: gift (x) & R (John, x)}]
\]

R can be construed as filling the thematic role of agent, or of goal. In (89), which appears as (91):

\[
(91) \ [\text{the x: gift (x) & to (Mary, x) & R (John, x)}]
\]

goal is assigned through the preposition to, and R is in consequence deprived of that role. The consequence reflects principles of grammar, not inevitable semantic constraints, as we can see by reflecting that if R also expressed goal, then (89) could be well-formed with the interpretation, "the gift that John and Mary received."
The facts about (89), and many others of a similar nature noted in Rappaport (1980), show that grammatical constraints on selection of thematic role are critical even for subject of NP. Anderson (1979), (1982) makes further points of this sort, although she does not follow the present idea that nouns never assign thematic roles on their own. Taken together, these facts seem to limit the validity of Williams' (1980) view that the subject of NP can bear any contextually reasonable relation to N.

We have been assuming that thematic roles inside NP are assigned by fixing the interpretation of various relation-symbols R, which may be overt or tacit, in such a way as to make the interpretation of the whole N compatible with the thematic selection governed by the head. Where r interprets R, let us say that r is an R-∅ if it is a relation taken among those licensed by the head of N, and an R-non-∅ otherwise. Where r is an R-non-∅, we may still regard its assignment as subject to the conditions of the ∅-criterion of Chomsky (1981), which requires that every argument be assigned a unique ∅-role, and in one and only one way. Indeed, the ∅-criterion, so understood, already suffices to rule out the interpretation of R as goal in (91), because that role must already be filled by the argument Mary.

However, the ∅-criterion now suffices to block raising nouns as well. Thus in (92):

(92) *[John's [likelihood [t to leave]]]

John is already assigned a ∅-role through the VP to leave; hence it cannot also be assigned a role through interpretation of a relation-symbol R, as it is required to do if our discussion of the assignment of thematic roles in nominals is correct. Evidently, the same considerations block passives like (93):

(93) *[John's [belief [t to be a nice fellow]]]

Two further points, I believe, speak in favor of our diagnosis of the failure of raising and passive in nominals. The first is that, unlike Kayne (1981), we allow WH-movement freely in cases like (94):

(94) Who did you hear [NP talk about t].

The second is that our analysis squares with the intuitive fact that raising in nominals appears unrelated to the possibility or impossibility of raising in the corresponding adjectives or verbs. Thus, to my ear, (95) is crashingly ungrammatical, although (96)-(97) have the same status, and (98) is fully well-formed:
(95) John is necessary [t to leave]
(96) John's necessity [t to leave]
(97) John's likelihood [t to leave]
(98) John is likely [t to leave]

In this discussion I have considered a possible general view on the structure of linguistic representations at the level LF of logical form, and rules of grammar involving LF that, if I am on the right track, account for some complex and surprising phenomena. Further work, as always, remains to be done. It is my hope that, as research progresses, we shall obtain a more articulated notion of the structure of LF than we have at present.

*An earlier version of this paper was presented at the Cornell Conference on Government and Binding Theory, Cornell University, 9 July, 1982. Prior discussions with Noam Chomsky, Isabelle Haïk, Howard Lasnik, Malka Rappaport, and Edwin Williams were especially helpful to me in formulating the views there presented. I am further indebted to the Conference participants for discussion and criticism.

The reader should know that what follows is not self-contained; in particular, some familiarity with technical terminology and literature in Government and Binding Theory (Chomsky (1981)) is presupposed. In this respect, the paper reflects its origin in a specialized conference. Besides the references cited in the body of the paper, Radford (1981) provides an introduction to some of the technical notions employed below.

FOOTNOTES

1 I use the notion of a schema in the sense of Quine (1970); and by saying that schemata are generalized, I mean to allow for restricted quantifications of various sorts, as well as generalized quantifiers.

2 We could, for example, assign freely to each NP a set of indices, requiring that this set be a singleton if NP is singular. The notions of bondage and freedom of NP's then become more complicated, because the index sets may overlap without being identical. Furthermore, several questions of interpretation arise. I hope to discuss these matters in work in preparation.

3 Lasnik (1982).

4 Specifically, I would conjecture that (25) is weakened in Thai to something like (i):
(i) If X c-commands Y, and Y is an R-expression, then if Y is linked to X, Y agrees with X.

where repetition of a proper name, for instance, counts as "agreement." Some examples in Evans (1980) may be explicable along similar lines.

It has been pointed out to me that the Thai examples might be treated in a theory with coindexing, if that theory is supplemented by a statement such as (ii):

(ii) An R-expression is free, except for arguments with which it agrees.

The contrast between the Thai equivalents of (27) and (28), however, shows that the pertinent notion of agreement is asymmetric. This asymmetry then goes unexplained. In an account with linking, however, the basis of the asymmetry is clear: the epithet the nut can pick up its interpretation from the proper name John, but not vice versa; and since an element cannot have its antecedent within its own c-command domain, (28) is impossible, although (27) is permitted (in Thai). In a coindexing theory, of course, one could define the notion of "antecedent" such that the antecedent element of a coindexed pair is the element of that pair that c-commands the other; but this amounts to stipulating (26), and thus agrees with the conclusion reached here.

5Indeed, (35) is analytic, given that the relation of antecedence obtains between two items only when the interpretation of one is determined through the interpretation of the other.

6Notice that (59) will in fact violate the much better supported condition (26), which forbids a pronoun from c-commanding its own antecedent. But (56) represents a strengthening of (26), as well will see directly.

7I will not in this space attempt to do justice to Haãk's full discussion, or analyze the empirical differences between her approach and that advocated here.

REFERENCES


Fourier, Z. de (1980) "Remarks on the "extended" linear model," ms., MIT.


Reinhart, T. (1980) "Coreference and Bound Anaphora," ms., Tel Aviv University, Tel Aviv, Israel.


LF, ECP, AND NON-VACUOUS QUANTIFICATION

C. T. James Huang

University of Hawaii

For the purposes of this paper I will assume that at the linguistic level of Logical Form (LF), all quantificational expressions appear in operator positions and all non-quantificational expressions appear in argument positions. I will examine the syntax of LF on the basis of a range of data from English, Chinese, and Japanese. Within this model, quantifier-variable representations may be derived by the rule Move α, which may be applied in the component of Syntax or in the component of LF. Sentences with ordinary quantificational NPs are subject to QR (May 1977), an instance of Move α. Wh-movement is taken to be a rule that applies universally, if not in Syntax then in LF. Thus, Chinese and Japanese differ from English not in having or not having a rule of wh-movement, but in where the rule may apply. Furthermore, we will treat all focused constituents as quantificational in some real sense, even if they do not involve movement in Syntax, on a par with ordinary quantificational NPs and wh-phrases.1

An important principle of grammar that will figure prominently in our discussion is the Empty Category Principle (ECP) of Chomsky (1981):

(1) ECP (Chomsky 1981):
    A trace must be properly governed.
A category $\alpha$ is said to be properly governed if and only if (a) it is lexically governed, i.e. governed by a $V$, $N$, $A$, or $P$; or (b) it is "locally controlled", i.e. governed by its own antecedent. The notion government is defined, following Aoun and Sportiche (1981), as follows:

(2) Government (Aoun and Sportiche 1981):
$\alpha$ governs $\beta$ iff every maximal category containing $\alpha$ also contains $\beta$, and conversely.

As is pretty well known, the ECP is intended to account for a fairly wide range of subject/object asymmetries observed in language. For example, the asymmetries illustrated by the following pairs of sentences are assumed to fall under the ECP:

(3) a. This is the book that I said that the man bought t.
   b. *This is the man that I said that t bought the book.

(4) a. ?This is the book that I wonder how much you like t.
   b. *This is the man that I wonder how much t likes the book.

(5) a. Je n'ai exige que la police arrete personne.
    'I didn't require that the police arrest anybody.'
   b. *Je n'ai exige que personne soit arrete.
    'I didn't require that anybody be arrested.'

(6) a. Who remembers why we bought what?
   b. *Who remembers why who bought the book?

(3)-(4) illustrate a subject/object asymmetry under overt wh-movement, commonly known as the "COMP-trace effect". (5) illustrates the same asymmetry under the LF rule of QR, sometimes referred as Kayne's "ne personne effect" (Kayne 1979, 1981). And (6) illustrates the same asymmetry under abstract wh-movement in LF, the "superiority effect" in the sense of Chomsky (1973). In each of the (b) (but not (a)) sentences, a subject trace created either by Move $\alpha$ in Syntax, or abstract movement in LF is neither lexically governed, nor locally controlled, thus not properly governed, in violation of the ECP. In Chomsky (1981) it is further proposed that the asymmetry between legitimate raising with $\mathcal{S}$-deletion and illegitimate raising without $\mathcal{S}$-deletion, as shown below, may also be attributed to the ECP:

(7) a. John is likely [$_s^S$ t to win].
   b. *John is probable [$_s^S$ t to win]].

Let us refer to asymmetries of the sort just illustrated as the "standard ECP effects". Of interest to our discussion here is that the full range of familiar ECP effects are apparently lacking in
Chinese and Japanese. A simple kind of evidence for this observation comes from the fact that, in these languages, one can question either a subject or an object embedded in an indirect question. For example, the sentence (8) in Chinese, with shei 'who' and sheme 'what' appearing syntactically unmoved in the embedded clause under xiang-zhidao 'wonder', can be interpreted as a direct question on either 'who' or 'what'.

(8) ni xiang-zhidao [shei mai-le sheme]?
   you wonder who bought what
   a. 'Who is the person such that you wonder what he bought?'
   b. 'What is the thing such that you wonder who bought it?'

This is because the sentence can be uttered as a question to which either (9a) or (9b) may be an answer:

(9) a. wo xiang-zhidao [Lisi mai-le sheme].
   I wonder Lisi bought what
   'I wonder what Lisi bought.'

b. wo xiang-zhidao [shei mai-le shu].
   I wonder who bought book
   'I wonder who bought the book.'

The answer (9a) provides a value for the embedded 'who' in (8), and (9b) provides a value for the embedded 'what'. The LF representations of the two readings of (8) are given in (10a) and (10b):

(10) a. [shei_i [ni xiang-zhidao [sheme_j [t_i mai-le t_j]]]]
    who you wonder what bought

b. [sheme_j [ni xiang-zhidao [shei_i [t_i mai-le t_j]]]]
    what you wonder who bought

These representations are parallel to the structures of (11a) and (11b) in English:

(11) a. *[Who_i [do you wonder [what_j [t_i bought t_j]]]]?

b. *[What_j [do you wonder [who_i [t_i bought t_j]]]]?

But although there is a difference in degree of acceptability between (11a) and (11b) in English, as between the (a) and (b) members of (3)-(6), both (10a) and (10b) must be considered well-formed in Chinese, given that both (9a) and (9b) are acceptable answers to (8). A similar observation can be made for Japanese, as indicated in (12), which has two readings on a par with (8):
(12) [dare-ga nani-o katta ka] kangaeteru ndai?
    who-NOM what-ACC bought Q thinking Q
    a. 'Who is the person such that you wonder what he bought?'
    b. 'What is the thing such that you wonder who bought it?'

    I will not try to enumerate all available evidence for the
observation that Chinese and Japanese lack a full range of stand-
ard ECP effects. Let us now consider the following question. How
is the apparently systematic difference between Chinese and Japanese
on the one hand, and English on the other, with respect to the ECP
to be accounted for? One easy answer could be that the ECP should
be formulated as a parameter which may be a principle of grammar
for some languages but may not for others, and not formulated as
a principle of UG. However, this conception appears to me to be
extremely implausible on grounds of learnability, especially because
we are dealing with empty categories, whose properties can hardly be
determined inductively on the basis of overt phenomena or experience,
and therefore presumably reflect inner resources of the mind. A
more plausible conception, I think, is to take the ECP as a prin-
ciple of UG, and derive the typological differences concerned from
some other learnable parameter or parameters.

    In this paper I am not concerned with what the parameter or
parameters are and how they might derive the observed typological
differences. I will only mention that a plausible idea to derive
the lack of ECP effects in Chinese and Japanese is to assume that
subjects in these languages are always properly governed in their
own clauses. One way to execute this idea is to assume that in
these languages S is a projection of V, and not of INFL, so that
subjects, like objects are directly governed by their lexical verbs.
Another way to execute this idea is to assume that the INFL is lexica-
lar, say in Chinese, though not in English. Both alternatives are
plausible and each has non-trivial consequences, though I will not
go into them here. My present purpose is to show that the second
conception of the ECP, namely that of taking it as a principle of
UG, is not only methodologically, but also empirically supported.
I will show that although Chinese and Japanese do not exhibit stan-
dard subject/object asymmetries, they nevertheless exhibit certain
systematic argument vs. adjunct asymmetries which may most profita-
bly be attributed to the ECP. The correctness of this analysis
thus supports the idea that the ECP represents a facet of the human
language faculty.

    To begin with, consider again (10a) and (10b), the LF repre-
sentations of (8). The fact that they are equally well-formed in
Chinese shows the lack of subject/object superiority, as we have
seen. Note now that the mere fact that each of them is well-formed
shows, furthermore, that abstract wh-movement in LF must be allowed
to violate subjacency, in this case the wh island constraint. In
(10a), 'who' has been moved across a wh island headed by 'what', and in (10b), 'what' has been moved across a wh island headed by 'who'. The example given in (13) further shows the irrelevancy of subjacency, in this case the CNPC:

(13) ni zui xihuan [[wo piping shei] de wenzhang]?
you most like I criticize who DE article
'Who is the person such that you like the articles in which I criticize him?'

The Japanese example corresponding to (13) is (14), which like (12) also shows the irrelevancy of subjacency:

(14) boku-ga dare-o hihansi-ta ronbun-o yon-da no?
I-NOM who-ACC criticized paper-ACC read Q
'Who is the person such that you read the papers where I criticized him?'

There is extensive further evidence showing that subjacency must be prevented from blocking the application of QR across certain islands. For example, in (15) we have a quantificational NP appearing within a relative clause in Chinese. As the translation shows, this Q-NP may be interpreted as having scope over the entire matrix clause, violating the CNPC:

(15) [[meige ren xie] de shu] dou hen youqu.
every man write DE book all very interesting
'For every x, books that x wrote are very interesting.'

What we have observed in Chinese and Japanese questions also applies to English multiple questions, in particular to the interpretation of syntactically unmoved wh phrases in this language. For example, as Baker (1970) has observed, the syntactically unmoved what in (16) may be interpreted as being paired with the matrix who; the movement of what in LF violates the wh island constraint.

(16) Who remembers where we bought what?

The sentence (17), with everybody interpretable as having sentential scope, so that it means everybody is such that pictures of him are on sale, further shows that QR may violate subjacency, in particular the subject condition of Chomsky (1973):

(17) Pictures of everybody are on sale.

It is therefore quite safe to conclude that subjacency is a condition that obtains only in the syntactic component, but not in LF.
In all the examples we have given, the phrases involved are noun phrases like 'who' and 'what'. There are cases where the abstract movement of a wh phrase apparently cannot violate subjacency, however, in particular if what is involved is an adverbial wh phrase like 'why' and 'how'. Consider sentences like (18), (19) and (20):

(18) ni xiang-zhidao [shei weisheme mai-le shu]? you wonder who why bought book 'Who is the person such that you wonder why he bought the book?' NOT: 'What is the reason x such that you wonder who, for x, bought the book?'

(19) ni xiang-zhidao [shei zeme mai-le shu]? you wonder who how bought book 'Who is the person such that you wonder how he bought the book?' NOT: 'What is the means x by which you wonder who, by x, bought the book?'

(20) *ni zuì xihuan [[wo weisheme piping Lisi] de wenzhang]? you most like I why criticize Lisi DE article 'What is the reason x such that you like the articles in which I, for x, criticized Lisi?'

Compare the sentences (18)-(19) with (8). Like (8), (18) and (19) each contain two syntactically unmoved wh phrases in the embedded clause. In (8), we have 'who' and 'what'. In (18) we have 'who' and 'why' and in (19) we have 'who' and 'how'. (8) is ambiguous in being interpretable as a direct question on either 'who' or 'what'. However, (18) and (19) are not ambiguous; they each can be interpreted as a direct question on 'who', but not as a direct question on 'why' or 'how'. This shows that, although 'who' and 'what' may violate the wh island constraint in LF, 'why' and 'how' cannot. Compare also (20) with (13). Like (13), (20) contains a wh phrase within a complex NP. The well-formedness of (13) as a question, as opposed to the ill-formedness of (20), shows that while 'who' may violate the CNPC in LF, 'why' cannot. What we have seen in Chinese can also be easily seen to obtain in Japanese and English. Take the paradigm (21) in English for example:

(21) a. Who remembers where we bought what?
    b. Who remembers where we met who?
    c. *Who remembers what we bought why?
    d. *Who remembers what we bought how?

Also the paradigm in (22):

(22) a. Who bought what?
    b. Which man pleased which woman?
c. *Which bought the books why?  
d. *Who bought the books how?  

The *wh phrase at the end of each of these sentences is syntactically unmoved and is moved in LF. The asymmetry shown in these examples shows that LF movement of who, what, which woman etc., may violate the *wh island constraint, but not the movement of why and how.  

What is the relevant distinction between the two types of *wh phrases that we have seen that gives rise to the asymmetry observed? I would like to suggest that what is relevant is the distinction between argument and adjunct. Who, what, which woman, etc. are arguments of sentences, but why and how are not: they are adjuncts, or one might say they are operators or predicates. This distinction seems to me to be correct, and is further supported by the following observation in Chinese. The interpretation of the so-called A-not-A question, as well as the cleft sentences in Chinese also appears to obey a full range of island conditions. An example of the A-not-A question is given in (23):  

(23) ni xihuan-bu-xihuan ta?  
    you like-not-like him  
    'Do you like him or not?'  

There is good reason to assume that the formation of an A-not-A question involves a phonological reduplication applied on the basis of a question operator in preverbal adverbial position, which I will call the A-not-A operator. According to this view, the S-structure of (23) looks something like (24):  

    you Q-yes/no like him  

Furthermore, there is good reason to assume that (23) undergoes movement in LF, by which the A-not-A operator is raised to operator position, giving rise to the LF representation (25):  

(25) [[A-not-A]_i [ni t_i xihuan ta]]  
    you like him  

An example of a cleft sentence in Chinese is (26):  

(26) Zhangsan shi mingtian yao lai.  
    Zhangsan SHI tomorrow will come  
    'It is tomorrow that Zhangsan will come.'  

The formation of a cleft sentence involves placing a copulative adverb in adverbial position, the element shi in the example (26). I will assume that interpretation of cleft sentences involves the
movement of this emphatic operator to operator position, as in (27):

(27) \[SHI_i [Zhangsan \ t_i \ mingtian \ yao \ lai]]

Zhangsan \ tomorrow \ will \ come

The assumption that movement is involved here enables one to account for the fact, among others, indicated in (28) and (29):

(28) ni xiangxin [Zhangsan xihuan-bu-xihuan ni]?
you believe Zhangsan like-not-like you
'Do you believe that Zhangsan likes you, or do you believe that
Zhangsan does not like you?'

(29) Lisi shuo [Zhangsan shi mingtian yao lai].
Lisi said Zhangsan SHI tomorrow will come
'It is tomorrow that Lisi said that Zhangsan will come.'

In (28) and (29) we have an A-not-A operator and a focus operator, respectively, appearing in an embedded clause. As the translation shows, each of the embedded operators may be interpreted as having matrix scope. (28) may be a direct question on the choice between 'likes' and 'doesn't like'; and (29) may have the embedded focused constituent 'tomorrow' interpreted as indicating the speaker's emphasis. Now, consider (30):

(30) ni xiang-zhidao [shei xihuan-bu-xihuan ni]?
you wonder who like-not-like you
'Who is the person such that you wonder whether he likes you or not?'
NOT: 'Do you wonder who likes you or do you wonder who doesn't?'

This sentence is not ambiguous. It may be interpreted as a direct question on 'who' but not a direct question on the choice between 'likes' and 'doesn't like'. This shows that although 'who' may cross a wh island headed by the A-not-A operator, the A-not-A operator may not cross a wh island headed by 'who'. Observe also (31), where we have a 'who' and a focus operator shi embedded in the lower clause:

(31) *ni xiang-zhidao [shei shi mingtian yao lai]?
you wonder who SHI tomorrow will come

This sentence is ill-formed. In particular, the sentence cannot have the interpretation according to which the focus operator shi has matrix scope. This shows, among other things, that the focus operator cannot cross a wh island headed by 'who'.

The two operators we have just seen, the A-not-A operator and and the copulative adverb shi are clearly not arguments of sentences, but adjuncts. The fact that they cannot violate island constraints
shows that they are on a par with 'why' and 'how', rather than with 'who' and 'what'. This is in accordance with our hypothesis that what is responsible for the asymmetry is the distinction between adjuncts and arguments.

How should we account for this argument/adjunct asymmetry? One simple solution is to stipulate that subadjacency, although it does not apply to arguments in LF, must still obtain for adjuncts in LF. (On the other hand, subadjacency applies to both arguments and adjuncts in Syntax.) However, there are a number of problems associated with this hypothesis, of which I will mention only the most obvious one: that it is a pure stipulation. Although it is a stipulation of UG and presumably causes no problem in learning, and it is true that certain stipulations are plainly unaviodable, it remains desirable to see if such a stipulation can be dispensed with.

I would like to suggest that the argument/adjunct asymmetry follows, free, from the ECP, not from a stipulation about subadjacency. The idea underlying this suggestion is as follows. Recall that the standard ECP effects reflect subject/object asymmetries. But why are there such asymmetries? One plausible functional explanation is that there is a closer dependency relation between a verb and its object than between the verb and its subject. In terms of government, the subject/object asymmetry is reflected by the fact that objects are governed by the verb, but not subjects. This proximity between object and its lexical governor in terms of government also corresponds to the dependency of subcategorization. Since there is an intrinsic tight dependency between object and verb, movement of an object may often go long distance. On the other hand, since there is a looser inherent relation between a verb and its subject, the subject may not be moved too far away from the verb, in order for the relatively loose relationship to obtain. This is the reason why movement of a subject has to obey a more strict locality requirement. In terms of the ECP, this is reflected in the requirement that traces of subjects in English, not being lexically governed, must be governed by their own antecedents, namely locally controlled, though traces of objects, being lexically governed, need not be locally controlled. Now, concerning adjuncts like 'why', 'how', A-not-A, and the focus operator, the dependency between an adjunct and a verb is obviously quite loose. Therefore, if our functional explanation for observed subject/object asymmetries is correct, then we will expect that movement of an adjunct has to obey a more strict locality requirement than that of an object. It is quite commonplace to assume that adjuncts lie outside the maximal projection of a VP. Therefore, adjuncts, unlike objects, are not lexically governed, according to (2) above. Not being lexically governed, the traces of adjuncts must therefore be locally controlled, like the traces of subjects in English. This is, I claim, exactly what accounts for the asymmetry between adjuncts like 'why', 'how', etc., and arguments like 'who',
'what' etc. In other words, I argue that the ill-formedness of the (a) sentences in the following paradigms is completely on a par with the ill-dormedness of the (b) and (c) sentences:

(32) a. *Who remembers what who bought?
    b. *Who remembers what we bought why?
    c. *Who remembers what we bought how?

(33) a. *What did who buy?
    b. *What did you buy why?
    c. *What did you buy how?

    b. *Tell me what you bought why.
    c. *Tell me what you bought how.

Let us see how the (b) and (c) sentences can be excluded on a par with the (a) sentences. I will now concentrate on English examples only, but it will not be difficult to show that the strict locality requirement on movement of adjuncts in Chinese and Japanese can also be accounted for in the same fashion. As a way of execution, let us assume, along with Aoun, Hornstein, and Sportiche (1981), that there is a rule that indexes a COMP at S-structure on the basis of the index of the wh phrase it contains at that level. Thus, after COMP indexing applies at SS, the structure of (33a) is (35), with the COMP identified as \(_i\), the index of \(\text{what} \) which it contains.

(35) \[[\text{comp}_{i} \text{what}_{i}][\text{did who buy t}_{i}]\]?

After the syntactically unmoved \(\text{who} \) is moved to COMP in LF, we have:

(36) \[[\text{comp}_{i} \text{who}_{j} \text{what}_{i}][\text{did t}_{j} \text{buy t}_{i}]\]?

In this structure, the trace of \(\text{what} \), \(t_{i} \), is lexically governed by the verb, in accordance with the ECP. The trace of \(\text{who} \), \(t_{j} \), on the other hand, is not lexically governed by the verb; therefore, ECP requires it to be governed by its own antecedent, \(\text{who} \). However, \(\text{who}_{j} \) is properly contained in the branching COMP, which is identified with the index of \(\text{what}_{i} \), not \(j \), and COMP is itself a maximal node, blocking government of \(t_{i} \) by \(\text{who}_{j} \). The subject trace is therefore neither lexically governed nor locally controlled, hence the ill-formedness of (33a). In a similar fashion, (33b) and (33c) can be ruled out by the ECP. Take (33b) for example. At SS, the structure of (33b) is (37), after COMP-indexing applies:

(37) \[[\text{comp}_{i} \text{what}_{i}][\text{did you buy t}_{i} \text{why}]\]?
After *why* is moved into COMP in LF, we have (38):

(38) \[ [\text{comp}_i \, \text{why}_j \, \text{what}_i][\text{did you buy}_t \, \text{t}_j] \]

The trace of *what* is lexically governed and therefore properly governed, as before. The trace of *why*, *t*_j, however, lies outside of the maximal projection of the VP, and is therefore not lexically governed by the verb. As such, it must be locally controlled, governed by *why*. But, again, because *why* occurs within the maximal projection of COMP which is not identified with the index of *why*, *why* does not govern its trace. The trace of *why*, *t*_j, is thus not properly governed, in violation of the ECP. It should be relatively easy to see how the same assumption will enable one to rule out (33c), as well as all the sentences in (32) and (34).

Note that the assumption we make not only rules out the sentences in (32)-(34). It also accounts for contrasts of the sort indicated in (39)-(41) below:

(39) a. Who bought what?
    b. *What did who buy?

(40) a. Tell me why you bought what.
    b. *Tell me what you bought why.

(41) a. Tell me how you bought what.
    b. *Tell me what you bought how.

We have already indicated how the (b) sentences in (39)-(41) can be ruled out. Let us see how the (a) sentences can be ruled in, thus accounting for the contrasts. Consider first (39a). At SS, after COMP indexing, we have (42); at LF, after *what* is moved, we have (43):

(42) \[ [\text{comp}_i \, \text{who}_i[[\text{t}_i \, \text{bought}_t \, \text{what}]]] \]

(43) \[ [\text{comp}_i \, \text{what}_j \, \text{who}_i[[\text{t}_i \, \text{bought}_t \, \text{t}]]] \]

Note crucially that the COMP is identified with the index of *who*, not with *what*. The trace of *what*, *t*_j, is not directly governed by its own antecedent, but since it is already lexically governed in object position, it is already properly governed. The trace of *who* is not lexically governed, so must be locally controlled. Since COMP is identified with *who*, the antecedent of *t*_j, it is itself the antecedent too. Since the COMP (=*who*) does govern *t*_i, *t*_i is locally controlled, therefore also properly governed; hence the well-formedness of (39a). In a similar manner, (40a) and (41a) can be ruled in. In (40a), since *why* is moved in Syntax, its trace is locally
controlled in the same way that the trace of who in (39a) is locally controlled. Likewise for the trace of how in (41a).

What we have seen up to now is that the subject/object asymmetry formerly known as the superiority phenomenon is but a special case of a more general complement/non-complement asymmetry, namely the asymmetry between complements like objects on the one hand (which are lexically governed), and non-complements like subjects and adjuncts on the other (which are not lexically governed). Now I will show that the subject/object asymmetry formerly known as the COMP-trace phenomenon is also but a special case of the more general complement/non-complement asymmetry. Consider the paradigms below:

(44) a. ?This is the book which I wonder where you bought t₁.
b. ?This is the book which I wonder when you bought t₁.
c. ?This is the book which I wonder why you bought t₁.
d. ?This is the book which I wonder how you bought t₁.

(45) a. *This is the man who I wonder where t₁ bought the book.
b. *This is the man who I wonder when t₁ bought the book.
c. *This is the man who I wonder why t₁ bought the book.
d. *This is the man who I wonder how t₁ bought the book.

(46) a. *This is the place where I wonder [who bought the book t₁].
b. *This was the day when I wonder [who bought the book t₁].
c. *This was the reason why I wonder [who bought the book t₁].
d. *This was the means by which I wonder [who bought the book t₁].

All of the sentences in (44)–(46) violate subjacency. But, as is well known, the sentences in (44) are considerably better than the ones in (45). This illustrates the COMP-trace effect. The sentences in (44) violate only subjacency, but the sentences in (45) violate both subjacency and the ECP. Note that the sentences in (46) are on a par with those in (45), not with (44), in their degree of acceptability. Like the sentences in (45), the sentences in (46) are completely impossible with the construal indicated, i.e. with the traces of 'where', 'when', 'why', 'by which' construed as originating from the lower clauses. Note that the fact that (46) parallels (45) rather than (44) readily follows from our conception of the ECP. Since the trace of 'where', 'when', 'why', etc., is not lexically governed, it must be locally controlled. But in each of (46), local control is impossible, because the COMP immediately above each of the relevant traces is lexically filled with 'who', preventing proper government of the adjunct trace.³

We have seen that the ECP accounts for a full range of observed complement/non-complement asymmetries. These asymmetries are those that exist between various constituents within sentences. There is
also an important complement-adjunct asymmetry within noun phrases that has not been observed before, but is apparently quite systematic. Consider (47) and (48):

(47) Of which city did you witness [the destruction t]?

(48) *On which table did you like [the books t]?

In both these sentences, a PP has been extracted out of an NP. The PP in (47) is a complement of the noun destruction. The PP in (48) is not a complement of the noun books, but rather its modifier, i.e. an adjunct. The asymmetry between (47) and (48) is thus the same type of asymmetry that we have been concerned with all along. This asymmetry, furthermore, also readily follows from our conception of the ECP. It is standard practice to regard noun phrase complements as dominated as N, and noun phrase modifiers as adjoined to NPs. Thus, we have the D-structures below for the bracketed NPs in (47) and (48):

(49) [np the [- destruction of which city]]

(50) [np on which table]

The PP of which city is governed by the noun destruction, a case of lexical government. The PP on which table, on the other hand, is not lexically governed by the head books, since a maximal NP node intervenes between the head and the PP. Therefore, when the latter PP is moved, its trace must be locally controlled. In (48), however, local control is impossible, since the antecedent and the trace are separated by a maximal node, the outer NP node containing the trace. The ECP thus correctly rules out sentences like (48).4

In conclusion, we have shown in this paper that well known subject/object asymmetries should be seen as constituting a special case of a more general complement/non-complement asymmetry. This conception of the ECP eliminates an otherwise necessary stipulation on subjacency involving movement in LF. Our analysis, if correct, as a number of interesting implications. Some of these will be pointed out in the remainder of this paper.

First, although Chinese and Japanese lack a full range of "standard ECP effects", they do exhibit argument/non-argument asymmetries. If the latter are to be accounted for by the ECP, as I have argued here, then the ECP also applies in Chinese and Japanese. This provides support for the idea that the ECP should be formulated as a principle of UG rather than a parameter itself, an idea we have indicated to be also plausible on grounds of learnability.
Secondly, our analysis provides an interesting empirical argument for the reformulation of the superiority condition of Chomsky (1973) as a special case of the ECP. According to the definition of superiority given by Chomsky, \( \alpha \) is superior to \( \beta \) if and only if \( \alpha \) asymmetrically c-commands \( \beta \). The superiority condition says that movement must affect the superior of two terms first before it affects the inferior one. Thus, since subjects asymmetrically c-command objects, they must be moved before the latter. This accounts for the contrast between (39a) and (39b). The same superiority condition can be extended to account for the contrast between (40a) and (40b), as well as that between (41a) and (41b). This is because adjuncts like why and how, just as subjects, also asymmetrically c-command objects, and are superior to objects. On the other hand, if neither of two terms asymmetrically c-commands the other, as in the examples of (51) and (52), then no difference in grammaticality need result:

(51) a. To whom did you give which book?
    b. Which book did you give to whom?

(52) a. Tell me to whom I should give which book.
    b. Tell me which book I should give to whom.

All of (39)-(41) and (51)-(52) can be accounted for by either the ECP or the superiority condition. The crucial evidence in favor of the ECP, however, comes from (53) and (54):

(53) a. *Why did who buy the book?
    b. *Who bought the book why?

(54) a. *Tell me why who bought the book.

There are two logical possibilities concerning subjects and adjuncts with respect to superiority. Either that neither is superior to the other, or that one of them is superior to the other. If the former, then the superiority condition does not apply, and we should expect both the (a) and (b) sentences to be both good, which is contrary to fact. If the latter, then the condition applies, and we should expect the (a) sentences to be good and the (b) sentences bad, or conversely. This is, again, contrary to fact, since both (a) and (b) are equally bad. On the other hand, the ECP correctly predicts that all of these sentences are ill-formed, as one can see upon a moment's thought. Since both who and why are not lexically governed, both their traces must be locally controlled. But only one of the traces at most can be locally controlled, since there is only one COMP per sentence, and each COMP can be given one unique index only. Both (a) and (b) are thus excluded by the ECP, regardless of whether why or who gets moved first. 5
A similar consequence of our analysis is that it provides an interesting piece of evidence for the ECP as an empirically more adequate principle underlying the COMP-trace phenomenon, than the earlier formulation proposed in Chomsky and Lasnik (1977). In Chomsky and Lasnik's work, the COMP-trace configuration is ruled out by a filter which states that a trace immediately following a COMP is ill-formed. However, this filter applies only to subject traces, which directly follow their COMPs, but not adjunct traces, which do not immediately follow the COMP. Chomsky (1981) has indicated a conceptual advantage in eliminating the filter in favor of the ECP. Now we also have an empirical argument for the elimination of the filter.

Note also that, if our analysis is correct, we have argued for a more generalized version of the ECP than is proposed in previous work. For example, Jaeggli (1980) formulates the ECP as a principle specifically applying to NP-traces, thus exempting traces of PPs and adverbial phrases from the effect of the principle. Our analysis argues for an empirically more adequate and conceptually simpler (and more general) version of the principle.

Finally, a corollary of the above is that movement of a category, be it an argument or an adjunct, must leave a trace. Note that this is beyond the requirement of the Projection Principle. If our analysis is correct, it might be taken as an argument for a strong version of the principle of non-vacuous quantification (cf. May 1977, Chomsky 1982), so that everything appearing in operator position, regardless of whether it originates as an argument or an adjunct, must bind a trace at LF. As this does not seem to be a requirement of any logical language, this is indication for the assumption that LF exists as a linguistic level of representation, distinct from the level of real semantics.

FOOTNOTES

1This follows the proposal made in Chomsky (1976). See Huang (1982) for some speculation on why focused constituents are quantificational in nature.

2There is also a reading according to which the sentence (8) is taken as a declarative containing an indirect multiple question, i.e. "you wonder who bought what". We exclude this reading from our discussion.

3Note that there is a difference between overt movement and abstract movement with respect to wh-phrases like 'where' and 'when'. Compare (46) in the text with the following:
(i) a. Tell me what you bought where.
b. Tell me what you bought when.
c. *Tell me what you bought why.
d. *Tell me what you bought how.

The difference is that overt movement of where, when, why, how cannot cross a wh island, abstract movement of where and when may, though abstract movement of why, and how still cannot. We have indicated how (ic) and (id) are excluded. Now the problem is to explain why (ia) and (ib) are well-formed, and why (46a-b) should differ in status from (ia-b). As for the well-formedness of (ia)-(ib), I submit that this stems from the fact that where and when may be analyzed as NP arguments of prepositions while why and how cannot be so analyzed: from where, since when, etc., but *for why,*by how, etc. We may assume that where and when are dominated by NP in the structure [PP P NP], where the P may or may not be phonetically realized. Thus the well-formedness of (ia(ib) may be derived from the fact that movement of where and when need affect only the NP node, thus stranding a (phonetically null) preposition. The trace of where and when is thus properly governed (by the P). This mode of movement is ruled out in the case of (ic-d), however, because why and how are directly dominated by PP or AP, not by NP after P. Hence the difference between (ia-b) and (ic-d). As for the difference between (ia-b) and (46a-b), this is due to the existence of a principle blocking movement out of a non-properly governed domain which, unlike the ECP, applies only at SS but not at LF. (For extensive defense of the existence of this condition, see Huang 1982.) Among other things, this condition excludes preposition stranding in English from a non-subcategorized (non-properly governed) PP at SS, though allowing it at LF:

(ii) a. *Which class did you fall asleep during t?
b. Who fell asleep during which class?

(46a-b) are then excluded as follows. If movement of where and when affects the PP node dominating them, then this is excluded by the ECP, on a par with (46c-d) and (ic-d). If movement affects only the NP node, stranding a (phonetically null) P, this is ruled out by the CED, on a par with (iia). Since the CED applies only at SS, but not at LF, the difference between (46a-b) and (ia-b) follows, on a par with the distinction between (iia) and (iib).

The sentence (i) below is well-formed, as is well known:

(i) Who did you say [t[ t will win]]?

This is because the original trace of who is governed by the trace in COMP; a case of local control. The sentence (48) in the text is ill-formed, on the other hand, because NP has no COMP, thus providing no
intermediate trace as a local controller to satisfy the ECP.

5. The following two sentences seem to be as good as each other.

(i) a. Whom did you give which book to?
b. Which book did you give to whom?

The two terms under consideration are whom (not to whom) and which book. Here which book is superior to whom, but we do not have a superiority effect. This may also be taken to indicate that the superiority condition expresses a spurious generalization, a point already indicated by the sentences (53) and (54) in the text. Within an ECP account, these sentences do not pose any problem, since both whom and which book are lexically governed (by to and give respectively), it does not matter which of the two gets moved first. The point indicated by (i) here may not be so clear, however, because there may be some reason to believe that the notion c-command must be somehow sharpened so that embedding an NP within a dative PP would not decrease the NP's prominence as far as the notion of c-command is concerned. This move seems to be independently needed, anyway, as shown below, in which himself is not strictly c-commanded by its antecedent.

(ii) I talked to John about himself.

REFERENCES


LATIN PASSIVE WITHOUT NP MOVEMENT*

John T. Jensen
University of Ottawa

0. Introduction

Three difficulties arise for the rule Move-α and the interpretive account of morphology found in Chomsky (1981) when applied to an analysis of passive in Latin. The first concerns the source of passive morphology. Chomsky assumes that morphology is part of the phonological interpretive component, but this is difficult to maintain in view of the lexical irregularities of Latin inflections. The second problem is that passive morphology is required in cases not directly related to the passive construction, most strikingly in the case of the deponent verbs. The final problem arises in 'raising' constructions, which the government-binding theory handles using the device of 'exceptional Case marking,' where a verb such as believe exceptionally assigns accusative Case to the subject of its infinitival complement after the removal of the intervening S node. Exceptional Case marking cannot be the correct device for Latin, which allows accusative and infinitive constructions in subject and predicate positions, where exceptional Case marking could not apply. The solution to these problems lies in consigning passive morphology (in fact all morphology) to the lexicon and in abandoning relation-changing movement rules. The central core of government-binding theory, including Case theory and θ theory, remains, however, since it provides an explanation of all the facts of Latin passive, with the modifications indicated incorporated into it. The result is a more constrained version of the theory, which can probably be adapted to English and other languages that have been studied from this point of view.

1. The morphological problem

There are four assumptions in the government-binding theory which bear on problems of morphology, listed in (1).
Government-binding assumptions bearing on morphology

a. Lexical insertion occurs at D-structure.
b. Case is assigned (or checked) at S-structure.
c. Verb morphology is determined by the node INFL (=AUX) by a rule ('Rule R' = affix hopping).
d. Move-α can change the grammatical relations of NPs.

I propose to develop an alternative model, which is also suggested by Chomsky (1981, 91), in which there is no Move-α, and in which lexical insertion occurs in S-structure. This model is shown schematically in (2).

Lexicon
(including morphology and lexical phonology)

Phrase-structure rules

Syntactic rules
(feature specifications for agreement and government; insertion of specific morphemes; cliticization; WH-movement)

Surface Structure

Deletions; scrambling; post-lexical phonology

Construal, etc.

Phonetic form Logical form

The revised assumptions of the model illustrated in (2) are given as (3).

a. Lexical insertion occurs at S-structure.
b. Case is assigned at S-structure before lexical insertion in each Case-marked position.
c. Verb morphology is generated in the lexicon and correctly-inflected verbs are inserted in S-structure.
d. There is no Move-α.

The reasons for adopting (2) and (3) are based on the way in which syntactic features are assigned to lexical items. The rule Move-α may change the grammatical relations of an element that it moves. Because of this, a language with any inflectional markers of agreement or government cannot insert fully-specified lexical items until all such movements have been accomplished. This
is clear even in English, which has only minimal inflection. The sentences of (4) illustrate this.

(4) a. John reads many books.
    b. Many books are read by John.

The verb in (4a) is third person singular to agree with its subject, John. In (4b), the passive version of (4a), the verb is third person plural to agree with its subject, many books, moved to subject position from the position of the trace. In Latin, the point is illustrated even more vividly, since not only the verb, but also the noun phrases, have their morphological properties determined by the S-structure configuration of the sentence, as illustrated in (5).

(5) a. Ioannes legit multos libros 'John reads many books'
    (nom) (3 sg) (acc p1)
    b. Multi libri leguntur a Ioanne 'Many books are read
    (nom p1) (3 pl pass) (abl) by John'

In (5a), Ioannes is nominative singular, as befits its subjecthood, legit is third person plural, in agreement with Ioannes, and multos libros is accusative plural, by virtue of being governed by the verb lego, which is not marked lexically for taking any special Case. In (5b), multi libri is the subject, hence nominative, and also plural, leguntur is third person plural in agreement with the subject, while Ioannes is ablative singular under government of the preposition a, which is lexically marked for ablative complements.

The D-structures of (5a) and (5b), in Chomsky's system, are (6a) and (6b) respectively.


I have omitted the desinences from these structures, since they are not determined until after Move-α. The rules that insert (or spell out) the desinences will clearly have to be of a rather complex sort, referring to the structural position of each lexeme, its declensional or conjugational class, and perhaps other properties, as in the case of forms with suppletive inflections. A priori, it seems desirable to substitute lexical generation of inflections for spell-out rules. In the next section we will provide additional evidence for this position, by showing that passive morphology, if assigned after lexical insertion, must make reference to four unrelated constructions.
2. The problem of deponent verbs.

According to Chomsky (1981, 124 et passim), passive is characterized by the cluster of properties (7), somewhat paraphrased.

(7)  
   a. Passive morphology  
   b. [NP, S] receives a θ-role, if at all, only from a coindexed trace  
   c. [NP, VP] does not receive Case within VP, for some choice of NP in VP.

We need (7b) to prevent dual θ-role assignment to the subject of a passive verb: since the subject receives its θ-role from its trace, it cannot receive another role from the verb without violating the θ criterion. Intransitive passives, in languages like German and Latin, also conform to (7b), in the sense that, with no trace to transmit a θ-role to the subject, the subject is necessarily empty. In German, word order requires something before the verb; in (8) it is the dummy es; it can also be a dative object (13c).

(8) Es wurde getanzt 'dancing took place' (Lit.: 'it was danced')

In Latin, overt subjects are not required. As in German, the verb is third person singular, and if a participle appears as part of the verb, it is neuter singular nominative. Some examples of Latin impersonal passives are given in (9).

(9)  
   a. sic itur ad astra 'this is the way to the stars' Aen ix, 641  
   b. postquam ventum est 'after (our) arrival' after come is  
   c. diu atque acriter pugnatum est 'the fighting was fierce long and fiercely fought is and long' BG I, 26

Because the verbs of (8) and (9) are intransitive, we cannot understand the failure of their subjects to receive a θ-role in terms of avoiding dual θ-role assignment, as we did in the case of transitive verbs. Chomsky assumes the uniformity principle to account for this.

(10) **Uniformity principle:** Each morphological process  
     either (i) transmits θ-role uniformly  
     (ii) blocks θ-role uniformly  
     or (iii) assigns a new θ-role uniformly

By virtue of (10), since passive verbs fail to assign a θ-role to their subjects in transitive cases, they likewise fail to do so in intransitive cases.

However, the uniformity principle cannot be part of universal grammar. Latin presents a clear counterexample in the case of
deponent verbs, those verbs described by traditional grammars as having passive forms with active meaning.\footnote{In government-binding terminology, these are verbs with passive morphology that assign a θ-role to their subjects (violating 7b) and, if transitive, assign Case (accusative) to their direct objects (violating 7c).} Significantly, these verbs can never be used passively (except in the gerundive, which is always passive).

(11) a. Cicero Caesarem sequitur 'Cicero follows Caesar'  
    (nom) (acc)  

b. *Caesar sequitur t a Cicerone  
    (nom) (abl)

To accommodate these facts within the government-binding framework, we must assume that the choice of passive morphology is one lexical property of verbs and the ability to assign a θ-role to the subject is another. Deponents are therefore lexically marked for both properties. If we regard deponents as a marked case, we must consider the uniformity principle as a strong tendency, not an absolute constraint. The disappearance of the deponent verbs in the Romance languages attests to their highly marked status.

Rejecting the uniformity principle as an absolute constraint does not oblige us to abandon the explanatory possibilities of Case theory and θ theory, as these principles correctly account for the three ways in which Latin passive diverges from standard, transitive passivization. These are (a) intransitive passives (9), (b) passives of verbs whose complements have lexical Case (12), and (c) the deponent verbs (11).

The passive of intransitive verbs is impersonal, as in (8) and (9), because verbs with passive morphology (except deponents) do not assign a θ-role to their subjects. An intransitive verb has no object; if passive it has no trace to transmit a θ-role to the subject. Therefore, the subject is not a θ position, and must be empty (as in Latin) or a dummy (as in German).

The passive of verbs with non-accusative objects is impersonal. This is really a subcase of intransitive verbs, but may be mentioned separately because their English translation is transitive. Some examples appear in (12).

(12) a. Amici nostri nobis persuaserunt 'our friends persuaded us' friends our us (dat) persuaded

b. Nobis ab amicis persuasum est 'we were persuaded by us (dat) by friends persuaded (neut) is our friends'
    (adapted from Colebourn 1948, 81)

As is usual with passives, the agent phrase, if expressed, appears as a prepositional phrase headed by ab 'by' and the ablative. As in the impersonal verbs of (9), these verbs in the passive are third person singular, with a neuter participle if a participle is
German too uses verbs with non-accusative objects impersonally in the passive (Borgert and Nyhan 1976, 141).

(13) a. Ich helfe ihnen 'I help them'  
I help them (dat)
b. Es wurde ihnen geholfen 'they were helped'  
it was them (dat) helped
c. Ihnen wurde geholfen 'they were helped'  
them (dat) was helped

Following Chomsky, we assume that the Case of the complement of these verbs is part of their lexical entry, since such Case is not predictable. (This notion is formalized in §4.) Such a Case-marked NP cannot be moved to a governed position without incurring Case conflict. With no movement, there is no trace to assign a θ-role to the subject. Since the verb is passive, it assigns no θ-role to the subject either, so the subject position can only be empty (Latin, e.g. 12b, German, e.g. 13c) or a dummy (German, e.g. 13b). Since there is no overt subject to assign person, number, gender, and Case to the verb, it takes the least marked value of the features for these categories, namely third person singular neuter nominative. In terms of a feature analysis, given in (14), this means that it is marked with the minus value of all the relevant features, under the assumption that minus is always the unmarked value of morphological features.

(14) [-I, -II, -pl, -animate, -feminine, -oblique, -theme, -source]

Finally, the impossibility of deponent verbs to passivize is also predicted by government-binding theory. We have assumed that deponent verbs carry two independent lexical marks: (a) that they require passive morphology (with the exceptions listed in footnote 1), and (b) that they assign a θ-role to their subjects, contrary to (7) and (10). In a sentence such as (11b), the subject would receive two θ-role assignments: one from the verb (by b) and one from the trace of the object, as in normal passives. This violates the θ criterion; hence sentences such as (11b) are ungrammatical.

As we have noted, Chomsky's assumptions (1) entail that morphology is interpretive, in the sense that inflected forms are spelled out as part of phonological interpretation. On this view, the assignment of passive morphology must be split into four distinct, construction-specific devices, listed in (15).
(15) Assignment of passive morphology under the assumptions of (1)

a. Assign passive morphology to a verb with a trace following it in S-structure coindexed with the subject (4b, 5b).

b. Assign passive morphology to a verb with an unspecified subject (depending on lexical marking that varies from language to language), e.g. rain is active, seem is active in English but passive in Latin (videtur), etc. (8, 9, 12b, 13b, 13c).

c. Assign passive morphology to a verb followed by an agent by-phrase in S-structure (4b, 5b, 12b).

d. Assign passive morphology to deponent verbs (lexically marked as such).

In order to unify the assignment of passive morphology, I suggest that all morphological processes are performed in the lexicon. This idea receives additional support from the theory of lexical phonology (Kiparsky 1982, Mohanan 1981), which shows that often the same morphological function is performed by elements derived at distinct levels in the lexicon, which could not be done by post-syntactic interpretive rules. To follow through on this proposal, we must abandon the idea that passive is part of a more general movement process (Move-α). Instead, we formulate passive as a syntactic rule that merely specifies an accusative NP following a passive (non-deponent) verb as empty if the verb c-commands the NP and the NP directly follows the verb. Let us use the feature [passive] to refer only to the morphology of the verb, and let us introduce a feature [T] which relates to the verb's role in assigning a θ-role to its subject. A verb that assigns a θ-role to its subject is [+T], and a verb which does not is [-T]. We can write passive as (16).

(16) [V
    [+pass]    N³
    [-T]      [-ob1]
    +theme     [-source]
    +

e
All possible combinations of the features [T] and [passive] are actually realized. The feature specification [-pass, +T] belongs to normal active transitive (e.g. amo 'love') and intransitive (e.g. dormio 'sleep') verbs. The combination [+pass, -T] represents normal passive (e.g. amor 'I am loved') and raising-to-subject (e.g. videor 'I seem') verbs. The combination [+pass, +T] belongs to deponent verbs, e.g. sequor 'follow'. Finally, the combination [-pass, -T] belongs to idiomatic impersonal verbs like taedet.
'tire' (cf. 24). Rule (16) applies to normal passives and to raising-to-subject verbs, but not to deponents, and its application has the effect of preventing lexical insertion in the NP position specified as empty. This empty node has the same properties as trace in theories with movement: it is governed, its antecedent cannot be in a θ position, the relation between e and its antecedent satisfies the subadjacency condition, and e must be bound. This version of passive retains the explanations available in government-binding theory for the facts of Latin passive, reviewed already in terms of the movement analysis. Intransitive passives are generated in the lexicon, and are [+pass, -T], but they have no following NP to be marked e by rule (16). Since they are [-T] and have no e in V1 to assign a θ-role to the subject, they are necessarily impersonal. A similar argument holds for dative-object verbs, since the object of such verbs fails to meet the Case requirement for the NP following the verb in (16). They are [-T] and have no e to assign a θ-role to the subject, and hence they are impersonal. Finally, deponents are lexically represented as [+pass, +T], hence the NP following such a verb cannot be marked e.

3. The interaction of lexical insertion with syntax

In this model, the interaction of lexical insertion with phrase structure and syntactic rules is slightly more complex than in Chomsky's model, so perhaps an example will clarify the interaction of these three components as given in the diagram (2). In the generation of a simple sentence, say (5a), the phrase-structure rules first determine the basic form of the sentence, then the lexicon inserts the subject Ioannes (nominative singular in form, since it is in subject position), then a syntactic rule marks the VP node third person singular (with appropriate features; cf. footnote 4) in agreement with the subject. Lexical insertion is constrained by the non-distinctness criterion: a lexical item can be inserted only where it does not contradict any features determined by syntactic rules (Chomsky 1965, 164). Therefore, lexical insertion must insert a verb specified as third person singular, along with its subcategorization frame (see §4). The verb lego is subcategorized for a single direct object unmarked for Case, and so legit is inserted with a following NP position, which by government from the verb is assigned features appropriate to accusative Case. (I assume throughout that, in Latin, morphological Case is identical to abstract Case, in view of the lack of evidence to the contrary.) Into this object NP position an accusative NP can be inserted, such as multos libros. The ordering of these processes in intrinsic phrase structure necessarily precedes lexical insertion, and lexical nodes cannot be filled until all the features for their particular position in the tree are determined by syntactic rules. Within sentences, these operations proceed from left to right (in languages with SVO word order). This ensures that the subject controls verb agreement and that the verb controls its governed complements. In complex sentences, these rules proceed anti-cyclical-
ly (i.e. from top to bottom), which ensures that elements of higher sentences govern aspects of lower sentences, as is required in sequence-of-tense phenomena. Finally, the head of a phrase is inserted before the rest of the phrase, to ensure that the head controls agreement within the phrase, as the head noun governs agreement in gender, number, and Case of modifiers within its syntactic phrase. I further assume that syntactic rules cannot change features specified on constituents, but merely add features.

A given constituent may receive features from several sources. A verb may have its tense determined by a higher verb, and its person and number determined by its own subject. The ordering of lexical insertion ensures that no verb is inserted in any given verb position until all necessary features have been appropriately specified.

4. Subcategorization

Chomsky notes (1981, 31) that there is a certain redundancy in having subcategorization frames in the lexicon as well as phrase-structure rules that generate these same frames. One effect of the projection principle is that each lexical item projects its subcategorization frame into the phrase structure into which it is inserted. Taking this idea a step further, using the conventions of Jackendoff's (1977) X-bar syntax, we can say that lexical entries are actually \(X^1(X')\) phrases, rather than simply lexical items \(X^0\). To illustrate with some examples from English, instead of specifying that the verb put can be inserted into a \(V^1\) phrase which contains after \(V\) a \(N^3(N')\) and a \(P^3\), we can write the lexical entry for put as the \(V^1\) phrase itself, as in (17).

\[
\begin{align*}
(17) & \quad \text{put} \\
& \quad \text{V}^1 \\
& \quad \text{V} \\
& \quad \text{N}^3 \\
& \quad \text{P}^3 \\
& \quad \text{put}
\end{align*}
\]

Now, phrase-structure rules are not needed to expand \(X^1\) phrases. Lexical insertion inserts (17) into a \(V^1\) position. Then, phrase-structure rules come into play again to expand \(N^3\) and \(P^3\), until these dominate \(N^1\) and \(P^1\) phrases, respectively, when lexical insertion again operates, until there are no longer any lexical (now interpreted as \(X^1\)) slots remaining to be filled.

In addition to providing a more constrained representation of subcategorization frames, the inclusion of \(X^1\) phrases in the lexicon allows a more satisfactory representation of certain idioms. In earlier versions of transformational grammar, idioms had to be considered as exceptional in requiring phrasal structure in the lexicon. Under the view of the lexicon I am suggesting here, all lexical entries are phrases, and so these idioms need not be regarded as exceptional in this respect—they are exceptional only in that some of their strictly subcategorized complements are lexically filled. Some English examples are given in (18).
(18) a. take...to task

```
  V^1
     /
    V  N^3  p^3
   take  to task

b. crane one's neck

```

The designation (anaphoric) in (18b) is intended to suggest, informally, that this idiom can be inserted into a well-formed syntactic structure only if the N^3 so marked can be interpreted as anaphoric with a c-commanding N^3, subject to the usual bounding conditions. As an additional example, consider verb-particle constructions, which Emonds (1972) argues have discontinuous lexical entries. Such discontinuous expressions receive a natural lexical form, given as (19), just as the idioms of (18) do.

(19) a. look up (as in look the answer up)

```
  V^1
     /
    V  N^3  Prt
   look  up

b. send out (as in send a schedule out to the stockholders)

```

Some Latin examples are given as (20) through (24) and (26). The simplest example is that of an intransitive verb whose V^1 phrase contains no strictly subcategorized phrases. We can assume further that the lexical entry contains the root without inflections, given that lexical rules of the type developed in Jensen (1981) derive inflected forms from roots by affixation, lexical processes, or suppletion. Under these assumptions, the lexical representation of dormio 'sleep' is given in (20).

(20)  

```
     V^1
      /
     V   
   \   \  
  V^{-1}  (stem level)  
  \   \  
  V^{-2}  (root level)  
      
  dorm  

```
The next case is verbs which are subcategorized for a single noun-phrase object, such as lego 'read,' whose lexical representation is (21).

(21)

\[
\begin{array}{c}
V^1 \\
N^3 \\
V^{-1} \\
V^{-2} \\
\text{leg}
\end{array}
\]

The N^3 in the structure of (21) is assigned accusative Case, in features [-oblique, +theme, -source], since this node is governed by the verb and has no inherent Case. These features percolate down to the N node resulting from the expansion of N^3 by phrase-structure rules and lexical insertion of N^1, which ensures that only a noun with accusative Case features can be inserted in this position.

When the object of a verb is not accusative, its Case cannot be predicted by government, and so must be specified in the lexical subcategorization of the verb. A simple case occurs when the verb takes a dative indirect object in addition to an accusative direct object. For example, the verb do 'give' has the lexical representation (22).

(22)

\[
\begin{array}{c}
V^1 \\
N^3 \\
\text{do} \\
\left[ \begin{array}{c}
\text{oblique} \\
\text{theme} \\
\text{source} \\
\text{dative}
\end{array} \right]
\end{array}
\]

The first noun-phrase object in (22) is governed by the verb, and has no inherent Case features. It is therefore marked accusative by the rule of government. The second noun phrase object, although governed by the verb, has an inherent Case marking. According to the principle that syntactic rules cannot change features, it cannot be marked for Case again, and so remains dative. When such verbs occur in passive sentences, only the accusative object can become the subject.

A number of verbs have no accusative object, only an object in an oblique Case. We will consider here verbs that require dative objects, such as noceo 'harm.' Although there is some semantic similarity among the verbs that take dative objects, the Case required by such verbs cannot be predicted because of the existence of transitive synomynms for some of them, such as laedo, which also means 'harm' but takes the accusative. We represent noceo lexically as in (23).
As we have noted, the passive of such verbs cannot have the dative object as a derived subject—the object remains dative and the subject is unspecified. This follows from the requirement that syntactic rules cannot change feature specification: since the Case of the object noun phrase is fully specified in the lexicon, it must appear in that Case in all syntactic permutations.

As an example of an idiom in Latin, we may give the phrase eum taedet vitae 'he is tired of life', which has a subcategorization frame similar to (22), with the additional specification that the verb is impersonal. In the government-binding framework, this specification is given by stating that no θ-role is assigned to the subject of such verbs; in terms of our features, the verb is specified [-T]. The idiom has the representation (24).

\[\text{(24)}\]

5. The problem of 'raising' verbs

As a final example of subcategorization, we will consider verbs that take sentential complements. These can be divided into two general categories, infinitival and finite, and infinitival clauses can be further subdivided into those with and without overt subjects. For example, volo 'want, be willing' takes an infinitival complement with PRO subject, iubeo 'order' takes an infinitival complement with an overt subject, and moneo 'advise, warn' takes a finite clause introduced by ut if positive, ne if negative. Examples of these are found in (25).

\[\text{(25) a. volo ire 'I want to go'}\]
\[\text{b. te ire iubeo 'I order you to go'}\]
\[\text{c. monuit ut nova urbs conderetur 'He proposed that a new city be founded'}\]

(Colebourn 1948, 9, 31)
We can give the subcategorization of these verbs by specifying their complements as sentential (V³ in Jackendoff's X-bar framework), and as finite or infinitival represented as a feature [finite] on the V³ node,⁷ and directly specifying the subject of the complement of volo as PRO. These subcategorizations are shown in (26).

(26) a. \[
\begin{array}{c}
\text{volo} \\
\downarrow \\
\text{V}^1 \\
\downarrow \\
\text{V}^2 \\
\text{vol} \\
\text{N} \\
\text{PRO}
\end{array}
\]

b. \[
\begin{array}{c}
\text{iubeo} \\
\downarrow \\
\text{V}^1 \\
\downarrow \\
\text{V}^2 \\
iub
\end{array}
\]

c. \[
\begin{array}{c}
\text{moneo} \\
\downarrow \\
\text{V}^1 \\
\downarrow \\
\text{V}^2 \\
\text{mon}
\end{array}
\]

Since the subcategorizations given the verbs in (26) are characteristic of intransitive verbs, we might expect them to have only impersonal passives. On the contrary, iubeo and similar verbs have only personal passives, and volo has no passive uses at all. Some passive forms are illustrated in (27).

(27) a. consules iubentur scribere exercitum (Livy 3, 30, 3) 'the consuls are bidden to enroll an army'

b. Nolani muros adire vetiti sunt (Id. 23, 16, 9) 'the people of Nola were forbidden to approach the walls'

c. Sestius accusare Clodium non est situs (Cic. Sest. 95) 'Sestius was not allowed to prosecute Clodius'

(Examples and translations from Woodcock 1959, 102)

The active version of (27a) would be represented as (28).
In the case of these 'raising' verbs, Chomsky invokes the device of 'exceptional Case marking' for English. On the basis of arguments given in §5, I consider exceptional Case marking to be unavailable for the analysis of such verbs in Latin. The subject of the embedded clause in (28) is assigned accusative Case by a rule given in §5. If the higher verb is lexically inserted in its passive form, the subject position in the embedded clause is marked by (16). Thus, (27a) is derived without NP movement or exceptional Case marking. The failure of volo (and similar verbs) to appear in passive structures is a consequence of the lexical fact that such verbs are not provided with passive forms.

5. Case assignment

We have discussed two mechanisms for Case assignment so far: Case assignment by government and Case assignment by strict subcategorization (inherent Case). In order to account for Case assignment to the subject of infinitives (as in 28) without 'exceptional Case marking,' the facts of Latin require a third mechanism, suggested by Williams (1981): Case assignment by phrase-structure rules. We will restrict Case assignment by government to cover only those instances of Case assignment to phrases which are strictly subcategorized by lexical items. Lexical items whose complements bear inherent Case in those items' lexical entries do not assign Case to those complements. Where no lexical Case is specified in subcategorization frames, Case is determined by government according to the principles of (29).

(29)  a. Assign [-oblique, +theme, -source] (=acc.) to the complement of [-N] (=V, P).

   b. Assign [+oblique, -theme, -source] to the complement of [+N, αV] (i.e. assign genitive to the complement of N and dative to the complement of A.
The question of assigning Case to the subject of embedded clauses such as (28) is part of the general problem of assigning Case to subjects, which Chomsky subsumes under the notion of government, using INFL to govern the subject of finite clauses and assuming exceptional Case marking to assign Case to the subjects of infinitives. This analysis seems to be tied to certain idiosyncratic facts about English, and cannot work for Latin, where the subject of an infinitive is always accusative, whether or not that subject is governed by a higher verb in Chomsky's sense. Infinitive clauses appear not only in object position, the normal position of such clauses in English, as in (30a), but also as the subject of the matrix verb, as in (30b, c, and d) and as predicate nominative (as in 30e).

(30) a. Caesarem adesse nuntiavit 'he reported that Caesar was present' Caesar (acc) to be present he reported

b. Caesarem adesse nuntiatum est 'it was reported that Caesar was present' Caesar to be present reported is

c. Hoc facere illum mihi quam prosit nescio 'I don't know how his doing this benefits me' Att ii, 1.6
(Lit.: 'I don't know how him to do this benefits me')

d. quae tandem Ausonia Teucros considere terra invidia est? 'why grudge the Trojans their settling on Ausonian land' Aen. iv. 349-50.
(Lit.: 'what malice is it Trojans to settle on Ausonian land?')

e. Rumor erat Caesarem adesse 'there was a report that Caesar was present' report was Caesar (acc) to be present

(30a, b, e from Allan and Greenough 1901, 287)

In (30a), Caesarem is in a position to be governed by the matrix verb, if exceptional Case marking is assumed. But in none of the other sentences of (30) is it possible to conceive of the subject of the infinitive being governed by the matrix verb.

Another argument for not considering government as the mechanism for assigning Case to the accusative subject of infinitives comes from certain verbs which take non-accusative objects or accusative plus infinitive complements. The clearest example is mememmi 'remember,' which is subcategorized to take genitive objects, as shown in (31).
(31) Animus meminit praeterorum (C. Div. 1, 30)  
    (gen pl)  
    'the soul remembers the past'  

But when memini occurs with an infinitival complement, the subject  
is invariably accusative, as shown in (32).  

(32) Memini Catonem...mecum et cum Scipione disserere (C. Am. 3, 11)  
    (acc)  
    'I remember that Cato...maintained a discourse with Scipio  
and me'  

We might expect at least a few examples of genitive 'subjects' of  
infinitives in the complement of memini if the Case of infinitive  
subjects were determined by government.  

Similar considerations apply to sentences like (30c) if the  
notion of 'subject' clause is interpreted as in Jackendoff (1977),  
following Emonds (1970), as the complement of V². The structure  
of (30c) is (33), under these assumptions.  

(33)  

There is no way to consider illum as governed in this structure,  
even by a radical extension of the notion of 'exceptional Case  
marking.' An alternative method of assigning Case to subjects is  
suggested by Williams (1981, 251), where Case is assigned to sub- 
jects directly in the phrase-structure rule that expands S. Using  
the feature system and Jackendoff's X-bar system, we can give this  
phrase-structure rule as (34).
(34) \[V^3 \rightarrow [\text{finite}] \rightarrow N^3 \rightarrow \begin{cases} \text{-oblique} & \text{[finite]} \\ \text{-source} & \text{[finite]} \\ \text{-atheme} \end{cases} \]

6. Conclusion

We have shown that NP movement is both unnecessary for passive in Latin and a cause of needless complications in the morphological component. By dispensing with NP movement, we lose nothing in the way of explanation of the facts, since these fall out of the theory of \(\theta\)-roles, the theory of Case, and bounding theory, with or without NP movement. Syntactic movements can be confined to operations which do not change grammatical relations, such as WH-movement and scrambling. A unified theory of morphology then becomes possible by confining all morphological operations to the lexicon. For Latin, we do not need the node INFL or the device of exceptional Case marking—Case marking of subjects is done by phrase-structure rules, as we showed to be necessary on independent grounds. The uniformity principle must be relaxed to allow for deponent verbs in Latin.

No doubt many of these conclusions have wider applicability, but this analysis is not directly applicable to English, for example. The distinction between trace and PRO is crucial to an understanding of empty categories in English. This distinction does not seem important for Latin, as empty NPs are generally permitted anywhere, even in strictly subcategorized positions with no controller. Further work is necessary before it is possible to determine whether the distribution of empty categories can be reduced to a single principle in the two languages or in UG.

FOOTNOTES

*I wish to thank Heles Contreras, Paul Hirschbühler, Louis Kelly, Marfa-Luisa Rivero, Donca Steriade, Margaret Stong-Jensen, and Douglas Walker for comments and criticisms on earlier versions of this paper, and Wayles Browne, Lori Davis, Wayne Harbert, James Higginbotham, and Annie Zaenen for comments following the presentation of this material at the Conference on Government and Binding. My apologies for any omissions; all errors are my own. This work was supported in part by a leave fellowship from the Social Sciences and Humanities Research Council of Canada, Grant number 451-82-0133.

\(^1\)There are certain active forms of deponent verbs: the future infinitive, the present and future participles, the gerund, and the supine. Also the gerundive of such verbs is passive in meaning, as is the passive participle of some deponents.
According to the criteria of Cole et al. (1980), there is no possibility of considering these passivized non-accusative objects as syntactic subjects.

Here, I am assuming a movement analysis for expository purposes. The same result holds in a non-movement analysis, as we will show directly.

I am assuming the following feature system in this representation and throughout:

(i) Case features (based on Bierwisch 1967)

Cases

<table>
<thead>
<tr>
<th>Features</th>
<th>nom</th>
<th>acc</th>
<th>dat</th>
<th>gen</th>
<th>abl</th>
<th>voc</th>
</tr>
</thead>
<tbody>
<tr>
<td>oblique</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>theme</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>source</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

(ii) Gender features

<table>
<thead>
<tr>
<th>Features</th>
<th>neuter</th>
<th>masculine</th>
<th>feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>animate</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>feminine</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

(iii) Number features

<table>
<thead>
<tr>
<th>Features</th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>plural</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

(iv) Person features

<table>
<thead>
<tr>
<th>Features</th>
<th>first</th>
<th>second</th>
<th>third</th>
</tr>
</thead>
<tbody>
<tr>
<td>[I]</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[II]</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Wayles Brown pointed out to me that this holds only for transitive deponents; nothing in the theory prevents intransitive deponents from being used as impersonal passives, analogous to (9). As I have found no convincing examples of this use, nor any principle disallowing it, I leave the matter open.

The θ criterion may also prove to be relative, holding absolutely in English and Latin, but allowing violations in the case of the Classical Greek middle voice, where in (i) Κύρος evidently has both the agent role and the benefactive role. See also Lappin (1982).

(i) Κύρος λουτάι 'Cyrus takes a bath (i.e. washes himself)'
I am using [finite] to replace [tense] as the appropriate feature to represent the distinction between finite and infinitival, since, in Latin, infinitives have tense distinctions.

REFERENCES


Bierwisch, Manfred (1967) "Syntactic features in morphology: general problems of so-called pronominal inflection in German," To Honor Roman Jakobson 1, 239-270.


Emonds, Joseph E. (1972) "Evidence that indirect object movement is a structure-preserving rule," Foundations of Language 8, 546-61.


Williams, Edwin (1981) "On the notions 'lexically related' and 'head of a word,'" Linguistic Inquiry 12, 245-274.

Free Relatives and the Null-Head Parameter*

Juliette Levin
MIT

1. Introduction

The syntax of free relatives has been the subject of much linguistic research in recent years. Issues of particular theoretical interest have been the position occupied by the wh-phrase or relative pronoun introducing the free relative, and possible restrictions on both the case and category of this wh-phrase, referred to in earlier analyses as "matching effects" (Bresnan and Grimshaw(1978), Groos and van Riemsdijk(1979)). Of the various analyses proposed, that with the most cross-linguistic evidence is the COMP-hypothesis.\(^1\) The COMP-hypothesis, as proposed by Groos and van Riemsdijk (1979), asserts that the wh-phrase is in the complementizer position at surface-structure, and that the head of the free relative is filled by an empty category. Under this analysis, both matching and non-matching free relatives have the same internal structure. Matching effects are accounted for under a parametrized version of the COMP Accessibility Hypothesis.\(^2\)

In this paper we will investigate the syntax of free relatives within the current Government and Binding framework (Chomsky(1981), (1982)). We will argue that while various cross-linguistic evidence supports the claim that the wh-phrase in free relative constructions is, in fact, in the complementizer position, evidence of an empty category in the head of the free relative is only found in languages which exhibit non-matching effects. We will also argue on purely theoretical grounds, that positing an empty category in the heads of free relatives for matching languages such as English, French, and
German, raises serious problems within the present theory of empty categories.

Instead, we will propose that matching and non-matching free relatives have different syntactic structures, as illustrated in Figure I.

Figure I

a. Matching FR

```
NP
  S
  COMP
  [..wh..] ..e..
```

b. Non-matching FR

```
NP
  S
  COMP
  e
  [..wh..] ..e..
```

Matching free relatives, which show no evidence of an empty category in the head position, are analyzed as literally "headless". Matching effects are derived from general properties of case assignment and/or case checking. Non-matching free relatives in languages such as Morrocan Arabic and Finnish are characterized by the existence of an empty category in the head. Thus, they act like normal headed restrictive relatives, where case assigned to the entire clause percolates to the head, and not to the wh-phrase in COMP which is co-indexed with the gap or resumptive pronoun in the lower clause.

We are thus able to characterize matching phenomena in terms of the existence or non-existence of a null-element in a specific structural position, i.e. the head of a relative clause. The similarities between this empty category and the 'pro' of null-subject languages, leads us to propose the null-head parameter as an extension of the null-subject parameter. Indeed, all the languages in which we have examined examples of non-matching free relatives are null subject languages.

2. Towards a typology of free relatives

Before examining the different types of free relatives, it is necessary to specify what is meant by the term free relative (FR). A FR is a restrictive relative clause with no apparent lexical head. Since the majority of headed restrictive relative clauses are limited to the category NP, this discussion will be limited for the most part to FRs of the category NP. Cross-linguistic study of FRs has led to a distinction between two basic types of FRs: matching and non-matching. In the following sections we will give various examples of matching and non-matching FRs and attempt to extract from them a precise characterization of the matching effects.
2.1 Matching free relatives. The term "matching effects" was first used to refer to observed categorial matching in English FRs:

(1) a. The rabbit eats [\textit{NP} \textit{whatever} it finds \textit{e}]
   b. He votes for [\textit{NP} \textit{whoever} she votes for \textit{e}]
   c. The rabbit followed me [\textit{PP} \textit{wherever} I went \textit{e}]
   d. Nicole will run [\textit{AdvP} \textit{however fast} her mom \textit{runs} \textit{e}]
   e. The future will be [\textit{AP} \textit{however bright} we make it \textit{e}]

In (1) the category of the \textit{wh}-phrase is identical to the category required by the matrix verb's subcategorization frame.\textsuperscript{5} In French we see that FR matching effects also include the strict PP subcategorization frames of certain verbs:\textsuperscript{6}

(2) a. Je le dis \textit{pour qui} je dois le dire \textit{e}
   'I say it for whom I must say it'
   b. J'ai donné le livre à \textit{qui} tu avais fait allusion \textit{e}
   'I gave the book to whom you alluded'
   c. *J'ai rencontré à \textit{qui} tu m'as dit de parler \textit{e}
   'I met to whom you told me to speak'
   d. *Je l'ai donné \textit{pour qui} je dois le dire
   'I gave it for whom I must say it'

The particular PPs selected by the matrix verbs in (2)a,b are also those selected by the embedded verb. (2)c is ruled out by the general subcategorization matching seen to operate in English, since the verb \textit{rencontrer} subcategorizes for an NP, not a PP. The sentence (2)d is illformed because the matrix verb donner subcategorizes for a PP of the type [\textit{PP} a NP], unlike the lower verb which may optionally take [\textit{PP} pour \textit{NP}].

The term matching effects was further generalized to include case matching in case-marking languages such as German:\textsuperscript{7}

(3) a. [Wer nicht stark ist], muss klug sein
   'Who(nom.) isn't strong must be clever'
   b. *[Wen/wer Gott schwach geschaffen hat], muss klug sein
   'Who(acc)/(nom) God has created weak must be clever'
   c. *[Wem/wer Gott keine Kraft geschenkt hat] muss klug sein
   'Who(dat)/(nom) God has given no strength must be clever'

(4) a. Ich nehme,[wen du mir empfiehlt]
   'I take whom(acc) you recommend to me'
   b. *Ich nehme,[wem/wen du vertraust]
   'I take whom(dat)/(acc) you trust'
   c. *Ich nehme,[wer/wen einen guten Eindruck macht]
   'I take who(nom)/(acc) makes a good impression'

As we can see in (3) and (4), the FRs in German are only grammatical when the relative clause and the matrix clause require identical cases on the relative pronoun. A further example will show that this phenomena of case matching is morphological in nature. Within
the case paradigm of wh-pronouns in German, one instance of syncretism occurs: the form was functions as the nominative and the accusative neuter singular form. In (5) we see that though case non-matching does occur, the sentences are grammatical.

(5)  
a. [Was du mir gegeben hast], ist prächtig  
'what(nom/acc) you have given to me is wonderful'

b. [Was du vorschlägst], kommt nicht in Frage  
'what(nom/acc) you propose is out of the question'

c. Ich habe gegessen [was noch übrig war]  
'I ate what(acc/nom) was left'

d. Hast du [was im Programm war] schon kopiert?  
'have you already copied what(acc/nom) was on the program?'

This might lead us to propose that case-matching does not refer to abstract case. The English FRs in (6) offer further evidence suggesting that case matching effects in FRs refer only to case structurally encoded in particular lexical items.6

(6)  
a. [Whoever you invited] will come  
b. *[Whomever you were talking to] has a loud voice  
c. I spoke to [whoever telephoned]  
d. *I spoke to [whomever telephoned]

Though sentences (6)a,c are both examples of non-matching abstract case, they are grammatical in dialects of English where the form whom is optional, if not obsolete. However, in the same dialects, or in others where whom is used regularly for various objective cases, non-matching is ungrammatical in sentences such as (6)b,d. We will return to this issue at a later point as evidence for case checking, but let us now extract from the available data a formal description of matching effects.

As we have seen, categorial matching effects refer to the fact that in certain languages, if a verb subcategorizes for an NP, then the FR will be of the type [\[NP \ Y \ [NP \ ... \ wh \ ...\] S] where Y is yet to be determined. Likewise, if a verb subcategorizes for a PP, the FR will be of the type [\[PP \ Y \ [PP \ ... \ wh \ ...\] S]. We may generalize our description over any category XP.

(7)  
\textbf{Categorial Matching}

Where a matrix verb subcategorizes for a category XP, a matching FR in the subordinated clause will be of the type [\[XP \ Y \ [XP \ ... \ wh \ ...\] Z ]], Y and Z variables.

We now turn to the issue of case matching. If we adopt a theory which incorporates the notion of case assignment, we might argue that sentences such as (3)b,c, (4)b,c are ungrammatical due to the fact that the structural position occupied by the relative pronoun is doubly case-marked. However, as we noted above, in the case of the relative pronoun \underline{was} in German, as well as the
relevant English sentences in (6), this generalization cannot be upheld. This suggests that a process of case-checking, operative at the PF-component, is at work, checking whether Case has been properly assigned at previous levels of structure. In sentences like (3)b, the case-checking process will check that the lexically realized NP in subject position has the proper case (nominative), while the wh-phrase is checked for accusative case. Because the case-checking process in matching FRs is apparently checking the same structural position twice, the lexical form in this position will only pass successive case checkings if it has multiple interpretations at PF. Thus, the sentences in (5) pass this case-checking process and are grammatical. The same holds for the English examples (6)a,c for the relevant dialects. Note that we have been assuming all along that in matching FRs, the case assigned to the higher NP somehow percolates down into the wh-phrase. We will see how this occurs when we examine the structural position occupied by the wh-phrase in the FR, and the existence or non-existence of an empty category in the NP head (cf. sections 3.1, 3.2).

We are now ready to describe the observed phenomena of case matching in FRs.

(8) Case Matching

In a matching FR of the type NP, case checking requirements of the maximal NP, and of the chain (wh-, e) are satisfied by the same lexical item at PF.

Let us now turn to examples of non-matching FRs and attempt to determine in what ways these relative clauses are exceptions to the matching effects described in (7) and (8).

2.2 Non-matching free relatives. Though languages with non-matching FRs are not numerous, they have received a good deal of attention in the current literature. Some of these languages are: Basque, Finnish, Gothic, and Moroccan Arabic.

The sentences in (9) from spoken Moroccan Arabic illustrate categorial non-matching in FRs (from Fassi-Fehri(1982)):

(9) a. ja mCam an dwiti came with-whom spoke-you 'whomever you spoke with came'
    b. skit bmCam an dwiti complained-I of-with-whom spoke-you 'I complained about whomever you spoke with'

Further examples show that Finnish is a language which exhibits case non-matching: (from Carlson(1977)):
(10) a. Valitsen mitä sinäkin valitset
    choose-I what(part) you-too choose-you
    'I choose what you choose'

b. Valitsen mistä sinä pidät
    choose-I what(elat) you like-you
    'I choose what you like'

In (10)a the verb valita 'to choose' takes a partitive NP object, while in (10)b pitää 'to like' takes an elative NP object. The case of the relative pronoun is that assigned to it within the subordinate clause. Unlike matching FRs, case assigned (or checked) to the maximal NP clearly does not percolate down to the wh-phrase, as evidenced by the ungrammaticality of such sentences as (11).

(11) a. *Valitsen mitä sinä pidät
    choose-I what(part) you like-you

b. *Pitän mistä sinä valitset
    like-I what(elat) you choose-you

Non-matching FRs in Finnish are interesting in that they may only occur when the entire NP is assigned nominative, accusative or partitive case from the matrix verb (or from INFL).

(12) a. Valitsen mistä sinä pidät
    choose-I what(elat) you like-you

b. *Pitän mitä sinä valitset
    like-I what(part) you choose-you

Sentence (12)b is thus ill-formed, since the matrix verb pitän subcategorizes for an elative NP object. With cases such as this one, matching must occur, as shown in (12)a. Though non-matching is also characteristic of headed restrictive relatives in Finnish, no such case restrictions as those observed above occur.

These facts from Morrocan and Finnish lead us to propose that non-matching FRs have the same properties as headed relatives, but that given the structure in (13), Finnish poses certain restrictions on the possible case of eₖ.

(13)

```
NPᵢ
  /\  \
NP_j  S
  /\  /
 e_j  COMP
     /\  /
    [...wh...k] [...ek...]
```

If the case assigned to NPᵢ is nominative, accusative, or partitive, then case will percolate from NPᵢ to NPⱼ. However, if case assigned to NPᵢ is some other case, such as elative, case seems to percolate into COMP. One possible analysis of this is explored in section 3.2.
3. The internal structure of free relatives: the COMP-hypothesis

In this section we will review the arguments in support of the COMP-hypothesis proposed by Groos and van Riemsdijk (GVR), (1978). This hypothesis attributes to FRs, both matching and non-matching, the structure in (14):

(14)

\[
\begin{array}{c}
\text{XP} \\
\downarrow \\
\text{S} \\
\text{COMP} \\
\downarrow \\
[... \text{wh} ...]_j [... e_i ...]
\end{array}
\]

As we can see from (14), the COMP-hypothesis involves more than positing that the wh-phrase or relative pronoun in FRs is in COMP. It also incorporates the presence of an empty category in the head. In 3.1 we will review arguments for the wh-phrase in COMP. In 3.2 it will be demonstrated that the only evidence provided by GVR for an empty head in matching languages such as German, Dutch, and English, may be accounted for independently under the Empty Category Principle. Finally we will argue that only in non-matching FRs is there evidence for an empty category in the head.

3.1 Wh-phrase in COMP. Recent analyses of FRs in Dutch, German, English, French, Catalan, Arabic, Hungarian, and Hebrew have all included convincing arguments that the wh-phrase or relative pronoun is in its canonical COMP position. This in fact does seem to be the null hypothesis, given the nature of wh-movement and variable binding within the present theory. We will therefore assume the null hypothesis, that the wh-phrase is in COMP, and will only briefly review evidence supporting this position.

GVR observe that S extraposition in German and Dutch may apply to FRs as well as to sentential complements. They argue that if, in fact, the wh-phrase was in the head position and not in COMP, one would expect extraposition of FRs to be ungrammatical for the same reasons that NP extraposition is ungrammatical. In addition, one would also expect that it would be possible to leave the relative pronoun in place, and to extrapose the rest of the clause. However, we can see from (15) and (16) that this is not so.

(15) a. Der Hans hat das Geld zurückgegeben [das er gestohlen hat]
   'Hans has returned the money that he has stolen'
b. Der Hans hat zurückgegeben [was er gestohlen hat]
   'Hans has returned what he has stolen'
c. *Der Hans hat was zurückgegeben er gestohlen hat
(16) a. Ik heb de vis opgegeten [die over was]
   I have the fish eaten that was left
   (Dutch) 'I have eaten the fish that was left'
b. Ik heb opgegeten [wat (er) over was]
c. *Ik heb wat opgegeten (er) over was

The FRs in the b. sentences of (15) and (16) behave in the same way as the extraposed Șs in the a. examples. GVR conclude that the rule deriving these b. sentences must be the same rule of relative clause extraposition (extraposition from NP) in operation in the a. sentences. Thus, the conclusion that the relative pronoun is in COMP.

From a more general point of view, we may observe that if the wh-phrase is in the head position in languages such as English and French (as proposed by Bresnan and Grimshaw(1978)), an argument for the ungrammaticality of the following sentences must be constructed:

(17) a. *I saw [whoever [who [that]]] you saw e
    b. *Je le dis pour qui[que] je dois le dire e

These sentences are ungrammatical precisely because of the position of the wh-phrase in COMP. As such, they are ruled out by the doubly-filled COMP Filter which states: *[(e) \( \infty \) COMP] (Chomsky and Lasnik, 1977).

For further arguments supporting the position of the wh-phrase in COMP in both headed restrictive relatives and FRs, I refer the reader to the sources cited in Footnote 9.

3.2 Evidence for an empty category in the head. Though GVR rely implicitly on the assumption that an empty category occupies the head of a FR, they do not provide much evidence supporting this hypothesis. The only empirical evidence brought to bear on this question again concerns Ș extraposition. If we reexamine the examples of Ș extraposition in (15) and (16), we see that they both involve extraposition of a FR NP from object position. (18) and (19) illustrate that extraposition from other positions such as NP/S or object of a preposition is ungrammatical in German.

(18) Extrapolation from NP/S
   a. [Was du darlegst] ist unsinnig
      'what you expound is nonsensical'
b. *Ist unsinnig, [was du darlegst]

(19) Extrapolation from object of P
   a. Ich habe mich über was er zurückbrachte, sehr gefreut
      I have myself about what he brought back very rejoiced
      'I was very pleased about whatever he brought back'
Extrapolation of a FR from subject position in this case is ruled out by the verb-second principle in German. Thus, if another element occupies sentence initial position, the FR may extrapose, as seen in (20), (21).

(20) a. Morgen wird e verurteilt [was heute noch für richtig gilt] tomorrow is condemned what today still as correct counts 'What is considered correct today will be condemned tomorrow'
b. *e wird morgen verurteilt [was heute noch für richtig gilt]

(21) a. Es ist e nun endlich getan worden [was getan werden musste] there has now at last done been what done been had to 'what had to be done has now at last been done'
b. *e ist nun endlich getan worden [was getan werden musste]

However, the ungrammaticality of FR extrapolation out of PP’s as in (19) cannot be accounted for in this way. GVR claim that the inability to extrapolate from a PP is due to the fact that the rules of co-indexing, which co-index the head of the relative clause with the wh-phrase in COMP, are sensitive to the Head Constraint which states:

(22) Head Constraint

No rule may involve $X_i$ ($X_j$) and $Y$ in the structure:

$$...X_i...Y...X_j$$

if $Y$ is c-commanded by the lexical head $H$ (N, V, A, P) of $X$, where $X = N", V", A", P"$. Therefore, the rule of co-indexing may not relate the relative pronoun external to $P"$ to an element c-commanded by the preposition. But this is clearly wrong in the case of lexically headed restrictive relatives, and must therefore be supplemented by the following statement:

(23) The Head Constraint operates at the level of co-indexing in logical form, but it only affects null anaphors (non-lexical elements), and no lexicalized positions.

This is equivalent to saying that $Y = [e]$ in (22). GVR conclude:

This qualification yields the desired contrast: the relative pronoun of an extrapolated normal relative can, but the relative pronoun of an extrapolated FR cannot be coindexed to their respective heads. As a consequence, the empty head of a FR which has been extrapolated from a PP does not receive an index and the structure is excluded as unwellformed. (p.19)
However, the argument above is reduced quite simply to the Empty Category Principle, if we assume that the category left by extraposition is an NP trace. As will be shown further on, the S of matching FRs can be viewed as exhaustively dominated by NP, so that S extraposition will in fact leave behind a category \([_\text{NP}_e]\). Now, it follows from the ECP that an NP trace must be properly governed. Because F is not a proper governor in German, extraposition from PP is ruled out. With this, we eliminate the sole evidence provided by GVR for an empty category in the head of FRs.

For matching FRs, we are therefore left with the structure in (24):

(24) 

\[ \begin{array}{c}
  \text{NP} \\
  \text{S} \\
  \text{COMP} \\
  \text{[...wh...]} \\
  \text{[...e_1...]} \\
\end{array} \]

This is the structure we propose in Section 4. for FRs in languages which exhibit matching effects.

If we turn now to non-matching languages, we may start out by observing that the phenomena of non-matching itself is characteristic of headed relative constructions (with the exception perhaps of a language such as Basque). From the data already noted in Moroccan Arabic and Finnish, we are led to the null hypothesis again, that these FRs do in fact have heads, and that the heads are empty. We could then propose the following syntactic structure:

(25) 

\[ \begin{array}{c}
  \text{NP} \\
  \text{e} \\
  \text{NP} \\
  \text{S} \\
  \text{COMP} \\
  \text{[...wh...]}_i \\
  \text{[...e_i...]} \\
\end{array} \]

Under this analysis, the case restrictions on Finnish FRs discussed in 2.2 may be restated as: Case of \([_\text{e}_e]\) = nominative, accusative, or partitive. These cases have been regarded as the unmarked cases in Finnish (Carlson, 1978), and it is not surprising to find that a base generated empty category may only be interpreted as having unmarked case.

Case restrictions on this empty category, which we will later define as a type of 'pro', are also apparent in Basque FRs, and show interesting dialectal variations (Baker, 1982).
4. Matching vs. non-matching FRs: structurally defined

From the discussion presented above, we arrive at the conclusion that the difference between matching and non-matching FRs is a structural one. Matching FRs are literally headless, while non-matching FRs are headed.

(26) Matching FR

(27) Non-matching FR

This structural difference accounts for matching and non-matching effects in a simple way. In truly headless relatives (26), case assigned to the entire NP percolates down to $\hat{S}$ and then to the structural head of $\hat{S}$ which we will assume is COMP. However, the wh-phrase also receives case from the case-marked position of the variable in the lower clause. From this it follows that in (26), the element in COMP is doubly case-marked. Assuming again a process of case-checking, the FR will be well-formed if and only if the lexical item in COMP checks out correctly for each case assigned to it. This explains why non-matching in abstract or structural case may only occur in phrases where the wh-word in COMP has multiple interpretations at PF, as we saw in the case of German was, and English who. In the headed FRs, case assigned to the higher NP percolates to its head $[NP, e]$. The wh-word again receives the case of the variable in the lower clause (or that assigned to it by the relevant preposition, should pied-piping have occurred). Non-matching occurs as a natural result of the headed construction.

4.1 The Null-head parameter. The null-head parameter may now be redefined as the existence or non-existence of a base generated empty category in the head of a relative clause. Languages with this empty category are non-matching, while those without it exhibit headless FRs with the resulting matching effects. This empty category cannot be an NP trace, since it is not in a chain with an NP in a $\hat{S}$-position. It cannot be a variable since it is not locally $\hat{A}$-bound. It cannot be PRO because it occurs in a governed position. Within the present typology of empty categories, it can only be some type of 'pro', (small pro) of null-subject languages.

Let us take a short aside here to reconsider the implications of the COMP-hypothesis as proposed by GVR, given the nature of this
empty category. This would entail positing 'pro' in the head of matching and non-matching free relatives alike. Languages like English, French, and German which are not pro-drop languages would have to be analyzed as having 'pro', but only in heads of free relative clauses. In addition, only this empty category in this structural position would trigger COMP accessibility. Clearly these results would not only weaken a theory of empty categories, but would also parametrize the notion of a 'null-subject parameter'. We should therefore not be surprised that the present study has led us to reject the COMP-hypothesis employing the notion of an empty head in accounting for matching phenomena.

Returning to the nature of the empty category in the head of non-matching FRs, we see that it differs from 'pro' of the null-subject languages in that it may occur in positions other than subject of a tensed sentence. This might lead us to expect that languages with non-matching FRs should be able to generate 'pro' in subject position, thus falling into the class of null-subject languages. Indeed, Moroccan Arabic, Finnish, Basque, and Gothic are all languages with non-matching free relatives, and all are null-subject languages. Because 'pro' can occur in positions other than NP/S in these languages, it is possible to view the null-head parameter as an extension of the null-subject parameter into positions other than NP/S. Not all of these languages allow pro-drop in object position however, leading us to hypothesize that non-matching FRs represent an intermediate level between null-subject and null-object languages as depicted in (28).

\[(28)\] Extensions of the null-subject parameter

\[
\begin{array}{c}
\text{(NP/S)} \\
\text{Italian: null-subject} \\
\text{(NP/NP)} \\
\text{Finnish: null-head} \\
\text{(NP/VP)} \\
\text{Basque: null-object}
\end{array}
\]

Should the null-head parameter prove to be an extension of the null-subject parameter, a comparison of null-subject languages with matching FRs to those without could provide finer delimitations within the set of properties commonly associated with null-subject languages.
5. Appendix: Free relatives and reanalysis

Proposing the structure in (1) to account for the observed properties of matching FRs might prove problematic for interpretations of X-theory, or for theories assuming a set of internalized generalizations about the lexicon expressed perhaps in the form of PS rules of the type: \[ \text{NP} \rightarrow \text{NP}(S) \].

(1)

Indeed, previous analyses of FRs all assumed a structure of the type:

Using the formal devices discussed by Goodall (1982), we will show how generalizations over the lexicon of this type might force reanalysis which would provide NP with the needed lexical head. Consider the FR in (2).

(2) I like [\text{NP}_S [\text{whoever you hate}_e]]

According to the above analysis, we may represent the FR as in (3).

(3)

This configuration may be represented by the abstract tree in (4), where C is the COMP node, b stands for a possible empty element in COMP, c for the string of terminal elements under S, and (5) is the Reduced Phrase Marker (RPM) associated with (4).

(4) \quad (5) \{A, B, Cc, abD, abc\}

Should a speaker of English have internalized a generalization of the form \[ \text{NP} \rightarrow \text{NP}(S) \] for relative clauses, reanalysis of (3) will be necessary,
taking the wh-word in COMP as the head of NP. Reanalysis will consist of adding the appropriate monostring to the RPM in (5). Adding the monostring aB will give us the new RPM in (6) and the new tree in (7). The translation of (7) back into a labeled tree yields (8).

(6) \{A, aB, abD, abc\}

(7)

(8)

\[ \text{S} \]

you hate e

Reanalysis of this type would therefore create an NP head for the headless relative, and uphold the phrase structure generalization \( \text{NP} \rightarrow \text{NP(S)} \) for relative clauses. The possible coexistence of the trees (3) and (8), in a three-dimensional tree-structure (as suggested by Goodall(1982)), could explain the long-time FR controversy over whether the relative pronoun is in the head, or in COMP. The tree in (3) is the only one which will account for matching effects and the proper relationship of control between the variable and its \( \tilde{\alpha} \)-binder. The tree in (8) is the only one which will account for generalizations about English phrase structure. Thus, the three-dimensional tree is optimal in satisfying conditions of well-formedness derived from both the set of principles we call Universal Grammar, and from internalized generalizations over the lexicon.

A similar process of reanalysis may be proposed for FRs of the category PP. The French sentence in (9) may have one of the two possible tree configurations shown in (10) and (11).

(9) \( \text{J'ai parlé à qui tu as parlé} \) \[ \text{PP} \]

(10)

(11)

If we propose the structure in (10), we must provide some preposition deletion rule, perhaps a rule dependent on processing limitations, which will delete one of two identical prepositions in this configuration. Though such a rule may exist, it is not clear how such a rule would be formulated. We leave this question for further study. In (11) we are faced with a figure similar in structure to (3). In this case, an
NP exhaustively dominated by PP might be justifiable in languages such as French and English due to the existence of other PPs with no overt heads, such as those in (12), (13).

(12)  
\begin{itemize}
  \item a. Je suis allé [PP chez elle]
  \item b. [PP Où] as-tu mis tes chaussettes [PP e]?
\end{itemize}

(13)  
\begin{itemize}
  \item a. I went [PP home]
  \item b. [PP Where] did you put the book [PP e]?
\end{itemize}

As we noted in the case of FR NPs, should a generalization of the type PP \(\rightarrow\) P NP be extracted from the grammar, reanalysis of structures like that of (11) would take place. The tree in (11) is represented by the abstract tree in (14), where c stands for the string of terminal elements under S. (15) is the RPM associated with (14).

(14)
\[ A \]
\[ B \]
\[ C \]
\[ D \]
\[ E \]

(15) \( \{A, B, C, Dc, abE, abc\} \)

Reanalysis will add the monostring ab to the RPM in (15) producing the new RPM in (16) and a new tree, (17). Filling in label values of (17) we get (18).

(16) \( \{A, ab, aC, abE, abc\} \)

(17)
\[ a \]
\[ B \]
\[ C \]
\[ b \]
\[ E \]
\[ c \]

(18)
\[ \text{PP} \]
\[ \text{NP} \]
\[ \text{qui} \]
\[ \text{tu as parlé [PP e]} \]

This account of reanalysis seems plausible, given the fact that sentences such as (10) are ungrammatical, while in (9), the PP à qui is clearly simultaneously satisfying the subcategorization frames of the higher and lower verbs. This reanalysis strategy is not possible in a language like English which allows preposition stranding.

(19)  
\begin{itemize}
  \item a. *I talked to whom(ever) you gave the book.
  \item b. The man to whom you gave the book smiled.
  \item c. I talked to whom(ever) you gave the book to.
\end{itemize}
Clearly pied-piping occurs in English, as exemplified by (19)b. This makes the question of why (19)a is ill-formed problematic. Under the present analysis, it might be the case that reanalysis only occurs when it is forced, as a last resort for mediation between principles of Universal Grammar and internalized generalizations over the possible strings of lexical items. In English, reanalysis of sentence (19)a is not possible since an optional strategy of preposition stranding exists, producing (19)c. Thus, where an alternate strategy is possible within the grammar, reanalysis does not occur.

FOOTNOTES

* Many thanks to N. Chomsky, G. Goodall, R. Manzini, D. Pesetsky, and K. Safir for much helpful discussion.

1 Though Borër(1979) adopts the Head-hypothesis to account for FRs in Hebrew, in more recent work (Borer(1981)), she argues convincingly for the COMP-hypothesis. Of the work done since 1979 on FRs, all seems to support the hypothesis that the relative pronoun is in COMP.

2 Two versions of the COMP accessibility hypothesis exist. In one, certain languages have the property that their COMP position is accessible to the "rules of regimen", which include case-marking and subcategorization. Another account is that COMP is universally accessible, and that in certain languages (those with matching FRs), empty categories may never be referred to by rules of regimen. See GVR for further details.

This characteristic has led some linguists to refer to FRs as headless relatives. We refrain from using this term, since a large part of the controversy over FRs concerns precisely whether they are headless or not. As we will discover later in this paper, not all FRs are "headless".

4 For one possible solution to the problems posed by FRs of the category PP, see the Appendix.

5 As mentioned earlier, it is not within the scope of this paper to discuss FRs of the categories A, Adv. For one possible treatment of these in English see Bresnan and Grimshaw(1978).

6 Examples are from Hirschbuhler(1976). Though many speakers do not accept PP FRs with different verbs in the matrix and subordinate clauses, some do. This fact does not alter our argument, but rather suggests that the reanalysis solution presented in the Appendix might involve V or VP identity for one class of speakers.

7 All German and Dutch examples in this paper are taken from Groos and van Riemsdijk(1979).
8I am assuming a spellout rule for wherever forms in COMP which not only distinguishes NP from S in sentences such as a. and b., but also adds a type of semantic indefiniteness, or universality which tends to be quite idiosyncratic from speaker to speaker. Thus, the fine semantic distinction(s) between sentences c. and d.

a. I know who you like.
b. I know whoever you like.
c. I ate what he ate.
d. I ate whatever he ate.

9Hirschbühler (1978); Groos and van Riemsdijk (1979); Hirschbühler and Rivero (1981); Horvath (1981); Borer (1981); Fassi Fehri (1982).

10 Basque exhibits two different relative clause configurations for both normal headed restrictive relatives and FRs. One configuration always shows matching effects, while the other does not. For more of the facts, see De Rijk (1972), Baker (1982).

11 COMP as head of S is discussed in Chomsky (1981), Stowell (1981).

12 Exceptions to this are languages which exhibit Case attraction. Case attraction is when the Case assigned to the higher NP is realized in the wh-phrase in COMP, though the gap in the lower clause is of a different Case. Attraction only occurs with certain combinations of Cases, and seems to reflect some sort of Case hierarchy. One recent discussion of Case attraction (Franks, 1980) attributes the phenomena to the result of different types of Case (Abstract, structural, surface) being assigned (or checked) at different levels of structure (deep structure, S-structure, and PF). See Harbert (this volume) for an analysis of case attraction in Gothic.

REFERENCES


Baker, M. (1982) "Relative Clauses in Basque" ms. MIT.

Borer, H. (1979) "Restrictive Relative Clauses in Modern Hebrew," ms. MIT.


Carlson, L. (1977) "Remarks on the Syntax of Free Relative Clauses," ms. MIT.

(1978) "Problems in Finnish Syntax," ms. MIT.


Fassi Fehri, A. (1978) "Comparatives and Free Relatives in Arabic," ms. MIT.


Goodall, G. (1982) "Notes on Reanalysis," ms. MIT.


A-BINDING*

JUDITH Mc A'NULTY

UQAM

This paper makes a tentative proposal to fill the gap left in Binding Theory by the absence of a hypothesis for A-binding which, like A-binding, is based on configurationally defined NP types. More specifically, I argue that (1) makes it possible to define anaphors, pronominals, and variables in this manner, then submit them to a theory of A-binding which parallels A-binding.

(1) Case Inheritance and Clitic Absorption apply on the left side of the grammar

Once A-binding and A-binding can thus be compared, I explore the possibility of reducing the two subtheories to a general theory of X-binding.

Minimally, A-binding includes (2):

(2) A variable is A-bound

more interestingly, it includes a set of three local conditions, informally described as (4), which correspond to those in (3):

(3) A-binding
   A. An anaphor is A-bound
   B. A pronoun is A-free
   C. A variable is A-free (non-local)
(4) \( \bar{A} \)-binding
   A. An anaphor is \( \bar{A} \)-free
   B. A pronominal is \( \bar{A} \)-free
   C. A variable is \( \bar{A} \)-bound

1. What is a variable?

   The first major problem which arises when moving from \( A \)-binding to \( \bar{A} \)-binding is that of defining the term variable, since we must now take into account elements such as intermediate traces in COMP, which were not crucial to \( A \)-binding. No configurational definition seems appropriate even if we make reference to \( A \) and \( \bar{A} \) positions.

   Locally \( \bar{A} \)-bound \([e] \)s include at least the categories in (5):

   (5) a. wh-traces and quantified NP-traces in their original Case-marked positions (\(+A, -Case\))
   b. empty A positions associated with clitics (\(+A, -Case\), at least if one accepts Jaeggli's 1980 rule of Clitic Absorption)
   c. intermediate traces in COMP (\(-A, \dagger\)Case depending on the status of Case Inheritance)

   Any definition based on the features in (5) will be disjunctive.

   Chomsky (1981) offers a solution to the disjunction problem by suggesting the definitions in (6) (p. 330):

   (6) a. \( \alpha \) is a variable iff it is locally \( \bar{A} \)-bound and in an A-position
   b. if \( \alpha \) is an empty category and not a variable, then it is an anaphor

   In (6), variables are partially defined in terms of their binders, though not exclusively. The content of a theory of \( \bar{A} \)-binding as a set of well-formedness conditions is reduced, but the A-position criterion still allows for excluding certain constructions: for instance, if an empty category is locally \( \bar{A} \)-bound but not in an A-position, then it is construed as an anaphor, in which case condition A of (3) intervenes, stipulating that it must be locally \( \bar{A} \)-bound, a contradiction in terms if the condition refers to local binding as defined in (7)

   (7) \( \alpha \) is locally bound by \( \beta \) iff \( \alpha \) is \( X \)-bound by \( \beta \)
   \[ \text{and if } \gamma \text{ Y-binds } \alpha \]
   \[ \text{then either } \alpha \text{ Y-binds } \beta \]
   \[ \text{or } \gamma = \beta \]

   (Chomsky, 1981, p. 185)
which leaves room for excluding what we would now call cliticization of adjuncts (see Gross (1968) and Ruwet (1972)):

(8) a. \([S_{\bar{A}} \text{ Discuter avec cet individu}] \text{ est difficile}\)
    b. \(\text{Il est difficile } [S_{\bar{A}} \text{ de discuter avec cet individu}]\)
    c. *\(\text{Il en est difficile}\)

(9) a. \(\text{Il semble que Paul soit parti}\)
    b. \(\text{Il le semble}\)

But notice that (5) analyzes intermediate traces in COMP as anaphors: although \(\bar{A}\)-bound, they do not occupy A-positions. Yet they are undeletable now that ECP is assumed to apply at LF.

More suitable definitions for empty categories can be found in (10):

(10) a. If \(\alpha\) is locally \(\bar{A}\)-bound by a t-antecedent,\(^2\) then it is a variable
    b. If \(\alpha\) is locally A-bound by a t-antecedent, then it is NP-trace, a non-pronominal anaphor
    c. In any other case, an empty category is PRO

(Chomsky, 1981, p. 328)

But now both A-binding and \(\bar{A}\)-binding cease to be well-formedness conditions on structures. Checking mechanisms formerly associated with Binding Theory must depend on other modules.

In what follows I pursue a different approach, where both A-binding and \(\bar{A}\)-binding remain well-formedness conditions and where variables are configurationally defined as in (11):

(11) variable = [e], +Case

Two further assumptions must be made:
    a) verbs may assign Case to traces in COMP;
    b) that is the head of \(S\) and is a Case assigning head.

The first is justified in Kayne (1980) and in the work of Julia Horvath on Hungarian; the second I assume to be plausible if \(S\) and \(S\) are different maximal projections, since \(S\) must then have a head, which given selectional restrictions between that and AGR must be the tense marker in COMP, a tense marker which is closely related to another tense marker (AGR) which does assign Case.

The rest depends on (1).

2. Predictions

2.1 Clitics are \(\bar{A}\)-binders for variables. By (1), the empty categories associated with clitics are construed as variables, not ana-
phors. These empty categories lose their Case on the left side of the grammar: on the right, they remain empty categories with Case.

Two arguments can be brought forth to support this claim:
a) Assuming condition A of A-Binding to be correct, if empty categories associated with clitics were anaphors, the clitics themselves would have to be arguments. This entails two subcategorization frames for languages where verbs take clitics, the first when a clitic is present, the second when the verb is followed by a full NP.
b) All empty NPs associated with clitics must have Case before Absorption takes it away from them: no clitic is associated with an NP which is not a position for Case-marking:

(12) a. $\Delta$ semble [$S$ lui être fou] → 
b. *Il le$_{i}$ semble [$S$ e$_{i}$ fou]

If the [NP e] associated with the clitic could be construed as an anaphor, that is, if Absorption applied in the syntax, there would be no difference on the right side of the grammar between (12b) and sentences where cliticization is possible, which puts the whole burden on the Case Filter (left). Although this is a possible solution, it seems premature to endorse it at this point because 1) it entails modifying the Case Filter to include categories which are generally agreed upon not to be NPs (Kayne, 1975); 2) it suggests that even if the Case Filter does apply to clitics, Case Inheritance is the only means for clitics to receive Case - creating new mysteries for ethical datives and other clitics for which there is no possible argument source; 3) it does not account for the contract between (13) and (14), where a clitic is deleted before being submitted to the Case Filter, a situation which is comparable to Chomsky's analysis of example (iii) in footnote 5.

(13) a. Jean le lui/lui le donne → (DO deletion)

b. Jean {lui}[[i]] donne (Quebec French)

(14) a. $\Delta$ semble à lui; [$S$ lui j fou]
b. *Il le$_{j}$ lui$_{i}$ semble [$S$ e$_{j}$ fou] → (Quebec French)
c. *Il lui semble fou (where ungrammaticality refers to the b. interpretation, not to the interpretation where il is the raised subject)

This last example rather supports our hypothesis that when an empty category associated with a clitic is not Case-marked, it is construed as an anaphor and the sentence is excluded by condition A of A-Binding.

2.2 The difference between adjectives and verbs. Given our previous assumptions, wh-traces in COMP preceded by adjectives and not followed by that have no Case on the right side of the grammar: they receive Case on the left and transmit it to the wh-phrase in
COMP to meet the Case Filter. For binding purposes then, they cannot be construed as variables.

This corresponds precisely to Chomsky's (1981) judgements of Kayne's (1980) examples:

(15) a. Who is it essential she talk to \( t_1 \)?
    b. Who is it essential \( t_1 \) talk to her?

    (16) Who is it essential \( t_1 \) that John see \( t_1 \)?

    (17) Who do you think \( t_1 \) came?

In (15), the circled \( t_1 \) is inappropriate as a bindee, having no Case. It must be treated as an anaphor, where it violates condition A of A-binding.

Similarly, a trace in subject position may be an inappropriate bindee, as in (18):

(18) *Who did you try \([S \quad t_1 [S \quad t_1 \text{ to win}]]\)??

One thus avoids having to assume that try is intransitive, like seem (LGB, p. 297), in order to explain that the trace in COMP receives no Case that it can transmit to the empty subject.

Despite accurate predictions for (15-17), our hypothesis fails to account for the difference between (19) and (20):

(19) *Jean croit \([S \quad [S \quad Paul \quad \text{être le plus intelligent de tous}]]\)
(20) Qui crois-tu \([S \quad t_1 \quad [S \quad t_2 \quad \text{être le plus intelligent de tous}]]\)?

In Chomsky (1981), the difference between the two sentences is explained as follows: along with Kayne (1980), Chomsky assumes that Case assignment may cross one major category boundary, but not two, and that Case is assigned to an index, then optionally transmitted to coindexed elements. In (20), \( t_2 \) receives Case from croire and \( t_1 \) inherits it from \( t_2 \). This solution is not available here, where Case Inheritance applies on the left.

What I suggest is based on a hypothesis proposed by Couquaux (1979). Couquaux noticed a surplus in the distribution of en (Ruétet's 1972 EN-AVANT), a gap in the distribution of se, as well as an interesting correlation between the two: contexts permitting EN-AVANT are precisely those which exclude se, that is NP - être - ADJ (or passive participle):
(21) a. La tour en est penchée
   b. *La tour en fumée

(22) a. Le sage s'est tué
   b. *Le sage s'est complètement indifférent

He concludes that the deep structure of (21a) and (21b) should be (23):

(23) être - NP - AP

EN-AVANT thus becomes unnecessary, while the gap in the distribution of se is accounted for by the independently motivated constraint that the antecedent of se must be a deep structure subject (Kayne, 1975).

Let us assume this to be correct. The surface structure for (19) then becomes (24):

(24) *Jean croit [₃₅ [₃₅ Paul₁ être t₁ le plus intelligent de tous]]

The trace of the subject can be Case-marked by être. If this happens, the sentence is excluded by conditions C of A-binding and of A-binding. If no Case-marking occurs, then Paul never receives Case (in contrast with tensed sentences such as Paul₁ est t₁ le plus intelligent de tous), and the sentence is excluded by the Case Filter. Example (20) is grammatical, however, because if the original trace of qui (t₁) receives Case in (25):

(25) Qui. crois-tu [₃₅ t₁ [₃₅ Δ être t₁ le plus intelligent de tous]]?

both traces have Case, t₁ by croire, t₁ by être, and local binding conditions are satisfied.

2.3 Traces of adjuncts are anaphors. In discussing (7) and (10) earlier, I referred to the possibility these definitions opened of treating traces of adjuncts as anaphors (see examples (8) and (9)). Having rejected (7) and (10) on the grounds that the first turned wh-traces in COMP into anaphors that did not meet condition A of A-binding while the second transferred well-formedness conditions traditionally associated with Binding Theory to other modules, I now show that (1) accounts for the same facts.

The more interesting case is that of Stylistic Inversion (SI) in French:

(26) Quand est parti Jean?

It is well known that SI leaves behind an unbound variable in sub-
ject position. For the construction to meet binding conditions, there must be a rule applying between S-structure and LF which creates an antecedent for the variable. In "ECP extensions" (1981), Kayne proposed such a rule, which I will call NP - to - S.³ NP - to - S adjoins the postposed subject to S, leaving behind the trace of an adjunct:

(27) a. Quand Jean est parti?
   b. Quand t₁ est parti Jean → (NP - to - S)
   c. [S Quandᵢ [S Jeanᵢ [S t₁ est parti t₁₂ tᵢᵢ]]]

Suppose that t₁₂ is a variable, as predicted by applying Case Inheritance in the syntax. After NP - to - S, (27c) is excluded because the variable is A-bound by t₁ (condition C of A-binding). Here, the optionality of Chomsky's Case Inheritance cannot save the construction since the surface structure is (27b), and not (27c): Jean must receive Case to meet the Case Filter. In contrast, (1) predicts that although Jean receives Case, this happens on the left side of the grammar: on the right, the NP, and therefore its trace t₁₂, remains Caseless and is construed as an anaphor. The sentence is therefore predicted to be grammatical: t₁₂ is A-bound by the variable t₁, which is itself A-bound by the NP adjoined to S, an explanation similar to what Chomsky proposes to account for (29).

(28) [S Quandᵢ [S Jeanᵢ [S t₁ est parti t₁₂ tᵢᵢ]]]?
    A-binding  A-binding

(29) [S WHOᵢ [S t₁₂ has been killed tᵢᵢ]]?
    A-binding  A-binding

A second prediction concerns cliticization of adjuncts. As the examples in (8) show, extrapoosed clauses cannot be cliticized; neither can NPs moved by Stylistic Inversion:

(30) a. Quand est parti Jean?
   b. *Quand l'est parti?

With Case Inheritance on the right, the trace (or empty category) associated with the clitic is construed as a variable. If one assumes Chomsky's (1981) definition of c-command to be correct, this variable meets condition C of A-binding, and the sentence is predicted to be grammatical. One again, the optionality of Case Inheritance is irrelevant: insofar as Case can be assigned, there exists one possible derivation for (8c) and (30b). Of course, if cliticization occurs, NP - to - S cannot take place under a movement analysis of Clitic Placement (the same constituent is moved to two different places) - which raises the question of how the original trace in subject position is bound. This new problem is however completely dependant on a movement analysis for clitics (if clitics are base-generated, NP - to - S may apply to the posposed UP, while
the base-generated clitic  \( \bar{A} \)-binds the empty category associated with it), and remains irrelevant in the case of extraposed constituents, where a designated element has replaced the offending trace.

2.4 PRO DROP Languages. With some verbs at least (see Burzio, 1981), PRO DROP Languages exhibit free in version of the subject.

(31) e ha telefonato Gianni

According to Chomsky (1981), e in (31) must be PRO: it cannot be trace because of the difference between PRO DROP and non-PRO DROP languages is to be defined as (32),

(32) R may apply in the syntax (where R is the rule which affixes AGR to the V, that is, inside VP, where it no longer c-commands the subject).

The trace of the subject is un govered when R applies in the syntax; when it does not apply, the trace, being both governed and Case-marked, becomes on unbound variable. In contrast, if e is construed as PRO, these constructions are grammatical under the option that R applies in the syntax.

My contention is that it is in fact possible to return to the trace analysis first proposed by Chomsky in the original version of the Pisa Lectures and further expanded in Sůněr (1982). Suppose that we extend Kayne's NP - to - S strategy to all cases where a variable is unbound. Under the option that R does not apply in the syntax, the trace in subject position, a variable, is correctly \( \bar{A} \)-bound by the preposed NP.

(33) \( \left[ S _{Gianni} _1 \left[ S _{t_1} _1 ha telefonato t_1 _2 \right] \right] \)

The problem is again the trace of the adjunct. If Case Inheritance applies on the right, no variable solution is possible: the NP Gianni obviously needs Case in order to meet the Case Filter, but Case assignment will turn t_1.2 into a variable which is incorrectly \( \bar{A} \)-bound by t_1. Given (1), however, the trace is construed as an \( \bar{A} \)-bound argument.

Consequently a moved NP leaves behind a trace, not a PRO (a PRO analysis depends on (10)), from which it inherits nominative Case in a straightforward manner.

Finally (1) has little effect on clitic doubling, the original motivation for Absorption. Most of Jaeggli's arguments rest on Case assignment, rather than government per se, Case assignment being impossible when the relevant NP is un govered. None of this is changed of Absorption applies on the left.

As for the right side of the grammar, our hypothesis states that if the empty categories associated with clitics are indeed PROs, then PROs are governed and Case-marked (variable), and appropriately \( \bar{A} \)-bound by the clitics.

Finally, \( \bar{A} \)-binding accounts for the impossibility of questioning on NP doubled by a clitic, which for Jaeggli does depend on the absence of government resulting from Absorption - the trace of the wh-phrase, if not a true PP, is un gov ered.
(34) a. ¿* A quién lo viste?
   b. ¿ A quién viste?

In (34a), the trace of a quién is governed, Case-marked and correctly bound by lo. However lo, as a pronoun, is subject to Disjoint Reference by Condition B of A-binding (see 4).13 (The sentence is also excluded by the Bijection Principle (Koopman & Sportiche, 1981), a principle which can perhaps be subsumed by A-binding.

3. Local domains

3.1 The Minimal Distance Principle and Superiority. In the preceding sections, I have referred rather loosely to "local" conditions, without defining what constitutes a local domain (D), a notion which I now attempt to define more precisely in order to explore further the predictions of (3) and (4).

Assuming the three conditions in (4) to be basically correct, local domains cannot be identified with governing categories, whether the relevant node is $S$ or $\overline{S}$:

(35) a. $\left[ \overline{S} \, \textit{Who} \!_{i} \left[ \textit{S}=\textit{D} \, \textit{t} \!_{i} \textit{l} \textit{i} \textit{likes himself} \right] \right]$
   b. $\left[ \overline{S}=\textit{D} \, \textit{Who} \!_{i} \left[ \textit{S} \, \textit{t} \!_{i} \textit{i} \textit{likes himself} \right] \right]$

if $D = S$, the variable $\textit{t} \!_{i} \textit{l}$ is not $\overline{A}$-bound in $D$; if $D = \overline{S}$, the anaphor himself is incorrectly $\overline{A}$-bound in $D$. Crossing lines and nested constructions also show that a local domain cannot be defined strictly in terms of the Minimal Distance Principle (MDP).

(36) a. $\overline{S} \, \textit{Who} \!_{i} \left[ \textit{S} \, \textit{as-tu donné} \!_{i} \textit{t} \!_{i} \textit{t} \!_{j} \right]$
   b. $\left[ \overline{S} \, \textit{Who} \!_{i} \left[ \textit{S} \, \textit{everybody} \!_{i} \textit{hates} \!_{j} \textit{t} \!_{i} \right] \right]$

(37) $\overline{A} \, \textit{qui} \!_{i} \left[ \textit{I} \!_{i} \textit{as-tu donné} \!_{i} \textit{t} \!_{i} \textit{t} \!_{j} \right]$

I will however make a tentative proposal based on a relativized version of the MDP, in conjunction with the Superiority Condition. By a "relativized" interpretation of the MDP, I mean that the condition holds when a rule may apply to two or more possible antecedents, much like Chomsky's, 1973 version of the A/A Principle and the Superiority Condition. Consequently, local domains are not fixed domains. In addition to ambiguous application, I borrow from the initial Superiority Condition the notion of asymmetrical c-command, as opposed to strict c-command, which may be mutual. Within the limits of Subjacency, a local domain is the domain of the first superior possible antecedent:

(38) Local domain $(D)$

$\alpha$ defines a local domain for $\beta$ if $\alpha$ c-commands $\beta$ and if there is no $\gamma$ such that $\alpha$ is superior to $\gamma$ and $\gamma$ is superior to $\beta$ ($\alpha, \gamma =$ possible antecedents for $\beta$).14
The general effect of (38) is illustrated in (39):

(39)

Only with respect to Superiority does (38) basically differ from a translation into the terminology of local domains of Chomsky's (1981) definition of local binding (see (10)), a difference which seems necessary to explain why, in the case of multiple complements, Disjoint Reference (DR) is not limited to elements in VP:

(40) a. Jean [VP le lui a donné]
   b. Jean [VP lui a donné une pomme]

The presence of another clitic in (40a) and of a full NP in (40b) does not exempt the clitics on the full NP from DR with the subject.

Turning now to the notion of possible antecedent, assume that all +wh constituents are possible antecedents for empty +wh constituents (interrogatives) and that all +Q constituents are possible antecedents for empty +Qs. Given non-distinctness, it follows that all NPs and clitics are possible antecedents for pronominals and anaphors, as well as for Case-marked empty categories that exhibit neither +wh nor +Q features.

3.2 Consequences. One of the consequences of (38) is that the domain of himself in (35) is S and the domain of its binder (t, l) is S. Both are appropriately bound in their respective domains, with no violation of either (3) or (4).

Another is that the anaphor each other in constructions such as (41) may be bound by the subject or the direct object if the "empty" preposition (at) does not block symmetrical c-command of V-complements:

(41) They shot the arrows at each other

both are possible antecedents for the anaphor, but the DO does not create a local domain.

Turning now to the more controversial area of variables, we see that an NP which is not marked +wh does not create a local domain for a wh-trace:
(42) a. [D What did John see t]?
   b. *What did [D who see t]?

(42a) simply restates the fact that wh-traces are not subject to the Specified Subject Condition. In the second case however, the implication is that the original Superiority Condition, which (38) also subsumes, still holds, accounting for the ungrammaticality of the sentence and replacing the current ECP analysis, a correct claim, I believe, since the ECP account for (42b), described in (43), does not carry over to sentences where the wh-phrases belong to different clauses (44),

(43) LF [Who_i [what_j [t_i saw t_j]]]

(44) *What did you accuse who of doing?

and relies heavily on ECP applying at LF, a perennial problem for QR (the prediction being that wide scope of object quantifiers is the unmarked case).

With respect to multiple quantifiers, (38) predicts that object quantifiers are in the scope of subject quantifiers - in fact, it derives Lakoff's 1970 global rule on the surface order of quantifiers:

(45) a. Many men read few books
   b. Few books are read by many men

(46) a. Everyone loves somebody
   b. Somebody loves everyone

the object quantifier cannot cross-over the +Q subject if it is within VP.

Finally, (38) accounts for the difference between cliticization and wh-movement in Faire-infinitive constructions. It is well known that in sentences such as (47), the NP preceded by à may be cliticized as well as fronted by Wh-Movement, while (48) exhibits only Wh-movement:

(47) a. Jean a fait boire le vin à NP
   b. Jean lui a fait boire le vin
   c. A qui Jean a-t-il fait boire le vin?

(48) a. Jean a fait téléphoner Paul à NP
   b. *Jean lui_i a fait téléphoner [D Paul e_i]
   c. A qui_i Jean a-t-il fait [D^t_i téléphoner Paul e_i]?

These facts are derived from our concept of local domains: in (48), the subject of the embedded clause defines the local domain of the e associated with the clitic; however, since this subject is not +wh, it does not define the local domain for the wh-trace. In (47), whatever the analysis of the FI construction, there is no NP asymmetrical ly c-commanding either empty categories in the embedded S.
4. **X-binding**

With this in mind, I now turn to the comparison of (3) and (4).

The two subtheories each contain a set of three conditions, but they cannot easily be compared, because condition C of A-binding is not a local condition.

4.1 **Condition C of A-binding.** In recent literature, condition C of A-binding has come under close scrutiny, especially with respect to "names" (Engdahl, 1982; Lasnik, 1982). Lasnik shows that the condition is subject to parametric variation and suggests that it may be pragmatic in nature. In addition to Lasnik's data, the claim is supported by the fact that we would expect a pragmatic condition to apply at S-structure rather than at LF, which accounts for the difference between QR and Wh-movement:

(49) a. SS: He liked every book that John read

(b) LF: [Every book x that John read] he liked x

(50) a. *He really likes which one of the people that John met

b. Which person that John met does he really like?

Finally, with the exception of names, it cannot be convincingly argued that variables must be free "in every governing category":

- Traces of non-interrogative wh-phrases in clefts, Tough-Movement and relatives are always bound by arguments outside their local domains, since their antecedent must be A-bound for these constructions to be interpretable. While it may be the case that these wh-phrases are subject to a specific local rule, there is no reason to expect a chain reaction where the effect of the rule carries over to the trace itself by neutralizing the violation of condition C.

- Clitics may be A-bound outside D. The empty categories associated with clitics therefore share this property.

- As for traces of interrogative wh-phrases, one can also claim that they may be A-bound outside D, though this possibility is never realized because of the nature of their antecedents (names):

(51) *John wonders [\[ which you prefer [t]]

thus in (51), coreference between John and the trace ti is impossible because who, the antecedent of ti, cannot itself be A-bound.

Two options can be then taken with respect to condition C: the first is to do away with the condition altogether, moving it to the level of interaction between syntax and pragmatic, the second is to keep the condition for [NP e], Case, but to exclude names and
relegate these to pragmatics. Both options are compatible with the preceding facts. The second option, which I tentatively adopt here (see 4c), entails that condition C of A-binding is a local condition, like all other binding conditions, making a comparison between (3) and (4) possible.

4.2 Condition C of X-binding. Given local domains, the most interesting characteristic of our local version of condition C for A-binding is that it does not seem to be very useful. Where the predictions made by the condition are not questionable, it is apparently subsumed either by other binding conditions or by general principles of grammar.

Clear cases of condition C are strong cross-over, empty categories associated with non-reflexive clitics, QR, and intermediate traces in COMP. Unclear cases include weak cross-over phenomena.

Examples (52) - (55), where 1 refers to the predictions made by condition C of A-binding and 2 to the principle which subsumes it, illustrate most of the clear cases:

(52) Strong cross-over
   a. *Who_i does John_i love t_i ?
      \[2\] \[1\]
      2 : surface pragmatic condition on names
   b. *Who_i does he_i love t_i ?
      \[2\] \[1\]
      2 : condition B of A-binding

(53) Empty categories associated with non-reflexive clitics
   *NP_i Cl_i e_i
      \[1\]
      2 : condition B of \(\overline{A}\)-binding

(54) QR
   a. [[Someone_i] he loves t_i
      \[2\] \[1\]
      2 : condition B of \(\overline{A}\)-binding
   b. (i) [Someone_i] John loves t_i
      \[1\]
   (ii) John loves someone
      \[2\]
      2 : surface pragmatic condition on names
As for intermediate traces in COMP, when they have Case and are construed as variables, they cannot be A-bound because, if they were, the argument binding them would create a local domain within which they could not meet condition C of A-binding (where in subject or object position NP_i asymmetrically c-commands t_i and exhibits the same features)

\[
(55) \quad *[S \ldots NP_i \ldots [S \quad t_i \ldots ]] \quad \text{[argument]} \quad +\text{Case}
\]

unless a clitic is also present in this domain, which brings us back to (53) [condition B of A-binding].

Notice finally that the case of weak cross-over is no more straightforward in its current account by condition C of A-binding than it is here: for exclusion by condition C, c-command restrictions must be relaxed, a situation which only occurs with respect to this condition (and not say, with condition B of A-binding). Our alternate explanation (exclusion of (56a) by condition B of A-binding) calls for a weakening of the NP boundary, a plausible solution, but one for which we have no convincing explanation:

\[
(56) \quad \text{a. } *\text{Who}_i \text{ does } [_{NP \text{ his}_i \text{ mother}}] \text{ love } t_i
\]

\[\begin{array}{c}
1 \\
2
\end{array} \quad \text{condition B of A-binding}
\]

\[
\text{b. } [_{NP \text{ His}_i \text{ mother}}] \text{ loves him}_i
\]

So it turns out that if we take condition C of A-binding to be a local condition, it need not be stated as such in the grammar. The only necessary condition for variables is condition C of A-binding, repeated here as (57):22

\[
(57) \quad \text{A variable must be A-bound}
\]

4.3 Conditions A and B. Local domains for conditions A and B of (3) and (4) are determined by the first superior NP.

In the case of condition A, if this NP is coindexed with the anaphor and in argument position, the sentence meets condition A of A-binding; if the NP coindexed is not in argument position, the sentence violates both A conditions: a violation of (4a) is a violation of (3a). Binding of anaphors within a local domain yields similar results: when a clitic binds an anaphor (*\text{Al}, 'a été tué e_i), (3a) is violated along with (4a) when no antecedent is available (condition B is violated in the presence of an antecedent).23 In contrast with condition C, here the condition from A-binding may subsume the condition from A-binding, which has nothing left to exclude.

Conditions B of A-binding and A-binding are different in that neither one can subsume the other. They can however, be construed as a single condition which does not mention the nature of the c-commanding NPs: from which pronominals must be disjoint in reference.
The three conditions for X-binding can then be stated as (58):

(58) X-binding (local)
   A. An anaphor is $A$-bound
   B. A pronominal is free
   C. A variable is $\bar{A}$-bound.

NOTES

* This research is part of a larger project supported by a grant from the Programme d'aide financière aux chercheurs (PAFAC) at UQAM. I thank Anne-Marie di Sciullo, who conceived the project with me, as well as José Bonneau and Pierre Layoie, for their comments and suggestions.

1. Used here in the broad sense, the term configurational includes inherent features and structurally assigned features but excludes reference to binders. Binding Theory this remains a set of well-formedness conditions on structures: as in Chomsky (1979) and in the first four chapters of Chomsky (1981), antecedents are not used to determine whether an element is an anaphor, a pronominal or a variable, therefore whether the well-formedness of the structure it occurs in depends on condition A, B, or C.

2. Where t-antecedent refers to the target position of a movement rule.

3. This hypothesis is justified in Mc A'Nulty (in preparation) and references cited there.

4. Note that if the relevant Case is the unmarked Case, that may not assign Case when there is Case assigner in the embedded clause,
   (i) *I think that he came
   (ii) *I believe him came

One may further speculate that if there exists such a minimal c-command strategy for Case assignment, this strategy could account for the absence of wh-movement from embedded clauses introduced by for, since for would automatically assign Case to the wh-phrase (or its trace) in COMP, leaving the subject Caseless.

5. I will not examine here the possibility that $\bar{A}$-binding is a sufficient condition for anaphors. Although this should not be rejected outright, it leads to many sentences having ambiguous derivations (with consequences that are unclear at present)

(i) a. Who was killed?
   b. [$_S$ Who$_i$ [$S$ t$_i$2 was killed t$_i$]]
   c. [$_S$ Who$_i$ [$S$ $\Delta$ was killed t$_i$]]
(ii) a. John seems to have left  
   b. \( S[\text{t}_i \text{ to have left}] \)  
   c. \( S[\text{t}_i \text{ to have left}] \)

and to predictions of grammaticality for (iii)

(iii) *The man \( S \) that \( S \) you tried \( S[\text{t}_i \text{ to win}] \)

I also take it that morephological binding can be reduced either to A-binding or to A-binding.

6. See also Burzio (1981).

7. Note that qui cannot be placed into the \( \Delta \) position before moving on to COMP: If it were and \( t_i \) had Case, the sentence would be excluded for the same reasons as (24); if \( t_i \) received no Case, the sentence would be excluded by condition A of A-binding: although \( t_i \) would be correctly A-bound, the additional trace occupying the \( \Delta \) position in (25), also an anaphor would not be bound by an argument.

8. The rule first appears in "Extensions on binding and Case-marking" (1980) under the name NP - to - COMP. In 1981, Kayne modifies it slightly so that it has the same target as May's (1977) QR. For present purposes, it is irrelevant which formulation we choose.

9. One consequence of this hypothesis is that anaphors can be A-bound at LF. For further details, see my "Multi-level binding" (ALNE XIII/NELS XIII Proceedings).

10. c-command (Chomsky, 1981):
    \[ \alpha \text{ c-commands } \beta \text{ if and only if} \]
    
    (i) \( \alpha \) does not contain \( \beta \)
    
    (ii) Suppose that \( \gamma_1, \ldots, \gamma_n \) is the maximal sequence such that
    
    (a) \( \gamma_n = \alpha \)
    
    (b) \( \gamma_i = \alpha^j \)
    
    (c) \( \gamma_i \) immediately dominates \( \gamma_i + 1 \)
    
    Then if \( \delta \) dominates \( \alpha \), then either (I) \( \delta \) dominates \( \beta \), or (II) \( \delta = \gamma_i \) dominates \( \beta \)

11. The argument does not hold, however, if the definition of c-command is changed to (i):
    
    (i) cmax command
        \[ \alpha \text{ cmax commands } \beta \text{ if neither } \alpha \text{ nor } \beta \text{ dominate each other} \]
        
        and if Emonds' rule of Finite Verb Raising applies on the left side of the grammar (see Mc A'Nulty, 1983a): the clitic simply does not c-command the empty category.

12. This will not save the ungrammatical (i):
    
    (i) *Est parti Jean

if we assume that the NP adjoined to S can only serve as a binder,
not a proper governor (this follows from the original hypothesis that
ECP applies at S-structure): a wh-phrase or its trace must be present
at S-structure for ECP to be met (see Mc A'Nulty, 1981, and the refe-
rences cited there).
13. The absence of Disjoint Reference between the clitic and the pro-
noun in lo vieste a él remains unexplained.
14. A second condition may turn out to be necessary to account for
the island behavior of reflexive clitics and NP associated with them.
15. See Mc A'Nulty (1983a) for a justification of having clitics
under VP rather than under INFL on the right side of the grammar.
16. The implication is that traces keep their features. Resumptive
pronoun strategy seems to indicate that this is so, as well as agree-
ment between a past participle preceded by avoir and a moved DO in
French: as Emonds has pointed out (p.c.), this local rule, which ap-
plies only to certain dialects, is most simply described as agreement
between a past participle and an adjacent trace with the feature
[+ FEM].
17. This was one of the characteristics of Chomsky's 1973 version of
the Superiority Condition:
   (i) a. Which books did John give to which students?
       b. To which students did John give which books?
18. Current literature abounds with counterexamples, which I made
no attempt to explain. The only point raised here is whether the un-
marked case is free ordering of quantifiers.
19. Given Subjacency, D = COMP in both b. and c. under a movement
analysis of clitics.
20. I take it that relative pronouns are A-bound by the NP that
contains them, as illustrated in (i):
   (i) [SN₁ L'homme [D à qui₁ tu as parlé t₁]]
   one could argue that the pronoun, and consequently its trace, is bound
not by an argument, but by an argument chunk (l'homme). However, I
see no reason to deny wh-phrases a type of indexation that has long
been accepted for other NPs in the Bach/Bouton paradoxes:
   (ii) [SN₁ Les enfants qui méritent les prix qu' [SN₁ ils]
       désirent] les auront.
21. Where names include full NPs and all constituents which can
truly be interpreted as operators (interrogative wh-phrases – quanti-
fied NPs).
22. With this unique condition, reflexive clitic constructions in
French, which were at odds with condition C of A-binding as stated,
cease to be counterexamples:
   (i) Les enfants se sont parlés e
23. This does not hold for sentences such as (i):
   (i) *Quand t₁₁ s'1'a été tué t₁₂

   These are excluded because the variable t₁₁ is unbound.

REFERENCES


EMONDS, J.E. (1978) "Le composé verbal V'V en français", Cahier de linguistique n° 8, Montréal, P.U.Q.


KAYNE, R.S. (1981a) "On certain differences between French and English" LI, vol. 12, n° 3.


MC A'NULTY, J. (in preparation) "Pourquoi INFL (AGR) est la tête de P"


SPANISH COMPARATIVES, DELETION OPERATIONS AND THE ECP

Carlos Piera
Cornell University

1. Introduction. This paper is concerned with the distribution of Spanish comparative clauses. Some of these exhibit overt wh movement, while others don't. The surface configurations in both cases are very similar, which makes this diversity highly non-functional, and therefore unlikely to derive from anything other than deep-seated properties of the language faculty. This possibility is made plausible, furthermore, by the fact that native speaker judgements on which of the two alternatives should be chosen in a given environment are very clear, while at the same time the phenomenon has never, to my knowledge, been described satisfactorily by grammarians -and hence cannot be taught (see, e.g., Gili Gaya, 1969, simply reproduced in Real Academia Española, 1973, or, in English, Spaulding's criticisms, of a reasonably good approximation by Ramsey (Spaulding, 1962).

The analysis proposed here is consistent with these expectations: the apparent arbitrariness of the distribution follows from independent properties of grammar. The crucial construct invoked is the Empty Category Principle (ECP) of Chomsky (1981):

(1) ECP: [e] must be properly governed

However, for the ECP to produce the desired results it must be allowed to apply at S-structure, not only, as is assumed in much recent literature, at the level of Logical Form (LF). Other principles are introduced in the course of the discussion: it is proposed that wh words in relatives and comparatives are anaphors, and an attempt is made to prevent deletion rules from creating configurations which would have been illegitimate S-structures.

2. Basic data. It is well-known that comparative constructions give rise in most languages to a bewildering variety of surface forms. The phenomenon discussed here appears overtly only in inequality comparisons, but even so we will have to confine ourselves to a drastically reduced sample of forms. They are, in my view, all that is required, but readers unfamiliar with Spanish must unfortunately take my word for it.
Our basic data will be the following:

(2) Julia tiene más libros que Marta
J. has more books than M.

(3) Julia tiene más libros aquí que Marta en casa
J. has more books here than M. at home

(4) Julia tiene más libros (aquí) de los que tiene Marta (en casa)
J. has more books (here) of wh-word has M. (at home)

Los que in (4) -where los, which takes gender and number, must be masculine plural in agreement with libros- is typically found in wh constructions such as headless and PP relatives (cf. footnote 1):

(5) Los que llegaron tarde no encontraron comida
Those who arrived late did not find any food

(6) Los invitados para los que preparé comida llegaron tarde
The guests for whom I prepared food arrived late

Notice now that it cannot appear in the context of (2)-(3):

(7) *Julia tiene más libros de los que Marta

(8) *Julia tiene más libros aquí de los que Marta en casa

On the other hand, its omission in (4) leads to ungrammaticality:

(9) *Julia tiene más libros (aquí) que tiene Marta (en casa)
The same pattern is observed with adjectival comparatives, where the 'neuter' form lo que appears, as it does with quantifiers such as mucho:

(10) a. Julia está más contenta aquí que María en casa

       b. *Julia está más contenta aquí de lo que María en casa
          J. is happier here than M. at home

(11) a. Julia está más contenta aquí de lo que estaba en casa

       b. *Julia está más contenta aquí que estaba en casa
          J. is happier here than she was at home

We now turn to the status of these basic data. There is a possibility that Spanish may lack phrasal comparatives, i.e. comparative constructions whose second term is not, underlyingly, a sentence. When a pronoun can be marked for Case, it appears in comparatives in the nominative form, not in the form expected after a preposition:
(12) a. Un libro para mí / *yo
   A book for me (non-nomitive)

b. Yo / *mí no tengo libros
   I (nomitive) have no books

c. Julia tiene más libros que yo / *mí
   J. has more books than I (nomitive)

Compare (12c) with its Italian translation, where the grammatical-
ity values are reversed:

(13) Giulia ha più libri di me / *io

The issue of phrasal comparatives is interesting in its own right, but we need not rely here on it. Consider again (3):

(14) (=3)) Julia tiene más libros aquí que Marta en casa]
The verb in α is missing, but α is unmistakably sentential, just as (15) would be:

(15) que Marta tiene en casa

Notice that Marta en casa in (14) cannot be a constituent, and hence not the complement NP of the PP que Marta en casa, as would be required in a non-sentential analysis. For one thing, extraposition from NP is impossible with non-restrictive relatives:

(16) a. Julia, que es mi amiga, llegó tarde
   J., who is my friend, arrived late

b.*Julia llegó tarde, que es mi amiga

If Marta en casa were a constituent, (17) below would be possible, since the relative clause would be adjacent to its head; it is excluded, however, for the same reason as (16b) is:

(17) *Julia tiene más libros aquí que Marta en casa, que es mi amiga

The example in (3) and (14), therefore, includes a configuration structurally equivalent to (15). Its complement is verbless either because the verb has not been inserted or because it has been deleted. What matters here is that its structure is identical to that of (4), except for the fact that wh movement does not seem to have applied in it. I will base my discussion on a comparison of these two types of structures.
3. The ECP and the distribution of wh movement. We will concentrate, then, in the pattern in (3), (4), (8) and (9), repeated here for convenience:

(18) Julia tiene más libros aquí que Marta en casa
(19) Julia tiene más libros aquí de los que tiene Marta en casa
(20) *Julia tiene más libros aquí de los que tiene Marta en casa
(21) *Julia tiene más libros aquí que tiene Marta en casa

That wh movement applies in (19) has been established by Rivero (1979) and Plann (1980). I will not repeat their arguments here. Notice that not only el (la, los, las) que, but also the wh word cuanto/a(s), can be used in these constructions -(22) -. This belongs to the somewhat literary style of (23):

(22) Julia tiene más libros aquí de cuantos tiene Marta en casa
(same gloss as (19))

(23) Tráeme cuantos libros tengas
    Bring me what books you have

As for wh movement not applying in (18), notice that by COMP-to-COMP movement (24) should be possible:

(24) *Julia tiene más libros aquí que cree Carmen que Marta
    en casa

J. has more books here than C. believes that M. at home

We will come back to this. At this point, the pattern in (18) -(21) suggests the following descriptive generalization:

(25) When the verb in the complement is missing, there is no
    overt wh movement

This is the generalization that we have to account for.

We could try to derive (25) from the proposal that deletion operations, or their interpretive equivalents, are somehow forbidden in the domain of a wh word, i.e. that wh movement creates an island for deletions as it does for extractions. But this would be wrong, since Gapping does apply in Spanish within wh islands:

(26) ¿A quién le doy el libro y a quién el caramelo?
    To whom (do I) give the book and to whom the candy

Alternatively, we might consider an explanation in terms of Case. Assume that what accounts for the absence of the verb in (18) is that it has not been inserted, i.e. that the category V dominates an empty node, as in the systems of Jackendoff (1972), Wasow (1979) and Williams (1977). A missing verb cannot assign Case. If variables must have Case, then the trace of wh movement out of
a verbless complement will induce ungrammaticality, since it will be caseless. This account, however, is also unsatisfactory. Comparatives can involve items which are clearly caseless, such as adverbial quantifiers, as in (27):

(27) Julia trabaja mucho
J. works a lot

(28) Julia trabaja más de lo que trabaja Marta
J. works more than M. works

There is no reason to require mucho to have Case in (27); notice, in particular, that the intransitive verb could not assign it, that mucho does not occupy a 0-position, and that similar quantifiers are generally unmarked in languages with overt Case systems. Therefore, a trace of this same category should not be excluded if not Case-marked.

I suggest that what lies behind the generalization (25) is the ECP. It would operate as follows: wh movement of the compared constituent leaves a trace, as it always does; this trace qualifies as the empty 5 of (1) above; if the original position of the compared constituent is in the VP, the trace is not governed, let alone properly governed, if the verb has not been inserted; therefore the ECP disallows traces of wh movement when the verb is absent. This still does not cover the case of subject traces, as in (29):

(29) Vinieron más mujeres de las que fueron invitadas
More women came than were invited

Subject extraction in languages with the general characteristics of Spanish can operate in one of two ways, according to recent studies. Rizzi (1982) proposes that subjects in these languages are never extracted from the preverbal position; it is always possible to move the subject to a position adjoined to the VP, where the verb governs it, thereby permitting extraction without violating the ECP. Torrego (1981) argues that many wh constructions in Spanish trigger a fronting of the verb: the verb moves to the left of S, from where it governs the preverbal subject position. Suppose that wh movement in Spanish does indeed trigger V-fronting, as seems likely (notice the word order in the examples given here). If the verb is present, it governs the trace of the subject; if it is not, the trace violates the ECP. An analogous account of subject extraction within Rizzi's system is straightforward.

One aspect of this account deserves some more discussion because of its theoretical implications. Recall that it relies on the assumption that the verb is absent at the level where the ECP applies. This level must follow 'Move $\alpha$', and hence is not D-structure. But it also cannot be the level of Logical Form. For
in LF the complement clause must have a verb. This follows from assumptions about this level which, as far as I know, are universally made: an S-level unit of LF shares with parallel well-formed formulas of predicate logic the property that it cannot consist of arguments without a predicate. Hence a structure like (30) would be ill-formed at LF:

\[(30) [S \text{ Marta} \ [V \text{ e}] \ [NP \text{ en casa}] ]\]

(30) is, in a simplified form, the S-structure of the complement of (18). At LF, it is interpreted as (31):

\[(31) [S \text{ Marta} \ [V \text{ tiene}] \ [NP \text{ x}] \ [en casa] ]\]

In other words, leaving now aside questions about the nature of \(x\), a copying rule of some sort must reproduce the relevant features of the main clause verb in \([V \text{ e}]\) of (30). Now, if the verb appears as in (31), the adjacent trace would be properly governed by it, just as it is in examples like (19) where the verb was never absent. The proper level for application of the ECP is here, therefore, the level at which this clause appears as (30), i.e. the level of S-structure.

4. Other regularities in the basic data. We have now explained why wh movement applies in (19) and does not apply in (18). It remains to account for the ungrammaticality of (21). This account will depend on what analysis is given of the compared constituent in (18), the parallel well-formed case. Suppose it is an unrealized wh phrase, in effect identical to the one in (19) except for its phonetic content. Then the problem reduces to the one posed by the equally ungrammatical (32) below, the counterpart to (22) where no movement has taken place:

\[(32) \text{*Julia tiene más libros (de) Marta tiene cuantos}\]

The violation in (32) is in turn comparable to the one occurring when wh words in relative clauses do not undergo 'Move \(\alpha\)'.

\[(33) \text{*El hombre que hablé a quién} \]

\[\text{The man that I talked to whom}\]

In Spanish, as in English, the reason cannot be that wh movement is somehow obligatory, since multiple questions are possible -(34)- and so are echo questions without movement -(35)-:

\[(34) \text{¿Quién ha visto qué?} \]
\[\text{Who has seen what}\]

\[(35) \text{¿(Que) has visto a quién?} \]
\[\text{(That) you have seen whom}\]

Notice also that the principle barring vacuous operators (Chomsky, 1982:11) does not help here, since it must apply at LF. What we
are asking is, in fact, why there cannot be LF movement in (21) -and, we assume, (32) and (33).

We could simply say that wh movement is irrelevant because the configurations in question are ill-formed at S-structure. Suppose 'relative pronouns' and, in general, non-interrogative wh phrases are anaphoric. If they are not moved to COMP -an Â position-, they are free in their governing or binding category, and therefore violate principle A of the binding theory. This applies to all of our examples. It will be noticed that this treatment is particularly appropriate for the operator-like quantified phrase involved in comparative formation, like cuantos in (22) and (32), which in no way can be construed as referential. Recall that the treatment proposed here is intended to be valid for adjective, adverb, and QP comparative constructions.

This treatment has consequences for the analysis of (18), reproduced below:

(36) (= (18)) Julia tiene más libros aquí que Marta en casa

Suppose (36) included the structure (37):

(37) que Marta [V e] [NP_i + wh] en casa

NP_i is phonologically unrealized, but contains the feature + wh. The same line of reasoning that led us to exclude (20) should exclude movement of NP_i: it cannot move, as we saw in section 3, without violating the ECP. But now we see that it cannot stay in place either, since we are claiming that it would then violate the binding theory. The structure in (37) would thus be ruled ungrammatical in any case. I conclude that the division of Spanish comparative complements into two classes -(18) and (19)- is not adequately characterized as a division between cases of overt wh movement -(19)- and latent LF wh movement -(18)-: LF wh movement cannot yield a grammatical output in the second class.

In view of the above, I propose that we treat (36) and similar constructions as containing a compared constituent not marked + wh; its determiner is presumably empty, and it:can be coindexed with its antecedent by a rule of construal at LF.

5. Interaction with deletion rules. The analysis just given for (36) creates problems elsewhere. Let us now assume that there exist rules of deletion 'under identity' (i.e. non-distinctnessness), as will be claimed below. Consider again the structure which motivated the preceding section:

(38) (= (21)*Julia tiene más libros aquí que tiene Marta en casa
Suppose the compared constituent in (38) does not have the structure of the one in (37), i.e. does not contain a + wh feature, as we are assuming is the case with the one in (36). In that case it may be possible for a deletion rule to erase it, creating a surface structure indistinguishable linearly from (38), without violating the principle of recoverability of deletion. In other words, if there are deletion rules, we have not excluded (38), but only a derivation of it from one out of two possible underlying structures.

At this point I will resort to stipulation, or rather to a descriptive generalization that seems to me exceptionless in languages like English and Spanish (but is prima facie invalid in, e.g., Japanese). Let us formulate it as (39):

(39) No deletion rule can change a string XYZ into X[e]Z,  
X, Z constant, where e = [+ N]_max in an A-position

For convenience, I will refer to (39) as the empty category formation restriction. Some version of it is clearly implicit in much contemporary linguistic research. Take, for instance, a sentence like (40):

(40) They know that they are late

Suppose we could delete the second instance of they under non-distinctness with the first. This would create an empty category as in (41):

(41) They know that e are late

This would circumvent the fact that English is not a Null Subject Language. Assuming that (41) is impossible because an e.c. in the position of e must be a subject pro, and English disallows them, implies that the e.c. in question cannot be the output of a deletion rule. The implication of this analysis is, in other words, that English obeys an e.c. formation restriction.

By formulating (39) we claim, of course, that all relative constructions leaving a gap arise through wh movement and that there are no rules of Complement Object Deletion (Lasnik and Fieno, 1974) and Comparative Deletion (Bresnan, 1974), in short that the program of Chomsky (1977) can be carried out successfully. We have seen that, as far as Spanish comparatives are concerned, this program does indeed make the right predictions.

It is sometimes claimed that deletions and interpretive operations on empty nodes are merely notational variants of one another. Note, however, that by virtue of their position in the grammar deletion operations could violate certain restrictions on the output of modules preceding them, whereas empty node
interpretation could not. If, as in Chomsky and Lasnik (1977),
deletions take place in the PF module, then they can produce
apparent violations of the ECP. Thus, recall example (26):

(42) (= (26)) ¿A quién le doy el libro y a quién le el caramelo?

If this were the output of an interpretive rule on [t, e] it
would be excluded as an ECP violation, since the trace of
a quién would not be governed (at S-structure, as with (20).)
One could argue that the indirect object trace in (42) might
not be governed by the verb, but this dubious move would not
help; similar configurations involve direct objects:

(43) a. Me preguntó qué debía darle a Julia y qué le a Marta
       S/he asked me what s/he should give to J. and what to M.

       b. Me preguntó qué comería el niño y qué el gato
       S/he asked me what the child would eat and what the cat

We are thus forced to attribute (42) and (43) to a deletion of
the verb. Equivalent English examples are given by Neijt
(1979: pp. 25, 164, 111 and 185).

Consider in this connection yet another class of Spanish wh
comparatives:

(44) Ha bebido más vino del que yo creía
       S/he has drunk more wine than I thought

In (44) the wh phrase does not come from a clause adjacent to the
matrix, but from a third one which is entirely absent from the
surface. There cannot be movement out of this clause if the verb,
together with all the other categories in it, is absent in
S-structure, since the trace would not be governed. As before,
we must attribute the absence of the most deeply embedded S'
in (44) to a deletion. (The obligatoryness of wh movement in
(44) is again due to the fact that a non-movement structure
would be ruled ungrammatical before reaching PF as a violation
of the binding theory.)

Let us now recapitulate what has been said here about the
non-movement cases. The relevant variables are the verb and the
compared constituent - an NP in our examples. Furthermore, the
compared constituent might conceivably be marked + wh or not.
The following alternatives are then possible:
(45) a. *Lexical V Empty NP NP is - wh 
b. *Lexical V Empty NP NP is + wh 
c. Empty V Empty NP NP is - wh 
d. *Empty V Empty NP NP is + wh

Case a is excluded by the e.c. formation restriction. Cases b and d violate principle A of the binding theory. Case d may also violate the ECP. Thus only case c gives a grammatical output. In non-movement comparatives, if the compared constituent is absent the verb must also be.

Movement comparatives, in principle, follow the opposite pattern: the ECP requires that the verb be present, whereas above it had to be absent. In (44) this regularity is masked by a later deletion rule. This brings us to the last issue that will be considered here. We still do not know why a verb deletion rule cannot apply to the well-formed (19) and form (20) out of it:

(46) a. (= 19) Julia tiene más libros aquí de los que tiene Marta en casa

b. (= 20) *Julia tiene más libros aquí de los que Marta en casa

I would exclude this particular derivation of the starred example by invoking a subcase of the general parallelism requirement under which deletion rules are known to operate. Specifically, in a structure S'₁ X S'₂, where S'₂ has a phrase in COMP but S'₁ does not, it seems that no rule can delete material internal to S'₂ under non-distinctness with material in S'₁. This covers many cases which Koster (1978) excludes by means of a filter prohibiting empty verbs after a lexical complementizer:

(47) (Koster's (220)) [S, C [S ... [V e] ...]]

(47), or its translation into a system with deletion as opposed to only interpretation, is too strong, since it would exclude (42) above. My own contention is simply that, in the unmarked case, the clause containing the deletion triggers must be symmetrical to the target clause, in some sense. Notice, incidentally, that many properties of gapped structures which are sometimes attributed to their obeying island constraints may simply follow from parallelism requirements. In any case, we can and must require that the two clauses be parallel in a stronger sense than is implied above; in particular, the fact that the lefthand clause contains a phrase in COMP does not suffice to license a deletion in the righthand one: the phrases in question must have the exact same source within their respective S(Fiengo, 1974; George, 1980). This condition is met
by (42), but not by (48b) below; (48a) is excluded by the weaker condition formulated first:

(48) a. *Julia knows that books should be lent to friends and which to Sue

b. *Julia knows who books should be lent to and which to Sue

Parallel examples can be given for Spanish.

6. Conclusion. I have assumed in this paper a model of grammar where rule optionality is maximized. This optionality extends to lexical insertion. Therefore, when a phenomenon P has been perceived to correlate with the absence of a certain element C, it has been necessary to argue that P followed both when C was an empty category at S-structure and when C was deleted in the phonological component.

Much of the complexity of the preceding sections arises from this fact. In particular, since deletions have been relatively neglected in recent syntactic research, the conditions proposed to restrict them are highly tentative; I would expect them to be derivable from more explanatory principles. In any case, I have argued that deletions of the type considered here -deletions under non-distinctness- must not be excluded from universal grammar (cf. Neijt, 1979, Sjoblom, 1980, for discussion of general properties of deletion rules, and Emonds, 1976, for the proposal that deletions cannot be structure-preserving, in his sense, which overlaps with the e.c. formation restriction.)

The complexity introduced by my discussion of deletions should not obscure the fact that, at pre-PF levels, the distribution of Spanish comparatives follows directly from the ECP, for the core aspects, and subsidiarily from the binding theory. If it didn't, I am sure that we speakers of Spanish would have found a simpler way of comparing things.

APPENDIX: ECP violations at LF

In the oral presentation of part of this material I speculated on the possibility that Spanish might not allow the ECP to operate at LF. This was rooted in a fallacy, since all that one can conclude from the data discussed is that it applies at S-structure. I did mention that this correlated perhaps with other apparent ECP violations, such as the following (noted also in Contreras, 1982):

(49) No quiero que nadie me interrumpa
The crucial fact here is that, contrary to what is predicted, *caeteris paribus*, in Rizzi (1982), Jaeggli (1980), (49) means exactly the same thing as its inverted subject counterpart:

(50) No quiero que me interrumpa nadie

(50), as expected, means 'For no x, I want that x interrupt me', but (49) is supposed to have this meaning blocked at LF by the ECP.

Pending a more detailed investigation of the data, to be carried out over the next five years or by someone else, some sort of treatment of these facts, however experimental, is owed to the participants in the conference. Suppose we keep the ECP also at LF in Spanish, as will be necessary if the ECP is not parametrizable. Then the long movement interpretation of (49) may arise as follows. Spanish systematically allows for the fronting of constituents. Suppose this freedom extends to subjects. The subject *nadie* in (49), if fronted, will properly govern its trace; extraction at LF may then apply to the fronted subject. This derivation would be parallel to the one in overt cases of long distance movement as in (51):

(51) a. Con esos amigos no quiero que vengas a mi casa
With those friends I do not want that you come to my house
'I don't want you to come to my place with those friends'
b. Para ver la televisión no quiero que vengas a mi casa
To watch TV I do not want that you come to my house

NOTES

1 Much of the confusion in the literature arises from the fact that grammarians do not distinguish between cases where the compared constituent is a free relative with *wh* movement -(i) below- and cases where *wh* movement is required by the comparative construction itself -(ii)-:

(i) Julia tiene más libros que [NP los que la critican]
J. has more books than those-who criticize her
(ii) Julia tiene más libros de [S, los que tenía]
J. has more books than she had

The *que* comparative in (i) is of course the same we find before simple NP's:
(iii) Julia tiene más libros que Marta

The surface similarity between (i) and (ii) has led many scholars to conclude that de and que relative complements are interchangeable. I deal later in the text with the properties of both classes.

2 Strictly speaking, what the argument shows is that, if the ECP applied only at LF, it would have to apply before the verb is copied. In current treatments of the role of the ECP at LF, this principle follows all the standard LF rules (R. May's Quantifier Rule, LF wh movement, etc.), in effect acting as a well-formedness condition on their output. The conclusion drawn in the text is therefore not warranted.

The ECP must follow copying processes if these are allowed to include the partial filling of layered traces for which the term 'reconstruction' is reserved in van Riemsdijk and Williams (1981). An argument analogous to the one here has been proposed independently by Huang (1982). Huang's argument, however, depends on reconstruction rules, whose status is somewhat more problematic than that of mere copying.

3 As for 'relative pronouns', it is well-known that in many languages, including Spanish, they are very similar to question words, a fact which no doubt reflects their common wh properties. But they are not identical to them: in particular, they are systematically stressless, the question word cuál appears as el cual in relatives, and there is no interrogative counterpart to the genitive relative pronoun cuyo. These differences presumably correlate with their different anaphoric status, if the suggestion made here is correct.

4 After writing this paper I have read Plann's (1982) very interesting treatment of this kind of cases. Plann starts from the fact that many of these 'comparative ellipsis' constructions do not exhibit an agreeing wh word such as el que (masculine singular like vino) in (44). Instead the 'neuter' lo que is used also in NP comparisons:

(i) Ha bebido más vino de lo que yo creía

Based on this fact, she claims that there is no deletion of the most deeply embedded S', but merely wh movement of the object of verbs like creía in the second S'. If this is right, then (44) must be ruled ungrammatical, and hence provides no argument for deletion rules. On the other hand, this makes the pattern entirely symmetrical: there is a direct and exceptionless correlation between the presence of a verb in the complement and wh movement of the entire compared constituent.
REFERENCES

Bresnan, J. (1975) 'Comparative deletion and constraints on transformations' Linguistic Analysis 1.25-74


and H. Lasnik (1977) 'Filters and control' Linguistic Inquiry 8.425-504

Contreras, H. (1982) 'Multiple questions in English and Spanish' Paper given at the 12th Linguistic Symposium on Romance Languages, Pennsylvania State University


Spaulding, R.K. (1962) 'De (1, la, lo, los, las) que vs. que (el, la, lo, los, las) que or the force of tradition' Hispania 45.309-314


Williams, E. (1977) 'Discourse and logical form' Linguistic Inquiry 8.101-139

Notes added in press

1/ The phenomenon exemplified by (49) does not occur in Catalan (C. Picallo, unpublished work). In this language, but not in Spanish, the negative NP subject must c-command a preverbal particle no within the same clause. Hence, as expected, the alternative derivation suggested for the Spanish version of (49) cannot yield in Catalan a grammatical output.

2/ Jessie E. Pinkham's important The formation of comparative clauses in French and English (distributed by the Indiana University Linguistics Club, 1982) was not available to me before the completion of this paper.

3/ The reservations expressed in the text with respect to the applicability of (39) to languages like Japanese turn out to be unjustified, given Huang's convincing recent analysis of zero pronouns in these languages ('On the distribution of zero pronouns', unpublished, University of Hawai.)
On anaphora and control in Russian

Gilbert Rappaport

The University of Texas at Austin

0. Introduction

One of the goals of current linguistic research has been to provide a Sentence Grammar theory of coreference. Such a theory would isolate the factors which determine when two Noun Phrases (NPs) in the same sentence must, can, or cannot refer to the same entity. In the program for a theory of language elaborated in Chomsky 1981, there are two sets of principles regulating coreference. First, Binding Theory provides a set of category definitions which classify NPs in accordance with their general requirements for an antecedent. These definitions are accompanied by a set of binding conditions which specify the precise range of antecedents permitted or excluded for NPs of each category. Second, a rule of control can select the antecedent(s) of a given NP. The focus of this paper is the nature of control and its interaction with Binding Theory.

The discussion here will proceed on two planes: descriptive and theoretical. On the descriptive plane, the interpretation of reciprocal and reflexive pronouns in Russian will be examined. It will be shown that while the coreference properties of the
reciprocal follow directly from Binding Theory, those of the reflexive pronouns do not. On the theoretical plane, it will be argued that the facts of the reflexive pronouns are best handled not by modifying Binding Theory, but by elaborating the notion of control. While control rules are usually considered only in connection with PRO (the assumed null subject of a nonfinite clause), it will be proposed here that this restriction is not appropriate. An explicit statement of the control principle for the Russian reflexives will be proposed. It will be suggested that Binding Theory represents the unmarked case, with control taking precedence when appropriate. These results will be generalized by considering parallel phenomena in certain other languages.

This paper is organized as follows. A brief review of Binding Theory (following Chomsky 1981) is provided in section 1. Section 2 is an introduction to the reciprocal and reflexive pronouns in Russian. The two types of pronoun are then taken up in detail, the reciprocal in section 3, and the reflexive in section 4. A summary and discussion of the results is found in section 5.

1. Binding Theory

1.1 Three categories of NP can be defined in terms of the requirement for an antecedent. An R-expression (‘referring expression) is an NP which is potentially referential, and therefore does not require an antecedent of any sort. An anaphor is an NP which can have no reference independent of the sentence containing it (no 'inherent' reference); that is, it must have an antecedent in the same sentence. More precisely, there must be another NP which a) has the same reference (extension) as the anaphor, and b) either is syntactically expressed in the same sentence as the anaphor, or is understood (via ellipsis) to function in that sentence. A third category, that of pronominals, lies, in some sense, between that of R-expressions and that of anaphors. A pronominal is not referential in the sense that an R-expression is, but neither is it restricted in the way that an anaphor is. A pronominal can a) have an NP antecedent (expressed or understood) in the same sentence, b) have an NP antecedent in another sentence of the discourse, or c) denote an entity fixed by some non-linguistic convention between the participants of the speech event ('deictic' usage).

The coreference properties of these three categories of NP are given by the following three binding conditions:

1. Binding conditions
   a. An anaphor is bound in its binding category.
   b. A pronominal is free in its binding category.
   c. An R-expression is free everywhere.
Chomsky (1981:220) defines a binding category as follows: 'B is a binding category for A if and only if B is the minimal category containing A and a SUBJECT accessible to A'. The terms 'SUBJECT' and 'accessible' are independently defined. The essence of this definition, phrased more transparently if less elegantly, is given by (2):

(2) The binding category of an NP A is the minimal S or NP containing A which also contains one of the following:
   a. a subject accessible to A;
   b. subject-verb agreement (AGR) which agrees with A.

The term 'subject' refers either to the grammatical subject of a sentence (including PRO), or to the overt lexical Specifier of an NP, as in 'my dog', or 'John's picture'. A subject is accessible to an anaphor if this subject is distinct from the anaphor and does not contain it. 'Bound' is defined as follows:

(3) An NP A is bound if and only if there is an NP B such that all of the following conditions are satisfied:
   a. A and B have the same reference;
   b. B c-commands A;
   c. B is in an argument position.

Finally, 'free' means 'not bound'.

1.2 Binding Theory is developed as a universal theory. This does not mean, of course, that the facts of any language will automatically fall out just as in English, even at the level of abstraction at which the binding conditions are defined. Superficial differences can be accounted for in several ways. There are marked phenomena, constructions which simply have to be learned. There is 'parametric variation', a difference in some abstract principle between languages which must be learned. The parameter approach is particularly powerful when several superficial differences can be traced to a single parametric option. It is also conceivable that the components of the theory interact in such a way that similar facts in two different languages follow from distinct principles. These possibilities must be explored before Binding Theory is summarily rejected. It is in this spirit that we now turn to some problems of binding in Russian.

2. **Reciprocal and reflexive pronouns in Russian**

The reciprocal pronoun in Russian, corresponding to English 'each other', is *drug druga*. This pronoun never occurs in the nominative case, but it takes the remaining five cases normally. The accusative is taken as the citation form. An example is given in (4):
(4) Roditeli dumali, <chto deti ljubjat drug drugu>. 3
   'The parents[i] thought <that the children[j] loved each other[*i,j]>.'

In this example, drug drugu behaves like the English anaphor each other. The embedded clause is the binding category of the Russian reciprocal, because it contains a subject (deti 'children') accessible to the reciprocal. (Here and below, angle brackets are used to mark the binding category of an indicated NP.) As indicated by the referential indices in the English gloss, coreference is impossible with a c-commanding NP outside the binding category. That is, (4) has only the following interpretation: the parents thought that each child loves the other children. There is no logical reason why (4) could not have a completely different interpretation: each parent thought that the children love the other parents. But this interpretation, which would be possible only if the reciprocal were bound with a c-commanding NP outside the binding category, is excluded. If we assume that the Russian reciprocal is an anaphor, whose reference is constrained by binding principle (1a), these facts follow directly.

Russian has two reflexive pronouns. The personal reflexive pronoun, like the reciprocal, never occurs in the nominative; the accusative form sebja is used as its citation form. An example is given in (5):

(5) Vanja znaet, <chto Volodja ochen' ljubit sebja>.
   'Vanja[i] knows <that Volodja[j] loves self[*i,j] very much>.'

The reflexive in (5), like the reciprocal in (4), must be bound in its binding category: sebja can refer only to Volodja. Unlike English, Russian also has a possessive reflexive pronoun: svoj. This pronominal modifier has the complete six case, three gender paradigm of modifiers; for example:

(6) Vanja slyshal, <kak Volodja rasskazyval o [svoej zhizni]>.
   'Vanja[i] heard <how Volodja[j] was talking about self's[*i,j] life>.'

Recall that according to the definition of an accessible subject, a NP is not accessible to itself. As a result, the Specifier svoj does not make the NP containing it its binding category. Rather, the binding category is the S above that NP, which is given in angle brackets; Volodja, then, is an antecedent for svoj. Reference outside this S, to Vanja, is excluded. The English reflexive cannot appear in Specifier position, which makes comparison with (6) difficult. The English reciprocal each other, on the other hand, can appear in the Specifier position. When it does, its coreference pattern is like that of svoj. For example,
(7a) and (7b) are parallel, with coreference permitted between the underlined constituents:

(7) 
   a. *Eti muzhchiny prodavali [svoi tovary].
      'These men were selling [self's wares].'
   b. *The men were selling [each other's wares].

In these constructions, then, the coreference of the possessive reflexive svoj is predicted by Binding Theory.

The Russian reciprocal and reflexive pronouns behave like the anaphors of Binding Theory in the simple constructions discussed so far. However, a more careful investigation of the facts reveals that there are some differences between the binding of these Russian pronouns and that of their English counterparts. First, it will be confirmed that the reciprocal pronoun is an anaphor in Binding Theory. Deviations from the theory as stated can be accounted for by making a small adjustment in the definition in the binding category (probably best viewed as a case of parametric variation). In contrast, it will be shown that the reflexive pronouns deviate from the theory more radically.

3. The reciprocal pronoun

The binding of the reciprocal pronoun is virtually identical in Russian and English. First consider cases in which the reciprocal pronoun is or is contained in the complement of a verb. Example (4) illustrated how a finite clause, containing an accessible subject and AGR, represents the binding category for a reciprocal in the Verb Phrase. If a clause is nonfinite, its PRO subject can serve as an accessible subject for an anaphor in the Verb Phrase just as well. In (8), PRO is the only possible antecedent of the reciprocal:

(8) My poprosili ix <PRO nalit' drugu drugu chajku>.
    'We[i] asked them[j] <PRO[j] to pour each other[*i,j] tea>.'

   Generaly prikazali soldatom <PRO ohranjat' druga druga>.
   'The generals[i] ordered the soldiers[j]
    <PRO[j] to protect each other[*i,j]>.'

Within the binding category, any c-commanding NP can be the antecedent. In examples (4) and (8), the antecedent was a subject, but this need not be the case. In (9), a direct object is the antecedent of the reciprocal:

(9) Neperyvnyaja stena vagonov tjanulas' mezhdu obeimi
    parami, i <ona zakryvala ix druga ot druga>.
    'A continuous wall of railroad cars stretched out
    between the two pairs, and it blocked them
from each other."

The c-command restriction is waived for certain prepositional phrase arguments of the verb, as in

(10) On ne smotrel na nix otdel'nno drug ot druga.
    'He did not look at them separately from each other.'

Otherwise, this restriction is in force. Consider, for example:

(11) Prepodavateli studentov znajut druga drugu.
    'The teachers[i] of the students[j] know each other[i,*j].'

(11) has the following interpretation: each of the teachers knows the other teachers. This sentence cannot express the perfectly coherent idea that for each student, his teacher knows all the other students.

Now consider cases in which the reciprocal is the complement of a noun. As in English, we find the importance of an accessible subject. Contrast (12a), in which the NP containing the anaphor has no overt Specifier, with (12b), in which the same NP has a Specifier:

(12)

a. <Oni chitali [zhaloby drug na druga].
   '<They[i] were reading [complaints against
each other[i]].'>

b. *Oni chitali <moi zhaloby drug na druga>.
   ('They[i] were reading <my[j] complaints against
   each other[i,*j]>.'

In (12a), there is no subject within the square-bracketed object NP, so that the matrix clause is the binding category. As a result, oni 'they' is eligible to be an antecedent for the reciprocal. In (12b), the object NP contains a subject accessible to the reciprocal: moi 'my'. The object NP, then, is the binding category of the reciprocal. Oni 'they', standing outside this binding category, is not an eligible antecedent for the reciprocal. Binding Theory permits moi 'my' to be an antecedent, but it is disqualified because it is singular. As a result, (12b) is ungrammatical.

Now consider cases in which the reciprocal either is the subject of a clause or is internal to the subject. When the reciprocal is the subject of a finite clause, there can be no lexical phrase in its binding category which c-commands it. Such a construction is ungrammatical:

(13) *My znali, <chto drug drug otsutsvovali na sobranii>.
    ('We knew <that each other were absent at the
meeting.

There are two cases in English of binding extending upward out of a clause. First, binding can extend upward from the subject of a nonfinite clause, in constructions such as the following:

(14)

a. <The men expect [each other to succeed]>.
   b. <The men would prefer [for each other to succeed]>.

Neither of these constructions is possible in Russian for reasons independent of binding. Second, anaphor binding in English can extend upward from inside the subject of a finite clause complement. This binding pattern is not possible in Russian:

(15)

a. <The dissidents knew [that articles about each other had appeared in the Western press]>.
   b. *Dissidenty znali, <chto stat' i drug o drug pojavils' v zapadnoj presse>.

The English construction (15a) is permitted by the definition of a binding category given in (2): AGR suffices to define the binding category only if it agrees with the subject. The ungrammaticality of the corresponding Russian construction (15b) follows from simply removing this agreement condition. The embedded clause then becomes the binding category (as indicated in 15b by angle brackets), and the binding condition for anaphors cannot be satisfied: there is no NP in the binding category which c-commands the anaphor.

The binding facts for Russian reciprocal pronouns, then, follow from the same principles as those which determine the binding of English anaphors, except that one condition in the definition of a binding category must be relaxed. This difference is probably a matter of parametric variation among languages. The binding category of an anaphor in Russian is defined, then, as follows:

(16) In Russian, the binding category of an anaphor \( A \) is the minimal S or NP containing \( A \) which also contains one of the following:

a. a subject accessible to \( A \);
   b. AGR.

In the following section, it will be shown that the coreference properties of reflexive pronouns are governed by rather different principles.
4. The reflexive pronoun

4.0 In this section, we consider the two reflexive pronouns in Russian, sebja and svoj. The coreference properties of the Russian reflexives can differ from those of the Russian reciprocal in several ways. These differences can be divided into three areas: syntactic conditions on binding, the possibility of arbitrary reference, and sensitivity to the semantic roles of NPs. These three areas will now be considered in turn. The purpose of this section is to describe the binding properties of the Russian reflexives; an explanation for these facts will be proposed in section 5. Although there are some differences between sebja and svoj (see the references in footnote 4), these differences affect the present discussion only when it is so stated.

4.1 Syntactic conditions on binding

Examples (5) and (6) from section 2 demonstrate that a Russian reflexive must be bound within its binding category, like the reciprocal. Also, binding from inside the subject to a higher clause is not possible for the reflexives, just as it is not possible for the reciprocal (cf. 15b):

(17) *Vanja znaet, <chto stat'ja o sebe pojavilas' v gazete>.  

('Vanja knows <that an article about self appeared in the paper.>')

*Vanja znaet, <chto stat'ja o svoej zhene pojavilas' v gazete>.  

('Vanja knows <that an article about self's wife appeared in the paper.>')</n
As we look more carefully, however, we find that the syntactic conditions on the coreference of reflexive pronouns are both more and less restricted than those on the interpretation of the reciprocal.

Reflexive pronouns are more restricted than the reciprocal in that, typically, the antecedent must be a grammatical subject (but see section 4.3). To see this, consider (18):

(18) 


While the English reflexive himself in (18a) can refer either to Vanja or to Volodja, the Russian reflexive svoj in the corresponding sentence (18b) can refer only to Vanja, denoted by the subject of the sentence. Similarly, the possessive reflexive in (19) refers unambiguously to the entity denoted by the subject, Sasha:
(19) Sasha beret im svoj xleb.
'Sasha[i] is taking to them[j] self's[i,*j] bread.'

For the time being, we will refer to this property as the subject control of the Russian reflexive. In section 5, we will motivate this use of the term 'control'.

We now turn to how Russian reflexives are less restricted than reciprocals. First, the null PRO subject of a tenseless clause is not sufficient to demarcate the binding category. That is, PRO is transparent to reflexive binding; e.g.,

(20)

a. <On ne razreshaet mne [PRO proizvodit' opyty nad soboj]>.
   (cited by Rozental' 1974)
   'He[i] does not allow me[j] [PRO{j} to perform experiments on self[i,j]]>.'

b. <Professor poprosil assistenta [PRO chitat' svoj doklad]>.
   (cited by Rozental' 1974)
   'The professor[i] asked his assistant[j]
   [PRO{j} to read self's[i,j] report]>.'

As indicated in the glosses, the reference of the reflexive in each of the examples in (20) is ambiguous.

Second, NP subjects are transparent as well. For example:

(21)

   'I[i] read [his[j] article about self[i,j]]>.'

   'I[i] read [his[j] article about self's[i,j] work]>.'

Again, the reference of the reflexive is ambiguous.

The reflexive facts discussed so far point to a simple generalization: a Russian reflexive pronoun is bound to any subject in the minimal finite clause containing the reflexive. The subjectionhood requirement will be investigated further below. What is important here is the domain within which the subject must be selected. Unlike the case of the reciprocal, both PRO and NP subjects are transparent. The Russian reflexives, then, need not be bound within the binding category defined by (2). This fact in itself might merely suggest that Russian defines the binding category differently from English. However, the relevance of the definition (2) for Russian is confirmed by the reciprocal pronoun, which behaves essentially like the English anaphors. Altering the definition of the binding category, then, would leave the reciprocals unaccounted for. We are left with the problem of accounting for why the reciprocal and reflexive pronouns in
Russian seem to define their binding category differently.

4.2 Arbitrary reference

When an anaphor is the subject of a finite clause, there can be no NP in its binding category which c-commands it. The binding condition for anaphors cannot be satisfied, and the result is ungrammaticality (e.g., (23)). The same problem arises when an anaphor is the complement of the subject noun; in this case, however, the c-command restriction on binding can be waived. This is observed for both the English anaphors (22a) and Russian reciprocal (22b):

(22)  
a. [The rumors about each other] annoyed them.  
[The rumors about himself] annoyed John.

b. [Spietni drug o druge] razdrazhali ix.  
'[The rumors about each other] annoyed them.'

The same problem arises when the anaphor is the Specifier of the subject of a finite clause. When an English anaphor, or the Russian reciprocal, is in this position, the result is ungrammaticality, just as when the anaphor is the subject itself:

(23)  
a. *[Each other's habits] irritated them.

b. *[Privychki drug druga] razdrazhali ix.

The Russian reflexive svoj, on the other hand, can often be assigned an interpretation in this position. In some cases, it is interpreted by a form of non-subject control discussed in the next section. If that fails, svoj can be assigned arbitrary reference. There are two forms of arbitrary reference. In (24), the antecedent is nonreferential, a form of universal quantification: the predication is true for any possessor:

(24) V xokkee, kak i v drugijx vidax sporta,  
<svoi steny igrajut znachitel'nuju rol'>.  
'In hockey, as in other sports, <one's own
walls play a significant role>.'

In (25), a referential antecedent is determined on the basis of the discourse context:

(25) On slushaet i otvechaet, i sam govorit, no <gde-to
vnutri idet svoja zhizn', svoi mysli>. (Kuznecov)
'He listens and answers, and speaks himself, but <somewhere
inside is going on his own life, his own thoughts>.'

These two forms of arbitrary reference are also possible when the reflexive is in a predicate nominal:
(26) U drugix, možhet, deti i poluchshe, i poumnee, a svoj vse ravno vsex dorozhe, potomu chto \((\text{on})\text{-svoj}\).
'Others, perhaps, have better and smarter children, but one’s own is nevertheless more dear, because it is one’s own.'

(27) \(<\text{Numeracija primerov v kazhdom paragrafe svoj}\>.
'\(<\text{The numbering of examples in each section is mine}\> [\text{literally, "self's"]}\>.'

Conversely, the arbitrary reference of \(\text{svoj}\) is restricted to these two positions (lexicalized expressions aside).

In permitting arbitrary reference, the Russian reflexives differ significantly from both the English anaphors and the Russian reciprocal. The range of attested examples excludes the possibility that arbitrary reference is only possible for lexicalized expressions. Assuming lexicalization in such cases would also not account for the fact that arbitrary reference is possible only in a clearly defined syntactic environment. The fact that only \(\text{svoj}\) (and not \(\text{sebja}\)) exhibits arbitrary reference follows from the fact that only \(\text{svoj}\) can appear in the required environment. There is no reason to see arbitrary reference as a property in some way peculiar to \(\text{svoj}\).

4.3 **The semantic roles of NPs: Theta control**

It will now be shown that the binding of Russian reflexive pronouns is sensitive to the semantic roles of NPs. There is no need to commit oneself to any particular theory of semantic analysis to see that this is true. The importance of semantic roles can be demonstrated for a given example by showing that a) binding cannot be attributed to the principle of control by the grammatical subject, and b) binding would not be possible to an NP bearing the same grammatical relation, but a different semantic role. We will now discuss several particular semantic roles which permit a non-subject NP to be the antecedent of a reflexive pronoun.

Russian has a group of predicates which take a dative case argument and express a range of modal, experiential, and quantificational predications. We will (arbitrarily) call the semantic role of the dative argument of such a predicate an \(\text{Experiencer}\). A dative \(\text{Experiencer}\) is not a surface subject \((\text{Perlmutter 1978})\). Nevertheless, it is a possible antecedent for a reflexive pronoun:

(28) Do slez zhalko mne stalo i \(\text{sebja}\) i Filimonova.
\(\text{(cited by Timberlake 1980a; hereafter, AT80)}\)
'I became sorry ['it became sorry to me'] to the point of tears (for) both \(\text{self}\) and Filimonov.'
Djia sebja, ej nihago ne nuzhno. (AT 80)  
'She doesn't need anything for herself  
["for self, to her nothing is necessary"].'  

Such examples contrast with a dative Goal, traditionally called an  
indirect object. The latter is not a possible antecedent for a  
reflexive:

(29) *Pered uxdom zadaju emu neskol'ko voprosov  
o samom sebe. (AT 80)  
('Before leaving I ask him a few questions about self.')  

Russian also has oblique Possessors, expressed by a normally  
locative prepositional phrase (PrepP). An oblique possessor is a  
possible antecedent for a reflexive pronoun; for example:

(30)  
a. U nix vsjakij zapas s soboj.(AT 80)  
'They have all kinds of supplies with them  
["By them there are all kinds of supplies with self"].'  
b. V zhenskix imenax tozhe est' svoi tonkosti.  
'Female names also have their own subtleties  
["In female names there are also self's subtleties"].'  
c. Mezdu nimi ustanovilis' svoi osobye otnoshenija.(AT 80)  
'Between them self's special relations  
were established.'

These examples contrast with the same locational PrepPs used in a  
on-possessive construction; in the latter case, binding is  
impossible:

(31) *Pervoe vremja u nas chasto byvali svoi sosedj. (AT 80)  
('At first self's neighbors were often at our place  
["by us"]').  
*Ja vdel v nem svoj nedostatki. (AT 80)  
('I saw in him self's faults.')  
*Mezdu nimi stojal svoj syn.  
('Between them stood self's son.')

There are several reasons to subsume the Experiencer and Possessor  
roles under a single rubric. In the absence of a more felicitous  
term, the semantic relation in question will be called  
Experiencer/ Possessor. There is no need to pursue this point  
here, and we will simply assume that a single semantic role is  
involved in such cases.

The agent of a passive construction in Russian is a possible  
antecedent for the reflexive pronoun:
(32) V pervoj (polovine izby) varitsja pishcha dlja sebja.
'In the first (half of the hut) is cooked food for self.'
Denezhnye dela Dar'i Aleksandrov by chuvstvovalis'
Levymi kak svoi sobstvennye. (L. Tolstoj)
'Dar'ya Aleksandrovna's financial matters were felt
by the Levins' like self's.'

It does not matter whether the passive in Russian is a syntactic
movement rule or a lexical rule (cf. Babby and Brecht 1975).
What is important is that a) a non-subject is the antecedent of a
reflexive, and b) this non-subject designates the Agent.

Finally, Timberlake (1980a) has discovered that the head of
a Benefactor prepositional phrase can be the antecedent of a
reflexive pronoun. Other PrepPs sharing the syntactic properties
and distribution of this PrepP cannot contain an antecedent;
contrast:

(33)

a. Vojna vse bol'she sosredoctochevalas' dlja Pekareva
   v krugu svoix del i ob"jazannostej. (AT 80)
   'More and more for Pekarev the war was concentrated
   in the realm of self's problems and responsibilities.'

b. *Vojna vse bol'she sosredoctochevalas' pri Petre
   Velikom v krugu svoix del i ob"jazannostej.
   ('More and more war was concentrated under Peter the
   Great in the realm of self's problems and
   responsibilities.')

We have, then, identified three semantic roles for NPs which
permit a non-subject NP to be the antecedent of a reflexive
pronoun. These roles are Experiencer/Possessor, Agent, and
Benefactor. When an oblique NP with one of these semantic roles
is a possible antecedent for a reflexive, we will speak of theta
control, to distinguish such cases from subject control. Both
forms of control are sensitive to the relative prominence of a
potential antecedent: subject control involves syntactic
prominence, and theta control involves the prominence of the NP's
syntactic role.

Theta control is an interpretive process which is distinct
from subject control and applies independently of it. In some
cases, theta control takes over when subject control fails because
there is no subject (e.g., 28a) or the anaphor is subject-internal
(e.g., 30b). In other cases, the simultaneous operation of both
control rules can result in ambiguity, if selectional restrictions
and context permit:

(34) Rebenok byl отправлен Annoj Pavlovnoj k svoim sestram.
    (cited by Klenin 1974)
'The child[i] was sent by Anna Pavlovnna[j] to self's[i,j] sisters.'

Thus, we see that the binding of reflexive pronouns in Russian is determined by two independent control processes functioning in parallel. The semantic roles of NPs do not figure in Binding Theory, nor do they affect the interpretation of the reciprocal pronoun in Russian.

5. Conclusion

5.0 To summarize the descriptive portion of this paper, we have discussed three properties of Russian reflexive pronouns. First, a linguistic antecedent of a Russian reflexive must be contained in the minimal finite clause containing the reflexive. Second, under certain well-defined conditions, arbitrary reference is possible. Third, binding is sensitive to both the subjecthood and the semantic role of a potential antecedent. These properties distinguish Russian reflexives not only from English anaphoric pronouns, but from Russian reciprocals as well.

We now turn to the question of how to best account for these properties of the Russian reflexives. The brute force method would be to assume that both the reciprocal and reflexive pronouns in Russian are anaphors, and to simply require that the grammar of Russian assign them different properties. Such a course would explain nothing; indeed, it would be no more than a statement of the facts. In what follows, we propose an alternative which a) accounts for the Russian facts, b) preserves Binding Theory, and c) is naturally extended to account for the facts of reflexives in other languages, as well as PRO. The essence of the proposal is more carefully articulated concept of control.

5.1 In section 4.1, it was shown that the Russian reflexives are subject controlled; in section 4.3, this principle was supplemented by a principle of theta control. Before proceeding, it is essential that we clarify the concept of control as it is applied here.

We assume that there are essentially two processes of 'coindexing', that is, of designating one NP as the antecedent of another NP in the same sentence. One process is essentially free coindexing between NPs satisfying certain configurationally defined criteria: an NP can be coindexed with any NP which c-commands it, within a locality constraint in the form of the binding category restriction. This form of coindexing, the one described by Binding Theory, will be called free binding. Any coindexing rule which (abstracting away from a locality requirement) is more restrictive than this represents a second category of coindexing processes; we define this second category as controlled binding, or simply control.
Work in the framework of generative grammar has generally assumed that the term 'control' refers only to the choice of antecedents for PRO (Chomsky 1982:7). The definition of control just proposed applies to NPs other than PRO, including the Russian reflexive. We will now counterpose the interpretation of the Russian reflexives with the reflexives of certain other languages and PRO, developing the concept of control as we proceed.

5.2 Recall that the antecedent of a Russian reflexive pronoun need not be in the binding category of the reflexive pronoun. There is, however, a locality constraint: the antecedent must be in the minimal finite clause containing the reflexive. We suggest that a rule of control can define its own locality domain; that is, the interpretive rule of control not only specifies the syntactic or semantic role which the antecedent must play, but it can also stipulate the locality domain which must contain the antecedent. We will call this locality domain a control category, distinguishing it from the binding category defined by Binding Theory (given in 2). The control rule for the Russian reflexive defines a control category as the minimal S containing the reflexive which also contains AGR. Within this locality domain, any suitably prominent NP is a possible antecedent of the reflexive; outside of this category, no NP is a possible antecedent. Non-controlled reflexives with a binding category significantly different from that defined in (2) are rare, if they are attested at all. On the other hand, there are numerous cases of controlled reflexives with a control category distinct from the binding category. We now consider the more representative possibilities for control categories.

In the most restrictive case, the control category for reflexives is the same as the binding category; this situation is attested in, e.g., New Testament Greek (Harbert 1982) and Georgian (Harris 1981). A less restrictive control category can be defined by relaxing either of the two defining properties of the binding category. In Mandarin Chinese, the control category for reflexives is the minimal S or NP with an accessible subject, regardless of AGR (Huang 1982). More common seems to be the relaxation of the accessible subject condition. The control category is then defined as the first S (or, vacuously, NP) with AGR, regardless of an intervening subject. This case is exemplified by the reflexive control rule in Gothic (Harbert 1982) and Hindi (Kachru and Bhatia 1977), as well as in Russian. The least restrictive control category is the maximal S containing the controlled reflexive. This is attested in Japanese (Kuno 1973), Korean (Kim 1976), and Malayalam (Mohanan 1981).

It is natural to conjecture that the binding conditions apply as the unmarked case, in the absence of a marked control rule which supercedes them. The interpretation of Icelandic reflexives demonstrates this markedness relation quite clearly.
In this language, the reflexive pronouns can be bound to any c-commanding NP in the same S, but only to the subject of higher S's (within the control category: the minimal indicative S). These facts follow from assuming that the subject control rule of reflexives in Icelandic is optional (as opposed to the obligatory rule in Russian). When the rule applies, the control category defines the range of eligible subject antecedents as going up to the first indicative finite clause. When the rule does not apply, the binding category defines the range of eligible NP antecedents as being any c-commanding argument NP within the clause.

The interpretation of PRO provides additional evidence that the range of admissible antecedents depends directly upon whether or not an NP is controlled. PRO is controlled in some configurational environments, and not in others. PRO is controlled when the S containing it is a) a VP-internal argument of a verb, and b) not associated with a complementizer (Manzini 1982). When PRO is controlled, its control category is its matrix clause: the antecedent of PRO must be syntactically present (if phonologically absent by ellipsis) in this clause. When PRO is not in a position of control, it is assigned arbitrary reference, and a linguistic antecedent need not be contained in the matrix clause:

(35) They thought ⟨that I believed [that [PRO getting to know each other] would require more time]⟩.

5.2 There is evidence that the reflexive pronouns of Russian are not anaphors, despite the fact that in simple constructions their reference appears to pattern like that of an anaphor. The most direct argument to this effect is based on the possibility of arbitrary reference (section 4.2). An anaphor, by definition, must have an antecedent in the same sentence. Arbitrary reference is a clear contradiction of this definition.

A more indirect argument is provided by the theta control of the Russian reflexive. A survey of control rules in various languages points to the following correlation:

(36) The control rule of an anaphor can be sensitive only to surface syntax.

Many examples of an anaphor which is controlled on a strictly syntactic basis have been reported in the literature. Subjecthood is a necessary and (within locality constraints) sufficient condition for an NP to be the antecedent of a reflexive pronoun in c-commands in such diverse languages as Classical Latin (se, suus), Icelandic sig, sinn), Mandarin Chinese (zija), Malayalam (taan, swa-), and Georgian (tav). On the other hand, while the reflexives in Japanese (zibun) and Korean (oaki) are essentially
subject–controlled, a variety of discourse and semantic factors (topicality, empathy, point of view) affect the necessity or possibility that a subject be the antecedent of the reflexive. Correspondingly, there are constructions in which the reflexive c-commands its antecedent, rather than vice versa. As a result, the reflexive can be the subject of a clause with no higher subject at all, giving structures of the form [that self[i] was a fool] saddened John[i]. In such cases, the reflexive is free in its locality domain, like a pronomial, not like an anaphor. These facts are consistent with (36): the reflexives in Japanese and Korean are controlled, but they are not anaphors. Correspondingly, their control need not be restricted to syntactic information.

The theta control of Russian reflexives is a clear case of a control rule referring to more than surface syntax. Interestingly, Russian reflexivization control is weakly sensitive to discourse factors as well, such as the topicality of the subject (see Nichols et al. 1980). Given (36), then, the sensitivity of the control rule for Russian reflexives to non-syntactic information indicates that these reflexives are not anaphors. Our conclusion, then, is that a) either anaphors or proninals can be controlled, and b) the Russian reflexive pronouns are controlled pronouns.

The control properties of PRO are perfectly consistent with our conclusions. PRO shares the two properties pointing to the pronominal status of Russian reflexives. Arbitrary reference is possible for PRO when it is not in a position of control:

(37) John asked Bill [how PRO to behave oneself under such circumstances]. (Chomsky 1981)

Also, the control of PRO is sensitive to the semantic roles of NPs. Contrast, for example, the interpretation of PRO in the following two sentences:

(38)

a. John[i] told Bill[j] [PRO[*i,j] to go to the store].

b. John[i] promised Bill[j] [PRO[i,*j] to go to the store].

The grammatical relation of John in both (38a) and (38b) is the same (that of the subject); what differs is the matrix verbs and the semantic roles they assign to their arguments. Control is obviously determined by these semantic factors. It has been observed that PRO exhibits some properties of an anaphor, and others — of a pronominal (Chomsky 1981); the control properties of PRO fall into the latter class.

Above, it was shown that the control category specified in a rule of control is marked, in the sense that it overrides the
(unmarked) binding category of Binding Theory. Similarly, the rule of controlled binding is marked in that it overrides the binding conditions themselves. This clearly has to be true in the case of a controlled anaphor: if rules of controlled and free binding were both permitted to operate, the former rule would be vacuous, never assigning an antecedent not also assigned by the latter rule. Assuming such a markedness relation is necessary for a pronominal control rule as well, explaining why a controlled pronominal such as a Russian reflexive need not be free in its binding category. This assumption also accounts for a problem mentioned by Chomsky (1981:221): if PRO is a pronominal, then it should not be permitted to have an antecedent in its matrix clause, as in:

(39) <John tried [PRO to win]>.

The matrix clause is the binding category of PRO, and, by binding condition (1b), PRO should be free in this category. If we assume that the binding conditions are inoperative in such constructions because PRO is controlled, then the problem disappears.

5.3 Our conclusions, then, may be summarized as follows:

(40)

a. Control is possible for any pronoun, that is, for anaphors and pronominals.
b. A control rule can specify its own locality domain, a control category.
c. The binding conditions and binding category defined by Binding Theory represent the unmarked case, that of free binding; they are overridden by the binding conditions and locality domain specified by a language-particular rule of control.
d. A control rule for anaphors can refer only to surface syntax; a control rule for pronominals is not restricted in this way.

Finally, assuming (40), the control rule for the Russian reflexive pronouns can be stated as follows:

(41) A pronominal B lexically marked as [+reflexive] is bound to an NP A if and only if both of the following conditions (a) and (b) are satisfied:
   a) binding condition: A is 'prominent', either syntactically or semantically; that is, A either
      i) is a grammatical subject; or
      ii) is assigned one of the following semantic roles by its governing verb or preposition:
      Experiencer/Possessor, Agent, or Benefactor.
   b) locality condition: A is contained in the minimal category also containing both AGR and B.
FOOTNOTES

1 This is a condensed version of a longer manuscript. I would like to thank Lee Baker, Sandra Chung, Samuel Gutmann, and Carlota Smith for their comments on earlier drafts of that paper. Portions of this work were presented in a lecture at the University of Texas at Austin, as well as at the Cornell Government-Binding Conference. The research reported here has been partially supported by the Research Institute of the University of Texas at Austin.

1 The discussion in this section is restricted to non-empty NP categories: PRO and trace are not considered.

2 We assume the following definition of c-command (there are others): B c-commands A if and only if A is dominated by the first branching node that dominates B, B is not identical to A, and neither A nor B dominates the other.

3 For typographical convenience, we deviate from the traditional transcription of Cyrillic letters in the following correspondences: щ -> sh; ћ -> zh; ё -> ch; є -> shch.


5 Actually, the range of possible antecedents can be more restricted when the reflexive is in a complex Adjective or Adverb Phrase. We leave these constructions aside here. See Yokoyama 1980 for some discussion.

6 Arbitrary reference cannot be assigned in all constructions fitting the syntactic conditions described here. On the other hand, examples such as those cited are too common and varied to be attributed to lexicalization. The precise conditions under which arbitrary reference is possible remain unclear.

7 To my knowledge, the first observation that the binding of reflexives in Russian is sensitive to both semantic roles and subjecthood belongs to Klenin (1974), who discussed agents. More recently, Timberlake has investigated this question in considerable detail (1980a, 1980c).

8 On how possession is expressed in Russian, see especially Chvany 1975, Issatschenko 1974, and the references cited there.

9 See Timberlake 1980a for detailed arguments against subsuming theta control under subject control.
The principle of theta control is discussed in greater detail in the full version of this paper.

On the Icelandic facts and various analyses, see, e.g., Thráinsson 1976, Maling 1981, and Anderson 1982.

For discussion and further references, see, e.g., Milner 1978 on Latin, the references in note 11 on Icelandic, Huang 1982 on Mandarin, Mohanan 1981 on Malayalam, and Harris 1981 on Georgian. The control category for reflexives in these languages varies.


REFERENCES


Manzini, Maria Rita. 1982. On control and control theory. Unpublished paper, MIT.


Nichols, Johanna; Gilbert Rappaport; and Alan Timberlake. 1980. Subject, topic, and control in Russian. Proceedings of Sixth Annual meeting of the Berkeley Linguistics Society, ed. by Bruce Caron et al., 372-86. University of California, Berkeley.


Timberlake, Alan. 1980c. Objects as controllers (Russian reflexivization). Unpublished paper, UCLA.

FREE RELATIVES AND THE pro-HEAD HYPOTHESIS

Margarita Suñer
Cornell University

It is the purpose of this paper to examine tensed and infinitival free relatives (FRs) in Spanish. Tensed FRs observe the matching condition whenever they appear in subcategorized positions; infinitival FRs do not. This difference in behavior is correlated with the [+ tense] character of the embedded clause part of the relative clause structures. It is theorized that this correlation is due to the fact that in relative clauses INFL is accessible. The general hypothesis developed in this study is that the head position of FRs is occupied by the empty category pro. Pro is the [+ pronominal, -anaphoric] element which was posited from the evidence provided by null subjects in pro-drop languages (cf. Suñer 1982a; Chomsky 1982). The pro-head hypothesis implies that the only difference between headed and headless relative clauses in outer structure resides in the [+ lexical content] of its head. The matching condition obeyed by tensed FRs in subcategorized positions is seen as a stratagem used to "determine" pro (cf. Chomsky 1982); the determining factor is provided by Case Theory.

This paper is structured as follows: After briefly presenting the two main competing analyses for the structure of FRs—the head proposal and the COMP analysis—in Section 1, Section 2 explores some of the problems faced by the COMP Accessibility (CA) hypothesis vis-à-vis the Projection Principle and the \$\theta\$-criterion. Section 3 makes explicit some of the assumptions made in this study. Section 4 discusses the matching condition and pro determination. Section 5 is devoted to infinitival FRs, since they provide strong proof against CA. Section 6 constitutes the conclusion where the principal findings are summarized. Speculations about the parametric variation with respect to matching vs. non-matching languages are also included in this section.
1. Free relatives

Free relatives are relative clauses which have no surface antecedent. Traditional grammars of Spanish state that the relative pronoun "includes" its antecedent (Ramsey 1956:196), or that there is an "implicit antecedent" (RAE 1974:220).¹

Two principal competing hypotheses have been advanced for the analysis of FRs: the Head proposal and the COMP proposal.

Bresnan and Grimshaw (1978) contains the most explicit analysis under the first hypothesis. According to this analysis, the WH-phrase is generated as head of its clause in antecedent position, and a rule different from WH-movement interprets the necessary gap. Thus, sentence (1) would be diagrammed as in (2).

(1) Busco quienes me ayuden.
I-am-looking-for who(pl) me(acc) help (3 pl) 'I am looking for people who can help me.'

(2) Under the X-theory, this hypothesis automatically explains the categorial "matching effect" between the sentential phrase and its head, since in this analysis free relatives are headed. Moreover, it accounts for the absence of a complementizer by not providing a COMP node.

By contrast, in the COMP proposal the WH-phrase is in COMP and the head remains empty.³ This alternative makes the claim that the only difference between regular relatives and FRs is the presence or absence of lexical material at the N level; their constituent structure is the same. Under this analysis, sentence (1) would be represented as follows:
Groos and van Riemsdijk (1979) reject the Head analysis on the basis of empirical and theoretical grounds (see also Hirschbühl and Rivero 1982). The Head hypothesis requires an expansion for NP (NP → N S) which unduly complicates the base rules and which is able to differentiate between FRs and headed relatives only by stipulation (i.e., WH + S = FR, WH + S = headed RC). Furthermore, as they point out, there is nothing which would prevent the WH-phrase in head position to be fronted by WH-question formation. Thus, from the Spanish FR in (4a), the ungrammatical (4b) would be derived.

(4a. Susana conversa con quien Mimi se casó ayer.
'Susana is-talking to whom Mimi got married yesterday.'

b. *¿Con quién conversa Susana Mimi se casó ayer?

Nevertheless, acceptance of the COMP proposal implies that matching conditions must be accounted for by another mechanism. Groos and van Riemsdijk (1979:10) attempt to solve this problem by theorizing that the COMP position of FRs must be accessible to the subcategorization and Case requirements of the matrix verb. This CA hypothesis handsomely encompasses example (1) because quienes 'who (pl)' is a noun and hence fulfills the subcategorization requirement for the verb buscar 'to look for'. In contrast, sentence (5) is ungrammatical because the prepositional relative phrase in COMP does not meet the subcategorization needs of the matrix verb.

(5) *Busco con quien tú saliste.
I-am-looking-for with whom you went-out

Thus, CA predicts that Spanish FRs are straightforward instances of categorial matching.
2. Problematic aspects of COMP Accessibility

2.1 Its incompatibility with the Projection Principle. Chomsky (1981) proposes the Projection Principle as a guiding principle of Government and Binding (GB) theory. In essence, the Projection Principle maintains that subcategorization properties of lexical items must be satisfied at each syntactic level (i.e., D-structure, S-structure and LF). And since subcategorization entails Θ-marking, it follows that the required Θ-positions must also be represented at each of the three syntactic levels. But notice that the Projection Principle and CA are incompatible. The reason for this paradox becomes evident upon examining what in Spanish goes by the name of verbo de régimen. These verbs select a particular (and obligatory) preposition. For example, the verb soñar 'to dream' requires the preposition con 'with' in Spanish.5

(6a) Soñé con quien tú soñaste.
I-dreamt with whom you dreamt.

b. *Soñé por quien tú luchaste.
I-dreamt for whom you fought.

CA predicts that sentence (6a) is grammatical because the prepositional pronoun con quien in COMP is of the appropriate category (i.e., PP) and moreover, it contains the necessary preposition (i.e., con) to satisfy the lexical idiosyncracy of the verb soñar. On the other hand, (6b) is ungrammatical, even though the phrase in COMP is a prepositional phrase, because the preposition is not the required one.6 Now consider the possibility that under CA (6a) is analyzed as having the D-structure in (7) and the S-structure in (8) (details aside).

(7) ...[VP soñé[PP e[ NP [n e][S[COMP][S[tú soñaste
[PP con quien]]]]]]

(8) ...soñé[PP e[ NP [n e][S[COMP][PP con quien i]]
[S[tú soñaste [PP t i]]]]]]

The D-structure (7) violates the Projection Principle because one of the lexical properties of the matrix verb— the fact that it selects a specific preposition—is not being fulfilled at this level of representation. Note there is no preposition con within reach of the matrix verb in (7).7 It is not until S-structure (8) that the necessary preposition becomes available to matrix soñar through CA.
FREE RELATIVES AND THE pro-HEAD HYPOTHESIS

If on the other hand (6a) is assumed to have the D-structure in (9) and the S-structure in (10), the matrix con appears at all levels.

(9) ...soñé [PP con [N P e][S [COMP][S tú soñaste [PP con quien]])]]

(10) ...soñé [PP con [N P e][S [COMP][PP con quien]]] [S tú soñaste [PP t_i]]]

However, if the WH-phrase in COMP is supposed to fulfill the subcategorization needs of the matrix verb (soñar con), (10) does not, because the phrase in COMP is a PP and not an NP. Hence, the structures in (9) and (10) must be rejected even by advocates of CA, since there is no categorial matching.

2.2 Its incompatibility with the θ-criterion. Furthermore, on the assumption that subcategorization entails θ-marking, CA leads to the violation of the θ-criterion—which states "each argument bears one and only one θ-role [...]" (Chomsky 1981:36)—because the phrase in COMP in both (8) and (10) bears two θ-roles: one assigned to it by the embedded V (because of the chain formed by this phrase and its trace) as well as another one assigned to it by the matrix verb (through CA).

In short, CA is at variance with both the Projection Principle and the θ-criterion.

3. Assumptions

In what follows, I shall sketch a solution which will account for the matching effect in tensed free relatives without resorting to CA. This proposal will allow us to preserve the integrity of the Projection Principle. In order to achieve this aim I shall make several assumptions.

3.1 First, I accept the COMP hypothesis as my working hypothesis because I assume that headed and headless relative clauses have parallel structures. Note that for every headless relative it is possible to cite a corresponding headed relative (which conveys about the same meaning) by merely adding a lexical head (cf. note 1). For example, the free relative in (1) can be paraphrased by the headed relative in (11).
(11) Busco personas que me ayuden.
I-am-looking-for persons that me (acc) help
'I am looking for people who can help me.'

This assumption permits the positing of (12) as the structure for all relative clauses in Spanish.

(12) \[ NP \left[ \text{N \ S} \right] \]

Furthermore, (12) has the advantage of being neutral with respect to the issue of whether to consider relative clauses as structure of predication in which the S predicates something about the antecedent N (cf. Chomsky 1977; Williams 1980), or as instances of construal in which the head is indexed with the relative pronoun in COMP (cf. Hendrick, 1982).

3.2 Second, I assume that the ec which serves as the head of FRs can only be pro, the [+pronominal -anaphor] governed category. I will adopt the definition of government offered by Belletti and Rizzi (1981).

(13) \( \alpha \) governs \( \gamma \) in a configuration like \( [\theta \ldots \gamma \ldots \alpha \ldots \gamma \ldots] \)
where:
(i) \( \alpha = X^0 (= \text{a lexical element}) \)
(ii) where \( \mathcal{Q} \) is a maximal projection, if
\( \mathcal{Q} \) dominates \( \gamma \), then either \( \mathcal{Q} \) dominates \( \alpha \), or \( \mathcal{Q} \) is the maximal projection of \( \gamma \).
(iii) \( \alpha \)-commands \( \gamma \)

This definition makes it possible for the head of a maximal projection--\( N_{\text{max}} \) in our case--to be accessible to an external governor, hence pro is governed by the matrix verb or preposition. In a way definition (13) implies that government percolates in a similar fashion to Case percolation. Hence, it makes sense to assume that the ec head of FRs is also Case-marked since it is desirable to keep the difference between headed and FRs to a minimum. Furthermore, if one goes through the inventory of ec's (cf. Chomsky 1982), it is possible to arrive at pro through a process of simple elimination. The ec cannot be a variable because it is not \( \bar{A} \)-bound by an operator. It cannot be PRO because the head position in a RC is governed and Case-marked by the matrix verb or preposition. And finally, it cannot be t because the ec does not form part of a \( \Theta \)-chain (i.e., it does not have an antecedent in argument position). In short, the ec in (3) must be pro. This in turn means that the only relevant difference between headed relatives and free relatives rests on the [+ lexical content] of the head, an intuitively very satisfactory result.
3.3 The third assumption I shall make is that the subcategorization conditions imposed by a verb are fulfilled by the $N^\text{max}$ node in (3), i.e., by the node that contains the entire relative, and not by the $N$ which serves as head of the relative. This is a natural assumption which views subcategorization as sisterhood and which is comparatively easy to demonstrate. Thus, the subcategorization requirements of *buscar ‘to look for’ are met by the whole NP is (3) and not by the head of the relative clause structure. Consider now the slightly more complicated case of a "verbo de régimen" such as *soñar con ‘to dream of’. Even though this verb selects a specific preposition, its subcategorization requirements are fulfilled by the whole PP in (14e) and not just by the head of the relative.

(14)a. Soñé con las vacaciones que me prometiste.  
I-dreamed with the vacations that to-me you-promised  
'I dreamed of the vacation you promised me.'

b. *Soñé las vacaciones que me prometiste.

c. Soñé con Telma.  
'I-dreamed of Telma.'

d. *Soñé con pro.

e. $\ldots \text{VP}$
   $\text{PP}$
   $\text{V}$
   $\text{P}$
   $\text{N}$
   $\text{N}^\text{max}$
   $\text{soñé}$
   $\text{con}$
   $\text{las vacaciones}$
   $\text{que}_{i} \text{pro me}$
   $\text{prometiste t}_{i}$

As (14b) shows, omission of the preposition required by the verb leads to ungrammatical results. Sentence (14c) demonstrates that, as long as the correct preposition and the appropriate head are used, what is "inside" the PP is quite irrelevant; it could be a simple NP (14c), or a relative clause (14a). Nevertheless, (14d) is ungrammatical apparently because the NP head has no lexical content. Why should this be the case when the subcategorization
requirement of the verb is met by the preposition in conjunction with the governed pro? There is a very plausible explanation for this: the content of pro in (14d) is not sufficiently "determined" (cf. Chomsky, 1982). This conclusion contributes towards explaining the empirical fact that Spanish bans sentences with dangling prepositions from the set of well-formed sentences in the language. However, this ban is not a basic restriction but rather the consequence of conditions on empty categories. Thus, the sequence prep t is ruled out by the Empty Category Principle because t is not properly governed; prep PRO is out because prep is a governor (though not a proper governor) and PRO must be ungoverned; and prep pro is not allowed because of the lack of determination of pro.11

In short, I have argued that the structure of free relatives parallels that of headed relatives, that the head of free relatives is the ec pro, and that the subcategorization requirements of the matrix verb are met by the node that dominates the whole relative structure and not by the head.

4. Matching effects

With this much background, it is now possible to turn our attention to the matching effect that Spanish tensed free relatives in subcategorized positions display.

4.1 [tense] FRs. Section 2 pointed out some shortcomings of the CA hypothesis—its incompatibility with the Projection Principle and the Θ-criterion—which already provide some grounds for rejecting this proposal. However, the crucial data for its rejection surfaces upon comparing tensed FRs (15) with infinitival FRs (16); in subcategorized positions, only the former must observe the matching condition.

(15)a. No encuentro { quien me ayude. 
*con quien (tú) puedas salir.
'I can't find { who would-help-me.' 
*with whom you-could go-out.'

b. Busco { a quien (tú) invitas. 
*de quienes tú te burlaste. 
'I-am-looking-for { whom you invited.' 
*of whom you made fun.'
FREE RELATIVES AND THE pro-HEAD HYPOTHESIS

(16)a. No encuentro { con quien salir.
    { de quien fiarme.
    'I can't find { with whom to-go-out.'
    { whom to trust.'

b. Busco { de quienes burlarme.
    { con quien discutir esto.
    'I-am-looking-for { of whom to-make-fun.'
    { with whom to-discuss this.'

Even though both encontrar 'to find' and buscar 'to-look-for' obligatorily subcategorize for an NP--as evidenced by the tensed FRs in (15) where only the matching ones are grammatical--there is no matching in the infinitival FRs (16) where the phrase in COMP is a PP. Now note that under CA there is no straightforward way to explain this disparity, since both [+tense] and [-tense] FRs have a COMP position, thus, CA would be equally possible for the two kinds of FRs. On the other hand, if one rejects CA in favor of INFL accessibility, the different patterning of [+tn] FRs can be accounted for directly: the matching effect is correlated with the [+tn] character of the relative.

4.2 pro determination. The main issue that remains to be settled is how to capture the matching condition itself. Recall that the hypothesis being developed maintains that matrix subcategorization requirements are met by the node which dominates the whole relative clause structure, and that the head position of the relative is occupied by pro. However, examples such as the ones found in (17), which exemplify the S-structure schema in (18), appear to present problems for the pro head hypothesis.

(17)a. *Busco con quien tú saliste.
    I-am-looking-for with whom you went-out

b. *Quisiera con lo que tú estabas jugando.
    I-would-like with what you were playing

(18) \ldots [NP [N pro] [S[COMP[PP]]]_S]

In these sentences the matrix verbs subcategorize a direct object NP; the phrase in COMP is of the PP category. The sentences are ungrammatical because the head and the phrase in COMP do not "match". How can this ungrammaticality be explained within the pro hypothesis? Note that N^{max} (=NP in (18)) fulfills the subcategorization requirements of the matrix which would predict that the sentences in (17) should be grammatical. Nonetheless, there is a principled explanation for the ungrammaticality of (17). This explanation has to do with the stipulation that the ec pro needs to
be "determined" (Chomsky 1982). Two possibilities for achieving pro determination within the environment of tensed FRs come to mind. One claims that pro determination amounts to category matching between the head of the RC and the phrase in COMP. Hence, since pro is a NP, the phrase in COMP must be a NP too. In (18) this categorial matching condition is not met, therefore, the ungrammaticality of the sentences in (17) is explained. The disadvantage of this approach is that it will star grammatical sentences characterized by the schema in (19) where the matrix verb and the one in the relative take the same preposition (cf. (6a) and Section 4.3 below). They would be considered ungrammatical because the pro head (a NP) would not match in category with the PP phrase in COMP.

(19) \[ \ldots V[PP, prep \ [NP, N, \text{pro}, [S, \text{comp}, [PP], S]] S' \ldots \]

The second alternative to pro determination consists of maintaining that the head of the FR and the phrase in COMP must Case-match. This Case-matching condition that tensed FRs in subcategorized positions must meet can adequately be stated as the well-formedness condition (20).

(20) If the structure:

\[ \ldots V (\text{prep}) \ [NP, N, \text{pro}, [S, \text{comp}, \text{WH-phr}], S] \ldots +tn \ldots \]

then (i) pro WH-phr

\[ [\text{\&}, \text{Case}] = [\text{\&}, \text{Case}] \]

or (ii) pro is non-distinct in Case from WH-phr.

There are two points that (20) captures directly. By including V in the pertinent structure, it follows that this matching condition is operative in subcategorized positions only. Thus, (20) predicts that in non-subcategorized positions FRs may be non-matching (provided no other principles interfere). This possibility is realized in Spanish since FRs in subject (21) and TOP (22) positions may be non-matching.

(21) Con \{ quien \{ el que \} me quiero casar \{ está en la esquina. \}

\{ vive a la vuelta. \}

with \{ who me(acc) I-want-to-marry \{ is on the corner \}

\{ lives around the corner \}

' The one I want to get married to \{ is standing on the corner. \}

\{ lives around the corner. \}
(22) Con quien el que me quiero casar, ése ni me habla.

with who me(acc) -want to-marry, that-one doesn’t even talk to me.

Second, (20) specifies that the embedded S part of the FR must be tensed; hence, it allows for [-tn] FR to be non-matching. This possibility is borne out by the non-matching character of infinitival FRs (cf. examples in (16) and Section 5).

Now it is possible to go back to the ungrammatical examples in (17). In these, the matrix varb assigns objective Case to the RC structure (i.e., to $N^{\text{max}}$), this Case percolates to the head of the RC (i.e., pro). However, since the relative pronoun in COMP is oblique, the Case-matching requirement (20) is not met and the sentences are excluded. This explanation exploits the modularity of the grammar by showing how Case Theory interacts with constituent structure and subcategorization to determine the grammaticality of FRs.

Statement (20ii), Case non-distinctness, is necessary because of examples like (23).

(23)a. Briana no encuentra pro quien t la ayude.
   Briana can't find who her(acc) would-help 'Briana can't find anybody to help her.'

b. Soñaba con pro quien (tú) me presentaste ayer.
   I-was-dreaming with who you introduced yesterday.
   'I was dreaming of the one you introduced me to yesterday.'

In (23a) pro is objective while the relative pronoun is nominative; in (23b) pro is oblique but the relative pronoun is objective. Nevertheless, the examples are grammatical. Why? Because the relative pronoun quien/es 'who, sg./pl.' adopts the same morphological shape regardless of Case. Case-non-distinctness finds confirmation in data from other languages as well. For example, Groos and van Riemsdijk (1979:37) point out that although modern German FRs require Case-matching, (24) is grammatical despite the mismatch in Case-requirements. This is because the neuter singular relative pronoun was has the same form whether nominative or objective.
(24)a. Ich habe gegessen was (acc/nom) noch übrig war.
    I have eaten what still left was
    'I ate what was left.'

b. Was (nom/acc) du mir gegeben hast, ist prächtig.
    what you me given have is wonderful
    'What you have given to me is wonderful.'

Even English makes limited use of the non-distinctness criterion.

(25)a. I'll hit whoever (acc/nom) tries anything.

b. She'll eat whatever (acc/nom) comes her way.

In brief, condition (20) makes all the right predictions for Spanish FRs. Moreover, it is possible to further streamline this filter; since Case-matching (20i) is just another instance of Case-non-distinctness (20ii), it is feasible to re-state (20) as (26).

(26) If the structure:

\[ ...V (prep) [NP[N_{\text{NP}} [pro][S_{\text{S}} [COMP \ WH-phr][S_{\text{S}}]... +tn... \]

then: pro and WH-phr must be non-distinct in Case.

Note that the implicit claim made by (26) is that Case Theory plays the central role in the matching effects observed by FRs, Case-matching being seen as the way in which pro determination is achieved in this type of construction. As a matter of fact, this matching requirement between the pro head and the phrase in COMP can be conceived as the instantiation of an agreement rule.

The place in the grammar where condition (26) applies is contingent upon the resolution of other open questions in the GB theory. My own inclination would be to consider it a rather late condition which applies in PF, but the decision seems to rest to a great degree on the decision as to where the Case filter operates. Since (26) is strictly local, it must apply after Move-\(\alpha\) has transported the WH-phrase to COMP. Furthermore, the criterion of non-distinctness—which essentially looks at the morphological shape of words—further points to the lateness of this condition.

4.3 Prepositional verbs. In Section 4.2 it was claimed that pro determination amounted to Case-matching up to non-distinctness. This decision was preferred over one which considers pro determination in terms of category matching. In this section, I plan to justify this position in detail.
Situations arise in which both the matrix verbs and the verb in the relative clause select the same preposition.

(27a). Anoche soñé con ese tío con quien tú soñaste el otro día.
   'Last night I-dreamt of that guy of whom you dreamt the other day.'

b. Me acabo de burlar de esa muchacha de quien tú te burlaste hace unos días.
   'I've just made fun of that (young) woman whom you made fun of a few days ago.'

The relevant parts of the S-structures of (27a) and (27b) are (28a) and (28b), respectively:

(28a). ...soñé [PP con [NP con ese tío][S COMPO con quien_i] ...
        [S tú soñaste t_i ...]_S NP]PP

b. ...burlar [PP de [NP de esa muchacha][S COMPO de quien_i]
        [S ...burlaste t_i ...]_S NP]PP

Sentences like those in (27) can be turned into free relatives by merely leaving the lexical head out (29).

(29a). Anoche soñé con quien tú soñaste t el otro día.
   'Last-night I-dreamt of whom you dreamt the other day.'

b. Me acabo de burlar de quien tú te burlaste t hace unos días.
   'I've just made fun of whom you made fun of a few days ago.'

However, when leaving out the lexical head, the two identical prepositions (two con's in (27a), and two de's in (27b)) surface as just one in (29a) and (29b). I would like to explain this phenomenon as follows: in S-structure these sentences have the structure in (30) with two prepositions—one required by the matrix verb, the other one by the embedded verb—the head of the relative clause is pro:

(30a). ...soñé con pro [S con quien...

b. ...burlar de pro [S de quien...

When these structures reach PF (or perhaps the morphological sub-component), and as a consequence of the lack of phonetic matrix in pro, the two prepositions come together and if identical one of them
is deleted. This rule could be informally stated as in (31), where the index represents identity.

\[(31) \quad \text{[prep}_i \text{ prep}_i \text{]} \rightarrow \text{prep}_i\]

Consider next the ungrammatical sentences in (32).

\[(32)a. \quad \text{*Sóñé en quien tú pensabas.} \\
\text{I-dreamed in whom you were-thinking}\]

\[b. \quad \text{Sóñé de quienes yo no me fiaría.} \\
\text{I-dreamed of whom I wouldn't trust}\]

\[c. \quad \text{*Me burlé con quienes tú fuiste.} \\
\text{I-made-fun-of with whom you went}\]

\[d. \quad \text{*Me burlé en quien tú pensabas.} \\
\text{I-made-fun-of in whom you were-thinking}\]

According to my hypothesis the above sentences have the structures in (33) (only (32a) and (32c) are used as illustration).

\[(33)a. \quad \text{Sóñé con pro [en quien... \]}

\[b. \quad \text{...burlé de pro [con quienes... \]}

The preposition \textit{con} in (33a) and \textit{de} in (33c) must be present in S-structure because they are selected by \textit{sóñar} and \textit{burlarse} respectively. Otherwise, the Projection Principle would be violated and the sentences would be ungrammatical from the start. In PF, the two prepositions in each sequence are contiguous; since the two prepositions are not identical, rule (31) can not operate and the sentences are discarded.

\[\S 4.4\quad \text{Personal a. Next, I would like to examine another set of examples which serves to demonstrate the viability of the explanation developed in this paper. This set shows how the so-called 'personal a', an accusative case marking particle, interacts with the proposed explanation for FRs. The parameters for the use of personal a are less than crystal clear but there is some consensus in certain areas. Personal a appears before a human specific (or individual) direct object (cf. (34a) and (34b)) but it is not generally used if the object is non-specific (cf. (34c) and (34d)).}\]

\[(34)a. \quad \text{Julia llamó a/\text{\textasteriskcentered} Paco.} \\
\text{'Julia called Paco.'}\]
(34)b. *Yo *el/al médico.
'I-saw the doctor.'

c. Pilar anda buscando una criada.
'Pilar is looking-for a maid.'

d. Allí hallará amigos que le acojan bien. (Ramsey 1956)
'There he-will-find friends that will receive him kindly.'

Now consider the following free relatives in (35) and (36) where only the first example in each is grammatical.

(35)a. Invito a quien tú invitaste.¹⁷
I invite whom you-have invited

b. *Invito con quien tú irás.
I-invite with whom you will-leave

(36)a. Quiero a quien tú quieres.
I-love whom you love

b. *Quiero con quien tú vienes.
I-love with whom you are-coming

c. *Quiero de quien tú me hablaste.
I-love of whom you spoke to me

Both invitar 'to invite' and querer in the sense of 'to love' select a human direct object. Consequently, personal a should co-occur with this direct object whenever the right conditions arise.¹⁸ According to my hypothesis the sentences in (35) should have the structures in (36).

(37)a. invito a [NP pro [\(\Sigma\) a quien tú invitaste \(t_1\)]]

b. invito a [NP pro [\(\Sigma\) con quien tú irás \(t_1\)]]

The subcategorization for invitar is met by \(N_{\text{max}}\), pro is the head of the relative.¹⁹ The structure (37a) is grammatical because the two a's meet the conditions for the PF rule (31).²⁰ On the other hand, the structure in (37b) leads to an ungrammatical result because a + con cannot be reduced to just one preposition, thus the sentence is rejected.²¹ A similar reasoning explains the sequences in (36).

It should be the case that, if my explanation has merit, since the verbs in (34c) and (34d) permit human direct objects to be
read as non-specific (thus, no personal a), it also should be possible to find FRs without personal a. This prediction is borne out by the examples in (38).

(38)a. Pilar anda buscando quien la ayude.
    Pilar is looking for who her (acc) can help
    'Pilar is looking for someone who can help her.'

b. Allí hallará quien le acoja bien.
    there he-will-find who him (acc) receive kindly
    'There he will find people who will receive him kindly.'

Thus, personal a in its interaction with direct object FRs totally parallels the behavior of this particle with lexical direct objects (cf. (34)). This fact supports the claim that ec's (in this case pro) mirror the behavior of overt elements.

5. Infinitival free relatives (IFRs).

I have already mentioned that one of the strongest arguments against CA is provided by IFRs. In contrast, if one takes INFL to be the accessible element, the disparity between [+tense] FRs reduces precisely to this feature, an intuitively very satisfactory outcome. Furthermore, the well-formedness condition in (26) readily captures this asymmetry because it specifies [+tn] as a condition for matching effects; since IFRs are [-tn], (26) does not apply to them.

In the first place, it should be pointed out that just as was the case for tensed FRs, for every IFR with an empty head (39) there is a parallel one with a lexical head (40). Hence, once again the difference lies in the [+lexical] character of the head.

(39)a. No tiene con que secarse.
    s/he doesn't have with what to-get-dry

b. No hay por quien votar.
    there isn't for whom to-vote

c. Al fin eligió con quien ir.
    Finally s/he chose with whom to-go.

d. No encuentra a quien vender la casa.
    s/he doesn't find to whom to-sell the house

(40)a. No tiene toalla con que secarse.
    s/he doesn't have towel with which to-get-dry
(40)b. No hay persona por quien votar.

(41) ...V [NP, lexical head][S[COMP WH-phr][S[...[-tn]...t]]]

Second, the WH-phrases in COMP in (39) and (40) cannot be
nominate (42), but they can be objective (43) or oblique (39).
This follows from Case Theory: since the embedded S is [-tn], there
is no AGR in Spanish infinitivals, thus the trace in subject
position can not receive Case, hence it can not be a variable.

(42)a. *No encuentra quien llegar.
S/he doesn't find who to-arrive

b. No encuentra [NP pro [S[COMP quien]...t... llegar]]

(43)a. No habfa a quien elogiar. 22
There wasn’t whom to-praise

b. No tiene a quien invitar.
S/he doesn't have whom to-invite

Third, IFRs occur not only in subcategorized positions (i.e.,
the positions where tensed ones must be matching) but also in non-
subcategorized places: subject (44a) and TOP positions (44b).

(44)a. En quien confrar no deja de ser un problema.
In whom to-trust remains to-be a problem

b. Por quien votar, ése es el asunto.
For whom to-vote, that is the thing.

In short, IFRs partake of the same characteristics of tensed
FRs provided no other principles enter the picture (cf. (42)). However, perhaps because there is no straightforward way to "determine" the content of the pro in head position (recall the matching condition (26) does not apply), IFRs occur with a rather limited class of matrix verbs (cf. Plann's 1981 comments). Intuitively, these verbs are limited to those which most readily take indefinite/unmodified objects. Related to this last fact seems to be the observation that IFRs are more likely to occur in negative contexts (cf. examples in (39) and (43)).

Before closing this section, I again want to highlight the fact that the hypothesis developed in this paper solves the mystery of tensed FRs vs. infinitival ones, a mystery which has puzzled researchers for quite a while. For example, Hirschbühlher and Rivero (1981:121) state:

...the question that must be solved in why infinitival relatives can be non-matching even in languages that do not exhibit this type of free relatives such as French:

(18) Il n'a avec qui parler.
   He not has with who talk

(19)a. *J'ai acheté où il habite.
   I have bought where he lives

   b. Il n'a où dormir.
   He not has where sleep

Although we have no explanation for this state of affairs, the comparison between Catalan and French indicates that a separate account of infinitival relatives is required, in addition to our proposal.23

It should be obvious that once CA is discarded in favor of INFL accessibility there is no need for a separate account of IFRs, both [ttn] FRs become unified under the same hypothesis supplemented by condition (26) which is seen as a device which aims at pro determination.

6. Conclusion

In sum, the hypothesis that pro [+pronominal -anaphor] occupies the head position of FRs together with the natural assumption that subcategorization requirements are met by the node that dominates the entire RC, renders the CA hypothesis irrelevant in
the context of relative clauses. INFL accessibility and the well-formedness condition (26) allows us not only to unify the behavior of [+tn] FRs but also permits us to keep the Projection Principle and the 6-criterion intact as guiding principles of GB theory.

The matching effect observed by tensed FRs in subcategorized positions is seen as a stratagem used to determine the content of the ec pro. This determination is achieved through Case Theory. This conclusion provides another example of the modularity of the grammar.

The implicit claim made by the pro head hypothesis is that it is the head of the FR (as opposed to the head in COMP), the constituent that is crucial to this type of construction. Furthermore, it also reduces the differences between lexically headed and pro headed RCs to the [+lexical content] of the head, these RCs are completely parallel in all other aspects.

The pro head hypothesis makes use of two pieces of evidence which are independently attested in the grammar of Spanish: the fact that Spanish derivatively prohibits dangling prepositions and the fact that human specific (or individualized) direct objects demand the personal a particle. The only additional machinery required is a rule in PF (cf. (31)), which in essence deletes one of two contiguous identical prepositions. Whether this type of rule is exclusive to FRs or not, is left open to future research.

Of general interest to GB theory is that the governed empty pronominal pro is the viable ec for the head of FRs. This fact shows that pro should not be restricted to the subject position of pro-drop languages (in the core cases). Moreover, the interaction of pro with personal a gives support to the claim that ec's parallel the behavior of overt elements: the features of pro are relevant in the explanation of matching effects in FRs in Spanish.

One issue that remains to be settled is that of the difference between languages which adhere to the matching effect in tensed FRs and those which do not. If this difference could be ascribed to a property of the head of the FR, then it will give more strength to my contention that the head is the crucial element for the well-formedness of FRs. I think it can. I would like to maintain that this parameter has to do with pro "determination". Matching languages require that the ec pro be sufficiently determined—thus, Case matching up to non-distinctness—maybe in an effort to avoid possible ambiguities. Non-matching languages do away with any need for further determination. One can find confirmation for this claim even within the history of the Spanish language. Hirschbülcher (1979:166) cites the following examples.
from 16th-17th century Spanish.

(45)a. Se rebeló...contra el servicio de a quien se había ofrecido.
    'He rebelled...against the service of (the one) to whom he had offered himself.'

b. No quería ver más a con quien estuviera.
    'He no longer wanted to see (the one) with whom he was.'

c. No quiero ver al ceño/Vainamente poco severe/De a quien la sangre ensalza o el dinero.
    'I do not want to see the frown/Vainly severe/of (the one) to whom blood or money exalts.'

The above examples show that in earlier centuries Spanish did not observe the matching condition;\(^{26}\) as a consequence, notice that a matrix preposition can come together with the preposition of the phrase in COMP, which in turn means that at that stage in the development of the language the pro head did not need any extra "determination".

Acknowledgments

My special thanks go to Wayne Harbert for our discussions on free relatives, and to M. Rivero and B. Freidin for their comments on an earlier version of this paper. I would also like to thank C. Piera and D. Wheeler for their feedback. As usual, I am responsible for all shortcomings.

Footnotes

\(^{1}\) More explicitly, the RAE (1974:526) says:

Frequently, the relatives que and quien are used without an explicit antecedent, either because it is unknown or indeterminant, or because it is of no interest to the speaker, or because the words causa 'cause', razón 'reason', motivo 'motive', hombre 'man', or similar ones are easily understood...[my translation, M.S.]
Although it is a recognized fact that the boundaries between the relative and the interrogative pronouns are quite hazy in some Spanish sentences, example (i) should be read as a FR parallel in interpretation to the headed relative in (i) and not as an indirect question. (For a recent treatment of Spanish indirect questions, see Plann 1982.)

(i) Busco personas que me ayuden.
    I-am-looking-for people that me (acc) help (3 pl)
    'I am looking for people who can help me.'

Tellingly, buscar does not appear to subcategorize for a sentential complement.

(ii) *Buscaba que María lo hiciera.
    'S/he was-looking-for that María would-do it.'

(iii) *Buscaba de quién (tú) podías depender.
    'S/he was-looking-for whom you could depend on.'
    (Cf. Sabía de quién (tú) podías depender.
    'S/he knew whom you could depend on. ')

Supposedly, the head could also be absent. In a footnote, Borer (1981:180) refers to one such proposal due to Fassi Fehri. According to this analysis FRs are instances of $\overline{S}$ marked with the feature [+N]; COMP is posited to be the head of $\overline{S}$. This hypothesis will not be entertained for reasons which will become clear in Sections 4.1 and 5.

In contraposition the infinitival FR in (i) is fully grammatical.

(i) Busco con quien salir.
    I-am-looking-for with whom to-go-out

This difference in grammaticality will be addressed in Sections 4.1 and 5 where it will be shown that it falls out from the hypothesis sponsored in this paper.

All Spanish verbos de régimen have the transitive structure $V_{PP}$ prep NP where a specific preposition is selected by the verb. Some other verbos de régimen are: pensar en 'to think of', acordarse de 'to remember', reírse de 'to laugh at', apiadarse de 'to have pity on', consistir en 'to consist of', empeñarse en 'to insist on', vanagloriarse de 'to boast of'.
6 I do not mean to imply that preposition selection is a specific problem of FRs, rather, what is problematic for FRs with respect to CA and the Projection Principle is the structure in (7).

7 The structures in (7) and (8) are modelled after Groos and van Riemsdijk example (51)—which in turn is a re-interpretation of Bresnan and Grimshaw's (107).

(51) I'll move [\text{pp} e [\text{pp} to whatever town}] you move

8 For example, Hirschbühl and Rivero (1981) claim that the matching condition for Catalan FRs in subcategorized positions amounts to categorial matching. Given the diachronic and synchronic closeness of Catalan to Spanish, it seems safe to speculate that the same hypothesis would be valid for Spanish (but see discussion below).

9 Although the correctness of the Projection Principle is not obvious, it seems to be methodologically sound to keep it in its strongest form until such a time when it is conclusively shown to be inadequate or incorrect. My position is that FRs do not provide an argument against the Projection Principle.

10 See Chomsky 1965 for a lengthy discussion on strict subcategorization as sisterhood.

11 Spanish does not allow re-analysis of the type which permits English to have stranded prepositions in certain structures. See Hornstein and Weinberg (1981) for discussion of the English phenomenon; and Kayne (1981) for a comparison between English and French; also Jaeggli (1980) for related issues in Romance. For the various ways to pro "determination" I refer the reader to Suñer (1982d, Section 3).

12 This is not an ad hoc requirement to save the pro head analysis. Case matching in FRs is a well-documented fact for languages with overt Case. For example, for German see Groos and van Riemsdijk (1979). It is also known that Case matching is not required for headed relatives, thus it makes sense to hypothesize that Case matching is one device grammars use to "determine" the empty head of FRs.

13 In contrast, for a language like French which requires matching in all positions (TOP included) the V specification should
be left out. Note that even CA must postulate a filter which guarantees matching in TOP position.

In 'Free Relatives, the matching parameter, and INFL Accessibility', a paper in preparation, I show that the parametrization of filter (20) allows one to capture the behavior of FR in other languages.

Just in passing, I should mention that non-matching FRs of the type found in (21) and (22) provide strong support for the COMP analysis of this construction (see also Hirschbülher and Rivero 1982).

14 I am grateful to S. McConnell-Ginet for this observation.

15 It could not be claimed that one of the prepositions is not inserted. This would violate a lexical property of the verb(s).

16 The hedge "generally" is justified by examples such as the ones in (i) and (ii). Pronouns, even indefinite ones, take this a.

(i) Busco a alguien que sepa portugués. 'I-am-looking-for somebody who knows (subjunctive) Portuguese.'

(ii) No conozco a nadie que sepa portugués. 'I do not know anybody who knows (subjunctive) Portuguese.'

17 Hirschbülher and Rivero (1981) use the parallel Catalan sentences in (35) to substantiate the CA hypothesis. They say that the Catalan equivalent of (35a) is grammatical because a quien is a NP, and invitar subcategorizes for a NP, while (35b) is deviant because the PP con quien does not satisfy the requirements of the matrix verb. My purpose is to account for the same data without resorting to CA.

18 Note that given the context, in both (35a) and (36a) the speaker must have a specific referent in mind; hence, personal a is mandatory.

19 Personal a does not change the category of the NP it precedes, i.e., the free relatives in (35a) and (36a) are NPs and not PPs. Cf. Jaeggli, 1980.
In this instance, however, there is an alternative explanation. As suggested by Borer (1981) for the Rumanian particle pe (the equivalent of Spanish personal a), it is conceivable that one of the a's is never inserted in these FRs. The crucial difference between this case and the one illustrated by (30) is that in (30) the prepositions are required by the verb whereas in (37a) a is not; rather it is a Case-assigner particle which could be inserted even in PF. However, regardless of where a is inserted in a structure like (37a) the output will be grammatical.

Alternatively, if a is not inserted, pro (objective) and the phrase in COMP (oblique) do not Case-match.

Since haber never takes personal a (cf. Suñer 1982b, Chapter 1), (43) unequivocally shows that that a is in COMP.

See also the comments to this effect in Hirschbülher 1979: 170.

INFL Accessibility through verb selection is independently needed in Spanish to account for mood distinctions in embedded clauses. See Suñer (in preparation) for details.

Nevertheless, it should be pointed out that this type of rule has been widely attested in morphology, for example see Stemberger 1981.

Keep in mind that Spanish developed from Latin, a language which had active overt Case and which did not adhere to the matching condition. See also Groos and van Riemsdijk (1979) for discussion showing that Modern High German allowed for non-matching FRs. In general, it seems to be the case that languages move away from non-matching FRs; there might be a correlation here with the decline of overt Case in the history of these languages.

References


Hirschbülher, P. 1979. The syntax and semantics of WH-constructions. IULC.


______. (in preparation). Free relatives, the matching parameter, and INFL accessibility.