THE DEVELOPMENT OF THE SPECIFIED SUBJECT CONDITION AND THE TENSED S-CONDITION / PROPOSITIONAL ISLAND CONDITION

Melinda Sinclair
University of Stellenbosch
ACKNOWLEDGMENTS

I would like to thank various people for their contribution to this study: Rudie Botha, who subjected the first draft of the manuscript to his customary detailed criticism; Hester Waher for her useful comments on §5; Cecile le Roux, who made numerous suggestions that led to considerable improvements in the formulation of this study; Marina Savini for valuable editorial assistance; and Mrs. L. Gildenhuys for her skill and patience in typing the manuscript. The responsibility for the numerous flaws that undoubtedly remain is mine alone.
CONTENTS

Acknowledgments ................................................................. (i)

1 Introduction ........................................................................ 1

2 The first stage in the development of the conditions .............. 4

2.1 General remarks ............................................................. 4

2.2 The SSC and TSC in (Chomsky 1973) .................................. 4

2.2.1 The final formulations of the conditions ......................... 4

2.2.2 The introduction of the conditions and the fundamental empirical problem of linguistics ................................. 7

2.2.3 The general nature of the motivation for the SSC and TSC .. 10

2.2.4 Dealing with potential empirical problems for the SSC and TSC .......................................................... 19

2.2.4.1 Wh-Movement, Strict Cyclicity and the COMP-escape hatch .. 19

2.2.4.2 The SSC and traces ................................................... 20

2.2.4.3 The SSC and the feature [+ definite] .......................... 27

2.2.4.4 The SSC and the notion 'agency' .............................. 28

2.2.4.5 The unsolved problem of Coreference Assignment ......... 30

2.2.5 The relative interpretation of conditions on rules and the possibility of parametric variation ................................. 30

2.2.6 The naturalness of the SSC ............................................ 31

2.3 Other relevant aspects of the SSC and TSC in the first stage of their development ......................................................... 32

2.3.1 Additional motivation for trace theory ............................ 33

2.3.2 The idealization of sentence grammar ............................ 34

2.3.3 The notion 'involve' ..................................................... 38

2.3.4 Further restrictions on the expressive power of transforma-

tions ..................................................................................... 42

2.3.5 Some changes in the formulation of the SSC .................... 43

2.4 Summary of main points ..................................................... 45

3 The second stage in the development of the conditions .......... 47
5.2.2.6  The asymmetry between pronouns and other anaphors ........ 98
5.2.3  Some empirical problems of the OB binding theory ............ 98
5.3  The GB binding theory as an alternative to the OB binding theory ................................................................. 101
5.3.1  General remarks .......................................................... 101
5.3.2.  The GB binding theory .............................................. 103
5.3.2.1  The formulation of the theory .................................. 103
5.3.2.2  Definitions of central notions ................................ 104
5.3.2.2.1  'Bound' and 'free' ............................................. 104
5.3.2.2.2  'Anaphor', 'pronominal', 'R-expression' .................... 105
5.3.2.2.3  'Governing category' ....................................... 106
5.3.2.3  Indexing in the GB theory ..................................... 109
5.3.2.4  An illustration of the GB binding theory .................... 110
5.3.2.4.1  General remarks ............................................. 110
5.3.2.4.2  The application of the GB binding theory in clauses ... 110
5.3.2.4.3  The application of the GB binding theory in NPs .......... 118
5.3.3  The GB binding theory and the conceptual and empirical problems of the OB binding theory ................................. 127
5.4  The 1979 versus the 1981 version of the GB binding theory ... 131
5.4.1  General remarks .......................................................... 131
5.4.2  Three different formulations of the GB binding theory ...... 132
5.4.3  The 1979 vs. the 1981 interpretation of the empirical differences between the OB and GB binding theories .......... 136
5.5  Some proposed modifications to the GB binding theory .......... 144
5.5.1  General remarks .......................................................... 144
5.5.2  Reformulating the GB binding theory in terms of the notion 'accessible SUBJECT' .............................................. 145
5.5.3  Further modifications to the GB-SUBJECT binding theory ... 156
5.6  Some problems for the GB binding theory .......................... 158
5.7  The elimination of structure building rules in the GB-framework ................................................................. 166
5.8  The Empty Category Principle ......................................... 169
5.9  Chomsky's view of the development of the GB binding theory ... 170
5.10  Summary of main points ................................................ 171

Footnotes ............................................................................ 174

References ........................................................................... 190

Previous issues of SPIL ......................................................... 194
Introduction

The Specified Subject Condition (henceforth SSC) and the Tensed S-Condition (henceforth TSC) were proposed in the early seventies by Chomsky as principles of universal grammar, restricting the application of both syntactic movement transformations and rules of semantic interpretation. An early informal reference to these conditions can be found in (Chomsky 1971:34-40). The first detailed technical discussion of the SSC and TSC appeared in "Conditions on transformations" (henceforth (Chomsky 1973)). Since the publication of the latter work, these two conditions have played a fundamental role in the theory of universal grammar (henceforth UG) within the Chomskyan approach to the study of language. The developmental history of these two conditions in fact provides us with helpful insight into important aspects of both the substantive and the methodological developments which Chomsky's general linguistic theory has undergone within the past decade.

The aim of the present study is to provide a comprehensive overview of the development of the SSC and TSC within Chomsky's syntactic theory since 1973. In particular, the study is an attempt to provide answers to the following questions:

(1) a. What are the changes which the SSC and TSC have undergone since 1973?
   b. What circumstances led to each change?
   c. What justification was provided for each change?
   d. What causal relations are there between the development of the SSC and TSC on the one hand, and developments in other components --- substantive and methodological --- of the Chomskyan approach to the study of language, on the other hand?

This study forms part of an in depth inquiry into the nature of progress in Chomskyan linguistics. The overview presented here provides
the data for this inquiry. However, I believe that this study of the development of the SSC and TSC can be useful in its own right. It provides a detailed account of some important recent developments in Chomsky's syntactic theory. Anyone who is interested in what has been happening in Chomskyan syntax during the past ten years will therefore find it informative. The study can also be used as the basis of an advanced course on Chomskyan syntax.

The changes which the SSC and TSC have undergone since 1973 are discussed in chronological order below. By distinguishing four different stages in the development of these conditions, I will impose some further structure on the discussion. These four stages are distinguished with reference to the question "What types of rules and/or representations are restricted by the SSC and TSC and the various conditions which have replaced them?". In the first stage of their development, the SSC and TSC were interpreted as conditions which restrict both syntactic movement transformations and rules of semantic interpretation. This stage is discussed in §2 below. In the second stage of their development, the two conditions were interpreted as conditions restricting rules of semantic interpretation only. See the discussion in §3 below. In the third stage of their development, the SSC and TSC (reformulated as the Opacity Condition and Nominative Island Condition, respectively) were no longer formulated as conditions on rules, but as conditions that restrict representations at a certain level. See the discussion in §4 below. In the fourth stage of their development, the two conditions follow, as special cases, from a more general binding theory, which also restricts representations at some level. Note that, strictly speaking, the SSC and TSC no longer form part of UG in this fourth stage. Instead, they are subsumed under a more general theory. See the discussion in §5 below. In the discussions that follow, the main emphasis is naturally on various works by Chomsky that deal with the SSC and TSC, and subsequent versions of these conditions. Works by Chomsky's associates or critics that bear on the development of these conditions will be discussed where appropriate.

This introduction is concluded with a brief, informal statement of the basic content of the SSC and TSC. The SSC stipulates that no rule can involve \( X, Y \) in a structure such as (2), where \( \circ \) contains a specified subject.\(^1\)
Thus, the SSC prohibits the Reciprocal Rule from associating the men and each other in (3), because the embedded clause contains a specified subject, namely John.

(3) *The men want [John to like each other]
2  The first stage in the development of the conditions

2.1  General remarks

(Chomsky 1973) is the primary work in which the SSC and TSC are presented as conditions that restrict both syntactic transformations and rules of semantic interpretation. In §2.2 below the introduction of the two conditions in this work will be discussed in detail. Other works by Chomsky in which the SSC and TSC are presented as conditions on syntactic transformations and rules of semantic interpretation include "Conditions on rules of grammar" (henceforth Chomsky 1976a), Reflections on language (henceforth Chomsky 1975a), and "On the nature of language" (henceforth Chomsky 1976b). The relevance of these works for the development of the SSC and TSC is discussed in §2.3 below.

2.2  The SSC and TSC in (Chomsky 1973)

2.2.1  The final formulations of the conditions

The formulations of the SSC and TSC adopted in (Chomsky 1973:257) are as follows.

(5)  The SSC

No rule can involve X, Y (X superior to Y) in the structure

\[ ... X \ldots [\omega \ldots Z \ldots - WYV \ldots] \ldots \]

where Z is the subject of WYV and is not controlled by a category containing X. 6)

(6)  The TSC

No rule can involve X, Y (X superior to Y) in the structure

\[ ... X \ldots [\omega \ldots Z \ldots - WYV \ldots] \ldots \]

where Y is not in COMP and \( \omega \) is a tensed S.

No definition of the notion 'involve' is provided in (Chomsky 1973). It is clear, however, that the notion must cover both syntactic movement
rules and rules of semantic interpretation. Both these types of rules are claimed to be constrained by the SSC and TSC in (Chomsky 1973) --- see the discussion in §2.2.3 below.

Although the formulations (5) and (6) imply that X must be to the left of Y, Chomsky (1973:272) suggests that the conditions should be generalized, eliminating the left-right asymmetry.

The SSC, as formulated in (5), has two subcases: (i) where Z is not controlled at all, i.e., where Z is a lexical subject, and (ii) where Z is controlled by a category which does not contain X. The various components of the SSC are illustrated by the following sentences.

(7)

a. The men each expected \([_S \text{the soldier to shoot the other}]\) \(25a\)
b. *The men expected the soldier to shoot each other \(25b\)

(8)

a. The men each saw \([_{NP \text{pictures of the other}}]\) \(28a\)
b. The men saw pictures of each other \(28b\)

(9)

a. The men each saw \([_{NP \text{John's pictures of the other}}]\) \(29a\)
b. *The men saw John's pictures of each other \(29b\)

(10)

a. The candidates each expected \([_{S \text{PRO to defeat the other}}]\) \(24a\)
b. The candidates expected to defeat each other \(24b\)

(11)

a. We each persuaded Bill \([_{\text{COMP PRO to kill the other(s)}}}\) \(111\)
b. *We persuaded Bill to kill each other \(112\)
Chomsky (1973:238) assumes that the (b)-sentences in (7)-(11) are all derived from the (a)-sentences by a rule of each-Movement, which moves each into the determiner position of the other. In each case $X = \text{each}$, and $Y = \text{the other}$. 

In (7a) the lexical subject $Z =$ the soldier intervenes between $X$ and $Y$. Movement of $X$ to $Y$ to derive (7b) is thus prohibited by the SSC. In (7), $\alpha\epsilon$ is S. In (8) and (9), $\alpha\epsilon$ is NP. In (8a) there is no subject $Z$ (subject being optional in NP), and each-Movement may apply to derive (8b). In (9a) a lexical subject John's intervenes between $X$ and $Y$. The SSC thus prohibits the derivation of (9b).

In (10) and (11) there are no lexical subjects in the embedded clauses, but PRO-subjects controlled by some category. In (10a) the subject PRO is controlled by a category containing $X$, namely the candidates each. Each-Movement can therefore apply to derive (10b). In (11a) PRO is controlled not by a category containing each, but by Bill. The SSC thus prohibits the application of each-Movement to derive (11b).

The TSC stipulates that no rule can involve $X$, $Y$ when $Y$ is in a tensed sentence. This is illustrated by the following sentences. (The case where $Y$ is in COMP will be discussed in §2.2.4 below.)

(12) a. The candidates each expected the other(s) to win

b. The candidates expected each other to win

(13) a. The candidates each expected that the other(s) would win

b. *The candidates expected that each other would win
In (12) and (13) the (b)-sentence is derived by the rule of each-Movement. In (13) each-Movement moves X (= each) to the position Y (= the other(s)), which is in a tensed clause. Consequently, the derivation of (13b) is prohibited by the TSC. In (12) Y is in a nontensed clause, and so the TSC does not prohibit the derivation of (12b).

The SSC and TSC replace the Insertion Prohibition, formulated by Chomsky (1965:146). The Insertion Prohibition prohibits the introduction of morphological material into a configuration dominated by S once the complete transformational cycle has applied to this configuration. Chomsky proposed this condition in order to explain the difference between I kept it near me and I aimed it at myself. The first sentence, but not the second, contains an embedded sentence. The application of the reflexivization rule in the first sentence is thus blocked by the Insertion Prohibition. The SSC and TSC specify the exact conditions that block the introduction of material into a cyclic node once it has been passed by the cycle of transformational rules.

2.2.2 The introduction of the conditions and the fundamental empirical problem of linguistics

Chomsky (1973:232f) explicitly relates the introduction of conditions such as the SSC and the TSC to the problem of accounting for the acquisition of knowledge of language, and to the objective of solving this problem by restricting the class of possible grammars. Chomsky has stated, and defended, his position on the acquisition of knowledge of language in a number of works. The following points represent the essence of his position.
(i) The system of grammar for a human language is a rich and complex system.

(ii) The acquisition of this system by individuals is remarkably rapid and uniform.

(iii) The evidence on the basis of which the system can be attained is both limited and degenerate.

(iv) Given (i), (ii) and (iii), the acquisition of knowledge of a human language can only be explained on the assumption that there is a fixed and highly restrictive initial state that determines the general framework of each natural language.

The following remarks by Chomsky (1980a:233) give a clearer indication of how the assumption (iv) makes it possible to explain the acquisition of knowledge of grammar.

(14) "The child's initial state, it seems, must lay down the general principles of language structure in fair detail, providing a rich and intricate schematism that determines (1) the content of linguistic experience and (2) the specific language that develops under the boundary conditions given by this experience. If the initial restriction is sufficiently severe, it will be possible for the child to attain a system of great intricacy on the basis of limited data, data sufficient to rule out all possibilities but one or a few. Then he will know the language compatible with his limited experience, though there will be no relation of generalization, abstraction, induction, habit formation, or the like that relates the system attained at the final state to the data of experience. The relation between experience and knowledge will be quite abstract. The principles of language structure incorporated in the initial state express this relationship."

The initial state of the language learner (and the linguist's characterization of this state) is known as "universal grammar" (UG). UG permits
only a limited number of final states, each of which corresponds to the grammar of a specific language.

The assumptions about the nature of language acquisition discussed above, and the view of UG as a characterization of the initial state of the language faculty, form part of the Chomskyan approach to linguistics. In particular, the assumption that the initial state of the language faculty is highly restrictive forms part of the hard core of substantive assumptions made about the object of inquiry in this approach. It is from this substantive assumption that it follows that UG must be as restrictive as possible, i.e., that it must restrict the class of available grammars as narrowly as possible.

Although the objective of reducing the class of grammars has served as a "guiding principle" in the study of generative grammar virtually since its outset, according to Chomsky (1981a:13), linguists have not always been successful in attaining this objective. In his overview of the development of generative linguistics Chomsky (1978:13) notes that the emphasis was initially on descriptive adequacy. Descriptive adequacy, in contrast to explanatory adequacy, often seems to require elaborating the available theoretical mechanisms, and thus extending the class of available grammars. So, for instance, the concept of a transformational rule was introduced precisely because of its strong expressive power. However, the initial concept of a transformational rule made available a very broad class of potential grammars. The basic goal of explanatory adequacy was therefore "left remote", as Chomsky (1978:14) puts it. Much work in transformational generative grammar has been aimed at showing how the richness of descriptive devices, including transformational rules, can be reduced without a loss of descriptive adequacy.

Newmeyer (1980:175–6) observes that the need for restrictions on the power of transformational rules became particularly acute around 1970, as a result of Peters and Ritchie's work on the weak generative capacity of transformational grammars. Peters and Ritchie showed that transformational grammar, as formulated then, made only one weak claim about human language, namely that its sentences could be generated by some set of rules. According to Chomsky (1978:15), a fruitful approach
to the problem of limiting the expressive power of transformational rules, and thus increasing explanatory adequacy, has been the study of general conditions on the functioning of transformations. Conditions on the functioning of such rules do not lead directly to a reduction in the class of grammars. They do, however, indirectly lead to such a reduction, by permitting the class of possible rules to be restricted. In particular, such conditions make it unnecessary for individual rules to be richly articulated. Rule-specific conditions can, for example, be eliminated.

The introduction of the SSC and TSC in (Chomsky 1973) must be seen against the background sketched above. The SSC and TSC (and the other conditions discussed in (Chomsky 1973)) are all conditions on the functioning of transformational and/or interpretive rules --- see Chomsky 1973:234. If successful, these conditions were expected to contribute to explanatory adequacy by making it possible to restrict the class of transformational rules and the class of interpretive rules.

2.2.3 The general nature of the motivation for the SSC and TSC

From what has been said in §2.2.2, it follows that the justification for adopting conditions such as the SSC and TSC must take the form of an illustration that the incorporation of these conditions in UG makes it possible to restrict the expressive power of transformational rules (and rules of semantic interpretation). The first steps towards such an illustration are taken in Chomsky (1973). Firstly, it is shown that a large number of transformational rules and rules of semantic interpretation obey the SSC and TSC. Secondly, a limited number of claims are made about reductions in the expressive power of transformations that can be effected if the SSC/TSC is adopted, or increases in expressive power that can be avoided.

The evidence that a large number of rules obey the SSC and TSC supports the claim that these conditions are general conditions on rules. If it can be maintained that the SSC and TSC (together with additional conditions) are general --- i.e., universal --- conditions on rules, then it may become possible to eliminate all rule-specific conditions, and
so to reduce the expressive power of transformations. The rules of which it is claimed in (Chomsky 1973) that they obey the SSC and TSC are the following.

(i) The Passive transformation

Chomsky (1973:237) proposes that Passive has the structural description (X, NP, V, NP, Y), and that it rearranges the NPs. The examples in (15) illustrate that Passive obeys the TSC.

(15) a. I believe the dog is hungry
    b. *The dog is believed is hungry (by me)

In (15a) the NP the dog is extracted from a tensed sentence in order to derive (15b). Consequently, (15b) is ruled out by the TSC.

(ii) each-Movement/each-Insertion

Chomsky (1973:238) follows Dougherty in adopting a rule which derives (16b) from (16a), by moving each into the determiner position of the other(s).

(16) a. The men each hated the other(s).
    b. The men hated each other.

(In fn. 17 Chomsky (1973:238) notes that if a rule of each-Interpretation were adopted instead of a rule of each-Movement, then the relevant conditions would apply to this interpretive rule.)

The sentences in (17) and (18) illustrate that each-Movement obeys the TSC and SSC, respectively.

(17) a. The candidates each expected \( s \) that the other(s) would win
    b. *The candidates expected that each other would win
(18) a. The men each expected [\(S\) the soldier to shoot the other] \[25a\]
b. *The men expected the soldier to shoot each other \[25b\]

In (17a) \(Y (= \text{the other(s)})\) is in a tensed clause. In (18a) a lexical subject \( (= \text{the soldier})\) intervenes between \(X (= \text{each})\) and \(Y (= \text{the other(s)})\).

(iii) it-Replacement

Chomsky (1973:239) adopts a rule of it-Replacement, which derives sentences such as (19b) by moving the object of the embedded clause to the position of it.

(19) a. It is easy to please John.
b. John is easy to please.

The sentences in (20) are presented by Chomsky to illustrate that it-Replacement obeys the SSC.

(20) a. It is a waste of time for us \(S\) for them to teach us Latin] \[32b\]
b. *Latin is a waste of time for us for them to teach us \[35b\]

In (20a) the lexical subject \(\text{them}\) intervenes between \(X (= \text{it})\) and \(Y (= \text{Latin})\).

(iv) Disjoint reference

Chomsky (1973:241) adopts a rule of interpretation which, when applied to the structure NP-V-NP, seeks to interpret the NPs as nonintersecting in reference. Where this is impossible --- e.g. in the case of
first and second person pronouns --- it assigns "strangeness". The sentences in (21) and (22) illustrate that this rule obeys the SSC and TSC, respectively. (In the case of some of the examples discussed below, I indicate more structure than Chomsky does.)

(21) \( \text{We expected } [S_{thems} \text{ to visit me}] \) \( [45a] \)

(22) \( \text{We believe } [S_{I} \text{ may still win}] \) \( [45d] \)

In (21) the application of the rule is blocked by the presence of the lexical subject them. In (22) the rule is blocked because \( Y (= I) \) is in a tensed clause.

(v) **The rule associating not and many**

Chomsky (1973:242) leaves open the question of whether the scope of negation in sentences such as (23) is determined by a syntactic transformation that extracts not from the NP object, or by an interpretive rule.

(23) a. I didn't see many of the pictures \( [46a] \)

b. I didn't see pictures of many of the children \( [46b] \)

Chomsky (1973:242) claims that, whatever the nature of the relevant rule, it obeys the SSC, as illustrated by (24).

(24) \( \text{I didn't see } [NP_{John's \text{ pictures of many of the children} }] \) \( [47] \)

In (24) the lexical subject John prevents the rule from associating not and many.
(vi) The rule associating not and enough

Chomsky (1973:242) tentatively adopts a rule which associates enough and not in sentences such as (25).

(25) You didn't understand the proofs of enough of the theorems (for me to be justified in giving you an A) [48a]

The sentence in (26) illustrates that this rule obeys the SSC.

(26) You didn't understand Euclid's proofs of enough of the theorems (for me to be justified in giving you an A) [48b]

The lexical subject (= Euclid) of the NP prevents the rule from associating not and enough. (26) thus receives no direct interpretation, according to Chomsky (1973:242).

(vii) The respectively-Interpretation rule

Chomsky (1973:261) briefly refers to the respectively-Interpretation rule, which associates respective with the matrix subject in sentences such as (27).

(27) We will obey any request to kiss our respective wives [154a]

The sentence in (28) illustrates that this rule obeys the SSC.

(28) *We will okay any request to kiss our respective wives [154b]

The embedded sentence in (28) has a PRO subject, which is not controlled by X (= we).
A large part of the discussion in (Chomsky 1973) is devoted to wh-Movement. The applicability of the SSC and TSC to this rule will be discussed in detail in §§2.2.4.1 and 2.2.4.2 below.

Let us now briefly consider the specific claims made in (Chomsky 1973) about the expressive power of transformations. The first such claim is made in connection with the Passive transformation. Chomsky (1973:237) claims that the TSC makes it possible to adopt the principle of blind application of transformational rules, at least for Passive. A transformational rule is said to apply "blindly" if it applies without regard to the semantic and grammatical relations holding between the constituents of a sentence. If Passive --- with the structural description (X, NP, V, NP, Y) --- applies in accordance with the principle of blind application, then Passive will overgenerate. In particular, ungrammatical sentences such as (15b) will be generated. The TSC, however, blocks the derivation of such sentences. The adoption of the principle of blind application entails a considerable reduction in the expressive power of transformations in comparison with the assumptions made in the literature at that stage. Postal (1976:151, fn. 7), for example, provides a long list of works in which transformational rules are presented that refer to the grammatical notion of 'subject', and so violate the principle of blind application. If the TSC does indeed make it possible to adopt this principle, then this fact provides strong support for the TSC.

Chomsky (1973:237, fn. 15) briefly considers, and rejects, two alternative solutions to the problem of overgeneration by Passive. Both these alternatives are primarily rejected because, unlike the TSC-solution, they do not lead to a reduction in the expressive power of transformations. The first alternative is to retain the notion of blind application of transformational rules, and to add a rule-specific condition to the Passive transformation. This solution is rejected because, according to Chomsky, it represents precisely the kind of enrichment of transformational theory that must be eliminated if explanatory adequacy is to be attained.
A second alternative solution to the problem of overgeneration by Passive lies in giving up the principle of blind application, and defining Passive in terms of relational notions such as 'subject' and 'object'. Passive will then fail to apply in (15a), because the subject of the embedded clause has not been raised to the object position of the matrix clause.

There are two reasons for Chomsky's rejection of this second alternative solution. Firstly, allowing structural descriptions to refer to relational terms entails that transformational rules have greater expressive power than would be the case under blind application. This second solution is thus less attractive in view of the guiding principle of reducing the class of grammars. In the case under discussion, Chomsky claims that there is no empirical motivation for the proposed revision.

The second reason why Chomsky rejects the alternative with Passive defined in relational terms, is that there are empirical objections to such a reformulation. Chomsky (1973:237, fn. 15) claims that this reformulation of Passive would be "ill-advised in the case of Passive because of pseudo-passives (see the discussion following (14)), double passives such as (14), indirect object constructions, and so on". These constructions indicate that constituents other than direct objects can also be moved by Passive. They consequently constitute potential counterexamples to a reanalysis of Passive in relational terms. The second solution to the problem of overgeneration by Passive thus fails in two respects. It not only leads to an increase in the expressive power of transformational rules (thus conflicting with the guiding principle of reducing the class of grammars). It also fails to attain descriptive adequacy.

Chomsky (1973:88) makes a further specific claim about the expressive power of transformations, this time in connection with an alternative to the SSC. In §8 (and §9) Chomsky argues for the extension of the SSC to include the case where the subject of the embedded clause is controlled by a category not containing X. One of Chomsky's arguments for such an extension of the SSC concerns the sentences presented in (29).
(29)  a. We each persuaded Bill [COMP PRO to kill the others(s)]

b. *We persuaded Bill to kill each other

(29b) is blocked by the SSC as formulated in (5) above. PRO is controlled by Bill, i.e., it is not controlled by the category containing X (= each). Chomsky (1973:255) also considers an alternative to the extended SSC to cover the control case. The alternative is to restrict each-Movement to a single clause. One would then in fact be assuming that (30b) is derived from (30a), and (31b) from (31a).

(30)  a. We promised Bill [COMP PRO each to kill the other(s)]

b. We promised Bill to kill each other

(31)  a. We wanted [COMP PRO each to kill the other(s)]

b. We wanted to kill each other

Chomsky (1973:255) points out two empirical problems arising from the adoption of this alternative to the extended SSC. Firstly, if (31a) must be derived from (31b), then it becomes impossible to block the sentences in (32).

(32)  a. *We each wanted to kill each other

b. *We would have both wanted to kill each other

Secondly, if each-Movement is restricted to a single clause, then it becomes impossible to derive the sentences in (33).

(33)  a. We like [S [NP pictures of each other] to be on sale]

b. They expect [S [NP each other] to win]
Apart from these empirical problems, there is a more fundamental reason why Chomsky rejects the relevant alternative solution to the problem posed by (29). Chomsky (1973:255) states this point as follows:

(34) "Furthermore, it would be highly undesirable to extend the general theory of transformations so as to permit transformations to be restricted to a single clause, and so far as I can see, there are no strong empirical reasons motivating such an elaboration of the theory, given the general framework that we are exploring here."

In fn. 34 Chomsky clearly indicates how much weight he attaches to the consideration of restricting the theory of transformations.

(35) "In the absence of other considerations, the general point that the theory of transformations should not be extended to permit this option is compelling, if not decisive."

Chomsky's rejection of the restriction of each-Movement to a single clause as a possible solution to the problem raised by (29), is thus based on two types of considerations. Firstly, the restriction poses some empirical problems. Secondly, such a restriction leads to an extension of the expressive power of transformations. Such an extension conflicts with the guiding principle of restricting the class of grammars. The second consideration also applies in the case of disjoint reference rule RI of (Chomsky 1973) discussed by Chomsky on p. 256-7.

In sum: The justification for the SSC and TSC presented in (Chomsky 1973) takes the form of an illustration that a large number of transformational and interpretive rules obey these conditions. Claims are also made that these conditions make it possible to restrict the expressive power of transformations, and that they make it possible to avoid alternatives that either lead to an increase in the expressive power, or that fail to lead to a reduction in this power.
2.2.4 Dealing with potential empirical problems for the SSC and TSC

2.2.4.1 Wh-Movement, Strict Cyclicity and the COMP-escape hatch

Chomsky (1973:243) points out that wh-Movement in (36b) poses a problem for both the SSC and TSC.

(36) a. COMP you told me \[ S \text{ COMP Bill saw something} \] \[50\]
b. What did you tell me that Bill saw \[49\]

Movement of the wh-phrase from the embedded sentence into the COMP position of the matrix clause violates both the SSC (because the embedded clause has a specified subject Bill) and the TSC (because the embedded sentence is tensed). wh-Movement in (36) thus constitutes a potential counterexample to the SSC and TSC.

On p. 243 Chomsky argues that if it is assumed that wh-Movement applies cyclically, then wh-Movement in (36) no longer represents a problem for the SSC. Chomsky (1973:243) formulates the Strict Cycle Condition as follows:

(37) "No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node."

It follows from the Strict Cycle Condition that wh-Movement must be a cyclic rule, since it applies in indirect questions and relative clauses. Given the cyclic nature of wh-Movement, (36b) will be derived from (36a) only via the intermediate stage (38), with wh-Movement on the innermost cycle.

(38) COMP you told me \[ S \left[ \text{COMP what} \right] \text{ Bill saw} \] \[52\]

(36b) is derived by movement of what from the COMP position of the innermost cycle to the COMP position of the matrix clause. The latter movement does not
violate the SSC as formulated in (Chomsky 1973) --- see (5) above. The specified subject in θ --- i.e. Bill --- no longer intervenes between X and Y. Thus, cyclic application of wh-Movement overcomes the problem which wh-Movement in (36b) poses for the SSC.

The derivation of (36b) from (38) is a case where a Y which is in COMP is moved out of a tensed S. Chomsky (1973:243) has the following to say about the conditions under which an item can be moved out of a tensed S.

(39) "An investigation of the conditions of the violation indicates that they are quite narrow: an item can 'escape' from a tensed sentence if it has been moved into the COMP position on an earlier cycle and is moving into the COMP position on the present cycle. Furthermore, in no case does an item in COMP position move to anything other than the COMP position. These specific properties of COMP may be considered alongside the property formulated as the Complementizer Substitution Universal."

The Complementizer Substitution Universal stipulates that only languages with clause-initial COMP permit COMP-substitution transformations. (See Chomsky 1973:234). Chomsky (1973:244) adopts the following base rules for English.

(40) S \rightarrow COMP S'

[54]

S' \rightarrow NP AUX VP

He takes S, and not S', as the domain of cyclic operations. The stipulation in the TSC that Y is not in COMP --- see (6) above --- thus makes it possible for an element in COMP to escape from a tensed clause. This possibility has become known as the "COMP-escape hatch."

2.2.4.2 The SSC and traces

Chomsky (1973:§10) considers various auxiliary assumptions to supplement the SSC, in order to increase the descriptive adequacy of the grammar of English. In doing so, Chomsky takes the important step of introducing
the notion 'trace' to deal with potential empirical problems to his proposed conditions.

The crucial cases in §10 are those in which X (in the structure \[\ldots \ X \ldots [\ldots \ Z \ldots - WYV \ldots \ldots \ldots ] \ldots \) is not a possible controller. Two such cases are distinguished: X = \textit{it}, as in the case of \textit{it}-Replacement, and X = COMP, as in the case of \textit{wh}-Movement.

Consider the following sentences, with X = \textit{it}.

(41) a. It is pleasant for the rich [\textit{COMP PRO} to do the hard work] [164a]
   b. The hard work is pleasant for the rich to do [164b]

(42) a. It is tough for me [\textit{COMP PRO} to stop [\textit{COMP PRO looking at Harriet}]] [166b]
   b. Harriet is tough for me to stop looking at [167b]

In (41a) X = \textit{it}, Y = the hard work, and Z (= PRO) is controlled by the rich. The SSC, as formulated in (5) above, wrongly predicts that (41b) is unacceptable, because Z is not controlled by a category containing X. The same is true in (42a), with X = \textit{it}, Y = Harriet, and Z (= PRO) controlled by me.

Chomsky (1973:262f) considers two possible solutions to the problem posed by sentences such as (41) and (42). A first solution is to supplement subcase [160b] of the SSC --- i.e., where Z is controlled by a category not containing X --- with the provision [161], presented as (43) below.

(43) "where the minimal major category containing X (i.e., MMC (X)) is a possible controller." 23) [161]

The addition of the provision (43)/[161] to subcase [160b] of the SSC leads to the prediction that (41b) and (42b) are acceptable. In
both cases the minimal major category containing X is not a possible controller.

A second possible solution is to adopt a rule of PRO-Replacement, which moves the NP the hard work in (41a) to the position of PRO on the internal cycle. The structure (44) will be derived.

\[
\text{(44) It is pleasant for the rich } [S \text{ COMP PRO the hard work to do}]
\]

It-Replacement can then extract the NP the hard work from the embedded clause because the structure in (44) does not satisfy the conditions of application of the SSC. A similar analysis can be made in the case of (42b).

Chomsky (1973:264) adopts this second possible solution to the problem posed by sentences such as (41) and (42). There are two arguments against adding provision (43) to subcase [160b] of the SSC. Firstly, this would lead to wrong predictions in the case of sentences such as (45a, b).

\[
\text{(45) a. } *\text{Who did John make a fortune by cheating}\]
\[
\text{b. } *\text{Where did John make a fortune while living}
\]

In (45a) and (45b) X is COMP, which is not a possible controller. If (43) were added to subcase [160b] of the SSC, the SSC would thus wrongly predict that (45a) and (45b) are acceptable.

Secondly, if (41) were added to subcase [160b] of the SSC, then the extraction of Harriet in (42a) and (46a) would violate the Subjacency Condition on extraction rules.

\[
\text{(46) a. It is tough for me } [\text{COMP PRO to stop Bill from looking at Harriet}]
\]
\[
\text{b. Harriet is tough for me to stop Bill from looking at}
\]
The rule of PRO-Replacement overcomes both the problems faced by provision (43)/[161]. Chomsky (1973:264f.) mentions three additional advantages of incorporating a rule of PRO-Replacement in the grammar of English. Firstly, addition of the rule of PRO-Replacement permits, at very little extra cost, the generalization of an obligatory rule already needed in the grammar, namely the rule that derives (47) from (48).

\[(47)\]
\begin{align*}
  a. & \text{John is likely to leave.} \\
  b. & \text{John seems to be a nice fellow.}
\end{align*}

\[(48)\]
\begin{align*}
  a. & \text{It is likely } \text{[S COMP John to leave]} \\
  b. & \text{It seems } \text{[S COMP John to be a nice fellow]}
\end{align*}

Thus, the addition of PRO-Replacement to the grammar of English contributes very little to the overall complexity of this grammar.

Secondly, adopting a rule of PRO-Replacement, with the restriction that it is obligatory under certain circumstances (see Chomsky 1973:265), allows some of the restrictions on it-Replacement to be explained. In particular, these restrictions can be explained in terms of the ordering of Passive and PRO-Replacement. Thirdly, adding PRO-Replacement can account for certain observations by Bresnan about stress contours.

However, as Chomsky (1973:265f.) shows, adopting a rule of PRO-Replacement also poses some problems. In particular, a grammar incorporating a rule of PRO-Replacement has difficulties in making the correct predictions about the sentences in (49), which are derived from (50).

\[(49)\]
\begin{align*}
  a. & \text{*The men are easy for each other to please} \\
  b. & \text{*John seems to the men to like each other} \\
  c. & \text{Toys are fun for the kids to give each other (?)}
\end{align*}

\[(50)\]
\begin{align*}
  a. & \text{It is easy for the others } \text{[COMP PRO to please each of the men]}
\end{align*}
b. It seems to each of the men [COMP John to like the others] [174b]

c. It is fun for each of the kids [COMP PRO to give toys to the others] [174c]

Chomsky claims that (49b) is the worst, with (49c) better than (49a). He (1973:266) points out that if each-Movement were ordered before it-Replacement, the correct predictions would be made about (49a) and (49b). Problems arise in connection with (49c). Suppose that the provision (43)/[161] is added to the SSC. The derivation of (49c) would then proceed as follows. On the innermost cycle, Indirect Object Movement gives (51).

(51) It is fun for each of the kids [COMP PRO to give the others toys] [175]

Each-Movement, followed by it-Replacement, applies on the external cycle to derive (49c). Application of each-Movement is permitted because the phrase each of the kids, which contains X (= each) controls PRO. In the case of it-Replacement X (= it) is not a possible controller. The provision (43)/[161] thus permits the application of the latter rule.

Suppose now that the rule of PRO-Replacement is adopted instead of the provision (43)/[161]. On the innermost cycle, PRO-Replacement gives (52).

(52) It is fun for each of the kids [COMP toys to give the others] [176]

Assuming still that each-Movement precedes it-Replacement, then each-Movement must apply to (52) at this point if (49c) is to be derived. However, each-Movement will apply only if the position of PRO in (51), now occupied by the complex structure [toys, PRO] (= toys) in (52),
is still controlled by the phrase each of the kids which appears in the matrix sentence of (52). If this position is no longer controlled by each of the kids, then the SSC --- now without the provision (43)/[161] --- blocks the application of each-Movement. In order to derive (49c) control must thus be regarded as an enduring property of the paired positions in cases like (52), where PRO-Replacement has created a complex structure consisting of PRO plus lexical item in the embedded subject position. Also, the SSC must be reformulated slightly so as to ensure that a position is not considered to be lexically specified if it is controlled. Chomsky (1973:266) remarks that "these consequences, while not intolerable, nevertheless do not seem to me particularly desirable".

In view of these undesirable consequences, Chomsky rejects the assumption that each-Movement precedes it-Replacement. No problem then arises in connection with (49c). In order to account for (49b), Chomsky assumes that when the NP John replaces it in (49b), it leaves behind a "trace" which it controls. This controlled trace will then block the application of each-Movement, so that (49b) is unacceptable. However, as Chomsky (1973:267) points out, the trace approach does not work in the case of (49a). He (1973:267) sums up the preceding discussion as follows:

(53) "I have explored the interconnections among various assumptions, reaching no firm conclusion. It seems reasonable to make the tentative assumption that PRO-Replacement operates and that it-Replacement in (174b) [= (49b) --- M.S.] leaves a 'trace', and, finally, that we can dispense with the qualification (161) [= (43) --- M.S.] and preserve the principle (81) of Subjacency. On this assumption we leave unexplained the ungrammaticalness of (173a) [= (49a) --- M.S.] .... but all of the other cases examined fall into place."

Chomsky then proceeds to extend the trace approach to wh-Movement, which exemplifies the second case where X (= COMP) is not a possible controller. Consider the sentences in (54) and (55).
(54)  a. COMP they expected [COMP Bill to kill who] [180]
    b. Who did they expect Bill to kill

(55)  a. COMP they expected [COMP PRO to kill who] [181]
    b. Who did they expect to kill.

Cyclic application of wh-Movement in (54a) and (55a) will yield the correct results, without the addition of provision (43)/[161] to the SSC. Consider now the following sentences.

(56)  a. COMP they each expected [COMP who to kill the
    others] [182]
    b. Who they each expected to kill the others [187]
    c. *Who did they expect to kill each other [188]

Cyclic application of wh-Movement in (56a) gives (56b). each-Movement can then apply to give (56c). However, (56c) does not have the interpretation of (56b). To overcome this problem, Chomsky (1973:268) first considers ordering each-Movement before wh-Movement. This assumption will block the derivation of (56c) from (56b), with the subject position of the embedded clause filled by who. The assumption does not, however, suffice in general. Consider the derivation in (57)/[189].

(57)  a. COMP Bill wanted [COMP they each to expect [COMP
    who to kill the others]]
    b. COMP Bill wanted [COMP they each to expect [who to
    kill the others]]
    c. COMP Bill wanted [COMP they to expect who to kill
    each other]
    d. COMP Bill wanted [who they to expect to kill each
    other]
    e. *Who did Bill want them to expect to kill each other
On the innermost cycle \( \text{wh-Movement} \) applies to (57a) to give (57b). On the next cycle, assuming that \( \text{each-Movement} \) precedes \( \text{wh-Movement} \), (57c) is first derived by applying \( \text{each-Movement} \), and then (57d) by \( \text{wh-Movement} \). On the last cycle, \( \text{wh-Movement} \), together with the obligatory rules of Auxiliary Inversion and Case Assignment, derive (57e). The assumption that \( \text{each-Movement} \) precedes \( \text{wh-Movement} \) therefore does not suffice to rule out the derivation of sentences such as (57e).

Chomsky (1973:269) consequently rejects the assumption that \( \text{each-Movement} \) precedes \( \text{wh-Movement} \). Instead, he assumes that \( \text{wh-Movement} \) --- like \( \text{it-Replacement} \) --- leaves behind a trace which, in the case of \( \text{wh-Movement} \), is controlled by the moved \( \text{wh-phrase} \). In (57b) \( \text{who} \) will thus control its trace in the subject position of the lowest embedded clause. Because of the presence of this controlled subject, the SSC will prohibit \( \text{each-Movement} \) from moving \( \text{each} \) into this embedded clause. The derivation of (57e) will then be blocked. Similarly, in the case of (56b) the trace of \( \text{who} \) will prevent the application of \( \text{each-Movement} \) to give (56c). The assumption that \( \text{wh-Movement} \) leaves behind a controlled trace thus makes it possible for the SSC to block the derivation of (56c) and (57e).

Chomsky (1973:269) claims that the assumption that \( \text{wh-Movement} \) leaves behind a trace permits a fairly simple rule of interpretation for \( \text{wh-questions} \). (See the discussion in Chomsky 1973:282f.) In fn. 49 he also claims that Emonds' observation concerning \( \text{NP-Preposing} \) --- viz. that it is obligatory in sentences but optional in noun phrases --- can be explained by assuming that all rules which move items from obligatory categories leave traces.

### 2.2.4.3 The SSC and the feature \([+\text{definite}]\)

Chomsky (1973:239, fn. 19) proposes a possible refinement of the SSC. In terms of this proposal the SSC must also incorporate the feature \([+\text{definite}]\). The incorporation of this feature makes it possible to explain a three-way gradation of acceptability with respect to \( \text{wh-Movement} \) from NPs. The data are as follows:
(58)  a. COMP you saw [\text{NP pictures of who}] [30a] 
b. Who did you see pictures of [30b]

(59)  a. COMP you saw [\text{NP the pictures of who}] 
b. Who did you see the pictures of

(60)  a. COMP you saw [\text{NP John's pictures of who}] [31a] 
b. *Who did you see John's pictures of [31b]

(58b) is perfectly acceptable, (59b) less acceptable, and (60b) completely unacceptable. The SSC, formulated as in (5) above, cannot account for the fact that (59b) is less acceptable than (58b), but more acceptable than (60b). Chomsky proposes that the incorporation of the feature [+ definite] in the SSC can solve this problem. If the SSC were to include the specification [+ definite], then (60b) would involve a double violation of the SSC (given that lexical subjects are [+ definite]). The distinction between a double violation, a single violation and no violation of the SSC could then account for the differences in acceptability exhibited by (58b), (59b), and (60b).

2.2.4.4 The SSC and the notion 'agency'

Consider the following sentences from (Chomsky 1973:261).

(61)  a. Why are John and Mary letting the honey drip on each other's feet [155] 
b. *Why are John and Mary letting Bill drip honey on each other's feet [156] 
c. Why are they letting the baby fall on each other's laps [157] 
d. *Why are they letting Bill drop the baby on each other's laps [158] 

The sentences in (61) are the result of the application of each-Movement. The unacceptability of (61b) and (61d) follows from the SSC. Both
these sentences contain a specified subject --- Bill --- which blocks each-Movement. In (61a) and (61c) there are also specified subjects: the honey in (61a) and the baby in (61c). Nevertheless the application of each-Movement is not blocked in (61a) and (61c). These applications of each-Movement thus constitute apparent violations of the SSC.

Chomsky (1973:261) proposes a possible modification of the SSC that will overcome the problem raised by each-Movement in (61a) and (61c). In terms of Chomsky's proposal, the notion 'specified agent' must replace the notion of a formal subject in the SSC. If the SSC is re-formulated in terms of the notion 'specified agent', then the application of each-Movement in (61a) and (61c) no longer constitutes an apparent violation of the SSC. The honey in (61a) and the baby in (61c) are not agents, even though they are specified subjects.

In fn. 37 Chomsky (1973:257) hints that the reformulation of the SSC in terms of the notion 'agency' would also make it possible to account for the difference in acceptability between (62a) and (62b).

(62) a. The men wanted to tell stories about killing each other.
   b. The men wanted to hear stories about killing each other.

According to Chomsky (62a), with the men the understood subject of kill and a relation similar to semantic agency between the men and stories, is more "natural" than (62b). In (62b) the men is not in a relation of agency to stories. Chomsky does not spell out exactly how the proposed reformulation of the SSC can account for the difference between (62a) and (62b). Presumably the explanation runs more or less as follows. In (62a), but not in (62b), the phrase containing each in the underlying structure --- the men each --- is also the agent of the NP [stories about killing the other(s)]. The agent of the latter NP in (62b) thus qualifies as a specified agent. each-Movement is thus blocked by the modified SSC in (62b), but not in (62a).

Chomsky (1973:261) briefly refers to the relevance of the proposed modification of the SSC in terms of the notion 'agency' for "the hypothesis
that transformations do not refer to semantic relations but only to the bracketing of phrase markers". He claims that if the SSC were so modified, the relevant hypothesis would remain unaffected. Similarly, the incorporation of the semantic notion 'agency' in the SSC would not affect this hypothesis.

2.2.4.5 The unsolved problem of Coreference Assignment

Chomsky (1973:238, fn. 16) notes that Coreference Assignment does not obey the TSC. Coreference Assignment is the rule that relates the NP John and the pronoun he in sentences like (63) with John and he interpreted as coreferential.

(63) John said [that he would leave]

In (63) the pronoun he is within a tensed clause. Coreference Assignment is thus a potential counterexample to the TSC.

Chomsky also notes that Coreference Assignment applies in other structures, e.g. coordinate structures, in which various other types of rules are blocked. The application of Coreference Assignment in (64) conflicts with the Coordinate Structure Constraint.

(64) John said that he and Bill would leave

Coreference Assignment is thus problematic with respect to conditions other than the TSC as well. Chomsky (1973) does not take any steps to solve the problem which Coreference Assignment poses for the TSC.

2.2.5 The relative interpretation of conditions on rules and the possibility of parametric variation

Chomsky (1973:235-6) distinguishes two possible interpretations of conditions on rules. On the absolute interpretation, a condition imposes an absolute restriction on rules of a certain type. On this interpre-
tation the TSC, for example, legislates against any rule that extracts
an element from a tensed clause, or that moves an element into a tensed
clause. On the other interpretation, a condition does not impose an
absolute restriction on rules of a certain type. Rather, a rule must
be interpreted in accordance with the condition unless otherwise speci-

died. Thus, one might formulate a rule that violates the TSC, but only
at a cost. The rule would have to explicitly mention that it may move
an element into or from a tensed clause. Chomsky (1973:236) remarks
that the logic of the second approach to the interpretation of condi-
tions on rules "is essentially that of the theory of markedness". The
significance of the adoption of the relative interpretation of condi-
tions such as the SSC and TSC is obvious. It now becomes possible
to maintain a condition as a general condition, while at the same
time admitting that certain rules violate the condition. In (Chomsky
1973) this idea does not yet play any special role. In later works
the notion of marked exceptions to general conditions plays an increa-
singly important role, as will be shown below.

Chomsky (1973:238, fn. 16) briefly mentions a second notion that has
come to play a very important role in more recent work. With reference
to the TSC, Chomsky remarks that one could make the weaker assumption
that \( \infty \) in the TSC is a language-specific parameter. That is, while
\( \infty \) has the value "tensed S" in English, it may have other values in
other languages. Again, this notion enables Chomsky to formulate
rules which violate a general condition on rules, without rejecting
the condition. The importance of this notion in more recent work,
specifically as it relates to the TSC and SSC, will be discussed below.

2.2.6 The naturalness of the SSC

Chomsky (1973:270) claims that the SSC has "a certain naturalness".
In particular, Chomsky observes that the SSC, in some cases, "has the
effect of reducing ambiguity, or, to put it differently, of increasing
the reliability of a reasonable perceptual strategy that seeks the
nearest NP to a verb (or the head noun of a nominal phrase) as its
subject". So, for example, the SSC implies that (65) must have the
interpretation indicated in (66a), but not that indicated in (66b).

(65) The men expected [the police to arrest each other] \[191\]

(66) a. The men expected [the police each to arrest the other(s)] \[192\]

b. The men each expected [the police to arrest the other(s)] \[193\]

(65) cannot be derived from (66b), because of the presence of the specified subject the police.

If, contrary to the assumption made above, the deep structure position of each plays no role in the interpretation of a sentence, then the SSC will guarantee a correspondence between deep structure position and scope as determined by surface structure interpretation rules. The latter consequence is characterized by Chomsky (1973:270) as "rather natural".

The requirement that general linguistic principles should be natural, is emphasized in recent works by Chomsky, such as (Chomsky 1978a). These will be discussed below.

2.3 Other relevant aspects of the SSC and TSC in the first stage of their development

Apart from (Chomsky 1973), there are a few other works by Chomsky in which the SSC and TSC are presented as conditions that restrict both syntactic transformations and rules of semantic interpretation. In §2.3 the various aspects of the two conditions dealt with in these works will be discussed.
2.3.1 Additional motivation for trace theory

Chomsky (1973:266 f.) introduces the notion of 'trace' to enable the SSC to apply to a wider class of cases. In works that follow (Chomsky 1973), it is also assumed that traces increase the generality of the SSC — see, e.g., Chomsky 1975a:102 f., Chomsky 1976a:320 f. Consider, for example, the following sentences presented in (Chomsky 1976a).

(67) The men like each other

(68) the men want [John to like each other]

(69) the men seem to John [t to like each other]

(70) John seems to the men [t to like each other]

(67) is analogous to (69), and (68) is analogous to (70). The rule of Reciprocal Interpretation \( \text{R} \) applies in (67). Similarly, it applies in (69), as if there is no specified subject in the embedded sentence, \( t \) being the trace of the men. The SSC blocks the application of the reciprocal rule in (68), because of the presence of the specified subject John in the embedded sentence. Similarly, the SSC blocks the application of the reciprocal rule in (70) with \( t \), the trace of John, acting as the specified subject. Chomsky (1976a:321) concludes on the basis of such examples that "the trace theory thus permits otherwise valid conditions to apply, again overcoming cases of misapplication of rules: overgeneration in the case of the reciprocal rule ...".

While the presentation in (Chomsky 1973) creates the impression that traces are merely mechanisms to extend the applicability of the SSC, this is not the case. In the works that follow (Chomsky 1973), trace theory forms an integral part of linguistic theory, with a number of important implications. Chomsky (1975a:93 f.), for example, mentions two independent considerations in support of trace theory. Firstly, under trace theory all semantic interpretation can take place at surface structure. Secondly, trace theory makes it possible to explain where downgrading rules are possible. Thus, there is independent confirmation of trace theory, apart from the fact that it allows a mere...
general statement of the SSC. \[30\) It will also be shown below --- see §3.2 --- that trace theory, in turn, has led to certain changes in the interpretation of the SSC and TSC.

2.3.2 The idealization of sentence grammar

Chomsky (1975a:105) distinguishes between sentence grammar, and non-sentence grammar. The SSC and TSC form part of sentence grammar. Consequently, rules that fall outside sentence grammar need not obey these conditions, and so cannot constitute potential counterexamples to them. Koster (1978:566) makes it clear exactly how the idealization of sentence grammar enables the linguist to deal with some potential counterexamples to the SSC and TSC.

(71) "The conditions proposed in Chomsky (1973) are seen as part of an idealization, sentence grammar. When the theory incorporating these conditions is confronted with problematic evidence, a crucial question is whether the rules that account for the problematic data have to fall within sentence grammar. Sometimes this question has been answered with 'no', and one of the major results has been that we are now able to make a principled distinction between sentence grammar and discourse grammar. \[27\]"

Koster relates the introduction of the idealization of sentence grammar to a general movement in Chomskyan linguistics, from what he calls "dataism" towards deeper explanations. In his recent works, Chomsky has repeatedly emphasized the great importance attached to depth of explanation as opposed to the relative unimportance of gross coverage of data. Chomsky regards the making of idealizations as an important tool in increasing explanatory depth. \[32\)

In the early works dealing with the SSC and TSC, Chomsky invokes the idealization of sentence grammar in three cases where rules violate the SSC and TSC. A first case concerns so-called "Picture Noun Reflexivization". Chomsky (1976a:316, fn. 23) refers to this rule as an "unsolved problem" for his conditions. Postal (1976:172) lists Picture Noun Reflexivization in sentences such as (72) as a potential counterexample to the SSC.
(72) Mike will not believe that this is a picture of himself.

The lexically specified subject *this* intervenes between Mike and himself.

Chomsky notes that Picture Noun Reflexivization "so far resists analysis under any general theory known to me". Fiengo and Lasnik (1976: 190) observe that Picture Noun Reflexivization violates a number of proposed conditions on rules, including the Complex NP Constraint, the Coordinate Structure Constraint, the Tensed-S Condition, and the Sentential Subject Constraint. Chomsky (1976a:316, fn. 23) tentatively suggests, following Helke (1971), that reflexivization in English consists of two parts: a process of bound anaphora subject to the conditions of sentence grammar, and another "more general" process that falls outside sentence grammar. The fact that Picture Noun Reflexivization resists analysis and violates a number of proposed conditions on rules, supports the hypothesis that this reflexivization process falls outside sentence grammar. If Picture Noun Reflexivization were outside sentence grammar, then it could no longer constitute a potential counterexample to the SSC or TSC (or any other condition of sentence grammar).

A second case where Chomsky invokes the idealization of sentence grammar to deal with an empirical problem for the SSC, concerns the rule that assigns an interpretation to the others. Chomsky (1976a:321 f.) compares the following two sets of sentences.

(73) a. the men like each other
b. the men want [John to like each other]
c. the men seems to John [t to like each other]
d. John seems to the men [t to like each other]

(74) a. each of the men likes the other(s)
b. each of the men wants [John to like the other(s)]
c. each of the men seems to John [t to like the other(s)]
c. John seems to each of the men [t to like the other(s)]
The pair (each of the men, the other(s)) is similar in meaning to the pair (the men, each other). The sentences in (73) thus correspond to the sentences in (74). While all the sentences in (74) are acceptable, only the (a) and (c) sentences of (73) are acceptable. The unacceptability of (73b) and (73d) can be explained on the basis of the SSC. The Reciprocal rule, which relates the men and each other, is blocked in these sentences because of the presence of a specified subject in the embedded clause: John in (73b) and the trace of John in (73d). However, the rule that relates each of the men and the other(s) in (74b) and (74d) is not blocked by the SSC. This fact thus constitutes potential counter-evidence to the claim that the SSC is a universal condition on rules.

As Chomsky (1976a:322) points out, it seems as if the difference between the Reciprocal rule and the rule assigning an interpretation to the other(s) forces one to formulate the SSC as a rule-particular principle. Chomsky avoids this undesirable reformulation by arguing, on the basis of the sentences (75)-(76), that there is a principled difference between the two cases.

(75)  
   a. Some of the men left today. The others will leave later.  
   b. *Some of the men left today. Each other will leave later.

(76)  
   a. Some of the articles are incomprehensible, but we expected John to understand the others.  
   b. *Some of the articles are incomprehensible, but we expected John to understand each other.

(75b) and (76b) are unacceptable. The unacceptability of (75b) shows that the Reciprocal rule is a rule of sentence grammar. Being a rule of sentence grammar, it is blocked by the SSC in (76b), because of the presence of the specified subject John. The acceptability of (75a) indicates that the rule relating the other(s) to a suitable NP is not a rule of sentence grammar. Consequently, it is not subject to the conditions of sentence grammar. In cases such as (76a) the SSC thus
does not block the rule. By arguing that the rule which assigns an interpretation to the other(s) is not a rule of sentence grammar, Chomsky avoids formulating the SSC as a rule-specific condition.

A third case where Chomsky uses the idealization of sentence grammar to overcome a problem for his conditions, concerns the rule of Coreference Assignment. Chomsky (1973:238, fn. 16) observes that this rule violates the TSC. In (Chomsky 1973) no steps are taken to solve this problem --- see the discussion in §2.4.5 above. Chomsky (1976a: 323) returns to Coreference Assignment, claiming that his observation that Coreference Assignment presents a problem for his theory "was simply an error". He now claims, following Lasnik (1976), that the rule of anaphora which (optionally) associates he/him and John in (77) is not a rule of sentence grammar.

(77) a. John thought that he would win.
    b. John thought that Bill liked him.

Notice that in (77b) the rule violates not only the TSC, but also the SSC. Because the rule of anaphora, applying in sentences such as (77), is not a rule of sentence grammar, it is not subject to conditions such as the SSC and TSC. Consequently this rule cannot present a real problem for Chomsky's theory.

In sum: The idealization of sentence grammar enables Chomsky to overcome three different problems faced by the SSC and TSC. In effect, this idealization enables Chomsky to avoid formulating these conditions as rule-particular conditions, and to retain them as general --- i.e., universal --- conditions. As a result, the theory attains a higher level of explanatory adequacy than would otherwise have been the case. The role of the idealization of sentence grammar in the development of the SSC and TSC thus clearly illustrates the role of idealizations in Chomskyan linguistics as a means of promoting explanatory depth rather than gross coverage of data.
2.3.3  The notion 'involve'

Chomsky (1976a:316, fn 22) provides the following definition of the notion 'involve' in the SSC and TSC.

(78) "In the case of a transformational rule, we may understand 'X is involved in the rule' to mean that X is changed by the rule or is a constant context for some change ....... Thus the terms involved in the rule are the factors that are not arbitrary strings, in accordance with the SD. In an interpretive rule, we may say that X and Y are involved if the rule establishes a relation of anaphora or control relating X and Y."

In the case of transformational rules there are two subcases:
(i) X is changed by the rule, and (ii) X is a constant context for some change. Chomsky refers to work by Fiengo and Lasnik (1976) for an example that falls under the second subcase. The rule in question is Q-float. While (Chomsky 1976a) does not provide any detail on the matter, (Chomsky 1977c:77 f.) contains a fairly detailed discussion of the nature of the problem posed by Q-float. Although, strictly speaking, the latter work belongs to the second stage of the development of the SSC and TSC, its discussion of Q-float will be considered in this section. There are two reasons for this. On the one hand, the problem in question is first raised in (Chomsky 1976a), and on the other hand, the issue is unaffected by the transition from the first to the second stage.

Fiengo and Lasnik (1976:188) formulate Q-float with the structural description $X, Q, NP, \{\text{NP, VP}\}, Y$. Q can then be moved to the position between the third and fourth factors. As Chomsky (1977c:78) points out, Q-float will then generate the acceptable sentences in (79) but not the unacceptable sentence (80).

(79) a. I gave the men all presents
    b. I persuaded the men all to leave
    c. I painted the houses all reddish-yellow.
Q-float, as formulated by Fiengo and Lasnik, will also generate (81).

(81) I promised the men all to leave.

The unacceptability of (81), in contrast to the acceptability of (79b), represents a potential counterexample to Fiengo and Lasnik's formulation of Q-float. Fiengo and Lasnik (1976:189 f.) show that the unacceptability of sentences such as (81) can be explained on the basis of a modified version of the SSC. They (1976:189) assume that the complements in cases such as (79b) and (81) are VPs. This assumption conflicts with Chomsky's assumption that the embedded clause in such cases has the form \([_{S \text{COMP}}_{S \text{PRO to VP}}] \). Chomsky (1977c:78) argues that the unacceptability of (81) can be accounted for by the SSC even if the latter assumption about the underlying structure of the embedded clauses is made.

The principal elements of Chomsky's proposed solution to the problem raised by the acceptability of (79b), as opposed to the unacceptability of (81), are the following.

(i) An assumption of the known control properties of promise and persuade.

(ii) An extension of the notion 'involvement' to cover adjacent constant terms, one of which is either an antecedent or anaphor and the other a constant category of the X-system. From this it follows that all and to leave in (79b) and (81) are involved in Q-float.

(iii) An assumption that PRO is a nonterminal node.

(iv) A modification of the notion 'specified subject', so that no rule can apply in the structure \(X \ldots [_{\alpha \ldots Z \ldots \text{WYV} \ldots}] \ldots \) if \(X \) and \(Y\) are involved in the rule and \(\alpha\) contains a subject not containing \(Y\) and not controlled by the category containing \(X\) or its trace. In terms of this modification the control of the subject of \(\alpha\) by the trace
of X will also permit the application of a rule involving X and Y in the relevant structure.

Given these assumptions, Chomsky is able to explain why (79b) is acceptable. Consider the following schematic representations of his analyses of (79b) and (81), respectively.

(82) \[ \text{I persuaded} \quad t \text{ the men, all } \left[ \text{PRO}\_i \quad \text{to leave} \right] \]
\[ \begin{array}{c}
\text{t} \\
\text{NP} \\
\text{Q} (= X) \\
\text{X}^n (= Y)
\end{array} \]
\[ \text{related} \quad \text{involved} \]

(83) \[ \text{I promised} \quad t \text{ the men, all } \left[ \text{PRO}\_i \quad \text{to leave} \right] \]
\[ \begin{array}{c}
\text{t} \\
\text{NP} \\
\text{Q} (= X) \\
\text{X}^n (= Y)
\end{array} \]
\[ \text{related} \quad \text{involved} \]

In (82) PRO is controlled by \text{t the men}, that is, by the category containing the trace of X. Since PRO in (82) is not specified in the appropriate sense, the SSC does not block the rule relating the NP the men and Q all. Consequently, (79b) is acceptable. In (83) PRO is controlled by I. PRO is thus not controlled by X (= Q), or by its trace. Therefore, PRO is a specified subject. The SSC thus prevents Q-float from associating the men and all in (83).

The proposed modification of the SSC, and the extension of the notion 'involve' have the status of tentative suggestions in (Chomsky 1977c). From the following remarks by Chomsky (1977c:78) it is clear that he himself is by no means convinced of the correctness of the proposed modifications.

(84) "The case is interesting in that the constant terms 'involved' are Q and VP, although the application of the rule related NP and Q. Judgments are unfortunately somewhat variable in the relevant cases and there are other possible analyses, but perhaps we can take this example at least as an illustration of the logic of the problem, and perhaps an actual illustration of the operative principles, though I am rather skeptical."
The aim of Chomsky's discussion of the problematic Q-float data is to show that these data can be handled without complicating the rule itself. Given the status of Q-float in the controversy between Postal on the one hand, and Fiengo and Lasnik on the other hand, it is quite important for Chomsky to be able to show this. Postal (1976) argues that the theory of transformations must be enriched to allow rules such as Q-float to refer to grammatical functions such as subject. That is, the principle of blind application must be rejected --- see §2.2.3 above. In particular, Postal (1976:161 f.) argues that if Q-float is formulated without reference to the notion 'subject', the rule faces numerous counterexamples. According to him, these examples are automatically accounted for if a formulation referring to the notion 'subject' is adopted.

Allowing transformational rules to refer to grammatical functions represents an undesirable enrichment of transformational theory from Chomsky's point of view. Fiengo and Lasnik (1976), who share Chomsky's views about the enrichment of transformational theory, reject Postal's argument. They argue that there is a "reasonably adequate analysis of Q-Floating" (p. 188), consistent with a more restrictive theory of transformations, that forbids reference to notions such as 'subject'. Fiengo and Lasnik's solution --- which also incorporates the SSC --- is based on the assumption that the complement of verbs like promise is a VP. Chomsky assumes that the relevant complement is an S. Consequently, he cannot simply adopt their solution to the problem posed by Q-float. Chomsky is therefore compelled to show that, within the framework of his own assumptions, the problem posed by Q-float can be handled without allowing reference to 'subject'.

Postal (1976:Appendix) actually admits that at least some of the counterexamples to a formulation of Q-float that does not refer to 'subject' can be handled by the SSC. On the basis of numerous potential counterexamples to the SSC --- see Postal 1976:172 f. --- Postal argues, however, that the SSC is neither a universal condition, nor a condition particular to English. Consequently, Postal claims that the SSC is not available as a means of ensuring that a formulation of Q-float that does not refer to 'subject' attains descriptive adequacy.
Fiengo and Lasnik, in turn, reject Postal's argument about the status of the SSC. In particular, they claim that some of the rules mentioned by Postal are counterexamples to almost all known conditions on rule applicability, and that some do not belong to sentence grammar. Because the SSC "correctly constrains the application of a wide variety of syntactic and semantic rules" (p. 190), Fiengo and Lasnik are unwilling to reject the SSC on the basis of the counterexamples cited by Postal. Chomsky obviously agrees with Fiengo and Lasnik.

2.3.4 Further restrictions on the expressive power of transformations

Recall that the aim with the introduction of conditions such as the SSC and TSC is the restriction of the expressive power of transformational (and interpretive) rules. Chomsky (1973) has claimed that the principle of blind application of transformational rules must be adopted. He (1976a:312) now proposes that this principle must be strengthened by a condition of minimal factorization. This condition disallows a structural description with two successive categorial terms unless one or the other is satisfied by a factor changed by the rule. For example, this condition rules out the structural description (85) for Passive, since only NP is changed by the rule.

(85) \((vbl, NP, AUX, V, NP, by, \#, vbl)\) [3]

According to Chomsky, Passive must now be formulated as (86a), or, equivalently, as (86b) (given Emonds' structure-preserving hypothesis).

(86) a. \((vbl, NP, vbl, NP, vbl)\)
b. Move NP.

The adoption of the condition of minimal factorization obviously leads to a drastic reduction in the expressive power of transformational rules. Clearly, however, a grammar limited to rules such as (86) will
overgenerate massively. Chomsky thus proceeds to show that this problem can, to a significant extent, be overcome by general (= universal) conditions on rules including, specifically, the SSC and TSC. Consider, for example, the following examples presented in (Chomsky 1976a).

(87)  

a. John₁ is believed [t₁ to be incompetent]  

b. John₁ is believed [t₁ is incompetent]  

c. John₁ seems [Bill to like t₁]  

The rule (86) generate all the sentences of (87), with t the trace of John. Only (87a) is acceptable, however. The rule (86), therefore, overgenerates quite drastically. The unacceptable sentences (87b) and (87c) are ruled out by the TSC and SSC, respectively. In (87b) Y (= t₁) is within a tensed S; in (87c) there is a specified subject, Bill, in the embedded clause.

While, according to Chomsky himself, universal conditions on rules constitute "the best case", he (1976a:315) points out that it is not only universal conditions that make it possible to reduce expressive power. Language-particular, or even rule-particular, conditions may also lead to a reduction in expressive power, if these conditions are regarded as parameters that must be fixed. It follows, for example, that if a language X does not obey the SSC and/or TSC, one need not abandon these conditions, or the resultant reduction in expressive power that follows from them. This point is taken up again in §3.4 below.

2.3.5 Some changes in the formulation of the SSC

Chomsky (1976a:316) formulates the SSC as follows:

(88)  

"Consider a structure of the form:

(11) ... X ... [\alpha \cdots Y \cdots ] ... X ...

Then no rule can involve X and Y in (11) .... where \alpha contains a specified subject distinct from Y and not controlled by X .... "

The formulation (88) differs from the formulation adopted in (Chomsky 1973), and presented in (5) above, in two respects. Firstly, in (5) it is stipulated that the specified subject must be to the left of Y within $\infty$, i.e., the specified subject must intervene between $X$ and $Y$. In (88) this stipulation is omitted. In (88) reference is made only to a subject distinct from $Y$. The formulation of (88) is obviously more general than that of (5). For example, while (5) allows the association of $X$ and $Y$ in (89), (88) forbids it (under the assumption that $S$ and NP are the cyclic nodes).

\[(89) \ldots X \ldots [S [\text{COMP } Y] [S \ldots Z \ldots]]\]

Chomsky (1976a) does not consider the empirical consequences of this change in the formulation of the SSC. The issue is taken up again in (Chomsky 1977c). See also the discussion in §3.5 below.

Secondly, in (5) it is stipulated that $Y$ is "not controlled by a category containing $X"$, while in (88) it is stipulated that $Y$ is "not controlled by $X"$. The stipulation in (5) concerning a category containing $X$ was needed in (Chomsky 1973) for the purpose of accommodating each-Movement.

\[(90) \begin{align*}
  a. & \text{ We each persuaded Bill [COMP PRO to kill the} \\
  & \quad \text{other(s)}] \\
  b. & \text{ We persuaded Bill to kill each other}
\end{align*}\]

\[(91) \begin{align*}
  a. & \text{ We each promised Bill [COMP PRO to kill the} \\
  & \quad \text{other(s)}] \\
  b. & \text{ We promised Bill to kill each other}
\end{align*}\]

each-Movement derives sentences such as (90b) and (91b) from structures such as (90a) and (91a), respectively. In both cases $X = \text{each}$, $Y = \text{the other(s)}$. In (90a) PRO is controlled by Bill, i.e. not by a category containing $X$. The derivation of (90b) thus violates the SSC. In (91a) PRO is not controlled by $X$. It is, however, controlled by a cate-
gory containing X, namely we each. (91b) can thus be derived. In works that follow (Chomsky 1973) --- e.g. (Chomsky 1976a) --- it is assumed that an interpretive rule associates we and each other in sentences such as (90b) and (91b). In (91b) X (= we) is then the controller of PRO. There is no longer any need to distinguish between X and a category containing X in the SSC.

2.4 Summary of main points

§§2.1-2.3 above contain a detailed description of the first stage of the development of the SSC and TSC, as reconstructed from the relevant works by Chomsky. In the present section I provide a brief summary of the main points of this description.

(i) In the first stage of their development the SSC and TSC restrict the application of both syntactic transformations and rules of semantic interpretation --- see §2.2.3.

(ii) The introduction of the SSC and TSC as principles of UC forms part of the attempt to solve the fundamental empirical problem of linguistics --- that of language acquisition --- by restricting the class of available grammars. The SSC and TSC contribute to the latter aim insofar as they make it possible to restrict the expressive power of one or more types of grammatical rules --- see §2.2.2.

(iii) Chomsky claims that the SSC and TSC make it possible to impose significant restrictions on the expressive power of transformational rules --- see §§2.2.3 and 2.3.4.

(iv) Various proposals are made about how to deal with potential empirical problems for the SSC and TSC.

a. Chomsky (1973) claims that the Strict Cycle Condition and the COMP escape hatch make it possible to overcome the problem posed by wh-Movement --- see §2.2.4.1.
b. A tentative proposal is made in (Chomsky 1973) that the SSC should incorporate the feature [+ definite] --- see §2.2.4.3.

c. A tentative proposal is made in (Chomsky 1973) that the semantic notion 'agent' should replace the syntactic notion 'subject' in the SSC --- see §2.2.4.4.

d. The idealization of sentence grammar is used in (Chomsky 1976a) to overcome the problems posed by three different rules which violate the conditions --- see §2.3.2. Earlier, one of these rules is mentioned in (Chomsky 1973) as an unsolved problem.

e. Chomsky (1976a) considers a definition of 'involve' that makes it possible to overcome the problem posed by Q-float --- see §2.3.3.

f. The assumption that movement transformations leave behind traces makes it possible to extend the applicability of the SSC --- see §§2.2.4.2 and 2.3.1.

(v) Chomsky (1973) proposes that a relative interpretation of universal conditions such as the SSC and TSC be adopted --- §2.2.5.

(vi) Chomsky (1973) allows for the possibility of parametric variation in universal conditions --- see §2.2.5.

(vii) Chomsky (1973) claims that the SSC is a "natural" condition --- see §2.2.6.

(viii) The formulation of the SSC adopted in (Chomsky 1976a) differs from the formulation adopted in (Chomsky 1973) on two points. One of these is quite minor, while the other point of difference has important consequences, which are considered in works that follow (Chomsky 1976a) --- §2.3.5.
3 The second stage in the development of the conditions

3.1 General remarks

The proposal that the SSC and TSC restrict only interpretive rules (rather than interpretive and transformational rules), is first made in (Chomsky 1976a). The proposed interpretation of the two conditions is finally adopted in "On wh-Movement" (henceforth (Chomsky 1977c)). This reinterpretation of the SSC and TSC is discussed in §3.2. In §§3.3-3.7 I discuss various other aspects of the conditions dealt with in (Chomsky 1977c). (Other aspects of the conditions dealt with in (Chomsky 1976a) have already been discussed in §2.3 above.)

From (Chomsky 1977c) onwards, Chomsky uses the term "Propositional Island Condition"/PIC to refer to the TSC. I will follow his example in the discussion below.

3.2 The reinterpretation of the SSC and PIC as conditions that restrict only interpretive rules

Chomsky (1976a:314) distinguishes two general approaches to the problem of overgeneration that results from the radical reduction in the expressive power of transformations proposed by him --- see §2.3.4 above for an illustration of this problem.

(92) "There are two general approaches to the problem of overgeneration in such cases as these: we may try to impose (I) conditions on the application of rules or (II) conditions on the output of rules, i.e., on surface structures. The latter will generally be related to rules of semantic interpretation that determine LF, under the assumptions of EST. As we will see, (I) and (II) may fall together."

The two approaches distinguished above can be illustrated with the aid of the sentences in (93). Both sentences are derived by the rule "Move NP". In both cases _ is the trace of John.
(93) a. John seems [t to like Bill]
   b. *John seems [Bill to like t]

(93b) is blocked by the SSC. Chomsky (1976a:319) points out that this can be interpreted in two ways. Assume that (93a) is derived by NP-Movement from the underlying structure "X seems [Bill to like John]", with X some kind of place-holder for NP. On one interpretation, the SSC prevents the NP-Movement rule from preposing John in "X seems [Bill to like John]" to derive (93b). On this interpretation, the ill-formed structure (93b) --- which could only result if the SSC were ignored --- is not generated at all. This interpretation represents the first general approach distinguished in (92), and is the interpretation adopted in the works discussed in §2 above.

On the second interpretation, the SSC is regarded as a condition on surface structure interpretation, but not on transformations. The NP-Movement rule then applies freely, giving both (93a) and (93b). The SSC must then filter out the ill-formed structure in some way. This can be done if the relation between an NP and its trace is regarded as a special case of bound anaphora. The SSC will then block the rule of bound anaphora in just those cases where movement leads to an antecedent-anaphor relation which violates the SSC. The SSC will thus filter out (93b), but not (93a). This interpretation represents the second approach distinguished in (92) above.

No clear choice between the two possible interpretations of the SSC (and PIC) is made in (Chomsky 1976a). Chomsky (1976a:320) states that "in principle, the two interpretations of SSC have distinct empirical consequences, but the issue is complex and it is not easy to sort out the consequences". In fn. 32 Chomsky briefly refers to a case discussed by Fiengo and Lasnik (1973), which suggests that the SSC must also govern transformational rules. He does not, however, discuss the matter.

The question arises why Chomsky would want to reinterpret the SSC and PIC as conditions on surface structure, in the apparent absence of clear empirical evidence supporting this reinterpretation. He (1976a:
345) mentions the following consideration.

(94) "Under EST, principles of semantic interpretation, at least those of SI - 1, would be expected to relate very closely, in many cases, to conditions on base and surface structures; particularly, surface structure, if the revision of EST suggested here proves to be essentially correct. It may be that conditions on transformations can be assimilated to conditions on surface structures and (in the best case) to independently motivated rules of semantic interpretation, under the trace theory."

Chomsky (1976a:336) proposes the following general structure for EST.

(95) \[ B \rightarrow \text{base structures} \rightarrow T \rightarrow \text{surface structures} \rightarrow \text{SI - 1} \rightarrow \text{LF} \rightarrow \text{SI - 2} \rightarrow \text{semantic representation} \]

The revision of EST referred to in (94), is the adoption of the principle that surface structure alone determines LF.

The assimilation of conditions on transformations to independently motivated rules of semantic interpretation would provide support for the revised EST. It is for this reason that Chomsky wants to show that conditions on movement rules can, under trace theory, be assimilated to the rule of bound anaphora.

The question of the reinterpretation of the SSC and PIC is taken up again in (Chomsky 1977c). He (1977c:74) formulates the SSC and PIC as follows.

(96) "The conditions (4) and (5) (PIC and SSC) refer to structures of the form (11), where \( \infty \) is a cyclic node:

\[
\text{(11)} \quad X \ldots [ \infty \ldots Y \ldots ] \ldots X \ldots
\]

As in the case of subjacency, I will take \( S \) and NP to be the cyclic nodes, delaying the discussion of other choices until later. PIC (the 'tensed-S condition' of the refe-
rences cited) asserts that no rule can 'involve' X and Y where $\alpha C$ is a finite clause (tensed-S). SSC asserts that no rule can 'involve' X and Y where $\alpha C$ contains a specified subject, i.e., a subject not containing Y and not controlled by X (I modify an earlier formulation here; I assume that Y contains Y). If $\alpha C$ contains a subject, then only the subject is accessible to rule, if the subject is specified in the defined sense."

Chomsky (1977c:75) provides the following explication of the notion 'involve'.

(97) "We now say that a transformational rule involves X and Y when it moves a phrase from position X to position Y and a rule of construal involves X and Y when it assigns Y the feature \([-\text{anaphoric to } i]\), where X has the index $i$ (or conversely, in both cases)."

In the case of transformational rules, the notion 'involve' is now restricted to movement transformations. In the case of interpretive rules, the notion 'involve' is now restricted to rules of construal. Suppose that the relation between a moved phrase and its trace is regarded as one of bound anaphora. It then becomes possible to provide a principled explanation for the fact that certain rules are not permissible. If the relation between a moved phrase and its trace is that of bound anaphora, it follows that any movement rule which would lead to a violation of one of the conditions on bound anaphora will be excluded. For instance, any downgrading rule which would lead to a violation of the requirement that an antecedent is superior to its anaphor, will be excluded --- see for instance the discussion in (Chomsky 1975a:109).

If the relation between a moved phrase and its trace is regarded as that of bound anaphora, it also becomes possible to unify the two cases of involvement defined in (97). Chomsky (1977c:76) notes that the notion 'involved in' as defined for rules of construal can be extended to movement rules by permitting the latter to apply freely, and applying the conditions to the moved phrase (the antecedent) and its trace (the anaphor). The SSC and PIC are then, in effect, interpreted as applying to transformational rules as filters. The result of applying
a transformational rule may or may not yield an appropriate case of bound anaphora.

Apart from the advantage of unifying conditions on transformations and conditions on interpretive rules discussed on p. 49 above, Chomsky (1977c:76) mentions another consideration which makes it desirable to restrict the conditions to rules of construal. By so restricting these conditions, it becomes possible to formulate "a somewhat stronger condition of autonomy of syntax". In particular, the semantic conditions that enter into the SSC —— see the notion of 'control' in the definition of 'specified subject' —— would then no longer affect syntactic transformations, but only semantic interpretive rules.

Chomsky (1975b:92) defines the absolute autonomy thesis as follows:

(98) "The absolute autonomy thesis implies that the formal conditions on 'possible grammars' and a formal property of 'optimality' are so narrow and restrictive that a formal grammar can in principle be selected (and its structures generated) on the basis of a preliminary analysis of data in terms of formal primitives excluding the core notions of semantics, and that the systematic connections between formal grammar and semantics are determined on the basis of this independently selected system and the analysis of data in terms of the full range of semantic primitives."

Chomsky (1975b) in fact argues for a weaker version of the autonomy thesis, the so-called parametrized autonomy thesis. The following remarks by Chomsky (1975b:92) will serve to illustrate the point.

(99) "... we might construct a linguistic theory in which formal grammar is independent in its structure but 'open' at certain designated points with respect to the full range of semantic primitives. The problem, then, will be to determine the specific ways in which semantic information enters into the determination of a formal grammar ... The theory of linguistic form may still be a theory with significant internal structure, but it will be constructed with 'semantic parameters'. The actual choice of formal grammar will be determined by fixing these parameters ... Note that the significant question with regard to the autonomy thesis may not be a question of 'yes' or 'no', but rather of 'more' or 'less', or more correctly, 'where' and 'how much'."
It is also important to note that the autonomy thesis (absolute or parametrized) bears on specific grammars, and not on universal grammar/UG. Chomsky (1975b:96) explicitly states that where a property of universal grammar is involved, "the matter is irrelevant to the absolute autonomy thesis". (This obviously also holds for the parametrized autonomy thesis.) It is against this background, that the following remarks by Lightfoot (1976:570) about the notion 'control' in the SSC must be interpreted.

(100) "The SSC is a constraint on syntactic (and other) rules and is sensitive to control properties, which in turn are determined in the semantic component. This raises a question about the hypothesis of the autonomy of syntax. It is consistent, of course, with the notion of a parameterized autonomy thesis (Chomsky (1974)), according to which there are certain prescribed areas where syntactic and semantic information may be confounded, such as the lexicon or conditions on rules. It is also consistent with the notion that conditions on rules may be part of the theory of grammar, and not part of specific grammars."

Lightfoot nevertheless says that "we could construct a plausibility argument for some version of the autonomy thesis, if we could show that syntactic rules are subject only to the syntactic aspect of the SSC and that only semantic rules are sensitive to the notion of control". Chomsky achieves the same result as Lightfoot --- viz. the possibility of a stronger version of the autonomy thesis, which presumably rules out semantic conditions on syntactic rules --- by stipulating that only semantic rules are constrained by the SSC.

Strictly speaking, the consideration of a stronger version of the autonomy thesis bears only on the SSC. The PIC, as formulated in (96), has no semantic content. Nevertheless, the PIC is also restricted to semantic rules. On the other hand, this increases the unity of the theory. On the other hand, the consideration mentioned on p. 49 above, concerning the desirability of unifying conditions on syntactic rules and rules of semantic interpretation, clearly holds for both the SSC and PIC.

In sum: In terms of the reinterpretation of the SSC and PIC proposed in (Chomsky 1976a) and finally adopted in (Chomsky 1977c), the SSC and
PIC restrict only rules of construal. Transformational rules, specifically movement rules, apply freely. The relation between a moved phrase and its trace is regarded as a special case of bound anaphora and, as such, subject to the SSC and PIC. The SSC and PIC, therefore, no longer prevent the derivation of ill-formed structures by transformations, as was the case in the previous stage of the development of these conditions. They rather act as filters on the output of transformations that apply quite freely.

3.3 Other implications of the formulations adopted in (Chomsky 1977c).

In this section I will consider various other implications of the formulations of the SSC and PIC adopted in (Chomsky 1977c) --- see (96) above --- and of the notion 'involve' --- see (97) above.

It has already been pointed out --- see p. 50 --- that, in terms of the definition (97) of 'involve', the SSC and PIC only constrain a subclass of interpretive rules: the rules of construal. This contrasts with the position adopted in (Chomsky 1973), where it was implicitly assumed that all interpretive rules are constrained by the conditions. This change in the definition of 'involve' enables Chomsky to deal with what would otherwise have constituted potential counterexamples to the SSC and PIC.

For instance, Chomsky (1973:242) previously claimed that the interpretive rule associating not ... many in sentences such as (101), giving the meaning "few", was subject to the SSC.

(101)  we didn't see pictures of many of the children [210a]
of (Chomsky 1977c)

The SSC would thus block the association of not and many in (102), because of the presence of the specified subject John.
(102) *we didn't see John's pictures of many of the children (*on the relevant interpretation)

Chomsky (1977c:116) provides the following example, in which the rule associating not and many violates both the SSC and PIC.

(103) we didn't believe that Bill had seen pictures of many of the children

Chomsky (1977c:116) claims that "in our present framework, there is no reason to suppose that" the rule associating not ... many is a rule of construal. Consequently, its application in cases such as (102) does not present any problem to the SSC and/or PIC. Chomsky (1977c:116) suggests that the unacceptability of (102) follows from quite a different principle: not and many cannot be associated when many is within a "specific" NP, where the NP [John's pictures of many of the children] is a "specific" NP.

Chomsky (1977c:80) briefly mentions other interpretive rules that are not rules of construal, and that violate the SSC and PIC. One example is a relativization process that does not involve movement, but only interpretation of a base-generated pronoun in the relative clause. Consider, for example, the following Hebrew sentences.

(104) a. ze ha- iš [še (oto) ra'iti etmol]  

  (this-is the-man [that (him) I-saw yesterday])

b. ra'iti et ha-iš [še natata li et ha-sefer  

  (I saw the-man [that you gave me the-book

  [še hu katav oto ] ]

  [that he wrote it ] ]

The association of ha-sefer and oto "it" in (104b) violates both the SSC and PIC. However, since the relevant rule is not a rule of construal
it does not represent a problem for these conditions. Chomsky points out that in the "rather artificial" English such that construction the SSC and PIC are also violated. Again, the relevant rule represents no real problem for these conditions, since it is not a rule of construal. The same is true for the rule involved in left-dislocation in structures such as (105).

(105) as far as John is concerned, I will never believe the claims that have been made about him

In (105) John and him are coreferential, apparently in violation of the SSC and PIC. Chomsky (1977c:81) claims that the relevant rule is not a rule of construal, and is thus not subject to the conditions in question. In the case of all these rules, the restriction of the SSC and PIC to rules of construal enables Chomsky to deal with what would otherwise have been potential counterexamples to the conditions.

Notice that Chomsky (1977c:81) also briefly refers to the role of the idealization of sentence grammar in eliminating potential problems for the SSC and PIC --- see §2.3.2 above for a detailed discussion of this point. The rule specifically mentioned by Chomsky (1977c:81) is VP-deletion. Following Sag and Hankamer, he claims that it is not a rule of sentence grammar. Consequently, VP-deletion is not subject to the SSC and PIC. The fact that it applies in sentences such as (106), apparently in violation of the SSC and PIC, thus poses no threat to these conditions.

(106) that John didn't hit a home run is not surprising, but that Bill knows that John didn't --- is a real shock

Chomsky (1977c:78) considers extending the notion of 'involvement' "to relate also adjacent constraint terms, one of which is either antecedent or anaphor and the other a constant category of the X-bar system". Chomsky considers this extension of the notion of 'involvement' in order to accommodate Quantifier Movement/Quantifier Construal. The details of the case were discussed in §2.3.3 above, and will not be repeated here.
Chomsky (1977c:75) adopts a certain modification to the PIC proposed by Vergnaud. In terms of this modification, a stipulation is to be added to the PIC, stating that $\alpha$ is the cyclic node which immediately dominates the category of $Y$. This stipulation is needed to overcome a problem posed by sentences such as (107).

(107) the men expected $[s$ that $[s [np$ pictures of each other$]]$ would be on sale$] ]$ [8]

The reciprocal rule, which associates the men and each other in (104), violates the PIC, as formulated in (96). each other (= Y) is in a tensed S. Nevertheless (107) is acceptable. Suppose now that the stipulation mentioned above is incorporated in the PIC. In (107) the cyclic node which immediately dominates Y, is NP. Consequently, the PIC will no longer prohibit the application of the reciprocal rule in (107).

The definition of 'specified subject' given in (96) --- "a subject not containing Y and not controlled by X", where Y contains Y --- is equivalent to the definition adopted in (Chomsky 1976a) --- "a subject distinct from Y and not controlled by X". The stipulation that the specified subject intervenes between X and Y is also absent from the formulation of the SSC adopted in (Chomsky 1977c), as well as from the formulation adopted in (Chomsky 1976a). The consequences of this omission are discussed in §3.5 below.

3.4 The SSC and PIC as part of core grammar

Two important points briefly mentioned in (Chomsky 1973) are taken up again in (Chomsky 1977c), where they are worked out in more detail. The first is that of the relative interpretation of conditions on rules --- see the discussion in §2.2.5 above. The second point is that general conditions on rules may contain language-specific parameters --- see the discussion in §2.2.5 above.
Under the relative interpretation of conditions, a condition does not impose an absolute restriction on rules of a certain type. Rather, rules are taken to obey the condition unless otherwise specified. Such a specification would lead to the rule being marked, in contrast to the unmarked rules that obey the condition. Thus, as Chomsky (1977c:76) puts it, "... the conditions become an integral part of an evaluation measure, rather than imposing absolute restrictions".

Chomsky (1977c:77) illustrates these general point about the relative interpretation of conditions on rules by means of a French rule that must handle "the peripheral Tous-Movement phenomena" of Kayne (1975:63-64). Kayne argues for a general rule L-Tous moving quantifiers to the left. Generally, this rule observes the SSC and PIC (and Subjacency). However, there are examples which apparently involve a violation of the PIC.

\[(108)\]

a. il faut toutes [qu'elles s'en aillent]  
b. il faut tous [qu'on se tire]

In (108), the quantifier tous is in each case construed with a pronoun that is within a tensed S. For reasons noted by Kayne, the L-Tous rule cannot be modified so as to derive (108). Chomsky proposes that the sentences of (108) are derived by a second rule with the structural description (109).

\[(109)\] \((vbl, V*, Q, que, \alpha , PRO, vbl)\)

In (109) Q is construed with PRO. V* is a certain class of verbs, including falloir, vouloir, Q is a quantifier and \(\alpha \) is either null or a "sufficiently short" NP. In (109) the antecedent --- Q --- is not adjacent to the anaphor --- PRO (or trace, if the rule in question is a movement rule). Two terms that are not variables intervene between Q and PRO, namely que and \(\alpha \). Chomsky (1977c:76) says the following about rules with structural descriptions such as (109), in which the antecedent and the anaphor are not adjacent.
"Let us say that the antecedent and the anaphor are involved in the rule if they are adjacent; otherwise not. Specification of constant terms intervening between antecedent and anaphor will then make the conditions inapplicable, at a cost, in accordance with the logic of markedness."

The derivation of (108), accordingly, does not involve a violation of the PIC. The rule (109) which is responsible for such cases is a marked rule, its marked status being due to its complexity. This then illustrates how, under the relative interpretation of conditions on rules, the rules of a specific language can differ with respect to a general -- i.e. universal -- condition on rules.

Chomsky (1977c:75) also rejects the view that the conditions must be invariant. In particular, he suggests that conditions might vary "within fixed limits". As regards the PIC, Chomsky refers to observations by Kim that the rules of anaphora in Korean meet a condition similar to the PIC, but with a somewhat different condition on $\alpha$. (See (96) above.) In Korean there is no formal distinction between tensed and tenseless clauses. There is, however, a category of embedded clauses that are not islands, as is the case with the infinitival clauses of English and the Romance languages. These "non-islands" in Korean are the complements of a certain class of "assertive" verbs. Chomsky notes that these verbs are very close in meaning to the verbs that in English take infinitival complements. Chomsky then suggests that a variant of the PIC can be formulated for Korean, with a different condition on $\alpha$. A more abstract formulation of the PIC can then be provided, with the English and Korean versions of the condition as special cases. Notice, however, that Chomsky does not make any specific proposals concerning the value of $\alpha$ in Korean or the more abstract formulation of the PIC. In fact, he (1977c:75) says that "in the absence of more extensive work on rule systems in other languages, I am reluctant to suggest anything further". The possibility that $\alpha$ in the PIC/TSC might have different values for different languages, was, of course, suggested as early as (Chomsky 1973:238, fn. 16). There it was suggested that $\alpha$ in the TSC might be a language-specific parameter.

In the case of the SSC, the application of this condition in a particu-
lar language depends on the characterization of the notion 'subject' in this language. Thus, while languages such as English and French seem to require a formal definition of 'subject', some case languages might require a characterization in terms of such notions as 'ergative', 'absolutive', or 'non-oblique'. Following a suggestion by Hale, that there are certain conditions on what can be taken as subject in the syntactically unmarked situation, Chomsky proposes that a language might characterize the notion 'subject' differently, "but at a cost in the grammar, in accordance with the logic of markedness". Chomsky (1977c:75-6) sums up his position on the status of conditions such as the SSC and PIC as follows.

(111) "For the moment, I would prefer to think of the conditions cited as instances of condition-schemata, part of the core grammar of English, pending further relevant work on rule systems that may provide evidence bearing on their viability and the more general formulation of the relevant schemata."

The core grammar of English, according to Chomsky (1977c:72-73), includes two transformational rules ("Move NP", "Move wh-phrase"), three interpretive rules (the Reciprocal Rule, the rule of Bound Anaphora, the rule of Disjoint Reference), and three conditions on rules (the Strict Cycle Condition, with the Subjacency Condition as part of the definition of the cycle, the PIC and SSC).

The elaboration of the two points mentioned on p. 56 above in (Chomsky 1977c), thus brings us to the theory of core grammar, which forms such an integral part of current Chomskyan linguistic theory. In (Chomsky 1977c) there are very few explicit remarks on the nature of core grammar. Consider, however, the following remarks from a publication that dates from the same year, namely (Chomsky and Lasnik 1977:430).

(112) "We will assume that UG is not an 'undifferentiated' system, but rather incorporates something analogous to a 'theory of markedness'. Specifically, there is a theory of core grammar with highly restricted options, limited expressive power, and a few parameters. Systems that fall within core grammar constitute 'the unmarked case'; we may think of them as optimal in terms of the evaluation metric. An actual language is determined by fixing the parameters of core grammar and then adding rules or rule conditions, using much richer resources, ...."
Rules which belong to the core grammar of a language are unmarked. Rules which belong to the non-core (or periphery) are marked. The rule of peripheral L-Tous Movement discussed above is an example of such a marked rule. The value of $\alpha$ in the PIC is an example of a parameter that must be fixed for each language.

The development of the theory of core grammar represents an attempt by Chomsky to overcome a well-known dilemma for linguistics, namely that of finding a means of developing a theory of UG that is sufficiently rich and highly structured to select descriptively adequate grammars, and at the same time sufficiently open to allow for the variety of languages. Given the theory of core grammar, a highly restrictive theory of UG, which defines only a small number of core grammars, can be proposed. The possibility of parametric variation, and the possibility of adding marked rules to the periphery, enable such a restrictive theory of UG to account for the variety of languages. The PIC and SSC play an important role in the development of the theory of core grammar. They (and the conditions that will replace them) form an integral part of the core grammars defined by UG.

The possibility of parametric variation in the conditions, and the possibility of marked rules in the periphery of a grammar, both have the effect of making it more difficult to test the SSC and PIC. In the case of a rule R₁ from a language L₁ which violates these conditions, two additional options are now available to the linguist who must interpret this violation. On the one hand, he has to consider the possibility that the two conditions must receive a different formulation in L₁, owing to parametric variation. On the other hand, he must consider the possibility that R₁ is a marked rule which belongs to the peripheral part of the grammar of L₁.

3.5 The case of wh-Movement

Chomsky (1977c:84-5) considers the consequences of the fact that in some cases wh-Movement appears to violate the SSC and PIC. The problem is that while the rules and conditions as formulated in (Chomsky 1977c) allow wh-Movement within a clause, they do not allow extraction of a wh-phrase from a clause. That is, they block COMP-COMP movement.
The derivation of (113a) from (113b) does not violate the SSC or the PIC. The {wh}-phrase is not moved out of {OC}. (See (96) above.) In the derivation of (114a) from its immediately underlying form (114b), however, both the SSC and the PIC are violated: the SSC because the embedded \( \exists \) (= \( \alpha \)) contains a specified subject she, and the PIC because the embedded \( \exists \) (= \( \alpha \)) is tensed.

As far as the SSC is concerned, the predictions made in (Chomsky 1977c) about {wh}-Movement differ from the predictions made in (Chomsky 1973). In (Chomsky 1973) it is stipulated that the specified subject intervenes between \( X \) and \( Y \). See (5) above. In (114b) \( Y \) (= who) is to the left of the specified subject. The SSC, as formulated in 1973, would then not block the derivation of (114a). In (Chomsky 1977c) the SSC merely stipulates that \( \alpha \) (= \( \exists \), NP) contains a specified subject. See (96) above. Consequently, the derivation of (114b) violates the SSC as formulated in (Chomsky 1977c). As far as the TSC/PIC is concerned, both the 1973- and 1977-formulations would have the effect of blocking (114a).42

Chomsky (1977c:85) points out two differences between clause internal {wh}-Movement and the extraction of a {wh}-phrase from a clause. Firstly, there are many languages (e.g. Russian, German) which allow movement of a {wh}-phrase within a clause, but not extraction of a {wh}-phrase from a clause. Secondly, while clause-internal {wh}-Movement in English is unconstrained, the extraction of a {wh}-phrase from a clause is lexically governed. Referring to the "bridge" character of certain matrix verbs that permit the escape of the {wh}-phrase from the embedded \( \exists \), Chomsky states that it is unclear just what property of matrix verbs allows them to function as "bridges".

Having formulated {wh}-Movement as "Move {wh}-phrase into COMP" (see Chomsky 1977c:85, [43]), Chomsky considers two possible solutions to the problem of extracting a {wh}-phrase from a clause. The first solution essen-
tially involves a language-specific COMP-COMP Movement rule.

\[(115)\] "move \textit{wh}-phrase from COMP to a higher COMP over a bridge" \[44\]

Chomsky suggests that the structural description of this rule must be approximately as in (116).

\[(116)\] "(COMP, X, \textit{wh}-phrase, vbl.), where X contains a VP" with certain special properties \[45\]

If the structural description of the COMP-COMP Movement rule incorporates the reference to "bridge" properties, as indicated in (116), then the rule does not satisfy the format proposed for transformational rules by Chomsky (1977c:74-75). According to the relative interpretation of conditions on rules, it can then be argued that the SSC and PIC are inapplicable to (115)/(116), the cost of this solution being the adoption of a complex and, hence, marked rule. Extraction of a \textit{wh}-phrase from a clause in a language such as English would then be the result of the application of a marked rule. No violation of the SSC and PIC would be involved.

The second possible solution to the problem of extracting a \textit{wh}-phrase from a clause dispenses with a language-specific COMP-COMP Movement rule. The "bridge" conditions are interpreted as conditions on rules of interpretation. To prevent the SSC and PIC from blocking COMP-COMP Movement, the language-specific proviso (117) \[46\] is incorporated in the SSC and PIC.

\[(117)\] "where Y is not in COMP"

Whenever Y is in COMP, the SSC and PIC would no longer be applicable. Consequently, the conditions would no longer block the extraction of a \textit{wh}-phrase (= Y) from the COMP-position of an embedded clause. The adoption of proviso (117) to permit COMP-COMP Movement was first proposed in (Chomsky 1973:144). See the discussion in §2.2.4.1 above. Chomsky (1977c:85) states that it is unclear which of the two approaches
to the problem of extracting a \( \text{wh} \)-phrase from a clause is preferable. He nevertheless adopts the second approach --- i.e., the addition of the language-specific proviso (114) to the SSC and PIC --- "without much reason".

What the two alternative approaches to the extraction of a \( \text{wh} \)-phrase from a clause have in common, is their reliance on the relative interpretation of conditions on rules, and the associated "logic of markedness". In the case of both approaches, COMP-COMP Movement disobeys the SSC and PIC at a cost. In the first case the cost is the addition of a complex, marked rule to the grammar of English. In the second case the cost is the addition of a language-specific proviso to the grammar of English. In both cases the grammar of English would be more highly marked than, e.g., the grammar of Russian which does not allow COMP-COMP Movement of \( \text{wh} \)-phrases.

Chomsky (1977c:99) mentions a potential problem for his analysis raised by COMP-COMP Movement in infinitival relatives such as (118).

(118) a. I found a book for you to insist that Bill should read t

   [106c]  

b. I found a book for you to insist that Bill tell Mary that Tom should read t

   [106d]  

Chomsky claims that, although he is not sure about the judgments, these sentences seem to him to be less acceptable than the comparable examples with \( \text{wh} \)-Movement in finite clauses. If this judgment is correct, then COMP-COMP Movement is less readily available in the case of infinitival relatives. In fn. 38 Chomsky briefly mentions a number of solutions to this problem. He does not, however, make a choice from among the available solutions. He clearly does not regard the problem as important, noting that "all that seems to be involved is a language-specific proviso and the precise formulation of a general principle for a domain of facts that are rather marginal".
3.6 The SSC and PIC and the notion 'cyclic node'

Chomsky (1977c:111 f.) considers the effect that it would have on the SSC and PIC (and Subjacency) if S, in addition to NP and $\overline{S}$, were to be regarded as a cyclic node. The PIC would only require a slight reformulation in order to ensure that movement from within a tensed S to the COMP position of the immediately dominating $\overline{S}$ is not blocked.

As far as the effect on the SSC is concerned, Chomsky (1977c:111) suggests that it would be in order to take S as a cyclic node in those languages in which there are many rules to which only subjects are accessible. Chomsky explains the consequences for the SSC if S were to be taken as a cyclic node as follows.

(119) "Given a structure of the form (176), no rule can now involve X and Y if S contains a subject not containing Y and not controlled by X:

(176) ..., X ..., [$\overline{S}$ ... Y ...] ..., X ...

Suppose in particular that Y is NP. Then a rule such as wh-movement, extracting an NP to the COMP position X outside of S, can apply to Y only if Y is the subject of S. In general, only subjects are accessible to movement rules involving an element outside of S, on this interpretation of SSC. It is well known that in many languages only subjects are accessible to many rules."

From these remarks about the SSC and the discussion on pp. 111-116 of (Chomsky 1977c) about the cyclic nodes for Subjacency, the following two points emerge:

(i) The class of cyclic nodes must be determined for each particular language. It is not the case that all languages have the same set of cyclic nodes. In other words, the class of cyclic nodes is a parameter that must be fixed for each language.

(ii) In the grammar of a particular language, the class of nodes that count as cyclic nodes for one condition may apparently differ from the class of nodes that count as cyclic nodes for
another condition. In particular, Chomsky proposes that S is a cyclic node for Subjacency in English, but not for the SSC.

3.7 Explaining the island conditions

Chomsky (1977c:89) discusses an important implication of his hypothesis that configurations derived by rules with the properties of (120)/[49J always result from the application of wh-Movement.

(120) "a. it leaves a gap
   b. where there is a bridge, there is an apparent violation of subjacency, PIC, and SSC
   c. it observes CNPC
   d. it observes wh-island constraints." 43)

If Chomsky's hypothesis is correct, then there is an explanation available for the island conditions, including the CNPC and the wh-island constraint. Chomsky (1977c:89) explicates this point as follows:

(121) "...we have some evidence that the island constraints of (50 ill, iv) [the latter should read (49 c, d) --- M.S.] can be explained in terms of general and quite reasonable 'computational' properties of formal grammar (i.e., subjacency, a property of cyclic rules that states, in effect, that transformational rules have a restricted domain of potential application; SSC, which states that only the most 'prominent' phrase in an embedded structure is accessible to rules relating it to phrases outside; PIC, which stipulates that clauses are islands, subject to the language specific 'escape hatch' (46)). If this conclusion can be sustained, it will be a significant result, since such conditions as CNPC and the independent wh-island constraint seem very curious and difficult to explain on other grounds." 44)

For the purpose of the present discussion of the development of the SSC and PIC, the important point about the remarks in (118) is the mention that is made of the role which these two conditions play in explaining the island constraints. The desirability of explaining the island con-
In terms of other properties of formal grammar is spelled out by Chomsky (1978:16). There Chomsky points out that the island constraints fail to meet two conditions which principles of UC must meet in order to qualify as deep, unifying principles. Firstly, they are not natural as principles of mental computation. Secondly, they constitute a descriptive catalogue; they are not genuinely explanatory in that they unify a variety of generalizations over observed structures and ground them in a system that has a certain degree of deductive structure. The remarks in (121) suggest that the SSC and PIC at least meet the first condition. Consider also the claims made in (Chomsky 1973) about the naturalness of the SSC --- see §2.2.6 above. In a later work, Chomsky (1980b:9-10) suggests that the SSC and PIC meet the second condition as well. This condition will play an important role in the further developmental history of the SSC and PIC, as outlined below.

3.8 Summary of main points

The main points that were made in the description of the second stage of the development of the SSC and PIC, in §§3.2-3.7, are as follows.

(i) In the second stage of their development, the SSC and PIC no longer restrict the application of transformational rules. Transformational rules now apply freely --- see §3.2.

(ii) It is proposed that the SSC and PIC apply to all rules of construal, including the rule of bound anaphora --- see §3.2.

(iii) The trace of a moved phrase is regarded as an anaphor bound by that phrase. The SSC and PIC thus filter out the ill-formed structures derived by the transformational rules via the rule of bound anaphora --- see §3.2.

(iv) In (Chomsky 1976a) and (Chomsky 1977c) two considerations in support of the above-mentioned reinterpretation of the SSC and PIC are presented --- see §3.2.
(v) The definition of 'involve' is modified in (Chomsky 1977c), so as to ensure that the SSC and PIC only apply to rules of construal, and not to other types of interpretive rules. This enables Chomsky to deal with what would otherwise have constituted potential counterexamples to these conditions --- see §3.3.

(vi) In (Chomsky 1977c) a stipulation is added to the PIC, in order to overcome an empirical problem faced by this condition --- see §3.3.

(vii) The proposals concerning a relative interpretation of conditions on rules, and of parametric variation are worked out in more detail, and the theory of core grammar is developed. The SSC and PIC are part of core grammar, with the possibility of parametric variation --- see §3.4.

(viii) Clause-external wh-Movement is analyzed as a marked phenomena, that falls outside the scope of core grammar. Hence, it is not subject to the SSC and PIC --- see §3.5.

(ix) Chomsky (1977c) suggests that it would be in order to take S (in addition to NP and S) as a cyclic node for the SSC in certain languages --- see §3.6.

(x) Together with the Subjacency Condition, the SSC and PIC provide an explanation for the island conditions --- see §3.7.
4 The third stage in the development of the conditions

4.1 General remarks

In the second stage of their development the SSC and PIC were interpreted as filters, or well-formedness conditions, on the output of transformational rules. They were, however, still interpreted as conditions on rules. In particular, they restricted the application of rules of construal. In "On binding" --- written in 1978 and first published in 1980 --- Chomsky reinterprets the SSC and PIC so that they no longer restrict the application of any rule. Instead, they form part of a binding theory, which sets limits on the domain within which an anaphor may or must find an antecedent. The binding theory presented in "On binding" (henceforth Chomsky 1980b) is known as the OB theory. The details of the OB-theory will form the subject matter of §4.

4.2 The SSC and PIC reformulated as the Opacity Condition and the Nominative Island Condition

The binding theory adopted in (Chomsky 1980b) comprises the Opacity Condition (122), and the Nominative Island Condition (123). The Opacity Condition replaces the SSC, and the Nominative Island Condition (henceforth NIC) replaces the PIC.

(122) Opacity Condition

\[ \text{If } \alpha \text{ is in the domain of the subject of } \beta \text{, } \beta \text{ minimal, then } \alpha \text{ cannot be free in } \beta \text{.} \]

(123) Nominative Island Condition

\[ \text{A nominative anaphor cannot be free in } S \text{.} \]
In (122) $\alpha$ is an anaphor. In English at least PRO, reflexives, reciprocals and pronouns in idioms such as John lost his way are anaphors.46) Lexical NPs are not anaphors. Chomsky (1980b:15) suggests that languages might vary as to what elements count as anaphors for the binding conditions (122) and (123). He proposes that $[\alpha e]$ (i.e., PRO and trace) fall under the binding conditions universally, while "more 'lexicalized' items" do so less freely. In this connection Chomsky mentions the equivalents of "reflexive" in Japanese and Korean, which apparently do not fall under the binding conditions, and in fact may be governed by conditions that fall outside sentence grammar. The fact that the notion 'anaphor' in the binding conditions (122) and (123) may vary from language to language again illustrates the possibility of parametric variation in conditions belonging to core grammar.

$\alpha$ is in the domain of $\beta$ if $\beta$ c-commands $\alpha$. $\beta$ c-commands $\alpha$ if $\beta$ does not contain $\alpha$ (and therefore $\beta \neq \alpha$) and $\alpha$ is dominated by the first branching category dominating $\beta$.

$\alpha$ is bound in $\beta$ if there is a category c-commanding it and co-indexed with it in $\beta$. Otherwise, $\alpha$ is free in $\beta$. NP_{arb} (i.e., (i.e., PRO with arbitrary reference) is thus always free. The Opacity Condition and the NIC are called "binding" conditions because they stipulate the domain in which an anaphor must find an antecedent, i.e., the domain in which it must be bound. In the terminology of (Chomsky 1980b:11), the domains of subject and Tense are "opaque" in the sense that anaphors that appear in these domains must be bound in the $\overline{S}$ or NP that immediately dominates subject or Tense.

Chomsky (1980b:10) assumes that the basic expansions of $\overline{S}$ and $S$ are (124), so that Tense c-commands both the subject and the predicate of $S$. He also assumes that NP is the subject of $\overline{S}$ in (124) and of NP_{1} in (125).

(124) $[\overline{S} \text{COMP} [S \text{NP Tense VP } ] ]$ [17a]

(125) $[\text{NP}_{1} \text{NP } \overline{N} ]$ [17b]
Nominative Case is assigned to the subject of a tensed clause, under the theory of Case Assignment presented by (Chomsky 1980b:25).

The Opacity Condition and the NIC can be illustrated with the following examples. 47)

(126)  

a. *The men expected \[S_1 \text{ the soldier to kill each other}\]  
b. *The candidates expected \[S_1 \text{ that each other would win}\]  
c. The candidates expected \[S_1 \text{ each other to win}\]

In (126a) \( \alpha \) (each other) is in the domain of the subject of \( \beta \) (= \( S_1 \)) namely the soldier. each other thus cannot be free in \( S_1 \), according to the Opacity Condition. It is, however, free in \( S_1 \), being coindexed with an NP (= the men) in the matrix \( S \). Consequently, (126a) is ill-formed. In (126b) \( \alpha \) (each other) is a nominative anaphor, since it appears in the subject position of a tensed clause. According to the NIC, each other thus cannot be free in \( S_1 \). It is, however, free in \( S_1 \), being coindexed with an NP (= the candidates) in the matrix \( S \). Consequently, (126b) is ill-formed. In (126c) \( \alpha \) (each other) is neither in the domain of the subject of \( S_1 \), nor a nominative anaphor. Consequently, each other can be free in the embedded \( S_1 \), and coindexed with an NP (= the candidates) in the matrix \( S \).

The qualification in the Opacity Condition, that \( \beta \) is minimal, is required for cases such as (127), with PRO\text{arb}.

(127)  

a. \[\beta_2 \text{ it is unclear } \beta_1 \text{ what PRO to do }\]  
b. \[\beta_2 \text{ their uncertainty as to } \beta_1 \text{ what PRO to do }\]

In \( \beta_1 \) PRO can be free, since it is not in the domain of the subject, and neither is it a nominative anaphor. The minimality qualification prevents (122) from taking \( \beta \) to be \( \beta_2 \). Consequently, PRO
can be indexed arb in (127a) and (127b), even though it is in the domain of the subject of $\beta_2$ (it in (127a) and their in (127b)).

Chomsky (1980b) in fact considers various reformulations of the binding conditions which he rejects in favour of (122) and (123). For ease of reference two other formulations considered by (Chomsky 1980b) are presented in (128) and (129).

(128) If $\alpha$ is an anaphor in the domain of the tense or the subject of $\beta$, $\beta$ minimal, then $\alpha$ cannot be free in $\beta$, $\beta = \text{NP or } \bar{S}$.

(129) A nominative anaphor in $S$ cannot be free in $\bar{S}$ containing $S$.

The condition (128) represents the first reformulation of the SSC and PIC considered in (Chomsky 1980b). (128) incorporates both the SSC and the PIC. (129) is the first formulation of the NIC presented in (Chomsky 1980b). (129) differs from (123) in that the former, but not the latter, contains a reference to $S$.

4.3 The justification for the reformulation of the SSC and PIC as the Opacity Condition and Nominative Island Condition

Chomsky (1980b) discusses three differences between the binding conditions and earlier formulations of the SSC and PIC. In this discussion Chomsky explicitly refers to the reformulation (128) of the SSC and PIC. However, the differences distinguished by Chomsky exist between earlier formulations of the SSC and PIC and any of the reformulations considered in (Chomsky 1980b). Consequently, any advantage which follows from these differences is, at the same time, an advantage of the Opacity condition and the NIC, relative to earlier formulations of the SSC and PIC.
The first difference mentioned by Chomsky is that the binding conditions are no longer conditions on some collection of rules in the grammar, as was the case with previous formulations of the SSC and PIC. Instead, the reformulated versions of the SSC and PIC adopted in (Chomsky 1980b) are conditions on some level of representation. Specifically, Chomsky assumes that the binding conditions are conditions on logical form (henceforth LF) (or on some late stage of interpretation within the rules giving LF). Chomsky does not mention an advantage that follows directly from this difference. One obvious advantage is that it is no longer necessary to stipulate what subclass of syntactic and/or semantic rules are subject to these conditions. That is, it is no longer necessary to define a notion 'X and Y involved in a rule'. Given the interpretation of the relation between a moved phrase and its trace as that of bound anaphora, the binding conditions formulated in (Chomsky 1980b) will automatically have an effect on the output of syntactic movement rules and rules of construal.

The second difference mentioned by Chomsky is that the binding conditions are conditions on anaphors, while earlier formulations of the SSC and PIC placed constraints on variables relating two positions involved in some rule. The significance of this difference, according to Chomsky (1980b:12), is that "it allows us to incorporate without specific mention the case of Arbitrary (uncontrolled) Reference ....". Arbitrary Reference has essentially the same properties as bound anaphora. Consider, for example, the sentences in (130).

(130) a. it is unclear \[s who \ t to visit \ PRO \]  \[22b\]
    b. it is unclear \[s who \ PRO \ visited \ t \]  
    c. it is unclear \[s who \ PRO \ to visit \ t \]  \[22a\]

In each case \(t\) is the trace of \(\text{who}\). (130a), (130b) and (130c) correspond to (126a), (126b), and (126c), respectively. In the ungrammatical (130a) PRO is in the domain of the subject \(t\). In the ungrammatical (130b) PRO is a nominative anaphor. In the grammatical (130c) PRO is neither in the domain of a subject (\(\not\) minimal), nor nominative. The binding conditions thus make exactly the right predictions about
Arbitrary Reference. Given the Opacity Condition, PRO cannot be free, and thus arbitrary in reference, in (130a). Given the NIC, PRO cannot be free, and thus arbitrary in reference in (130b). Neither condition prevents PRO from being free, and thus arbitrary in reference, in (130c). Since there is no question of two positions being involved, earlier formulations of the SSC and PIC could not cover Arbitrary Reference.

The third difference mentioned by Chomsky is the absence of the notion 'specified subject' from the reformulated versions of the SSC, i.e., the Opacity Condition (122) and the second part of (128). The absence of the notion 'specified subject' means that the reformulated versions of the SSC can overcome an empirical problem faced by earlier formulations. Consider the sentence in (131).

(131) which men did Tom think that Bill believed t saw each other

\(t\) is the trace of which men. The SSC would prevent the Reciprocal rule from associating which men and each other, since the specified subjects Tom and Bill intervene. Given the reformulation of the SSC as a binding condition, each other can be coindexed with the trace \(t\), so that it is not free in any opaque context. Chomsky (1980b:13) notes that it was in any event improper to relate the quantifier phrase which men to reciprocal each other, since the latter requires a "referring expression" as its antecedent. The reformulation of the SSC as a binding condition at LF also avoids this problem.

Chomsky (1980b:14) mentions a further advantage of the binding conditions over earlier versions of the SSC and PIC. Given the binding conditions, the COMP position of a tensed clause need no longer be stipulated as an "escape hatch" for movement. In the case of earlier formulations of the SSC and PIC the escape hatch status of COMP had to be stipulated --- see e.g., the discussions in §§2.2.4.1 and 3.5 above. To see how the status of COMP as an "escape hatch" follows from the binding conditions, consider the structure (132).
(132) who do they think \[ \text{COMP} t^1 \] \[ \text{Bill will see t}^2 \] \[ \text{[31]} \]

\( t^2 \) is not free in \( S \), since it is coindexed with \( t^1 \). \( t^1 \) is free in \( S \) (since it is not c-commanded by \( t^2 \)). Since \( t^1 \) is not in the domain of subject or nominative, it can, however, be free in \( S \). Thus, the binding conditions do not block (132). Consequently, under the binding conditions an element can escape from an opaque domain via COMP.

Let us now consider why Chomsky adopts a binding theory consisting of the Opacity Condition and the NIC, rather than condition (128) which incorporates both the SSC and PIC. The reasons for this choice are discussed by Chomsky (1980b:13-14). The first consideration that motivates the choice of the Opacity Condition and the NIC concerns a redundancy exhibited by (128), and in fact by all earlier formulations of the SSC and TSC/PIC. Consider the sentence in (133).

(133) they told me \[ \text{what I gave each other} \] \[ \text{[24]} \]

The sentence in (133) is blocked by both the SSC and PIC, in their earlier formulations as conditions on rules, as well as by the reformulation (128). In the terminology of the latter, the anaphor each other is in the domain of Tense and in the domain of the subject I.

The redundancy illustrated in connection with (133) can be eliminated if the PIC is restricted to the subject of a tensed clause, as is the case with the NIC. The NIC does not rule out sentences such as (133), because the anaphor each other is not in subject position. (133) is ruled out by the Opacity Condition, however, because each other is in the domain of the subject.

Chomsky regards the fact that the redundancy under discussion is avoided as an advantage of the binding theory consisting of the Opacity Condition and the NIC. He (1980b:13) characterizes the avoidance of the redundancy as an advantage "at the metatheoretic level".

The second consideration which motivates the choice of the Opacity Condition and the NIC concerns a difference "at the empirical level", in
Chomsky's (1980b:13) own words. Consider the sentence in (134).

(134) they expected [\( S \) that pictures of each other (each other's pictures) would be on sale]

The PIC, in its pre-1980 formulations and in the reformulation (128), can predict the grammaticality of (134), only if it is assumed that the PIC is constrained by Subjacency. The latter assumption is made in (Chomsky 1977c:75). The NIC, however, makes the correct prediction about (134) without recourse to the Subjacency stipulation. The reciprocal phrase each other in (134) is not nominative, and thus not subject to the NIC. Chomsky (1980b:14) claims that "now we have a much simpler explanation" for the ungrammaticality of (134).

A last point to be discussed in this section concerns the choice of the formulation (123) of the NIC over (129). Chomsky (1980b, fn. 19) explains that the reference to S in (129) is necessary to deal with the trace in COMP of a wh-moved subject, under the assumption that this trace is also nominative. Consider the structure in (135).

(135) who did they think [\( S \) [\( \text{COMP} \ t^1 \) [\( S \) \( t^2 \) would win]]]

Assuming that both \( t^1 \) and \( t^2 \) are nominative, (129) allows this construct. \( t^2 \) is bound in \( S \), while \( t^1 \), though free in \( S \), is not free in \( S \) contained in \( S \). In a discussion of Case Theory, Chomsky (1980b:36) makes the opposite assumption, namely that the trace in COMP of a wh-moved subject is not nominative. It then becomes possible to eliminate the reference to S in the NIC, i.e., to adopt the formulation (123).

Chomsky (1980b:13) claims that the reference to S in the formulation (129) of the NIC constitutes a disadvantage at the metatheoretic level. Neither condition (128), nor earlier formulations of the PIC, contain a reference to S. Chomsky (1980b:13) says that the reference to S in (129) represents "an undesirable complication", and on p. 36 he refers to it as an "inelegance of formulation". The advantage of
the formulation (123) of the NIC is then that it avoids this "undesirable complication" and "inelegance of formulation".

4.4 The role of the SSC and PIC in an optimal theory of construal

In the works discussed in §§2-3, much attention is given to the effect of the SSC and TSC/PIC on transformational rules. It is argued that the incorporation of these conditions in UG permits a dramatic simplification of transformational rules. This simplification, in turn, leads to a desirable reduction in the expressive power of transformations.

Right from the beginning, the SSC and PIC are interpreted as also restricting semantic interpretive rules, specifically the rules of construal. In (Chomsky 1976a) and (Chomsky 1977c) it is suggested that the SSC and PIC also permit a simplification of the rules of construal. Referring to the rules of construal, Chomsky (1976a:319) claims that "... the SSC function so as to permit a very simple formulation of rules". In (Chomsky 1977c:76), he proposes that the structural descriptions of rules of construal must conform to the same narrow format stipulated for transformational rules. In (Chomsky 1980b:6-10) he considers the effect of the SSC and PIC on the formulation of rules of construal in more detail.

Chomsky first examines the case of control, and explores the possibility of using "the simplest possible rule", namely Coindex. He argues that this approach will work for both the familiar cases of control, i.e., for indirect questions such as those in (136), and for sentences such as those in (137).

(136) a. \[ \left[ \infty \right. \theta \left[ \beta \right. \left. \text{NP} \right. \text{visited} \right] \]
   \(\text{(It is unclear who Bill visited)}\) \[13a\]

 b. \[ \left. \left[ \infty \right. \theta \left. \left[ \beta \right. t \right. \text{visited} \right. \right. \left. \text{NP} \right. \]
   \(\text{(It is unclear who visited Bill)}\) \[13b\]

 c. \[ \left. \left[ \infty \right. \theta \left. \left[ \beta \right. \text{NP}_1 \right. \text{to visit} \right. \right. \left. \text{NP}_2 \right. \]
   \(\text{(It is unclear who to visit)}\) \[13c\]
(137)  a. John promised (persuaded) Bill \( [\alpha \text{ that } \text{NP}_1 \text{ would (should) visit } \text{NP}_2 ] \) [14a]
b. John promised (persuaded) Bill \( [\alpha \text{ to visit } \text{NP} ] \) [14b]
c. John tried \( [\alpha \text{ to visit } \text{NP} ] \) [14c]
d. it is time \( [\alpha \text{ to visit } \text{NP} ] \) [14d]

Chomsky assumes that (137b)-(137d) have the embedded structure (138).

(138) \( [S \text{ COMP } [S \text{ NP}_1 \text{ to visit } \text{NP}_2 ] ] \) [15]

In (136a) and (136b) NP \# PRO, that is, control is impossible. In (136c) NP \# PRO. In (137a) neither NP\(_1\) nor NP\(_2\) can be PRO. In (137b)-(137d) NP\(_2\) cannot be PRO. NP\(_1\) in (136c) and (137b)-(137d) can be PRO, however. That is, only the subject of an infinitive is open to control. These properties of control follow automatically from the SSC and PIC. (What does not follow from the SSC and PIC, is that NP must be PRO in these cases, i.e., that control is obligatory in an infinitive. See Chomsky 1980b:18f. for an explanation of this property.) Control can thus be assigned by the simple rule Coindex.

Chomsky (1980b:9) proposes that the SSC and PIC also make it possible to adopt the simplest possible formulation of the rule that assigns an antecedent to each other.

(139) Each other is a reciprocal phrase. [16]

As Chomsky explains, the grammar of English can be reduced to (139) "for the core cases of reciprocals". Conventions belonging to UG will ensure that the coindexing of each other and some NP affected by (139) is correct. The SSC and PIC will ensure that only each other in the subject position of embedded infinitives can be coindexed with an NP in a higher clause.
It is claimed (Chomsky 1980b:9) that other cases of bound anaphora, including reflexives, can be dealt with in the same way as the control and reciprocal cases. Recall that traces are also regarded as bound anaphors. In addition, Chomsky claims that "essentially the same analysis carries over to disjoint reference". (In order to incorporate disjoint reference in his general approach, however, Chomsky must adopt a fairly complex indexing theory --- see the discussion in §4.5 below.) He (1980b:9-10) comments as follows on the advantages of the approach outlined above.

(140) "In this way, we considerably reduce the complexity of the required rules, approaching the potential limits. And we also have a highly unified theory, with a few abstract principles governing a wide range of phenomena."

Two points must be noted here. Firstly, the simplification of the rules of construal effected by the SSC and PIC derives its significance from the fact that it can lead to a reduction in the descriptive power of such rules. Secondly, in Chomsky's view, the theory of which the SSC and PIC form part has the property of being "highly unified". The SSC and PIC qualify as "abstract principles governing a wide range of phenomena". This means that the SSC and PIC meet at least one of the conditions which Chomsky (1978:16) imposes on the explanatory principles of linguistic theory: they must "unify a variety of generalizations and ground them in a system that has a certain degree of deductive structure".

4.5 The binding conditions and Disjoint Reference

Chomsky (1977c:72) formulates the rule of Disjoint Reference as follows:

(141) Assign to a pronoun the feature [- anaphoric to \( \xi \)] in a structure containing NP \( \xi \).
The rule of Disjoint Reference is a rule of construal. Given the formulations of the SSC and PIC as conditions on variables relating two positions in a rule (see e.g., (96) above) the SSC and PIC will also apply to the rule of Disjoint Reference. Chomsky (1977c:79) provides the following examples to illustrate that the PIC and SSC do indeed restrict this rule.

(142) a. PIC (i) they want [them to win] (they ≠ them)
         (ii) they prefer [that they win]

         b. SSC (i) they seem to me [t to like them] (they ≠ them)
         (ii) I seem to them [t to like them]
         (iii) what books do they expect [t to read t to them]
                (they ≠ them)
         (iv) what books do they expect [t to be read to them]
         (v) what books do they expect [Bill to read t to them]

In Chomsky (1980b) the SSC and PIC are reformulated as binding conditions on anaphors. Pronouns are nonanaphors. (See Chomsky 1980b:39). The question then arises of how the rule of Disjoint Reference can be incorporated under the binding conditions. Chomsky's solution to this problem is an indexing theory that incorporates anaphoric indices for nonanaphors. The essentials of this theory are set out below, specifically insofar as they relate to Disjoint Reference.

The indices of NPs are nonnegative integers. The index 1 is reserved for arbitrary reference. Some NPs receive indices via the movement rules, and others via the rules of construal. The latter indexing applies from "top to bottom" in the structure. An NP is assigned an index only when all NPs that c-command or dominate it have been indexed. The only NPs not assigned indices by the movement rules or rules of construal are the nonanaphors: lexical NPs, and pronouns (apart from the bound idioms, as in John lost his way).

The rule of Disjoint Reference assigns indices to the nonanaphors. Each nonanaphor is assigned a complex index (r, A), where r is the referential index and A the anaphoric index. The complex index is
assigned as follows:

(i) Indexing proceeds from top to bottom, until a nonanaphor \( \alpha \) is reached.

(ii) If \( \alpha \) has already been assigned an index \( \text{i} \) by a movement rule, then \( \text{i} \) is its referential index. If it has no index, it is assigned a new referential index \( \text{i} \geq 2 \).

(iii) The referential indices of all NPs that c-command \( \alpha \) are assigned to \( \alpha \) as its anaphoric index. If there is no c-commanding NP, then the anaphoric index of \( \alpha \) is empty.

The anaphoric index \( \{a_1, \ldots, a_n\} \) of \( \alpha \) means that \( \alpha \) is disjoint in reference from each NP with referential index \( a_i \). The binding conditions are taken as deleting certain indices from the anaphoric index of a pronoun, thus in effect blocking disjoint reference in some cases and permitting reference to be free. The binding conditions hold for pronouns, but not for lexical NPs.

In order to unify anaphors and pronouns for the purposes of the binding conditions, the notion 'designated index of \( \alpha \)' is introduced. In the case of an anaphor its referential index is its designated index. In the case of a pronoun its anaphoric index is its designated index. The notion "free" defined on p. 69 above is now generalized as follows.

(143) "Suppose that \( \alpha \) has the designated index \( j \) and \( i \) is an integer such that \( i = j \) or \( i \not\in j \). Then \( \alpha \) is free (\( i \)) in \( \beta \) if there is no \( \chi \) in \( \beta \) with the index \( i \) that \( c \)-commands \( \alpha \)."

The index \( \text{i} \) is necessarily referential. The case \( \text{i} = \text{j} \) is the case of an anaphor, and the case \( \text{i} \not\in \text{j} \) is the case of a pronoun. The binding conditions are then reformulated as rules that modify the designated index.
Suppose that \( \alpha \) has the designated index \( j \) and is free \((\hat{e})\) in \( \beta \) \((\beta = \text{NP or S})\)

where (a) \( \alpha \) is nominative 
or (b) \( \alpha \) is in the domain of the subject of \( \beta \), \( \beta \) minimal.

Then \( j \rightarrow 0 \) if \( j \) is an integer, and \( j \rightarrow (j - \{ i \}) \) if \( j \) is a set.

Case (a) of (144) is the NIC, and (b) the Opacity Condition. \( \text{NP}_o \) is not permitted in LF, where \( o \) is the referential index. \( \text{NP}_o \) is an inadmissible free variable, an anaphor that is not properly bound.
The effect of (144) on the rule of Disjoint Reference is illustrated by the structure in (145).

(145) \( \text{John}_2 \) told \( \text{Bill}_{(3,\{2\})} [\text{S PR0}_3 \text{ to visit him}] \)

The null anaphoric index of \( \text{John} \) is omitted in (143). \( \text{John} \) and \( \text{Bill} \) have been indexed by the assignment rule for nonanaphors; \( \text{PRO} \) by the rule of Control. \( \text{him} \), as a nonanaphor, is assigned the index \((4, \{2,3\})\). \( \text{him} \) is free \((2)\) in \( \text{S} \) but not free \((3)\) in \( \text{S} \), and is in the domain of the subject of \( \text{S} \). \( \text{him} \) thus undergoes rule (144), which removes 2 from its anaphoric index, leaving \( \text{him} \) with the index \((4, \{3\})\). \( \text{him} \) in (145) is thus understood as disjoint in reference from \( \text{PRO} \) and \( \text{Bill} \), but not necessarily disjoint in reference from \( \text{John} \).

In sum: By adopting a fairly complex indexing theory, which includes anaphoric indices, Chomsky is able to incorporate Disjoint Reference under the binding conditions, even though the rule of Disjoint Reference affects nonanaphors.
4.6 A problem for the Opacity Condition solved by structure building rules

Chomsky (1980b:16) argues that the notion 'subject' which figures in the Opacity Condition is a syntactic, and not a semantic, notion. In sentences such as (146), the phrases the books and John are not subjects of given and appear "in any semantically significant sense of the notion 'subject'". Nevertheless, they invoke Opacity, blocking the Reciprocal Rule and the rule of Disjoint Reference.

(146) a. They expect the books to be given to each other (to them) [36a]
b. They expected John to appear to each other (to them) to be qualified for the job [36b]

In (146a) and (146b) each other cannot be coindexed with they, and they and them can be coreferential.

In sentences such as (147) the rules of Reciprocal Interpretation and Disjoint Reference are blocked in the domain of the trace of the wh-phrase, indicating that overt subjects need not appear to invoke Opacity.

(147) a. what books did they expect t to be given to each other (to them) [37a]
b. who did they expect t to appear to each other (to them) to be qualified for the job [37b]

In (147a) and (147b) they and each other cannot be coindexed, and they and them can be coreferential. The Opacity Condition is thus analogous to the Specified Subject Condition, in that traces count as subjects in both cases. Chomsky (1980b:16) concludes that it is "the abstract syntactic subject that invokes Opacity, where 'syntactic subject' is a formal, configurational notion in English". The semantic relation between the subject that invokes Opacity and the elements in its domain is irrelevant.
Chomsky (1980b: 17) considers sentences which lack a syntactic subject at surface structure, but which nevertheless appear to be subject to the Opacity Condition.

(148)  
   a. They regard me as very much like each other (them) [38a]  
   b. I impress them as very much like each other (them) [38b]

Chomsky (1980b:17) observes that "there seems to be no syntactic motivation for assigning anything beyond the obvious surface structure to such sentences". Nevertheless, in (148a, b) the rules of Reciprocal Interpretation and Disjoint Reference are blocked, as if these sentences contained a subject which invokes Opacity. In (148a, b) they and each other cannot be coindexed, and they and them can be coreferential, exactly as in (146) and (147). If subject and object are inverted, as in (149), then neither rule is blocked.

(149)  
   a. I regard them as very much like each other (them)  
   b. They impress me as very much like each other (them)

In (149) they/them and each other can be coindexed, and they/them and them cannot be coreferential.

One possible solution to the problem posed by sentences such as (148) is to develop a semantic analogue to the Opacity Condition. This is rejected by Chomsky as "a dubious move", since it has already been shown that the Opacity Condition relates to the syntax, not the semantics, of LF. Given the similarity between the properties of (148) and those of sentences (such as (146) and (147)) that fall under the syntactic notion of Opacity, Chomsky states that "it seems natural to extend the Opacity Condition directly" to (148). In order to achieve such an extension, it must be assumed that sentences such as (148) are represented as (150) at the level of LF, where the binding conditions apply.

(150)  
   a. they regard me as \([_S \text{ PRO be very much like each other (them)}] \) [39a]
b. I impress them as \[
\frac{3}{5} \text{PRO} \text{be very much like each other (them)} \]

The verbs \textit{regard} and \textit{impress} have essentially the control properties of \textit{persuade} and \textit{promise}, respectively. \textit{regard} assigns object control, and \textit{impress} assigns subject control. In (150a) PRO is thus coindexed with \(\text{me}\), and in (150b) with \(\text{i}\). Thus PRO, coindexed with \(\text{me/1}\), then invokes Opacity in (148). each other must be bound in the embedded \(\exists\), but because it needs a plural antecedent it cannot be co-indexed with PRO. And because \textit{them} is free in the domain of a subject (PRO), it can be coreferential with the plural pronoun in the matrix clause. In (149), with subject and object inverted, the situation is reversed.

In order to extend the Opacity Condition to sentences such as (148), Chomsky must provide for a new class of rules among the rules that generate representations in LF: structure building rules that assign LF-representations such as (150) to sentences such as (148). The fact that the existence of such structure building rules enables the Opacity Condition to apply directly to (148) provides "positive, though indirect, evidence for such rules", according to Chomsky (1980b: 18). Chomsky defends his structure building rules from the potential charge that they exhibit all the weaknesses of earlier transformational rules involved in lexical decomposition. His main argument is that, since there are built-in restrictions on both the input of these rules (i.e. S-structure) and their output (i.e. LF-representations), it is unlikely that his structure building rules of interpretation "go beyond narrow limits". This, according to him, is in contrast to the arbitrary and varied nature of earlier lexical decomposition rules.

4.7 The elimination of the *[NP to VP] filter

One of the important topics of discussion in (Chomsky 1980b) is the elimination of the * [NP to VP] filter. In view of various problems with this filter, Chomsky (1980b:20f.) attempts to develop an alternative to it. The Opacity Condition/SSC and NIC/PIC are relevant to
this development, since they provide part of the motivation for the attempt to do away with the filter.

The \*[[NP to VP]] filter, proposed in (Chomsky and Lasnik 1977:458) and presented here as (151), explains the obligatory nature of control in structures such as (152).

(151) \*[[\(\omega\text{-NP to VP}\)], unless \(\omega\) is adjacent to and in the domain of Verb or for \([-N]\)].

(152) \(\ldots [S \text{wh-phrase } [S \text{NP to VP]}] \ldots\)

By convention, NP in filters is taken to be "lexical", i.e., containing lexical material or trace. In the immediate domain of a wh-phrase, as in (152), the \*[[NP to VP]] filter thus requires NP = PRO. In this way the filter explains the obligatory character of control in such structures, a property of control not explained by the binding conditions --- see the discussion in §4.4 above.

Chomsky (1980b:19f.) discusses a number of "metatheoretic" and "technical" problems raised by the \*[[NP to VP]] filter. One of the metatheoretic problems concerns a certain redundancy between the filter and the binding conditions. As Chomsky (1980b:19) puts it, "it [i.e., the \*[[NP to VP]] filter --- M.S.] in effect recapitulates the basic content of the PIC and SSC (NIC and Opacity), in that it explicitly stipulates a property of subjects of infinitives".

The alternative to the \*[[NP to VP]] filter proposed by Chomsky comprises a rule of obligatory deletion in COMP, up to recoverability (see p.21), and a Case theory. The latter consists of the general principles (153) and the filter (154).

(153) a. NP is oblique when governed by P and certain marked verbs;
    b. NP is objective when governed by V;
    c. NP is nominative when governed by Tense. 52)

(154) \*N, where N has no Case. 52)
The Case theory can account for the obligatory character of control in structures such as (152). No Case will be assigned to the subject NP position in such structures. Consequently, given the filter (154), no lexical NP can appear in this position, i.e., PRO must appear. Chomsky (1980b:27) claims that the Case theory manages to avoid the redundancy in the *[NP to VP] filter. He argues that "the principles (68) [= (153) --- M.S.] and (70) [= (154) --- M.S.] do single out the subject of an infinitive, but indirectly, without the explicit redundancy of the *[NP-to-VP] filter, and on principled grounds, if (68) and (70) prove to be of some generality".

Notice that the redundancy in the *[NP to VP] filter differs from the redundancy exhibited by the SSC and PIC, and the reformulated version (128). In the latter case, the problem is that the theory contains two mechanisms capable of doing the same job --- see in this connection the discussion in §4.3 above. In the case of the filter, the problem is not that of two mechanisms doing the same job. The filter and the binding conditions have distinct functions in the theory. The filter stipulates that only PRO can appear in the subject position of certain infinitives, while the binding conditions stipulate that the only position in which PRO and other anaphors can appear in embedded clauses, is in the subject position of infinitives. The problem is rather that both the filter and the binding conditions concern the same entity, an entity explicitly referred to by the filter, namely the subject position of infinitives.

4.8 Summary of main points

In §§4.2-4.7 a detailed description of the third stage in the development of the SSC and PIC is presented. The main points of this description are as follows.

(i) In (Chomsky 1980b) the SSC and PIC are reinterpreted as conditions that restrict representations at LF --- see §4.2.
(ii) The SSC is reformulated as the Opacity Condition, and the PIC as the NIC. The Opacity Condition and NIC are binding conditions that stipulate the domains within which anaphors must be bound —— see §4.2.

(iii) There are several considerations that justify the reformulation of the SSC and PIC as the Opacity Condition and NIC, respectively —— see §4.3.

a. It is no longer necessary to stipulate the subclass of syntactic and/or semantic rules that are subject to the conditions.

b. It is now possible to incorporate the case of Arbitrary Reference, without specific mention.

c. The Opacity Condition makes better predictions than the SSC in some cases.

d. The status of COMP as an "escape hatch" need no longer be stipulated, since it follows automatically from the binding conditions.

e. The Opacity Condition and NIC avoid a redundancy inherent in all earlier formulations of the SSC and PIC.

f. Unlike the PIC, the NIC does not need the stipulation that it is subject to Subjacency.

g. The formulation (123) of the NIC is chosen over (129), because the former formulation is simpler and more elegant.

(iv) The SSC and PIC, in some formulation, make it possible to adopt extremely simple formulations of the rules of construal —— see §4.4.

(v) In order to accommodate Disjoint Reference under the binding conditions, a complex indexing theory with anaphoric indices is adopted —— see §4.5.
(vi) In order to overcome an empirical problem for the Opacity Condition/SSC, the existence of structure building rules in the interpretive component giving LF is assumed —— see §4.6.

(vii) A redundancy in the binding conditions and the $[^{NP \text{ to } VP}]$ filter provides one of the arguments for the elimination of this filter —— see §4.7.
5 The fourth stage in the development of the conditions

5.1 General remarks

Recall that "On binding" --- referred to as (Chomsky 1980b) above --- was written in 1978. From 1979 onwards, Chomsky argued for the replacement of the binding theory presented in (Chomsky 1980b) by a new binding theory, called a "government binding (GB) theory". The relevant works by Chomsky are the following.

(i) "Markedness and core grammar" (henceforth (Chomsky 1981d)).

This paper was presented at the GLOW-conference in 1979. It contains a fairly detailed account of the reasons for developing an alternative to the binding theory of (Chomsky 1980b), as well as a brief outline of a possible alternative.

(ii) "Principles and parameters in syntactic theory" (henceforth (Chomsky 1981b)).

This paper, which dates from the same year --- 1979 --- as "Markedness and core grammar", contains a brief and informal exposition of the relevant ideas.

(iii) The Pisa lectures (henceforth (Chomsky 1979b)).

These lectures contain a detailed account of the proposed alternative binding theory, and of an important new principle of UG, the "empty category principle"/ECP.

(iv) Lectures on government and binding (henceforth (Chomsky 1981a)).

In this work, the ideas contained in the works mentioned above are brought together, and some proposals are made as to how these ideas can be modified. The work also provides a fairly detailed account of other principles of UG with which the binding theory (and the ECP) interacts.
The aim of §5 is to provide an account of the development of the new
government-binding theory, insofar as it can be reconstructed from
these works. The feature that distinguishes this fourth stage in the
development of the SSC/Opacity Condition and the PIC/NIC, is the fact
that these conditions are no longer stipulated as part of UG, but fol-
low from other principles.

In the discussions that follow I use the term "OB theory/framework"
to refer to the overall theory of UG assumed in (Chomsky 1980b). The
term "OB binding theory" refers to the binding theory which is incor-
porated in the OB theory. Other components of the OB theory are iden-
tified in the same manner, e.g. "OB Case theory". The term "GB binding
theory" is used to refer to the binding theory which replaces the OB
binding theory, and which forms part of the larger GB theory/framework.
The use of the term "GB binding theory" must not be taken to imply that
only one version of a government-binding theory is presented in the
literature. In fact, there are several versions of such a binding
theory. In the discussion that follows I will always explicitly indi-
cate which version of the GB binding theory it is that is being dis-
cussed.

The organization of §5 is as follows. In §5.2 I will discuss the
motivation for the development of an alternative to the OB binding
theory. §5.3 contains a detailed account of the GB binding theory
presented in (Chomsky 1981a:183-209). As Chomsky's exposition of this
version of the GB binding theory is much more detailed than that of
any other version, it constitutes a useful starting point for the pro-
posed exposition of the GB binding theory. In §5.3.2 the content of
the theory is outlined, and its application in clauses and NPs is
illustrated in detail. In §5.3.3 I will consider the degree of
success attained by this version of the GB binding theory in over-
coming the conceptual and empirical problems which triggered the search
for an alternative to the OB binding theory. Differences between this
version of the GB binding theory and two earlier versions are discussed
in §5.4. Further modifications to the GB binding theory proposed in
(Chomsky 1981a) are outlined in §5.5. In §5.6 some problems which
the GB binding theory faces are discussed. The consequences for the
binding theory of the elimination of structure building rules in the
GB framework are discussed in §5.7. §5.8 contains a brief exposition of the Empty Category Principle. §5.9 provides a meta-theoretical perspective by Chomsky on the development of the CB binding theory. A summary of the main points of §5 is presented in §5.10.

5.2 The motivation for the development of an alternative to the OB binding theory

5.2.1 Introduction

As in the case of the earlier versions of the SSC and PIC, the Opacity Condition and NIC must be seen within the context of the overall theory of UC of which they form part. In order to provide the necessary background to the discussion that follows, I will briefly outline the OB theory, which incorporates the OB binding conditions.

Chomsky (1980b:3) proposes that core grammar has the following structure.

(155) 1. Base rules
2. Transformational rules
3a. Deletion rules 3b. Construal rules
4a. Filters 4b. Interpretive rules
5a. Phonology and stylistic rules 5b. Conditions on binding rules

For our purposes, the following aspects of (155) are of special importance.

(i) The conditions on binding in (5b) are the Opacity Condition and the NIC.

(ii) Apart from the rule "Move OC", the transformational component includes the Case principles, discussed in §4.7 above.

(iii) The OB theory incorporates the indexing conventions discussed in §4.5 above.

(iv) Among the filters in (4a) is the $^[\text{that}-t]$ filter. Chomsky and Lasnik (1977:451) formulate this filter as follows:
(156) \[*[[S that [[NP e] ... ]}, unless \(S\) or its trace is in the context: [[NP NF \(\ldots\) ]]

The filter (156) will block structures such as (157), but not (157b-c).

(157) a. *who do you think \([[S that [[NP e] saw Bill]]\)]

b. the man \([[S that [[NP e] saw Bill]]\)]

c. a book \(t\) arrived \([[S that [[NP e] may interest you]]\)] (\(t\) the trace of \(S\))

Chomsky (1981a:18) uses the term "S-structure" to refer to the output of the transformational rules. The term "surface structure" is used to refer to the actual labelled bracketing of an expression at the level of phonetic form/Ph. I will adopt this terminological convention in the discussion below.

Chomsky (1981a:157-161 and 1981d:128-132) discusses six conceptual problems of this OB theory, the solution of which motivated the development of an alternative theory. These conceptual problems are all in some way related to the OB binding theory. In §5.2.2 I will discuss these problems. Since the discussion of these problems in (Chomsky 1981a) is virtually identical to that in (Chomsky 1981d), I will refer only to the former work.

From the discussion in (Chomsky 1981d:128ff.) and (Chomsky 1979b:7), it is clear that these conceptual problems, rather than any empirical problems of the OB theory, played the major role in the search for an alternative to this theory. This does not imply, however, that empirical problems played no role at that stage. Thus Chomsky (1981d:128) refers to the "inevitable problems of empirical adequacy" of the OB theory. The latter work in fact contains a fairly detailed examination of empirical differences between the OB theory and GB theory. For example, it is argued (Chomsky 1981d:140) that the GB binding theory overcomes one of the empirical problems of the OB binding theory. Chomsky (1981a:§3.1) contains a brief discussion of some
empirical problems of the OB theory, in addition to the discussion of the conceptual problems of the theory. In §5.2.3 I will briefly outline these empirical problems.

5.2.2 The conceptual problems of the OB theory

5.2.2.1 A redundancy between the theories of binding and Case

The first conceptual problem of the OB theory identified by Chomsky (1981a:157) concerns a redundancy in the theory, specifically between the theories of binding and Case. The three basic positions of NP in S are nominative Subject of Tense, subject of an infinitive, and complement of a verb. Both the OB binding theory and the OB Case theory single out the subject position of an infinitive. The binding theory singles out this position as the only transparent domain in S. Case theory singles it out as the only position in S in which no Case is assigned. This redundancy constitutes a conceptual problem for the OB theory. 56)

In a previous work, Chomsky (1980b:fn. 30) mentioned the possibility of characterizing the properties of PRO in indirect questions and other control structures in terms of Case theory, rather than binding theory, by stipulating that PRO cannot have Case. The appearance of PRO would then be restricted to the only transparent position, namely subject of an infinitive. In this way, the redundancy between Case theory and binding theory would be eliminated for PRO. However, this suggestion is rejected in (Chomsky 1980b) because it does not generalize to the other cases of binding, and also because, according to Chomsky (1981a:157), "the stipulation seemed rather ad hoc". Chomsky (1981a:157) nevertheless says that "the similarity between the structural properties of Case and binding seems more than fortuitous, and raises the question whether opacity cannot somehow be reduced to Case theory", just as the TSC/PIC was reformulated in terms of considerations of Case as the NIC.
5.2.2.2 Absence of an explanation for the two opaque domains

A second conceptual problem of the OB theory identified by Chomsky (1981a:158) is the failure of the theory to provide an explanation for the fact that the two opaque domains are the subject of a tensed sentence and the c-command domain of the subject of any category. Chomsky comments that in (Chomsky 1980b) (and earlier works) the two opaque domains are in no way related, and neither is particularly well-motivated (except in terms of "rather vague" functional considerations). He concedes that some of the principles of UG must simply be stipulated, at least if the language faculty is an independent cognitive system. Moreover, the binding principles "do serve to unify a fairly impressive range of observations and also interact as one would hope with other principles ...". He nevertheless considers it reasonable to ask "whether there are some more fundamental considerations from which the distribution of transparency and opacity derives".

5.2.2.3 An asymmetry between the NIC and SSC/Opacity Condition, and a contradiction regarding the NIC

The applicability of the binding conditions to the trace of wh-Movement leads to problems for the OB theory. Quoting Rizzi, Chomsky (1981a:158) provides the following Italian example to illustrate that in languages (such as Italian) that tolerate certain violations of the wh-Island Constraints, the SSC/Opacity Condition does not hold for the trace of wh-Movement.

(158) a. tuo fratello, a cui mi domando [che storie abbiano raccontato t], era molto preoccupato

b. ("your brother, to whom I wonder [which stories they told t], was very troubled")

In (158) the wh-phrase a cui moves in a single step to its S-structure position from the position marked by t. This trace is free in the domain of the subject abbiano of the embedded clause, in violation
of the SSC/Opacity Condition. According to Chomsky, the nonapplicability of the SSC/Opacity Condition to wh-traces appears "very natural", in the light of the similarity between variables and names, e.g. under the conditions of strong crossover. Freidin and Lasnik (1981) have pointed out that the similarity between variables and names with respect to strong crossover extends to the domain of Tense, i.e., to the NIC. Thus, in (159) the variable \( t \) and the pronoun he cannot be coindexed, i.e., they cannot be coreferential.

(159)  
\begin{enumerate}
\item \text{a. who did he say [Mary kissed } t \text{ ]}  
\item \text{b. who did he say [ } t \text{ kissed Mary ]}
\end{enumerate}

(159a) illustrates the nonapplication of the SSC/Opacity Condition to the wh-trace, and (159b) the nonapplicability of the NIC. The NIC and SSC/Opacity Condition are thus similar in that neither applies to variables, which behave like names in these constructions.

There are, however, examples which apparently indicate that the NIC does apply to wh-traces. The effect of the NIC on wh-Movement cannot be observed directly in Italian, because of an interaction with other principles. Referring to work by Sportiche, Chomsky (1981a:159) claims that the applicability of the NIC to wh-traces is clear in French. In English the effect of the NIC on wh-Movement is difficult to detect, since the relevant examples are all rendered unacceptable by the wh-Island Condition. There are, nevertheless, sentences which can be used to illustrate this effect. For example, compare (160) with the English translation (158b) of the Italian sentence (158a).

(160)  
the men, who I wonder [which stories \( t \) told to your 
brother], were very troubled

In (160) the wh-phrase who moves in one step from the position marked by \( t \), just as to whom moves from the position marked by \( t \) in (158b). In (158b) the trace of to whom is free in the domain of the subject they of the embedded clause. In (160) the trace of who is free in the subject position of a tensed clause. Chomsky (1981a:159) claims that
the status of the two sentences are clearly "quite different, even for dialects that mark (8) \[\text{M.S.}\] unacceptable because of a wh-island violation". The explanation for this difference in status appears to be that in (160) the NIC applies over and above the conditions that lead to wh-island violations (i.e., Subjacency with S and \(\bar{S}\) as bounding nodes), while in (158b) the SSC/Opacity Condition does not apply.62

The examples (158b) and (160) thus illustrate an asymmetry between the NIC and SSC/Opacity Condition: while the former appears to apply to wh-traces in some manner, the SSC/Opacity Condition does not. Moreover, the examples (159b) and (160) apparently indicate the existence of a contradiction with respect to the NIC in the OB theory: whereas (160) appears to indicate that the NIC applies to wh-traces, (159b) indicates that it does not.

Chomsky (1981a:160) interprets these facts as follows. The wh-Island Condition has two quite separate components. One of these relates to the choice of bounding nodes for Subjacency (and applies in the case of both (158) and (160)). The second component (which applies in the case of (160) but not of (158)) relates to something else, apparently the NIC. If this "something else" is in fact the NIC, then we have the unexplained asymmetry between the SSC/Opacity Condition and NIC. The problem is compounded by the fact that neither the SSC/Opacity Condition nor the NIC applies to wh-traces in strong crossover contexts. Chomsky concludes that the NIC expresses a spurious generalization, and that two distinct principles are involved in the phenomena that have been taken to fall under the NIC.

A possible solution to the problems sketched above is outlined by Chomsky (1981a:160). The NIC must be restricted to the category of phenomena in which there is complete symmetry between the NIC and SSC/Opacity Condition. Variables (including wh-traces) will thus be exempt from both conditions, and NP-traces will be subject to both. A distinct principle --- the residue of NIC/RES(NIC)" --- can then be formulated to account for the fact that wh-traces in sentences such as (160) are subject to something like the NIC.
5.2.2.4 The strangeness of the *[that - t ] filter

Chomsky (1981a:160) considers the "curious character of the *[that-t ] filter" to be a fourth conceptual problem of the OB theory. He again notes that some properties of UG must be stipulated. The filter, moreover, is attractive in that it serves "to unify many phenomena related to 'long movement' of nominative subjects in an enlightening way".64) Chomsky nevertheless claims that "the filter is so strange-looking that one would certainly want to derive it, if possible, from more natural principles".

There is a certain similarity between sentences such as (160) and (161), on the one hand, and sentences such as (157a) that fall under the *[that - t ] filter, on the other hand.

(161) John, I wonder how well understands this book [3.1 (13i)]

The similarity lies in the fact that in each case an unacceptable sentence contains an element X in COMP, followed directly by the trace of another element Y. The sentences in (160) and (161) are those in which some version of the NIC is applicable. These examples suggest that the NIC and the *[that - t ] filter are related. Chomsky mentions three attempts in the literature to eliminate the filter by explaining the relevant phenomena in terms of the NIC: those by Taraldsen, Pesetsky, and Kayne. As in the case of these attempts, Chomsky proposes to solve the problem posed by the *[that - t ] filter by relating the filter phenomena and the NIC phenomena. In particular, Chomsky derives the filter from RES(NIC).65)

5.2.2.5 The complexity of the OB-indexing conventions

The complexity of the indexing conventions incorporated in the OB theory raises a further conceptual problem for this theory. Chomsky (1981a: 161) particularly mentions the desirability of eliminating the concept of an 'anaphoric index' entirely "in terms of some more basic and simple notion". The OB-indexing conventions were discussed briefly in §4.5
above. It was shown that, within the OB theory, the complexity of the conventions (including the use of anaphoric indices) arises from the need to accommodate disjoint reference under the OB binding conditions.

5.2.2.6 The asymmetry between pronouns and other anaphors

The sixth conceptual problem of the OB theory identified by Chomsky (1981a:161) is closely related to the fifth. The conditions under which pronouns enter into disjoint reference are essentially the same as those under which anaphors enter into coreference, namely, the conditions stipulated in the binding conditions. It is this asymmetry between pronouns and other anaphors that gives rise to the complexity of the indexing conventions of the OB theory and of the complexity of the notion 'free (i)' defined in (Chomsky 1980b:Appendix), and explained in §4.5 above. The question arises why there is this asymmetry. Thus Chomsky asks, "why shouldn't pronouns have coreference, rather than disjoint reference, where, for example, reciprocals do?" Within the OB theory there is no explanation for this asymmetry between pronouns and anaphors.

5.2.3 Some empirical problems of the OB binding theory

In addition to the six conceptual problems of the OB binding theory discussed above, Chomsky (1981a:§3.1) mentions what he calls "certain technical problems" of the OB theory. Closer analysis reveals that these "technical" problems are in fact empirical problems. This interpretation is supported by Chomsky's (1981a:187) reference to the conceptual and empirical problems discussed in §3.1 of his work. While Chomsky (1981d:128) — written in 1979 — briefly refers to "the inevitable problems of empirical adequacy" of the OB framework, there is no systematic discussion of such empirical problems in the latter work. Also, no attempt is made to show that the GB binding theory overcomes a significant number of the empirical problems of the OB binding theory. In this respect the presentation of the OB binding theory in 1979 differs from the presentation in 1981.
The first empirical problem noted by Chomsky (1981a:155) concerns structures such as (162a), but with $\text{VP V-ing ... }$ in the place of $N$, i.e., (162b).

(162) [3.1 (2ii)]

a. $
\begin{array}{c}
\text{NP}^* \\
\text{N} \\
\text{\textcircled{\small \&}}
\end{array}$

b. $
\begin{array}{c}
\text{NP}^* \\
\text{VP} \\
\text{\textcircled{\small V-ing}}
\end{array}$

(163) is an example of a sentence with the structure (162b).

(163) *we preferred $[\text{NP}^* \text{ each other's reading the book}]$ [3.1 (3)]

each other (=\textcircled{\small \&}) is not free in the domain of the subject (=we) of $\chi$ (= the matrix S), since it is coindexed with we. The OB binding theory thus incorrectly predicts that (163) is acceptable.

With \textcircled{\small \&} in (162a) and (162b) a pronoun, the OB theory, once again, makes the wrong predictions.

(164) a. they read $[\text{their books}]$ [66]

b. John preferred $[\text{NP}^* \text{ his reading the book}]$ [3.2.3 (65)]

Their (=\textcircled{\small \&}) in (164a) is in the domain of the subject they. The SSC/Opacity Condition thus predicts that it must be free in (164a). Nevertheless, their and they in (164a) can be coindexed, as can his and John in (164b).

Chomsky (1981a:156) mentions some unexplained differences between PRO, on the one hand, and NP-trace and overt anaphors, on the other hand. The positions in which PRO appears are determined to a large extent by
the binding principles for anaphors. Like other anaphors, PRO may
appear as the subject of an infinitive or NP. However, in the case of
"long distance control" the antecedent-anaphor relation of PRO is not
subject to these binding conditions. Thus, consider the sentence (165),
in which PRO is controlled by they over more than one clause boundary.

(165) they thought I said that \[ \text{PRO} \{ \text{to feed} \} \text{each other} \]
would be difficult \[3.1 (5)\]

Chomsky explains that in (165) PRO is bound by they, but is free in the
domain of the subject I. The SSC/Opacity Condition thus wrongly pre-
dicts that (165) is unacceptable. Chomsky (1981a:222, fn. 3) claims
that (165) contrasts with the unacceptable (166), with the overt anaphor
each other in place of PRO.

(166) *they thought I said that pictures of each other were on sale

Chomsky admits that judgments in the case of examples such as (166) are
"not very solid". He nevertheless thinks that "there is a difference
of judgment in the assumed direction".

The OB theory fails to explain why PRO can appear in certain positions
from which trace is excluded.

(167) a. John tried \[ \text{PRO to win} \] \[3.1 (6)\]
b. *John tried \[ \text{t to win} \]
c. it is difficult \[ \text{PRO to win} \]
d. *John is difficult \[ \text{t to win} \]
e. John wanted \[ \text{PRO to win} \]
f. *John wanted \[ \text{t to win} \]
g. *John was wanted \[ \text{t to win} \]
h. it is unclear \[ \text{how PRO to solve the problem} \]
i. *John is unclear \[ \text{how t to solve the problem} \]

The structures (167b) and (167f) are excluded by the requirement that
the matrix subject position must be occupied by an argument.\(^{67}\)

Similarly, the OB theory cannot explain why PRO is excluded from certain positions in which overt anaphors can appear, as in (168).

(168) a. *they expected that pictures of PRO would be on sale

b. they expected that pictures of each other would be on sale

The unexplained asymmetries between PRO and other anaphors illustrated in (167) and (168) constitute further empirical problems for the OB binding theory.

5.3 The GB binding theory as an alternative to the OB binding theory

5.3.1 General remarks

Recall that §5 is primarily concerned with the binding conditions of the OB and GB frameworks. The main aim of §§5.3 and 5.4, accordingly, is to provide answers to the following two questions.

(169) a. What mechanisms replace the SSC/Opacity Condition and NIC within the GB theory?

b. How do these mechanisms solve the conceptual and empirical problems faced by the OB binding theory?

From our perspective, the major differences between the OB and GB theories can roughly be summarized as follows.

(170) a. The GB theory incorporates a new binding theory, in which the notion 'government' plays a central role. The OB binding conditions follow from this new binding theory.
b. The GB theory incorporates a new principle, the "empty category principle"/ECP.

c. The indexing theory of the GB theory consists of the simplest possible convention, namely free (or random) indexing, as opposed to the complex indexing conventions of the OB theory.

d. The *[that - t ] filter is not part of the GB theory.

As will become clear from the discussion below, it is these differences between the GB theory and the OB theory that are responsible for the fact that the GB theory manages to overcome (at least some of) the conceptual and empirical problems of the OB theory. The main emphasis in the following sections will be on the GB binding theory, since it is this component of the GB theory that directly corresponds to the OB binding conditions. The ECP will be discussed very briefly in §5.8.

There are two factors that complicate the attempt to provide an account of the GB binding theory and the ECP. The first is the highly modular nature of the GB theory as a whole. Chomsky (1981a:135) characterizes this modularity as follows.

(171) "The system that is emerging is highly modular, in the sense that the full complexity of observed phenomena is traced to the interaction of potentially independent subtheories, each with its own abstract structure."

For example, the GB binding theory interacts closely with Case theory and government theory. At least some aspects of these theories must be presented in order to explicate the binding theory itself. Moreover, some of the conceptual problems of the OB theory are solved, not by the GB binding theory itself, but by other components of the GB theory, or by the GB binding theory in conjunction with other components. The strategy that I will follow is to provide only as much information about these interacting theories and principles as is needed for an understanding of the GB binding theory, referring the reader to the relevant sections of (Chomsky 1981a) for more detailed information.
The second factor that complicates the attempt to provide an account of the GB binding theory is the fact that several versions of the GB binding theory are presented in the relevant works by Chomsky. In the discussion that follows I will explicitly indicate, where relevant, which version of the GB binding theory it is that is being discussed.

The organization of §5.3.2 is as follows. §§5.3.2.1-2 contain an exposition of the content of the GB binding theory of (Chomsky 1981a:183-209). In §5.3.2.3 the indexing theory incorporated in the GB theory is briefly discussed. In §5.3.2.4 the way in which the GB binding theory applies in clauses and NPs is illustrated in detail. The extent to which the GB binding theory manages to overcome the conceptual and empirical problems of the OB binding theory is considered in §5.3.3.

5.3.2 The GB binding theory

5.3.2.1 The formulation of the theory

Chomsky (1981a:188) formulates the binding principles of the GB binding theory as follows:

(172) "(A) An anaphor is bound in its governing category
(B) A pronominal is free in its governing category
(C) An R-expression is free."

In §§5.3.2.2 the various notions that feature in (172) will be explained. In §§5.3.2.2.1 I present Chomsky's definitions of the notions 'bound' and 'free'. In §§5.3.2.2.2 the three classes of nominal expressions referred to in the binding principles, viz. 'anaphor', 'pronominal', and 'R-expression', are defined. The notion 'governing category' is discussed in §§5.3.2.2.3.
5.3.2.2 Definitions of central notions

5.3.2.2.1 'Bound' and 'free'

Chomsky (1981a:184) distinguishes between two types of binding: A-binding and \( \overline{A} \)-binding. A-binding holds when the binder is in an A-position, while \( \overline{A} \)-binding holds when the binder is in an \( \overline{A} \)-position. He (1981a:184) provides the following formal definitions of 'X-bound' and X-free', where "X" can be replaced by "A" or "\( \overline{A} \)".

(173) " \( \alpha \) is X-bound by \( \beta \) if and only if \( \alpha \) and \( \beta \) are coindexed, \( \beta \) c-commands \( \alpha \), and \( \beta \) is in an X-position." (Fnt. 31 omitted)

(174) " \( \alpha \) is X-free if and only if it is not X-bound." [3.2.3 (5ii)]

A-positions are those positions in which arguments may appear in D-structure. Chomsky (1981a:35) uses the term "argument" to refer to expressions that are assigned \( \Theta \)-roles, that is, thematic roles such as agent-of-action. Arguments thus include names (e.g., John), variables (e.g., the trace of a wh-phrase), anaphors (e.g., each other), and pronouns (e.g., he). Excluded are idiom chunks (e.g., too much as in too much has been made of this problem), and elements inserted to occupy an obligatory position of syntactic structure (e.g., it as in it is certain that John will win, and existential there as in there are believed to be unicorns in the garden). A-positions include the subject position, and the complements to X.

Non-A-positions/\( \overline{A} \)-positions include the head of X, and adjuncts of any sort. If it is assumed that the operation performed by wh-Movement involves adjunction of the wh-phrase to COMP, then it follows that the position occupied by a wh-phrase in COMP is an \( \overline{A} \)-position.

The distinction between A-binding and \( \overline{A} \)-binding can be illustrated with the following structure.

(175) who \( [\_g t \text{ seemed} [\_g t' \text{ to have been killed} t']] \) [3.2.3 (4)]
t is $\overline{A}$-bound by who, since who is in an $\overline{A}$-position. t is in an $A$-position, namely that of subject. Thus, t' is $\overline{A}$-bound by who and A-bound by t. Similarly, t" is $\overline{A}$-bound by who, and A-bound by t and t'.

The GB theory is a theory of A-binding. That is, the terms "bound" and "free" in (172) are synonymous with "A-bound" and "A-free" respectively.

5.3.2.2 'Anaphor', 'pronominal', 'R-expression'

Principle (A) of the GB binding theory applies to anaphors. Chomsky (1981a:188) declares that "intuitively anaphors are NPs that have no capacity for 'inherent reference'". Two types of anaphors are distinguished: lexical anaphors, such as reciprocals, and NP-trace.

Principle (B) of the GB binding theory applies to pronominals. Chomsky (1981a:102) informally characterizes pronominals as elements that have "the features gender, number and person, and perhaps other grammatical features, but not those of overt anaphors or R-expressions". Two classes of pronominals are distinguished: pronouns, which have a phonological matrix, and PRO, which lacks a phonological matrix. 69)

Principle (C) of the GB binding theory applies to R-expressions. According to Chomsky (1981a:102), this category includes "noun phrases with heads that are in some intuitive sense 'potentially referential' (e.g., John, wood, sincerity, book, etc.) and variables ...". Chomsky (1981a:185) defines the notion 'variable' as follows:

(176) " $\infty$ is a variable if and only if

(i) $\infty = [\text{NP e}]$

(ii) $\infty$ is in an $A$-position (hence bears an $A$-GF)

(iii) there is a $\beta$ that locally $A$-binds $\infty$. " 70)

According to Chomsky (1981a:102, 185), the possible binders $\beta$ of $\infty$ include wh-phrases and their traces, quantifier expressions (e.g., everyone) and an empty element in COMP. Chomsky (1981a:185) provides the following definitions of the notions 'locally bound' and 'locally X-bound'.


106

(177) "\(\alpha\) is locally bound by \(\beta\) if and only if \(\alpha\) is \(\text{X-bound by } \beta\), and if \(\gamma\) \(Y\)-binds \(\alpha\) then either \(\gamma\) \(Y\)-binds \(\beta\) or \(\gamma = \beta\)."

(178) "\(\alpha\) is locally \(X\)-bound by \(\beta\) if \(\alpha\) is locally \(X\)-bound and \(\beta\) is \(X\)-bound by \(\beta\)."

In (177) "\(X\)" and "\(\gamma\)" may be independently replaced by "\(A\)" or "\(A\)". In (178) "\(X\)" may be replaced by "\(A\)" or "\(A\)". In the structure (175), \(\xi\) is a variable: it is \([\text{NP } e]\), it is in an \(A\)-position, and it is locally \(\overline{A}\)-bound by the wh-phrase in COMP.

5.3.2.2.3 'Governing category'

The notion 'government' plays an important role in the GB theory, and various definitions of this notion are considered in (Chomsky 1981a: §3.2.1). I will consider only the final definition presented by Chomsky (1981a: §3.2.1), namely the definition presented in [11]. Where it becomes relevant in subsequent sections, I will compare the consequences of this definition with those of earlier definitions.

According to Chomsky (1981a:165), \(\alpha\) governs \(\gamma\) in (179).

(179) "\(\begin{bmatrix} \beta \ldots \gamma \ldots \alpha \ldots \gamma \ldots \end{bmatrix}\) where \(\begin{bmatrix} \text{3.2.1 (11)} \end{bmatrix}\)

(i) \(\alpha = \overline{X}\)

(ii) where \(\phi\) is a maximal projection, if \(\phi\) dominates \(\gamma\) then \(\phi\) dominates \(\alpha\)

(iii) \(\alpha\) o-commands \(\gamma\)."

In terms of (179i), \(\alpha\) is N, V, A, or P. Chomsky (1981a:25) assumes that the expansion of S in English is "NP INFL VP". INFL may in principle be analyzed as \(\begin{bmatrix} \text{[ } \pm \text{Tense} \text{]} \end{bmatrix}\), (AGR). AGR represents the complex of features person, gender, and number. It is identified with PRO, and hence is a lexical category, N. AGR is thus also a proper choice for \(\alpha\) in (179). In terms of (179ii), the maximal projec-
tions ($\bar{S}$, NP, AP, PP, VP) are absolute barriers to government. 73)

No such maximal projection may dominate $\gamma$, unless it also dominates
$\alpha$, the governor. 'c-command' is defined as follows.

\[\text{(180)} \quad \alpha \text{ c-commands } \beta \text{ if and only if } \]

(i) $\alpha$ does not contain $\beta$

(ii) Suppose that $\gamma_1, ..., \gamma_n$ 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)
    $\delta = \gamma_i$ and $\gamma_i$ dominates $\beta$.

This notion 'c-command' can be illustrated with reference to the fol­
lowing structures.

(181)

\[
\begin{array}{c}
S \\
NP \\
\quad \xrightarrow{VP} (= \gamma_1 = \hat{\delta}) \\
\quad \xrightarrow{V} (= \alpha)
\end{array}
\]

(182)

\[
\begin{array}{c}
VP \\
\quad \xrightarrow{VP} (= \gamma_1 = \hat{\delta}) \\
\quad \xrightarrow{NP^*} (= \beta)
\end{array}
\]

In (181) $V$ does not c-command NP, because $VP (= \gamma_1 = \hat{\delta})$ does
not dominate NP. In (182) $V$ c-commands NP*, since $VP (= \gamma_1 = \hat{\delta})$
dominate this NP ($= \beta$). $V$ in (182) also c-commands NP.

Given these definitions, instances of government such as those in (183)
can be distinguished. 74)
In (183a) V governs both NP and NP'. In (183b) P governs NP. In (183c) the prepositional complementizer for governs NP₁. for does not govern NP₂, since the maximal projection VP dominates NP₂ but not for. NP₂ is governed by V.

The GB binding principles (172) stipulate the conditions for boundedness within governing categories. Chomsky (1981a:188) defines 'governing category' as follows.

(184) " \alpha \) is the governing category for \( \beta \) if and only if \( \alpha \) is the minimal category containing \( \beta \) and a governor of \( \beta \), where \( \alpha = \text{NP or S}.\)"

In (183c), for example, S is the governing category for NP₂, since S is the minimal NP or S that contains both NP₂ and its governor, V.

Like the GB binding theory, the GB Case theory is closely linked to government theory. Chomsky's (1981a:170) formulations of the fundamental properties of Case assignment clearly demonstrate the link between Case theory and government theory.

(185) "(i) NP is nominative if governed by AGR
(ii) NP is objective if governed by V with the subcategorization feature: ___ NP (i.e., transitive)
(iii) NP is oblique if governed by P
(iv) NP is genitive in [NP ___ X]
(v) NP is inherently Case-marked as determined by properties of its [-N] governor."

The Case theory also includes the following Case filter.75)
Chomsky (1981a:183) comments as follows on the interrelatedness between Case theory and government theory.

(187) "In its essentials, Case theory forms part of the theory of government. That is, the basic and central instances of Case-assignment are instances of government by a Case-assigner."

The fact that both the GB binding theory and the GB Case theory are closely linked to government theory, enables the GB theory to overcome some of the conceptual problems of the OB theory, as will be shown in §5.3.3 below.

5.3.2.3 Indexing in the GB theory

The GB theory incorporates a very simple indexing theory. It is assumed (Chomsky 1981a:185) that coindexing of a moved element and its trace is, by convention, part of the rule "Move OC". All other indexing is free. (According to Chomsky (1981a:186), it might even be assumed that traces and moved elements are freely indexed at S-structure.) All cases of improper indexing are ruled out by independent conditions, including the binding conditions, Subjacency, and the ECP.

Chomsky (1981a:186-187) claims that the same indexing theory applies to pronouns. Thus pronouns are "proximate" when coindexed with an antecedent (e.g., his in John lost his way), and "obviative" if not coindexed with an antecedent (e.g., him in John saw him). The GB indexing theory thus differs from the OB indexing theory with respect to the indexing of pronouns. In terms of the OB theory, anaphoric and referential indices are assigned to pronouns and names to account for the proximate-obviative distinction, and more generally, for disjoint reference. Only referential indices, in the sense of the OB theory, are assumed under the GB theory.
5.3.2.4 An illustration of the GB binding theory

5.3.2.4.1 General remarks

Having defined the notions that feature in the GB binding principles, it is now possible to take a closer look at the application of this theory. In §5.3.2.4.2 the application of this theory in clauses is illustrated and in §5.3.2.4.3 its application in NPs. Differences between the predictions made by the OB binding theory and those made by the GB binding theory will be noted where relevant. The main focus in both these sections is on principles (A) and (B), since it is from these principles that the SSC/Opacity Condition and NIC follow.

One further preliminary point: Chomsky (1981a:225, fn. 35) stresses that the binding principles as formulated in (172) do not presuppose that $\alpha$ has a governing category. That is, the principles in (172) must be understood as follows.

(188) "Let $\beta$ be a governing category for $\alpha$. Then

(A) if $\alpha$ is an anaphor, it is bound in $\beta$
(B) if $\alpha$ is a pronominal, it is free in $\beta$
(C) if $\alpha$ is an R-expression, it is free."

5.3.2.4.2 The application of the GB binding theory in clauses

The GB binding theory applies in clauses with the following basic structures, where $\alpha_n$ stands for any of the categories to which the binding principles apply.

(189) a. 

\[
S^* \\
| \downarrow \text{NP} \\
\alpha_1 \downarrow \text{INFL} \\
| \downarrow \text{VP} \\
V \downarrow \text{NP} \\
\alpha_2 \downarrow \text{PP} \\
P \downarrow \text{NP} \\
\alpha_3
\]
(i) The overt anaphors

Overt anaphors, such as each other, have phonetic content. They must therefore be assigned Case, by the Case filter (186). Overt anaphors must thus have a governing category in all the structures of (189). By binding principle (A), any overt anaphor must be bound in its governing category.

Suppose $\alpha_1 = \text{each other}$, INFL = [ [+ Tense] AGR]. Then each other is governed by INFL, and its governing category is $S^*$. 78)

(190) *we thought $[S^* \text{each other gave the books to Bill}]$ 79)

In violation of (A), each other is free in $S^*$. Consequently, (190) is unacceptable. Note that sentences such as (190) are also ruled out by the NIC.

Suppose that each other is the object of a verb or preposition in VP --- i.e., $\alpha_2$ or $\alpha_3$ in (189a).
(191) a. *they expected \[s^*_\text{me} \text{to introduce each other to Bill}\]
b. *they expected \[s^*_\text{me} \text{to point the gun at each other}\]

In (191a) each other is governed by the verb introduce, and in (191b) by the preposition at. In both cases, therefore, $S^*$ is the governing category. In both (191a) and (191b) each other must be bound by the subject of $S^*$, me. However, me is not a possible antecedent for each other. Consequently, (191a) and (191b) are ruled out by principle (A) of the binding theory. Note that (191a, b) are ruled out by the SSC/Opacity Condition as well.

Consider next sentences with each other in the subject position of an infinitive.

(192) a. \[s^*_\text{they'd prefer }[s \text{ for each other to win}]\]

b. *we expected \[s^*_\text{Bill to prefer }[s \text{ for each other to win}]\]

(193) a. \[s^*_\text{they believe }[s \text{ each other to be intelligent}]\]

b. *we expected \[s^*_\text{him to believe }[s \text{ each other to be incompetent}]\]

In (192) each other = $\infty_4$, and in (193) each other = $\infty_5$. In (192) each other is governed by for, so that its governing category is $S^*$. In (192a) each other is bound in $S^*$ (being coindexed with they). Principle (A) thus correctly predicts that (192a) will be acceptable. In (192b) each other is free in $S^*$, in violation of (A). Hence the unacceptability of (192b). (192b) is also ruled out by the SSC/Opacity Condition.

In (193) each other is governed and assigned Case by believe. In (193a) each other is bound in $S^*$ (being coindexed with they), and this sentence is acceptable by principle (A). In (193b) each other is free in $S^*$, in violation of principle (A). Thus
(193b) is unacceptable. Again this sentence is also ruled out by the SSC/Opacity Condition.

These examples illustrate that, in the case of overt anaphors in clauses, the NIC and SSC/Opacity Condition follow from principle (A) of the GB binding theory.

(ii) The non-overt anaphor, NP-trace

NP-trace lacks Case. 81) Nevertheless, NP-trace must be governed (Chomsky 1981a:136). 82) If governed, NP-trace has a governing category. By principle (A) of the binding theory, NP-trace must be bound in this governing category. NP-trace is excluded from positions $\alpha_3$ and $\alpha_4$ of (189) by other conditions. 83) Consider the following sentences, with $t = \alpha_1$ in (194), $t = \alpha_2$ in (195), and $t = \alpha_5$ in (196). 84)

(194) *they are believed [S* t are competent] 2.4.4 (17iv)

(195) a. [S* John was killed t] 2.4.4 (181a)
b. *they seem [S* Bill likes t] 2.4.4 (17 v)

(196) a. [S* they seem [t to be competent ]] 2.4.4 (17 i)
b. [S* they are believed [t to be competent ]] 2.4.4 (17 ii)

In (194) t is governed by INFL, with S* the governing category. In violation of principle (A) t is free in this governing category. Thus (194) is unacceptable. This sentence is also ruled out by the NIC.

In (195) S* is the governing category, t being governed by V in each case. In (195a) t is bound in S*, as required by principle (A). Thus (195a) is acceptable. In (195b) t is free in S*, in violation of principle (A). Thus (195b) is unacceptable. (195b) is also ruled out by the SSC/Opacity Condition.
In (196) \( t \) is governed by the matrix verb, thus \( S^* \) is the governing category. In both cases \( t \) is bound in \( S^* \), as required by principle (A). Thus (196a) and (196b) are acceptable.

The examples (194)-(196) illustrate that the NIC and SSC/Opacity Condition follow from principle (A) of the GB binding theory in the case of NP trace in clauses.

(iii) The overt pronominals, i.e., pronouns

A pronoun is necessarily Case-marked, because it has phonetic content. It thus has a governing category, in which it must be free by principle (B) of the GB binding theory. Consider the following examples.

(197) a. Mary thought \([_{S^*} \text{she gave the books to Bill}] \) [3.1 (4)]
   \( (\text{she} = \alpha_1) \)

   b. \([_{S^*} \text{John (everyone) introduced him to Bill}] \)
   \( (\text{he} = \alpha_2) \)

   c. Mary expected \([_{S^*} \text{me to introduce her to Bill}] \)
   \( (\text{her} = \alpha_2) \)

   d. \([_{S^*} \text{everyone introduced John to him}] \)
   \( (\text{him} = \alpha_3) \)

   e. Mary expected \([_{S^*} \text{me to point the gun at her}] \)
   \( (\text{her} = \alpha_3) \)

   f. \([_{S^*} \text{John (everyone) would prefer [for him to win}] } \)
   \( (\text{him} = \alpha_4) \)

   g. Mary expected \([_{S^*} \text{Bill to prefer [for her to win}] } \)
   \( (\text{her} = \alpha_4) \)

   h. \([_{S^*} \text{John (everyone) believed [him to be incompetent}] } \)
   \( (\text{him} = \alpha_5) \)

   i. Mary expected \([_{S^*} \text{Bill to believe [her to be incompetent}] } \)
   \( (\text{her} = \alpha_5) \)
In each case, \( S^* \) is the governing category for the pronoun. By principle (B) of the binding theory, the pronoun must be free in \( S^* \). Thus he/him must be disjoint in reference from John in (197b, d, f, h), and he/him cannot be in the scope of everyone in these sentences. The pronoun can, however, be bound by an element outside \( S^* \) without violating (B). Thus, she/her may refer to Mary in (197a, c, e, g, i). The NIC and SSC/Opacity Condition also make the correct predictions about these sentences. The NIC and SSC/Opacity Condition thus follow from principle (B) of the GB binding theory in the case of pronouns in clauses.

(iv) The non-overt pronominal PRO

Let us now consider the case of a pronominal without a phonetic matrix, that is, PRO. According to Chomsky (1981a:191), it is reasonable to regard PRO as a pronominal anaphor, since it is like overt pronouns in some respects, and like anaphors in others.\(^{86}\) PRO is like the overt pronouns in that it never has an antecedent within its own clause or NP. PRO is like the anaphors in that it has no intrinsic referential content, but is either assigned reference by an antecedent or is indefinite in interpretation, lacking specific reference. If PRO is indeed a pronominal anaphor, then it is subject to both binding principle (A) and binding principle (B): to (A), because it is an anaphor, and to (B), because it is a pronoun. By (A) PRO must be bound in its governing category. By (B) PRO must be free in its governing category. We thus have a contradiction if PRO has a governing category. Therefore, PRO cannot have a governing category, i.e., PRO must be ungoverned. The following principle, which expresses the "essential property" of PRO, is thus derived from the GB-binding theory.\(^{87}\)

(198) PRO is ungoverned

\[ 3.2.3 \text{ (20)} \]

The binding theory also determines that the positions of PRO are essentially those of the other anaphors. To see in more detail how the GB binding theory determines the basic properties of PRO, consider the following sentences with the structures of (189), where \( \alpha C = PRO \).
(199) a. *it is unclear [how \( S_A \) PRO solved the problem] \[2.4.2 (26i)\] 
    \( \text{PRO} = \alpha_1 \)

b. *it is unclear [how \( S_A \) to solve PRO] \[2.4.2 (26ii)\]
    \( \text{PRO} = \alpha_2 \)

c. *it is unclear [how \( S_A \) to give t to PRO] \[2.4.2 (26iv)\]
    \( \text{PRO} = \alpha_3 \)

d. *[\( S_A \) I'm eager [for [PRO to take part] ] ]
    \( \text{PRO} = \alpha_4 \) \[88\]

e. *[\( S_A \) I believe [PRO to be incompetent ] ]
    \( \text{PRO} = \alpha_5 \) \[89\]

f. John tried [PRO to win] \( \text{PRO} = \alpha_5 \) \[2.4.2 (8)\]

In (199a–e) PRO is governed: in (199a) by INFL, in (199b) by V, in
(199c) by the preposition to, in (199d) by the prepositional complemen-
tizer for, in (199e) by the matrix verb believe. In all these cases
PRO thus has a governing category. (199a–e) are thus ruled out by the
GB binding theory, since it follows from this theory that PRO cannot
have a governing category. In (199f) PRO is ungoverned. As predicted,
this sentence is acceptable.

The empirical problems which cases such as (165)–(168) above created
for the OB theory are now solved by the GB binding theory. Since the
relation of PRO to its antecedent (if there is one) is not deter-
mixed by the binding principles, we can have long-distance control, as
in (165) above. Since the relation of other anaphors to their antece-
dents is determined by the binding principles, the contrast between
(165) and (166) follows.

The distinction between PRO and trace in (167) can also be explained.
PRO can appear in these positions, since they are ungoverned. Trace
is excluded from these positions because it must be governed. \[90\]

The exclusion of PRO from the relevant position in (168a) is also
explained by the fact that it is a governed position. The GB binding
theory thus solves the empirical problems relating to PRO which the
OB binding theory failed to solve.
Principle (C) of the GB binding theory applies to R-expressions, i.e., to names and variables. Principle (C) stipulates that R-expressions must be free. The application of (C) to names is illustrated in (200), and its application to variables in (201).

\[(200)\]
\[
a. \text{ he said that John would win } [3.2.3 (25)] \\
b. \text{ John said that John would win }
\]

\[(201)\]
\[
a. \text{ who did he say Mary had kissed } [3.2.3 (26)] \\
\text{ (for which } \text{x, he said Mary had kissed } \text{x)} \\
b. \text{ who did he say had kissed Mary } \\
\text{ (for which } \text{x, he said } \text{x had kissed Mary)}
\]

If there is no emphatic stress, the embedded occurrence of John in (200) is understood as distinct in reference from the matrix subject. That is, this occurrence of John is free, as required by (C).

In (201), he cannot be replaced by the variable x in the associated LF-representation. That is, (202a, b) are not possible interpretations of (201a, b), respectively.

\[(202)\]
\[
a. \text{ for which } \text{x, x said Mary had kissed x } [3.2.3 (27)] \\
b. \text{ for which } \text{x, x said x had kissed Mary}
\]

This follows from principle (C), which stipulates that variables must be free.

The similarity between names and variables, e.g., in strong crossover contexts, is captured by the fact that principle (C) applies to both these categories. Since the NIC and SSC/Opacity Condition follow from principles (A) and (B), but not from (C), the inapplicability of the NIC and SSC/Opacity Condition to, e.g., wh-traces is explained --- see Chomsky 1981a:194-5 for discussion. Thus in (201a) the variable x is free in the domain of the subject Mary, and in (201b) it is free in
the domain of Tense. What remains to be explained, is the fact that, in some instances, bound variables do fall under something like the NIC. These cases do not fall under the GB binding theory (from which the NIC follows), but under an independent principle --- see the discussion in §5.8 below.

To summarize: The NIC and SSC/Opacity Condition follow from principles (A) and (B) of the GB binding theory in the case of arguments in clauses. The GB binding theory successfully overcomes the empirical problems of the OB binding theory with respect to PRO. The inapplicability of the NIC and SSC/Opacity Condition to variables (e.g., wh-traces) follows from the GB binding theory, since variables are subject only to principle (C) of the binding theory. The apparent applicability of something like the NIC to certain variables does not lead to any paradox, since these cases fall under an independent principle, viz. RES(NIC).

5.3.2.4.3 The application of the GB binding theory in NPs

The GB binding theory applies in NPs with the following basic structures.

(203) a.  

\[ \begin{array}{c}
\beta \\
\text{NP}^* \\
\text{N} \\
\text{PP} \\
\text{P} \\
\infty_6 
\end{array} \]

b.  

\[ \begin{array}{c}
\gamma \\
\text{NP}^* \\
\infty_7 \\
\text{N} 
\end{array} \]

According to Chomsky (1981a:207), principle (C) of the GB binding theory applies unproblematically to arguments in NP. The present dis-
cussion will be mainly concerned with the principles (A) and (B) of this binding theory, from which the OB binding conditions are supposed to follow. I will now illustrate the application of these two principles of the GB binding theory to arguments in NP. The predictions made by these principles will be compared with those of the OB binding theory. Notice that only the SSC/Opacity Condition is relevant here, since there is no Tense within NP.

(i) Anaphora

The non-overt anaphor, NP-trace, is excluded from the positions $\alpha_6$ and $\alpha_7$ by independent conditions. Only overt anaphors, such as each other, need therefore be considered.

Consider, firstly, sentences with each other in the position of $\alpha_6$.

(204) a. $[\text{NP}* \text{their stories about each other}]$ [3.2.3 (57i)]
    b. $[^*\text{we heard} [\text{NP}* \text{his stories about each other}]]$ [3.2.3 (57ii)]
    c. $[\text{we heard} [\text{NP}* \text{some stories about each other}]]$ [3.2.3 (57iii)]
    d. $[\text{we heard} [\text{NP}* \text{the stories about each other (that are being circulated)}]]$ [3.2.3 (57iv)]
    e. $[\text{we thought} [\text{that} [\text{NP}* \text{pictures of each other}]]$ would be on sale$]$ [3.2.3 (58)]

The SSC/Opacity Condition of the OB binding theory makes the following predictions about these sentences. In (204a) each other is bound in the domain of the subject $\beta$ (= their) of NP*. The SSC/Opacity Condition thus correctly predicts its acceptability. In (204b) each other is free in the domain of the subject $\beta$ (= his) of NP*. The SSC/Opacity Condition correctly predicts its unacceptability. In (204c) and (204d) NP* has no subject. each other can thus be free in NP*. The SSC/Opacity Condition correctly predicts the acceptability of these sentences. Similarly in the case of the more complex example (204e). NP* has no subject, and it is itself the subject of the embedded sentence. The
SSC/Opacity Condition thus correctly predicts the acceptability of (204e), with each other free in NP*.

Principle (A) of the GB binding theory makes rather different predictions about these sentences. In each case each other is governed by P. NP* is thus the governing category of each other. By principle (A) each other must be bound in NP*. In (204a) each other is bound in NP*. (A) thus correctly predicts the acceptability of this sentence. In (204b) each other is free in NP*. (A) correctly predicts the unacceptability of this sentence. In (204c-e) each other is free in its governing category NP*. Principle (A) thus incorrectly predicts that these sentences will be unacceptable. The GB binding theory in fact makes the right predictions only in those cases where NP contains a subject.

Chomsky (1981a:208) states that (204e) is "perhaps somewhat marginal and may be a marked construction, as consideration of some other languages suggests ...". If (204e) is marked, then the wrong prediction of the GB binding theory with regard to it does not represent a real problem for this theory. However, Chomsky claims that (204c) "in English it surely has a different status from such violations of the SSC as (59) [= (205) --- M.S.]".

| (205) | we thought [that [John's pictures of each other] would be on sale] [59] |

The predictions of the OB and GB binding theories about the sentences in (204) are presented schematically in (206), where "v" before a number indicates unacceptability of the relevant sentence, "v" indicates a correct prediction of (un)acceptability, and "x" indicates a wrong prediction).

| (204a) | v | v |
| (204b) | v | | v |
| (204c) | v | x |
| (204d) | v | x |
| (204e) | v | x |
Consider now sentences with each other in the position of $\alpha_7$ in (203).

\[(207) \begin{align*}
a. \quad & \text{we read } \left[ \text{NP* each other's books} \right] \quad \left[ 3.2.3 \ (60) \right] \\
b. \quad & \text{*they forced me} \ i \ \text{to read } \left[ \text{NP* each other's books} \right] \quad \left[ 3.2.3 \ (61) \right]
\end{align*}\]

The SSC/Opacity Condition of the OB binding theory makes the correct predictions in both cases. In (207a) each other is not in the domain of the subject of NP*, and it is bound in the domain of the subject we of the matrix clause. The SSC/Opacity Condition thus correctly predicts the acceptability of (207a). In (207b) each other is free in the domain of the subject PRO controlled by me. The SSC/Opacity Condition thus correctly predicts the unacceptability of (207b).

Let us turn now to the predictions of the GB binding theory about sentences such as (207). A central question is whether or not $\alpha_7$ (= each other) is governed by the head of $\bar{N}$ in these structures. The answer to this question depends on the precise definition of 'government'. In particular, it depends on whether or not the governor ($= \alpha$) of $\bar{X}$ in structures such as (208) must be an immediate constituent of $\beta$.

\[(208) \left[ \beta \ldots \ Y \ldots \ \alpha \ldots \ Y \ldots \right]\]

Chomsky does not resolve this question in §3.2.1. Definition [3.2.1 (11)], presented as (179) above, does not require that $\alpha$ must be an immediate constituent of $\beta$. The same holds for definition [3.2.1 (6)]\textsuperscript{94). Definition [3.2.1 (4)]\textsuperscript{95) however, does require that $\alpha$ must be an immediate constituent of $\beta$.

Suppose that $\alpha_7$ is ungoverned, and thus lacks a governing category (though it has Case). Principle (A) then makes the right prediction for (207a). NP* is not a governing category for each other, so that each other can be free in NP*. In (207b) each other has no governing category. Principle (A) thus incorrectly predicts that (207b) will be acceptable.

Suppose instead that \( c \gamma \) is governed by the head of \( \bar{N} \). Each other then has a governing category, \( \text{NP}^* \), in which it must be bound by principle (A). In both (207a) and (207b) each other is free in \( \text{NP}^* \). Principle (A) thus incorrectly predicts that (207a) will be unaccept­able, and correctly predicts that (207b) will be unacceptable, though, according to Chomsky (1981a:208), (207b) is barred "for what seem to be the wrong reasons".

Problems also arise when we have \( [\text{VP} \ V-ing \ldots] \) in place of \( \bar{N} \) in (203b).

(209) *they preferred \( [\text{NP}^* \ \text{each other's reading the} \ \text{book}] \)

As was pointed out above, the SSC/Opacity Condition makes the wrong prediction about sentences such as (209). Each other is not in the domain of the subject of \( \text{NP}^* \), and it is bound in the domain of the subject (= they) of the sentence. The SSC/Opacity Condition thus incorrectly predicts that (209) will be acceptable. The GB binding theory also makes a wrong prediction. Under none of the concepts of government considered by Chomsky, is each other governed in (209), VP being a barrier to government. Thus, each other has no governing category in (209). Principle (A) thus incorrectly predicts that (209) will be acceptable. Chomsky (1981a:208) suggests that examples such (209) "may not be crucial, since it might relate to the plurality requirement for reciprocals ...". In fn. 57 Chomsky (1981a:228) briefly discusses and illustrates the latter requirement in English. He points out that sentences such as (210) are unacceptable, because of a kind of plurality requirement elsewhere in the sentence imposed by each other.

(210) a. *they read each other's book
b. *they saw a picture of each other
c. *they turned the child against each other

All these become acceptable if book, picture, and child are made plural. The situation in English is more complex, however. Sentences such as
(210c), with other lexical material, can be acceptable.

(211) they kicked the ball towards each other

Apparently there is also the possibility of interlanguage variation with regard to the plurality requirement. Referring to personal communications by Lauri Carlson and Garald Garaldsen, Chomsky notes that this requirement does not hold for Finnish and Norwegian. Chomsky (1981a:228, fn. 57) states that "further investigation is necessary to determine the character and parameters of these constructions". In his §5.2 Chomsky makes some further comments about these constructions --- see §5.6 below for a discussion of these comments.

There is one further problem concerning reciprocals in the position of $\alpha 7$ in constructions such as (203b) that Chomsky (1981a:222, fn.3) notes. As we have seen, sentences such as (207b) are correctly ruled out by the SSC/Opacity Condition, and by the GB binding theory, under one possible definition of 'government'. In Dutch, however, the analogue to (207b) is acceptable.

The predictions of the OB and GB binding theories about the sentences (207) and (209) are presented schematically in (212).

(212)

<table>
<thead>
<tr>
<th></th>
<th>OB binding theory</th>
<th>GB binding theory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha 7$ governed</td>
<td>$\alpha 7$ ungegoverned</td>
</tr>
<tr>
<td>(207a)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>*(207b)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>*(209)</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

The SSC/Opacity Condition thus does not follow from the GB binding theory in the case of overt anaphors in NP. The GB binding theory also fails to solve the empirical problem of the OB binding theory regarding sentences like (163)/(209).
(ii) *Pronominae

Consider, firstly, the non-overt pronominal, PRO.

(213) a. *they expected that \([\text{NP pictures of PRO}]\) would be on sale (PRO = \( \propto_6 \))

b. *I like \([\text{NP PRO book}]\) (PRO = \( \propto_7 \))

c. I'd much prefer \([\text{NP PRO going to a movie}]\) (PRO = \( \propto_7 \), and \([\text{VP V-ing \ldots}]\) in place of \( \bar{N} \))

The OB binding theory, specifically the SSC/Opacity Condition, makes the wrong predictions about (213a) and (213b). In both sentences PRO is bound in the domain of the subject of the sentence (they and I, respectively), and not free in the domain of the subject of NP. The SSC/Opacity Condition thus incorrectly predicts that (213a) and (213b) will be acceptable. The SSC/Opacity Condition correctly predicts that (213c) will be acceptable. PRO is not in the domain of a subject of NP, and is bound in the domain of the subject I of the sentence.

The GB binding theory makes the correct prediction about (213a). In this sentence PRO is governed by of, and thus has a governing category. The GB binding theory also makes the right prediction about (213b), if it is assumed that the position \( \propto_7 \) in (203b) is governed. If, however, this position is ungoverned, the GB binding theory wrongly predicts the acceptability of (213b). The GB binding theory makes the right prediction about (213c). The position of PRO is ungoverned. PRO thus has no governing category and it is correctly predicted that (203c) will be acceptable. The predictions of the OB and GB binding theories about the sentences (213) are presented schematically in (214).

(214)

<table>
<thead>
<tr>
<th></th>
<th>OB binding theory</th>
<th>GB binding theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(213a)</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>*(213b)</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td>(213c)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
(* This is the prediction, if PRO is governed. If PRO is ungoverned, the GB binding theory also makes the wrong prediction. Chomsky himself does not mention this possibility. He (1981a:208) simply says that PRO is excluded "from the governed positions $\alpha_6$ and $\alpha_7$ ...")

Consider, secondly, the overt pronouns, i.e., pronouns. In (215), pronouns appear in the position $\alpha_6$ of (203a), with him proximate to John.

(215) a. John saw $\left[ NP^* \text{my picture of him} \right]$  

b. *I saw $\left[ NP^* \text{John's picture of him} \right]$

c. *John saw $\left[ NP^* \text{a picture of him} \right]$

d. John thought I saw $\left[ NP^* \text{a picture of him} \right]$

Repeating sentences like (215c), Chomsky (1981a:222, fn. 2) observes that "for some reason, disjoint reference seems less than obligatory in many such cases ...". However, it is clear from his remarks on p. 209, directly below the sentences $[63]$(215), that he considers (215) to be "ungrammatical".

The SSC/Opacity Condition makes the right predictions about all the sentences in (215). In (215a) him is free in the domain of the subject $\text{my}$ of NP*. In (215b) him is bound in the domain of the subject of NP*, given that it is proximate to John. In (215c) him is bound in the domain of the subject of the sentence. In (215d) him is free in the domain of the embedded subject I.

The GB binding theory makes the right predictions about (215a, b, d), but the wrong prediction about (215c). In all cases him is governed by of, with NP* as its governing category. By principle (B), him must be free in NP*. In (215a) and (215d) him is free in NP*, and the GB binding theory correctly predicts that these sentences will be acceptable. In (215b) him is bound in NP*, and the theory correctly predicts that this sentence will be unacceptable. In (215c) him is free in NP*, and the GB binding theory incorrectly predicts that (215c) will be acceptable with him proximate to John.
In (216) his appears in the position of \( \gamma \) of (203b), and his is proximate to John. In (216c) \( \bar{N} \) of (203b) is replaced by VP.

(216) a. John read \( [_{NP^*} \text{his book}] \)  
    b. John thought I saw \( [_{NP^*} \text{his book}] \)  
    c. *John preferred \( [_{NP^*} \text{his reading the book}] \)

The SSC/Opacity Condition incorrectly predicts that (216a) will be unacceptable in the relevant interpretation, since him is bound in the domain of the subject John. It correctly predicts the acceptability of (216b), with his free in the domain of the embedded subject I. In (216c) his is bound in the domain of the subject John, and the SSC/Opacity Condition correctly predicts that this sentence will be unacceptable.

Let us now consider the predictions which the GB binding theory makes about these sentences. Suppose that his is governed in (216a, b), i.e., that NP* is a governing category for his. By principle (B), his must then be free in NP*, which is the case. The GB binding theory thus correctly predicts the acceptability of (216a, b). If his is ungoverned, the same predictions are made. His will then have no governing category. The GB binding theory predicts that (216c) will be acceptable, since his is not governed in NP*, and thus has no governing category. However, (216c) poses no real problem for the GB binding theory, since its acceptability follows from another principle, the Avoid Pronoun Principle. 97)

The predictions of the Ob and GB binding theories about the sentences (215) and (216) are presented schematically in (217).
The SSC/Opacity Condition thus does not follow in full from the GB binding theory in the case of pronouns in NP. In some cases, viz. (213a), (213b), and (216a), where the OB binding theory makes the wrong predictions, the GB binding theory makes the right predictions. In the case of (215c), however, the situation is reversed.

5.3.3 The GB binding theory and the conceptual and empirical problems of the OB binding theory

In §5.2 a number of conceptual and empirical problems of the OB binding theory were outlined. The search for an alternative to the OB binding theory was to a large extent motivated by the desire to overcome these conceptual problems. The alternative proposed by Chomsky --- the GB binding theory --- was discussed in §5.3.2. In this section I will briefly indicate which of the conceptual problems of the OB binding theory are solved by the GB binding theory. A few remarks are also made about the solution of the empirical problems noted in §5.2.1 by the GB binding theory.

Chomsky (1981a:221-2) mentions three of the conceptual problems of the OB theory, that are solved.

(218) "... we have reached plausible answers to (1), (2) and (6): namely, (1) the problem of redundancies between Case and binding theory, now resolved in terms of their common component, the theory of government; (2) the problem of explaining why the two unrelated domains subject-of-AGR
and domain-of-subject should be opaque; (6) the problem of finding a more natural account of disjoint reference (but see note 39)."

Recall that within the OB framework both Case theory and the binding theory single out the subject position of an infinitive. Within the GB framework, the theories of Case and binding are both formulated in terms of the notion 'government'. The subject position of an infinitive is an ungoverned position in the unmarked cases (the marked cases being an infinitive introduced by the complementizer for, or an infinitive in the complement of an S-deletion verb like believe). In this way, then, the redundancy between Case theory and binding theory is "resolved in terms of their common component, the theory of government".

Within the GB framework the explanation for the fact that the two domains, subject-of-AGR and domain of subject, are opaque, is as follows. The PIC/NIC and SSC/Opacity Condition follow from principles (A) and (B) of the GB binding theory, formulated in terms of the notion 'governing category'. It is the latter notion which in fact unifies the PIC/NIC and SSC/Opacity Condition. The opaque positions are governed positions, i.e., elements that appear in these positions have governing categories.

Disjoint reference is treated as follows within the GB framework. Pronouns share with PRO the property of being pronominals. Hence they are subject to principle (B) of the GB binding theory. In terms of (B), pronouns must be free in their governing category. They thus differ from anaphors which, being subject to principle (A) of the GB binding theory, must be bound in their governing category. The indexing of pronouns is exactly like that of anaphors, according to Chomsky (1981a:186). That is, they are "proximate" if they are coindexed with some other element and "obviative" if not coindexed with another element. As Chomsky (1981a:222) notes, if the simple indexing theory proposed in his §3.2.3 could be adopted — i.e., random assignment of referential indices — then the fifth conceptual problem of the OB theory would also be solved. This problem — see §5.2.2.5 above — concerned the complexity of the indexing conventions of the OB theory.
While Chomsky (1981d:139) tentatively suggests that the GB theory can solve all the conceptual problems of the GB theory mentioned by him, he (1981a) admits that the GB theory does not manage to solve all of them. In particular, not all properties of disjoint reference can be accounted for by the simple indexing theory of the GB theory. Chomsky (1981a:226, fn. 39) points out that the examples given in §3.2.3 to illustrate the application of the GB binding theory to pronouns are restricted to distinct reference. Chomsky (1981a:§5.1) considers disjoint reference more generally, and discusses some cases that present problems for the simple GB indexing theory. Consider the following sentences.

(219)  
   a. *we lost my way  
   b. I lost my way  
   c. *I lost his way

(220)  
   a. *we expected me to like John  
   b. we expected John to like me

In (219a) the idiom requires coindexing of we and my. This is illustrated by the contrast in acceptability between (219b), which has coindexing, and (219c), which does not. The question is how coindexing of we and me in (219a) must be interpreted. According to Chomsky (1981a:286), this example "indicates that we must take coindexed elements to be strictly coreferential, not merely overlapping in reference ...". On this assumption, (219a) will be assigned its proper, ungrammatical, status.

Given this assumption about the interpretation of coindexing, consider now disjoint reference, as in (220a). Referring to Postal for the original observation, Chomsky (1981a:286) claims that (220a) clearly contrasts in status with (220b). Within the GB theory of indexing there are two options in the examples (220). The two pronouns may be either coindexed, or differently indexed.

Suppose that we and me are coindexed in (220a, b). The GB binding
theory will then correctly bar (220a) and admit (220b). In (220a) the matrix $S$ is the governing category in which $me$ must be free, and in (220b) the embedded $S$ is the governing category for $me$, thus allowing $me$ to be coindexed with $we$ in the matrix $S$. However, on this analysis $me$ in (220b) is understood as overlapping in reference with $we$ under coindexing. This is inconsistent with the analysis of (219a), where it is assumed that coindexed elements must not merely overlap in reference, but must be strictly coreferential.

Suppose then that $we$ and $me$ in (220a, b) are indexed differently. We then have consistency with the analysis of (219a). The GB binding theory now fails to explain the difference in status between (220a) and (220b), since $me$ is free in both sentences. This, according to Chomsky (1981a:286), is "surely an incorrect result". He points out that the relevant examples do not present any problem for the complex indexing theory of the OB theory. In the latter theory, the referential indices of the pronouns will differ, and the anaphoric indices will indicate the required properties of overlapping and disjoint reference in (219a), (220a, b). Chomsky (1981a:286) characterizes the status of the GB theory with respect to the fifth conceptual problem of the OB theory identified in §5.2 --- namely, the complexity of the OB indexing theory --- as follows.

"Clearly, then, the theory of indexing we have been using here is defective, and something more complex is required. The theory of anaphoric indices in the OB-framework over­comes these problems, with the exception of (11) and (2). In §3.1, I cited the complexity of this theory as one of the problems to be addressed in improving the OB-framework, and in the exposition above I have avoided all of these problems, but only by restricting myself to too narrow a class of examples. This problem, along with several others relating to the theory of indexing, therefore still stands, in contrast to the other problems raised, which receive a natural solution in the GB-framework."  

The two remaining conceptual problems discussed in §5.2 --- those relating to the RES(NIC) phenomena and the strange *[that - t]*filter --- are taken up in §5.8 below.

To conclude this section, let us briefly consider to what extent the
GB theory is successful in overcoming the empirical problems of the OB binding theory discussed in §5.2.3. Consider, firstly, the problems relating to PRO in the OB theory, as outlined in §5.2.3 above. Chomsky (1981a:191-2) claims that all these problems are in fact solved by the GB theory. For an explanation of how the GB theory manages to solve these problems, see §5.3.2.4.2 above.

Another empirical problem for the OB binding theory, mentioned in §5.2.3, is the wrong prediction which this theory makes about overt anaphors in structures of the form \([\text{NP} \_\_ \text{VP V-ing ...}]\) —— see e.g. (163) above. Chomsky (1981a:208) points out that this problem remains unsolved by the GB binding theory. He does, however, mention the possibility that the relevant example "may not be crucial, since it might relate to the plurality requirement for reciprocals ..." —— see the discussion in §5.3.2.4.3 above for the latter point. The OB binding theory also makes the wrong predictions about pronouns in structures of the form \([\text{NP} \_\_ \text{N}]\) —— see e.g. (164) above. This problem is solved by the GB binding theory, as explained by Chomsky (1981a:209).

The GB binding theory thus manages to solve all but one of the empirical problems of the OB binding theory discussed in §5.2.3. In §5.3.2.4.3 it has been shown that in some cases where the OB binding theory makes the right predictions, the GB binding theory makes the wrong predictions —— see e.g., (204c), (204d), (204e), (215c). In §5.4.3 I will consider the significance of these differences between the OB and GB binding theories.

5.4 The 1979 versus the 1981 version of the GB binding theory

5.4.1 General remarks

In §5.3 I have outlined and illustrated the version of the GB binding theory presented in (Chomsky 1981a:183-209). The aim of the present section is to compare this version of the GB binding theory with two earlier versions, namely those presented in (Chomsky 1979b) and (Chomsky 1981d). Both the latter versions date from 1979. In the discussion below I use the term "the GB (LGB) binding theory" to refer to the ver-
sion of the GB binding theory presented in (Chomsky 1981a:183-209). The term "the GB (Pisa) binding theory" refers to the version presented in (Chomsky 1979b), and "the GB (MCG) binding theory" refers to the version presented in (Chomsky 1981d). In §5.4.2 I compare the formulations of the GB (LCB), GB (Pisa), and GB (MCG) binding theories. A detailed exposition of the 1979 interpretation of the empirical differences between the OB and GB binding theories will be presented in §5.4.3. This interpretation differs from the interpretation adopted in (Chomsky 1981a). In §5.5 I discuss further developments in the GB binding theory as presented in (Chomsky 1981a). It will become clear that a change in Chomsky's interpretation of the empirical differences between the OB and GB binding theories played a significant role in these further developments.

5.4.2 Three different formulations of the GB binding theory

Chomsky (1981d:134) formulates the GB binding theory as follows:

(222) "A. If NP is lexical or a bound variable, then it is free
B. If NP is pronominal, it is free in its governing category
C. If NP is an anaphor, it is bound in its governing category."

He provides the following explications of the terms that appear in (222).

(i) "Pronominal" in (222B) refers to pronouns and PRO.

(ii) An argument is bound if it is c-commanded by a coindexed argument (where the argument positions are taken to be the NP positions within S or NP). If not bound, an argument is free.

(iii) \( \infty \) is the governing category for \( \beta \) if \( \infty \) is the minimal category in which \( \beta \) is governed (\( \infty = \text{NP or S} \)). \( \gamma \) governs \( \beta \) if \( \gamma \) minimally c-commands \( \beta \) (\( \gamma = a \))
lexical category or Tense); that is, $Y$ c-commands $\beta$ and there is no $\phi$ c-commanded by $Y$ and c-commanding $\beta$ but not $Y$. If $\beta$ has Case, then its governing category is the $S$ or NP in which it is marked for Case. $S$ and NP are absolute barriers to government.

Principle (222A) of the GB (MCG) binding theory corresponds to principle (172C) of the GB (LGB) binding theory; (222B) corresponds to (172B); (222C) corresponds to (172A). There are two obvious differences between the formulations of the GB (MCG) and GB (LGB) binding theories. Firstly, the binding principles of the GB (MCG) binding theory, but not of the GB (LGB) binding theory, are formulated as conditionals. Secondly, principle (222A) of the GB (MCG) binding theory refers to lexical NPs and bound variables, while the corresponding principle (172C) of the GB (LGB) binding theory refers to R-expressions. It appears to be the case that both these differences are without any significance --- empirical or otherwise. The discussion in §5.3.2.2.2 above makes it clear that the term "R-expression" used in the GB (LGB) binding theory refers to the same class of arguments that "lexical NP" and "bound variable" refer to in the GB (MCG) binding theory.

There is a third difference between the GB (MCG) and GB (LGB) binding theories that has empirical consequences. The two theories incorporate different notions of 'government'. Consequently, the class of governing categories of the GB (MCG) binding theory differs from that of the GB (LGB) binding theory. The empirical consequences of this difference between the two versions of the GB binding theory are discussed in §5.4.3 below. Chomsky's (1981a:§3.2.1) discussion of the concept of government makes it quite clear that the choice of a particular definition of this concept is an empirical matter.

The formulation of the GB binding theory presented in (Chomsky 1979b: 16) --- the GB (Pisa) binding theory --- differs from both the GB (MCG) and GB (LGB) binding theories. (The formulation of the GB (Pisa) binding theory presented in (223) is taken over from (Chomsky 1981a:225, fn. 35). This formulation merely omits the illustrative material incorporated in the formulation of (Chomsky 1979b:16)).
Chomsky (1979b:8) provides the following definitions for the notions 'governing category', and 'minimal governing category'.

(224) 
\[ \alpha \text{ is a governing category for } \beta = \text{def some } \gamma \text{ such that } \gamma \text{ governs } \beta \text{ and } \alpha \text{ contains } \gamma. \]

(225) 
\[ \alpha \text{ is a minimal governing category for } \beta = \text{def } \alpha \text{ is a governing category which properly contains no governing category.} \]

The notion 'minimal governing category' that features in the GB (Pisa) binding theory is equivalent to the notion 'governing category' in the GB (MCG) and GB (LGB) binding theories. The notion 'government' incorporated in the GB (Pisa) binding theory is the same as that incorporated in the GB (MCG) binding theory. Chomsky (1979b:16, 20) makes it clear that the conditions on the boundedness of the different types of arguments contained in the GB (Pisa) binding theory are in fact the same as those contained in the GB (MCG) binding theory --- see (222) above --- and the GB (LGB) binding theory --- see (172) above. (223) stipulates that anaphors must be bound in all governing categories, including their minimal governing category. Pronominals must be free in their minimal governing category. Lexical NPs and variables must be free. The question arises why the formulation of the GB (Pisa) binding theory differs so much from the formulations of the GB (MCG) and GB (LGB) binding theories. Having outlined the basic content of (223), Chomsky (1979b:20) provides the following answer to this question.
"That's roughly the content of that, although it is formulated in such a way as to make some other things follow. What follows is that PRO is ungoverned and that trace is a variable if it is case-marked. In particular it follows that the trace of NP movement is always not case-marked."

The ungoverned status of PRO follows from (223) in the following manner, according to Chomsky (1979b:17).

(227)  

a. Assume PRO is governed.

b. Then there is a minimal governing category \( \infty \) in which PRO is governed.

c. By (223C), PRO must be free in \( \infty \).

d. By (223A), since PRO is not a variable, it has to be bound in every governing category, including \( \infty \).

e. From the contradiction between (c) and (d) it follows that PRO has no governing category, i.e., PRO is ungoverned.

Chomsky (1981a:191) argues that the ungoverned status of PRO follows from the formulation (172) of the GB binding theory, given the assumption that PRO is a pronominal anaphor. The more complex formulation (223) is thus not needed to derive the principle that PRO is ungoverned. Chomsky (1981d:135) also shows that the conclusion about the ungoverned status of PRO follows from a simpler formulation of the GB binding theory, namely that of (222), the GB (MCG) binding theory.

The principle that trace is a variable if it is Case-marked follows from the GB (Pisa) binding theory in the following manner:

(228)  

a. Assume \( \left[ NP \ e \right] \) has Case.

b. Since \( \left[ NP \ e \right] \) lacks a phonetic matrix, it is subject to (223A).

c. Since \( \left[ NP \ e \right] \) has Case, it is subject to (223B).

d. By definition it is not an anaphor.
e. Therefore, \([\text{NP} e] \) with Case is a variable and free in every governing category.

(228) entails that the principle (229) follows from right to left in (Chomsky 1979b).

(229) "\([\text{NP} e] \) is a variable if and only if it has Case."

Since (229) follows from left to right from the Case filter, (229) holds in its full generality in (Chomsky 1979b). In (Chomsky 1981a) (229) is not assumed in its full generality. While it is assumed that variables have Case, it is not assumed that every \([\text{NP} e] \) with Case is a variable. The GB (LGB) binding theory thus differs from the GB (Pisa) binding theory in that it follows from the latter, but not from the former theory, that every \([\text{NP} e] \) with Case is a variable. Chomsky (1981a) drops the assumption that every Case-marked \([\text{NP} e] \) is a variable for empirical reasons. He (1981a:272) argues that in (230a), which is the S-structure of the Italian sentence (230b), \(\text{NP}^* \) is PRO with Case. This PRO is not a variable.

(230) a. \([S \ \text{NP}^* \ [\text{VP} \text{copula} - \text{AGR} \ \text{NP}] \] \ [4.5 \ (48i)] \n
b. sono io
   ("it's me")

5.4.3 The 1979 vs. the 1981 interpretation of the empirical differences between the OB and GB binding theories

Chomsky (1981a:207-209) discusses certain differences between the predictions of the OB binding theory and those of the GB binding theory. These predictions specifically concern the status of arguments in NP. In §5.3.2.4.3 above the differences between the predictions of the two theories are set out in detail. The cases where the two theories make different predictions are summarized in (231).
137

(231)  

a. Overt anaphor in the position of $\alpha_6$ in (203a)

<table>
<thead>
<tr>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(204c)</td>
<td>✓</td>
</tr>
<tr>
<td>(204d)</td>
<td>✓</td>
</tr>
<tr>
<td>(204e)</td>
<td>✓</td>
</tr>
</tbody>
</table>

b. Overt anaphor in the position of $\alpha_7$ in (203b)

<table>
<thead>
<tr>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(207a)</td>
<td>✓</td>
</tr>
<tr>
<td>* (207b)</td>
<td>✓</td>
</tr>
</tbody>
</table>

‼ The predictions of the GB binding theory depend on the exact definition of 'government'.

c. PRO in position of $\alpha_8$ in (203a)

<table>
<thead>
<tr>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(213a)</td>
<td>x</td>
</tr>
<tr>
<td>*(213b)</td>
<td>x</td>
</tr>
</tbody>
</table>

d. Pronoun in position of $\alpha_6$ in (203a)

<table>
<thead>
<tr>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(215c)</td>
<td>✓</td>
</tr>
</tbody>
</table>

e. Pronoun in position of $\alpha_7$ in (203b)

<table>
<thead>
<tr>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(216a)</td>
<td>x</td>
</tr>
</tbody>
</table>

f. Pronoun in position of $\alpha_7$ in (203b), with $[VP \quad V \quad - \quad ing]$ in place of $\overline{W}$

<table>
<thead>
<tr>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>*(216c)</td>
<td>✓</td>
</tr>
</tbody>
</table>

‼‼ This prediction of the GB binding theory does not present an actual problem, since sentences like (216c) are ruled out by an independent principle, the Avoid Pronoun Principle.

The summary in (231) shows that, while the GB binding theory makes
correct predictions in some cases where the OB binding theory makes wrong predictions, it also makes wrong predictions in some other cases where the OB binding theory makes the correct predictions. Chomsky (1981a:209) presents the GB binding theory as "a considerable improvement over OB on empirical and conceptual grounds ...". Cases where the GB binding theory, but not the OB binding theory, makes wrong predictions about the acceptability of sentences are clearly problematic for the GB theory. Let us consider Chomsky's reaction to these problematic cases.

That there are cases in which the OB binding theory, but not the GB binding theory, makes the correct predictions about the acceptability of sentences, is already acknowledged in (Chomsky 1981d) and (Chomsky 1979b). Chomsky (1981d:141f.), in particular, contains a fairly detailed discussion of the problems that each other in NP poses for the GB binding theory --- see case (231a) and (231b). In essence, the strategy adopted by Chomsky in 1979 with respect to these problems, is to claim that the relevant predictions of the GB binding theory are actually correct, and that it is the predictions of the OB binding theory that are wrong. Consider in this connection the introductory remarks to (Chomsky 1979b).

(232) "At the GLOW talk [= (Chomsky 1981d) --- M.S.] , I discussed some conceptual problems that arise in a theory of the OB type. I suggested another approach [= the GB theory --- M.S.] which I suggested would overcome to some degree, sometimes completely, sometimes not, these conceptual problems. I also mentioned that this approach has different empirical consequences. It's incomparable in empirical coverage in that it properly explains some things which were not explained in the other theory, but it doesn't cover some of the material in the other theory. I then suggested that that was a good result since the things covered in this theory seem very central whereas the things covered in the other theory and not in this one seem rather peripheral. We may look to the future Markedness theory to justify these differences."

The final remarks quoted in (232) make it clear that Chomsky appeals to the notion 'markedness' in order to reinterpret the apparently wrong predictions of the GB binding theory as actually correct, and the apparently correct predictions of the OB binding theory as actually
wrong. This point also emerges clearly from the discussion in (Chomsky 1981d:140-145). The role of the notion 'markedness' in Chomskyan syntax, and its relation to the notion 'core grammar', were outlined in §3.4 above. In order to understand exactly what Chomsky means in (232), it is necessary to consider in detail the role of the notion 'markedness' in Chomsky's (1981d) interpretation of the empirical differences between the OB and CB binding theories. The sentences discussed in (Chomsky 1981d:140-142) are the following:

(233) John read \([NP \text{his books}]\)

(234) *They'd prefer \([NP \text{each other's writing the book}]\)

(235) They read \([NP \text{each other's books}]\)

(236) *They found \([NP \text{some books for each other to read}]\)

(237) a. *They heard \([NP \text{my stories about each other}]\)
    b. They heard \([NP \text{the stories about each other} \text{ (that had been published last year)}]\)
    c. They heard \([NP \text{stories about each other}]\)

(238) a. *They expected that \([NP \text{my pictures of each other}] \text{ would be on sale}\)
    b. They expected that \([NP \text{the pictures of each other}] \text{ would be on sale}\)
    c. They expected that \([NP \text{several books about each other}] \text{ would be on sale}\)

(239) a. They think it is a pity that \([NP \text{pictures of each other}] \text{ are hanging on the wall}\)
    b. *They think he said that \([NP \text{pictures of each other}] \text{ are hanging on the wall}\)
In (233) there is a pronoun in the position of $\alpha_7$ in (203b). (233) thus corresponds with (216a) above. In (234) and (235) each other, an overt anaphor, appears in the position of $\alpha_7$ in (203b). (234) corresponds with (209), and (235) with (207a). In (237)-(241) each other appears in the position of $\alpha_6$ in (203a). These sentences correspond with those in (204c, d, e). In (236) each other appears in the position of $\alpha_4$ in (189). I will return to the status assigned to this sentence in (Chomsky 1981a) below. The differences between the predictions which the OB and GB binding theories make about these sentences, according to Chomsky (1981d:140f.), are summarized in (242).

(242)  

<table>
<thead>
<tr>
<th></th>
<th>OB</th>
<th>GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(233)</td>
<td>$\times$</td>
<td>✓</td>
</tr>
<tr>
<td>* (234)</td>
<td>$\times$</td>
<td>✓</td>
</tr>
<tr>
<td>(235)</td>
<td>✓</td>
<td>$\times$</td>
</tr>
<tr>
<td>* (236)</td>
<td>$\times$</td>
<td>✓</td>
</tr>
<tr>
<td>* (237a)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(237b, c)</td>
<td>✓</td>
<td>$\times$</td>
</tr>
<tr>
<td>* (238a)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(238b, c)</td>
<td>✓</td>
<td>$\times$</td>
</tr>
<tr>
<td>(239a)</td>
<td>✓</td>
<td>$\times$</td>
</tr>
<tr>
<td>* (239b)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(240a, b)</td>
<td>✓</td>
<td>$\times$</td>
</tr>
<tr>
<td>(241a)</td>
<td>✓</td>
<td>$\times$</td>
</tr>
<tr>
<td>* (241b)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Certain aspects of these predictions are in need of explication.

(i) The definition of 'government' adopted in 1979 in the GB (MCG) and GB (Pisa) binding theories differs from the definition adopted in 1981 in the GB (LGB) binding theory with respect to the class of categories that are considered to be barriers to government. On the 1979 definition $S$ and NP are taken as absolute barriers to government --- see §§4.2 above. On the 1981 definition all maximal projections are taken as absolute barriers to government. This difference entails that the predictions made about sentences such as (234)/[20ii] by the GB (MCG) and GB (Pisa) binding theories (the 1979 versions of the GB binding theory) differ from those made by the GB (LGB) binding theory (the 1981 version). As was noted in §§3.2.4.3 above, the 1981 version wrongly predicts that such sentences are acceptable since the VP internal to the NP blocks government --- see e.g., the discussion of (209) above. Thus, both the OB binding theory and the GB (LGB) binding theory make the wrong predictions about such sentences.

(ii) As regards sentences such as (235)/[21], corresponding to (207a) above, the GB (LGB) binding theory makes either the wrong or the correct prediction, depending on the exact definition of 'government' --- see the discussion following (207) above for details.

(iii) Sentences such as (236)/[22] are regarded as unacceptable by Chomsky (1981d), with the GB binding theory making the correct prediction and the OB theory making the wrong prediction. Chomsky (1981a:216) claims that, while such sentences are unacceptable to him, most speakers tend to regard them as acceptable. If the latter judgment is correct, then the advantage which Chomsky (1981d) claims the GB binding theory has over the OB binding theory, becomes a disadvantage. See §§5.5 below for a further discussion of such sentences.

(iv) Chomsky (1981d:142) claims that (237)/[23] and (238)/[24] illustrate that it is not a "definiteness restriction" that is involved in such sentences, as is the case in, e.g., (243)/[28].
(243) a. *Who did they hear \[NP \ \text{my stories about } t \]  
   b. *Who did they hear \[NP \ \text{the stories about } t \]  
   c. Who did they hear \[NP \ \text{stories about } t \]  

In (243b) the definiteness of the NP from which the \textit{wh}-phrase is moved is responsible for the unacceptability of the sentence. In (237b) and (238b) each other appears within a definite NP, and is bound outside this NP. Nevertheless, these sentences are acceptable.

(v) Chomsky (1981d:143) claims that (239)/[25], (240)/[26], and (241)/[27] illustrate that the subject that creates an opaque domain, i.e., that invokes the SSC/Opacity Condition, must be a possible argument. Thus, \textit{it} and \textit{there} do not invoke Opacity.

(vi) Chomsky (1981d:143) claims that (240)/[26] illustrates that the notion 'subject' is crucial, and not the notion 'possible antecedent'. In (240b) \textit{me} is in the position of a possible antecedent, but it does not create an opaque domain, so blocking coindexing of \textit{they} and each other.

Chomsky's interpretation of the data presented in (242) is as follows. The OB and GB binding theories make different predictions about the markedness of the sentences (233)-(241). The OB binding theory predicts that (233)/[18], (234)/[20ii] and (236)/[22] will be marked, and that all the others will be unmarked. The GB binding theory predicts that (235)/[21], (237b, c)/[23i, ii], (238b, c)/[24ii, iii], (239)/[25i], (240a, b)/[26i, ii] and (241a)/[27i] will be marked and that all the others will be unmarked. Chomsky (1981d:141) claims that the markedness predictions of the GB binding theory are correct, and that those of the OB binding theory are wrong. In support of this claim, Chomsky presents the following considerations. He (1981d:141) claims that (233)/[18] "is surely the normal case in the languages of the world", while "structures such as (21) [= (235) --- M.S.] appear to be rare". Moreover, the OB principle that permits (235)/[21] does not extend to similar structures such as (234)/[20ii], "which perhaps
represents a more general case across languages". As regards the un-starred sentences in (237)-(241)\textsuperscript{[23]-[27]}, Chomsky (1981d:143) claims that they are marked because (i) they "seem somewhat marginal", (ii) "judgments tend to vary", and (iii) "there appear to be differences in judgment depending on lexical choice".

In order to accommodate the cases that are marked under the GB binding theory, Chomsky (1981d:143) proposes that the grammar of English incorporate something like the following marked principle.

(244) "each other may be free in its governing NP if it is not free in the c-command domain of a lexical subject."

He notes that the exact formulation of the principle "depends on some rather questionable factual judgments involving not only each other, but also reflexives and pronouns". He also notes that (244) is in fact derivable from the GB binding theory, but that it is inoperative since a more restrictive condition --- namely, that each other must be bound in all its governing categories --- also follows. (244) applies only in special marked structures in which the general principles of the GB system are relaxed. Chomsky (1981d:145) claims that this may be "an example of the kind of 'analogic process' (in a rather abstract sense of the term) that we might expect to find outside of the central core of the system of grammar".

The approach adopted in (Chomsky 1981a) towards those cases in which the GB binding theory apparently makes the wrong predictions, differs substantially from the approach outlined above. Summarizing the relative merits of the OB and GB binding theories, Chomsky (1981a:209) says that "the GB-framework is a considerable improvement over OB on empirical and conceptual grounds, but there are still problems in the case of arguments within NPs ...". Apart from these empirical problems, Chomsky (1981a:207) also identifies a conceptual problem of the GB theory. Within this theory, there is no explanation of why NP and S are the two governing categories. Chomsky (1981a:207f.) attempts to modify the GB binding theory in such a way that it overcomes both this conceptual problem and the empirical problems relating to argu-
ments within NPs. Referring to his discussion of possible modifications to the GB binding theory, Chomsky (1981a:216) mentions that "it may be that this entire discussion properly belongs to the theory of markedness rather than of core grammar, and that the phenomena we have been discussing reflect marked properties of English". However, no mention is made of the specific considerations used in (Chomsky 1981d) to justify the claim that cases about which the GB binding theory apparently makes the wrong predictions are in fact marked.

The crucial difference between the 1979 and 1981 approaches to the data that are problematic for the GB binding theory can be summarized as follows. In the works that appeared in 1979, it is claimed that the relevant data represent marked phenomena, and that the mechanism responsible for them falls outside core grammar. In 1981, while the possibility that the data represent marked phenomena is not ruled out, an attempt is made to modify the GB binding theory to enable it to handle these data. The modified version of the GB binding theory must also overcome a conceptual problem of the earlier version of this theory, namely, the absence of an explanation of why NP and S are the governing categories. The proposed modifications to the GB binding theory are discussed in §5.5 below.

5.5 Some proposed modifications to the GB binding theory

5.5.1 General remarks

In §5.4.2 above I have outlined certain differences that exist among three different versions of the GB binding theory, namely GB (MCG), GB (Pisa), and GB (LGB). In spite of these differences, the three versions are similar in an important respect: the conditions under which the various types of arguments must be bound or free are the same for all three versions. In each case, the theory stipulates the boundedness, or otherwise, of an argument within the minimal NP or S that contains both the argument and its governor. Chomsky (1981a: 209-216) argues that the GB binding theory must be reformulated in terms of the notion 'accessible SUBJECT'. In order to qualify as a governing category, a category must not only contain a governor, but
also an accessible SUBJECT. Chomsky claims that this reformulation overcomes the conceptual problem noted in §5.4.3 above, as well as the majority of the empirical problems relating to arguments within NPs left unsolved by the three earlier versions of the GB binding theory. In the discussion that follows the term "the GB-SUBJECT binding theory" is used to refer to this reformulated version of the GB binding theory. The term "the GB-governor binding theory" is used to refer collectively to the three versions discussed in §5.4.2. The differences among these versions are irrelevant for the present discussion. An exposition of Chomsky's formulation of the GB-SUBJECT binding theory is presented in §5.5.2. In §5.5.3 I discuss some additional modifications to this binding theory proposed by Chomsky.

5.5.2 Reformulating the GB binding theory in terms of the notion "accessible SUBJECT"

Chomsky (1981a:209-210) first illustrates how the SSC/Opacity Condition and some version of the PIC/NIC can be unified in terms of the notion 'SUBJECT'. He (1981a:210) points out that the version of the PIC/NIC involved in this unification, is the one proposed by George and Kornfilt. This version takes agreement, rather than Tense, as the crucial element invoking opacity, since there is a one-to-one correlation between the two. In other languages --- e.g., Turkish --- where the two are disassociated, it is agreement, rather than Tense, that determines opacity. Chomsky remarks that "we have tacitly been assuming the accuracy of the George-Kornfilt theory all along, taking AGR --- the governor of the subject --- to be the crucial element determining opacity".

The SSC/Opacity Condition and the relevant version of the PIC/NIC can be unified in terms of the notion 'SUBJECT' in the following way. In an infinitive, NP, or small clause, the subject is the SUBJECT. In clauses where INFL contains AGR (as is the case in tensed clauses in English), AGR is the SUBJECT. In clauses (245a) AGR is thus the SUBJECT. In the embedded clause of (245b) John is the SUBJECT, and AGR is the SUBJECT of the matrix clause.
Suppose that the SSC/Opacity Condition were reformulated in terms of the notion 'SUBJECT', stipulating that no anaphor may be free in the domain of SUBJECT. The PIC/NIC would then be reduced to this reformulated version of the SSC/Opacity Condition, i.e., to the case where SUBJECT is AGR. There would be only one opaque domain, namely the domain of SUBJECT.

Such a reformulation of the SSC/Opacity Condition and PIC/NIC has a conceptual advantage over the GB binding theory. The reformulated version of the SSC/Opacity Condition and PIC/NIC provides an answer to the question of why NP and S are the categories in which anaphors cannot be free, i.e., why NP and S are the governing categories. In essence, the answer is that NP and S contain SUBJECTS, where a SUBJECT creates an opaque domain. The full answer is as follows. S is a governing category for $\infty$, since it always contains a SUBJECT. $^{110}$ NP is a governing category for $\infty$ only when it contains a SUBJECT $\neq \infty$. Other categories, e.g., adjective phrases, may also be governing categories if they contain subjects (hence SUBJECTS). $^{111}$ Given the distinction made in the reformulated SSC/Opacity Condition between NPs with subjects and NPs without subjects, it follows that the predictions of this condition about arguments in NPs will differ from the predictions made by the GB binding theory. They will, however, make the same predictions about arguments in clauses.

Chomsky (1981a:211) claims that both the GB-governor binding theory and the reformulated SSC/Opacity Condition have attractive features. The first solves many of the conceptual problems of the OB binding theory, and the second solves a conceptual problem of the GB-governor binding theory. Chomsky (1981a:211f.) therefore tries to amalgamate these two theories. The amalgamated theory, the GB-SUBJECT binding theory, has the following component's.

(i) The binding principles of the GB-governor binding theory --- see (172) above --- are taken over without any modification.
(ii) The principle (246) is adopted.

(246) "AGR is coindexed with the NP it governs" [3.2.3 (701)]

Given the general condition that a coindexed NP and pronominal (pronoun or PRO) must share the appropriate features, (246) accounts for the phenomenon of agreement. Recall that $AGR = PRO$. (246) thus reduces the phenomenon of obligatory subject-verb agreement to general properties of proximate pronouns. According to Chomsky (1981a:216), the principle (246) is required in some form in any theory.

(iii) The GB-SUBJECT binding theory incorporates a new definition of 'governing category'.

(247) "$\beta$ is a governing category for $\alpha$ if and only if $\beta$ is the minimal category containing $\alpha$, a governor of $\alpha$, and a SUBJECT accessible to $\alpha$.

It follows from (247) that $\beta$ is a governing category only if it has a SUBJECT. $S$ can thus always be a governing category, and NP can be a governing category when it has a subject. The same holds for $\Lambda^i$, i.e., small clauses. According to Chomsky (1981a:211), the choice of governing category receives "a rather natural characterization" in terms of (247). The conceptual problem faced by the earlier versions of the GB binding theory, namely, the problem of explaining why NP and $S$ are the governing categories, is now solved. The question of whether $S$ or $\overline{S}$ should be selected as the governing category no longer arises. Note that in (Chomsky 1979b:8) no justification is provided for the decision to take $S$ as the governing category.

(iv) The notion 'accessible' in (247) is defined in (249), in terms of the well-formedness condition (248).

(248) "$[\gamma \ldots \hat{\delta} \ldots ]$, where $\gamma$ and $\hat{\delta}$ bear the same index." [3.2.3 (73)]
" is accessible to $\beta$ if and only if $\beta$ is in the c-command domain of $\alpha$ and assignment to $\beta$ of the index of $\alpha$ would not violate (73) $= (248)$ --- M.S." 

Chomsky (1981a:212) claims that (248) holds for a variety of constructions apart from those directly relevant in the present case. In all the cases of (250) this well-formedness condition is violated.

(250)  
a. *[$N^p_i$ the friends of $[i$ each other's]
   parents]

b. *There is [$N^p_i$ a picture of $[N^p_i$ itself] ] on
   the mantelpiece

c. *[$N^p_i$ the owner of $[N^p_i$ his] boat]]

d. *[$N^p_i$ the friends of $[N^p_i$ their] parents]]

In (250b), for example, the NP *itself (= $x$) is contained in another NP the friends of each other's parents (= $y$), and $x$ and $y$ bear the same index. (248) thus marks (250b) as ill-formed. Note that (249) refers to possible, not actual indexing of $\beta$. That is, $\alpha$ becomes inaccessible to $\beta$ if coindexing of $\alpha$ and $\beta$ would lead to a violation of (248).

The GB-SUBJECT binding theory makes the same predictions about arguments in clauses as the GB-governor binding theory. The nominative subject of a clause always has an accessible SUBJECT, namely the AGR element of INFL, which also governs the subject. Hence the clause is a governing category. The nominative anaphor in such a clause must therefore be bound in it, which is impossible. Pronominals must be free in this clause. (PRO can in fact never appear in the relevant position, since it is governed.) Since clauses must have subjects, hence SUBJECTS, which are accessible, the governing categories in the other cases received remain the same.

However, different predictions are made about arguments within NPs. Consider firstly the case of overt anaphors, such as each other. These
fall under binding principle (A) of (172). In (251)-(260) I present
the examples with each other in the position of $\alpha_1$ discussed by
Chomsky (1981a:216, 217). Wherever an example is the same as, or
exactly analogous to, an example discussed in §§5.3.2.4.3 and 5.4.3
above, I will also provide the number of the latter example. These num-
bers appear directly below the relevant example.

(251)  
  a.  they heard $\left[_{\text{NP}} \text{my stories about each other} \right]$ [3.2.3 (78)]
         (= 204b, 237a)
  b.  they heard $\left[_{\text{NP}} \text{the stories about each other} \right]$ 
         (= 204c, 204d, 237b)
  c.  they heard $\left[_{\text{NP}} \text{stories about each other} \right]$ 
         (= 204e, 237c)

(252)  
  a.  *they expected $\left[_{\text{S*}} \text{me to hear}_{\text{NP*}} \text{ stories about each other} \right]$ [3.2.3 (79)]
  b.  they expected that $\left[_{\text{S*}} _{\text{NP*}} \text{pictures of each other} \right.$ 
         would be on sale$\left. \right]$ 
         (= 238c)
  c.  they expected that $\left[_{\text{S*}} _{\text{NP*}} \left\{ \text{PRO feeding each other} \right\} \right.$ 
         would be difficult$\left. \right]$ 

(253)  *they thought $\left[_{\text{S*}} \text{I expected that pictures of each other would be on sale} \right]$ 
         (= 239b)

(254)  they thought $\left[_{\text{S*}} \text{I expected that } \left\{ \text{feeding each other} \right\} \right.$ 
         would be difficult$\left. \right]$
(255)  

a.  they think it is a pity that pictures of each other are hanging on the wall  

(= 239a)  

b.  *they think he said that pictures of each other are hanging on the wall  

(= 239b)  

(256)  

a.  they think there are [some letters for each other] at the post office  

(= 241a)  

b.  *they think he saw [some letters for each other] at the post office  

(= 241b)  

(257)  

a.  I think it pleased them that pictures of each other are hanging on the wall  

(= 240a)  

b.  they think it pleased me that pictures of each other are hanging on the wall  

(= 240b)  

(258)  

they found [NP some books [S for each other to read]]  

(= 236)  

(259)  

*they preferred [NP each other's reading the book]  

(= 209, 234)  

(260)  

*they thought [S* I preferred each others reading the book]  

[3.2.3 (82)]  

[3.2.3 (83)]  

[3.2.3 (86)]  

[3.2.3 (88)]  

[3.2.3 (89)]  

[3.2.3 (87)]  

[3.2.3 (90)]  

[3.2.3 (91)]
Chomsky shows that, with the exception of (259), the GB-SUBJECT binding theory makes the correct predictions about all these cases. This contrasts with the GB-governor binding theory, which makes the wrong predictions about (251b, c), (252b), (255a), (256a), (257a, b) --- see §§5.3.2.4.3 and 5.4.3 above for details. Let me briefly explicate the predictions of the GB-SUBJECT binding theory about the sentences (251)-(260).

In (251a), but not (251b, c), the object NP contains a SUBJECT accessible to each other. The GB-SUBJECT binding theory thus requires that each other be bound in NP in (251a), but not in (251b, c). It thus correctly predicts the unacceptability of (251a) as opposed to the acceptability of (251b, c).

In (252a) NP* contains a governor for each other, but no accessible SUBJECT. Therefore, NP* is not a governing category for the anaphor. S* does contain an accessible SUBJECT, namely the subject me. S* is thus the governing category for each other in which it must be bound. The GB-SUBJECT binding theory thus correctly predicts the unacceptability of (252a). In (252b) NP* does not contain a SUBJECT accessible to each other. S* does contain a SUBJECT, namely AGR. AGR is coindexed with NP* by principle (246). It is thus not accessible to each other because of the well-formedness condition (248). It follows that S* is not a governing category for each other. The matrix clause, however, is a governing category for each other, since it contains an accessible AGR element. (252b) is thus acceptable with coindexing of they and each other.

(252b) contrasts with (253). In (253) S* does contain a SUBJECT accessible to each other, namely the subject I. The GB-SUBJECT binding theory thus correctly predicts the unacceptability of (253), with each other free in S*.

In (252c) NP* is a governing category for each other, since it contains the accessible SUBJECT PRO (coindexed with they). each other must therefore be bound in NP*. The GB-SUBJECT binding theory thus correctly predicts the acceptability of (252c), with they and each other coreferential. As regards PRO, the binding theory merely requires that it be ungoverned. This is the case in (252c), and also in (254).
Chomsky (1981a:214) remarks that the contrast between (255a) and (255b) may be attributed to a phenomenon "that has been frequently discussed in connection with the SSC, namely, that the nature of the subject that creates the opaque domain figures in determining the degree of violation of opacity, with agentive subjects inducing maximal violation and nonarguments minimal violation." He claims (1981a:214) that consideration of other examples, such as those in (261), suggests another approach.

(261)  a. *they think [it bothered each other that S] [3.2.3 (84)]
   b. *he thinks [it bothered himself that S]
   c. he thinks [it bothered him that S] (him proximate to he)

Chomsky argues that these examples indicate that it is not the agentivity of the subject it that explains the difference between (255a) and (255b). In (261) AGR in the embedded clause is a SUBJECT accessible to the italicized anaphor or pronominal. This clause is therefore the governing category for the anaphor or pronominal. By principle (A) the anaphor must be bound in it (hence the unacceptability of (261a, b). By principle (B), the pronoun must be free in it (hence the possibility of interpreting (261c) with him proximate to he). The examples (261) thus fall together with (255b), and contrast with (255a). The problem is then to explain the difference between (255a) and (255b) on the one hand, and between (255a) and (261) on the other hand.

According to Chomsky (1981a:215), the crucial difference between (255a) and (261) is the following. In (255a) the anaphor is internal to the extraposed clause associated with it, while in (261) the anaphor (or pronoun) is external to that clause. Suppose that the association between it and the extraposed clause is given by the standard device of coindexing. Then (255a) is of the form (262a), and (261) is of the form (262b).

(262)  a. they think [\textit{s}\textsuperscript{*} it \textit{AGR} \textit{VP} is a pity] [3.2.3 (85)]
   \[\textit{i} \text{that pictures of each other are hanging on the wall}\]
b. NP think(s) \[S^* \text{ it}_i \text{ AGR}_i \text{ VP bother } \alpha \] \[i \text{ that } S \]

In (262) it and AGR are coindexed by principle (246). In (262a), AGR of S* is not accessible to each other, since coindexing of the two would violate the well-formedness condition (248). Therefore S* is not the governing category for each other. The full clause is the governing category for each other, with AGR the accessible SUBJECT. In (262a) (\(=255a\)) each other is bound in its governing category (being coindexed with they). The GB-SUBJECT binding theory thus correctly predicts the acceptability of this sentence.

In (262b) AGR of S* is accessible to \(\alpha\), so that S* is its governing category. (261a, b) are thus unacceptable with the anaphor free in S*, and (261c) is acceptable with him free in S*, but bound by he in the matrix clause. The GB-SUBJECT binding theory makes the correct predictions in all the cases. Example (255b) falls together with (261), with an accessible SUBJECT (he) in S*. The distinction between (255a) and (255b) thus does not lie in a difference in the properties of the nonargument it and the argument he. Rather, the distinction lies in the accessibility of a SUBJECT in S* in (255b) as opposed to the inaccessibility of a SUBJECT in S* in (255a). The examples (256) illustrate that there has the same effect as it in (255a), on the assumption that there and the extraposed clause are coindexed.

The acceptability of (257a) follows from the GB-SUBJECT binding theory, in exactly the same way as the acceptability of (255a). This theory predicts that (257b) is also acceptable. According to Chomsky (1981a: 216), judgments vary concerning (257b). But, since it is much better than (255b), he takes the prediction of the GB-SUBJECT binding theory to be correct.

The OB binding theory predicts the acceptability of (258), as does the GB-SUBJECT binding theory. Only the matrix clause contains an accessible SUBJECT, and each other, being coindexed with they, is bound in this clause. The GB-governor binding theory predicts that (258) will be unacceptable. In 1979 Chomsky took the latter as the correct prediction --- see Chomsky 1981d. He (1981a:216) says that this accords with his own judgment. However, he recognizes that most
speakers regard (257) as acceptable, in which case (257) does not pose a problem for the GB-SUBJECT binding theory.

The GB-SUBJECT binding theory makes the correct predictions about all the examples presented in (251)-(260), with the exception of (259). The OB binding theory also makes the wrong predictions about (259). However, Chomsky notes that it is possible that the unacceptability of (259) follows from a different requirement, namely the plurality condition for reciprocals —— see §5.3.2.4.3 above for a brief discussion of this condition.

Chomsky (1981a:217) suggests that the GB-SUBJECT binding theory may nevertheless be an improvement over the GB-governor binding theory with regard to (259). The GB-governor binding theory does not prevent each other in (260) from being bound by they, since it is ungoverned in S*. In the GB-SUBJECT binding theory such coindexing is barred. AGR of S* is an accessible SUBJECT for each other, and each other must therefore be bound in S*. Under the GB-SUBJECT binding theory, (260) is thus ruled out by both the binding theory and the plurality requirement. On these assumptions, one would expect (260) to be more unacceptable than (259). However, Chomsky admits that "one can hardly rely on such judgments". He suggests that "comparative evidence should prove relevant". The existence of parametric variation with respect to these constructions is noted in (Chomsky 1981a:228, fn. 57).

Consider now examples with pronouns, which fall under principle (172) (B). In the examples of (263) he is proximate to John.

(263) a. John saw [NP* my picture of him] \[3.2.3 (92)\]
   (= 215a)

b. *I saw [NP* John's picture of him]
   (= 215b)

c. *John saw [NP* a picture of him]
   (= 215c)

d. John thought \[S* I saw [NP a picture of him] \]
   (= 215d)
The GB-governor binding theory makes the correct predictions about all these sentences, except (263c, e, g). (263g) is no real problem, since its status is determined by the Avoid Pronoun Principle. The GB-SUBJECT binding theory makes the correct predictions about all the examples in (263), except (263e, g), with (263g) accounted for by the Avoid Pronoun Principle, as noted. NP* and S* are the governing categories. In (263a, d, f) he is free in its governing category. Thus these examples are acceptable. In (263b, c) him is bound in its governing category. Hence the unacceptability of these examples. What remains a problem, then, is (263e). Chomsky (1981a:217) points out that either (263e) or (264) does not fall under the binding theory, since in these constructions the pronoun is not free in the position where the anaphor is bound.

(264) they read \[\text{NP } \text{each other's books}\] [3.2.3 (93)]

Chomsky assumes that (264) falls under the binding theory. Some other condition is then needed to permit (263e): a condition which would "override" the requirements of the binding theory. Chomsky points out that it has been suggested that his in (263e) is an obligatory variant of himself's, which is excluded from (263e), contrary to the prediction of the binding theory. He (1981a:218) remarks that the near complementary distribution between proximate pronouns and reflexives is only partially captured within any of the approaches investigated in (Chomsky 1981a). Given differences among languages as to whether the analogue of his in (263e) can be regarded as proximate to the matrix subject, it is possible that something other than the binding theory is involved in these cases.
In sum: the GB-SUBJECT binding theory overcomes a conceptual problem left unsolved by the GB-governor binding theory, i.e., it provides an explanation of why NP and S are the governing categories. The GB-SUBJECT binding theory makes the correct predictions in all those cases where the GB-governor binding theory makes the correct predictions. The GB-SUBJECT binding theory also makes the correct predictions about examples with anaphors as arguments in NPs, about which the GB-governor binding theory makes wrong predictions.

5.5.3 Further modifications to the GB-SUBJECT binding theory

Chomsky (1981a:219-221) considers two further modifications to the GB-SUBJECT binding theory: (i) the addition of another principle to the theory of government, and (ii) a redefinition of the notion 'governing category'.

Following a suggestion by Hornstein, Chomsky (1981a:220) proposes that the following principle be adopted as part of the theory of government.

(265) "A root sentence is a governing category for a governed element." [3.2.3 (99)]

The adoption of the principle (265) makes it possible to overcome a problem concerning anaphors noted by Rizzi. In sentences such as (266), anaphors are governed, but lack governing categories, because they do not have accessible SUBJECTS.

(266) a. [for each other to win] would be unfortunate [3.2.3 (97)]
   b. [for [i each other] to win] would be unfortunate for [them_i] [3.2.3 (98)]

Both (266a) and (266b) are in fact unacceptable. The GB-SUBJECT binding theory incorrectly predicts that they are acceptable, since each other has no governing category in which it must be bound. If (266) is to
barring because it has no interpretation, it will make it impossible to adopt the simplest rule for the interpretation of each other, namely, apply the rule to any coindexed pair (NP, each other). While this rule will fail to assign an interpretation to (266a), it will (wrongly) assign an interpretation to (266b). The adoption of the principle (265) solves this problem. each other in (266) now has a governing category, namely the matrix clause. In both (266a) and (266b) each other is free in this governing category. In (266a) each other is not coindexed with another category, and in (266b) it is not coindexed with a c-commanding category. --- see the definitions of 'bound' and 'free' in (174), (175) above.

Chomsky (1981a:220f.) considers the possibility of simplifying the definition of 'governing category'. He proposes the elimination of the reference to government. The term 'governing category' must then be replaced by something else. He proposes that it be replaced by 'binding category', defined as in (267).

(267) "β is a binding category for α if and only if β is the minimal category containing α and a SUBJECT accessible to α."

[3.2.3 (100)]

Binding principles (172) (A) and (B) are correspondingly reformulated as (268), and principle (265) is modified to (269).

(268) "(A) An anaphor is bound in its binding category

(B) A pronominal is free in its binding category."

[3.2.3 (101)]

(269) "A root sentence is a binding category for a governed element."

[3.2.3 (102)]

Chomsky (1981a:220-221) claims that this proposed revision has "no meaningful consequences" in the case of overt elements or NP trace. Consider the case of PRO. The basic property of PRO is that it is un-governed. As we have seen, this follows from the binding theory formulated in terms of 'government'. The ungoverned status of PRO no
longer follows from the binding theory revised as in (268). What follows from (268) is that PRO lacks a binding category, which does not imply that PRO is ungoverned. However, the conclusion that PRO is ungoverned does follow from (268) in conjunction with (269). Note that (269) is also required in the unrevised theory. If PRO is governed, then by (269) it always has a binding category, in which it must be both bound and free by (268). Hence, PRO is ungoverned.

Sentences like (270) are problematic for the proposed simplification of the GB-SUBJECT binding theory.

(270)  
a. John expected [him to win]  
b. John tried [ [PRO to win] ]  
c. John knows [how [PRO to win] ]

As Chomsky (1981a:221) points out, him cannot be coindexed with John, because then (268B) will be violated. But exactly the same argument shows that PRO cannot be indexed with John in (270b, c), since the matrix clause is the binding category for PRO. This is the wrong result, however. The unrevised theory --- i.e., with "governing category" in place of "binding category" --- gives the correct result. In (270b, c) PRO has no governing category, and no requirement of unboundedness is thus made by principle (B) of the binding theory. These examples indicate that it is necessary to introduce a crucial reference to government in the binding theory, i.e., that the simplification to (268) cannot be adopted. However, Chomsky (1981a:221) claims that the effects of such a reference to government "are so narrow as to suggest that an error may be lurking somewhere".

5.6 Some problems for the GB binding theory

Although the GB binding theory, and in particular the GB-SUBJECT binding theory, manages to solve many problems left unsolved by its predecessors, it is not itself without problems. Chomsky (1981a) points out various problems left unsolved by this theory. In some cases Chomsky merely mentions the existence of a problem. In other cases he provides a detailed exposition of the exact nature of the problem. In this sec-
tion I will briefly discuss the various problems pointed out by Chomsky. Suggestions by Chomsky about possible solutions to these problems will also be noted.

(i) The status of reflexives

Chomsky (1981a:218) notes that the concept 'anaphor' "has been left rather vague in the preceding discussion ...". Informally, an anaphor is characterized as "an NP with no intrinsic reference". Reflexives fall under this characterization. Their cross-linguistic status with respect to the binding conditions is problematic, however. Chomsky (1981a:229, fn. 62) points out that the analogue to the reflexive in languages such as Japanese and Korean does not obey the binding conditions. It has been argued that, in some languages at least, such elements can be subject to pragmatic control. That is, they can be without antecedents.118)

(ii) Pronouns in NP

As was noted in §5.5.2 above, no version of the GB binding theory, including GB-SUBJECT, makes the correct predictions about all cases of pronouns in NPs. For example, the GB-SUBJECT binding theory incorrectly predicts that his in (271) (presented as (263e) above) must be free.

(271) [s* John read [NP his book]]

(iii) Constructions to which the binding theory must be extended

Chomsky (1981a:225 fn. 36; 230:fn. 73) points out that there are cases to which the GB theory, including GB-SUBJECT, must be extended. He mentions constructions with left-dislocated items, topics, predicate nominals, heads of relatives, and subjunctives. Only in the case of the latter does he (1981a:230, fn. 73) give some indication of the nature of the problems involved for the GB binding theory. He claims
that, although they contain no overt AGR element, subjunctives behave like tensed clauses with respect to binding. (Unfortunately, Chomsky does not provide an example to illustrate this point.) For the present, he says, it must simply be stipulated that there is a null AGR in English subjunctives.

Chomsky (1981a:219) also points out that certain elements that do not function as anaphors in the narrow sense that applies to NP-trace, each other, PRO, etc., appear to obey the binding conditions. The sentence presented in (272) illustrates that the trace of extraposition is apparently subject to the binding conditions.

(272) *[John's novel t] arrived last week [that you ordered] [3.2.3(95i)]

In (272) the subject NP is the governing category for the trace of the extraposed clause. The unacceptability of (272) thus follows from the fact that this trace is free in its governing category. Chomsky refers to works by Fiengo and Lasnik and Quicoli, in which evidence is presented that the relation between an NP and a displaced quantifier (or maybe the relation between the trace of the quantifier and the quantifier) is also subject to opacity. Chomsky points out that these cases are relevant to "a proper understanding of the notion 'anaphor'."

(iv) The AUX-to-COMP case in Italian

Chomsky (1981a:225, fn. 36) briefly refers to an Italian construction derived by AUX-to-COMP movement with a nominative subject. This construction is discussed in an unpublished paper by Rizzi. Chomsky (1981c) also discusses the relevant construction, and the problems which it presents for the GB-SUBJECT binding theory. The details that follow are from the discussion in the latter work.

Chomsky (1981c:3) assumes that the gerundive construction is of the form (273).
G in (273) is a realization of INFL, with the property that it may or may not be a governor. Italian examples with the structure (273) are presented in (274).

\[(274)\]
\[
\begin{align*}
\text{a. Gianni avendo telefonato, ...} & \quad \text{[14i]} \\
\text{b. PRO avendo telefonato, ...} & \quad \text{[14ii]}
\end{align*}
\]

Italian has an AUX-to-COMP rule, that applies to structures such as (273), yielding structures such as (275).

\[(275)\]
\[
\begin{align*}
\text{AUX} & \quad G \quad [\text{NP} \ldots ] & \quad \text{[19]}
\end{align*}
\]

Although Chomsky does not spell this out, it follows that in (273) there appears a COMP to the left of NP, and an AUX in G. Chomsky (1981c:7) claims that G in (275) must be a governor, with the NP governed by AUX in COMP. PRO is thus excluded from the NP position in (275), since PRO must be governed. Consequently, the AUX-to-COMP rule derives (276a) form (274a), but not (276b) from (274b).

\[(276)\]
\[
\begin{align*}
\text{a. avendo Gianni telefonato, ...} & \quad \text{[21i]} \\
\text{b. *avendo PRO telefonato, ...} & \quad \text{[21ii]}
\end{align*}
\]

According to Rizzi, the nominative subject position in gerundive structures derived by the AUX-to-COMP rule is opaque. This follows quite straightforwardly from the NIC. Chomsky (1981a:225, fn. 36) states that the island properties of the relevant structure also follow from the GB-governor binding theory. If G is indeed a governor, as is claimed by Chomsky (1981c:7), then the clause dominating the structure (275) is a governing category for the subject. By principle (C) of (172) R-expressions must be free in this subject position.

Under the GB-SUBJECT binding theory, an additional assumption is needed. The relevant clause must not only be a governing category for the sub-
ject NP. It must also contain an accessible SUBJECT for this NP. Chomsky (1981c:8) states that the latter assumption implies that G in COMP must have the same property as AGR, that is, it must be an accessible SUBJECT. He states that "while this is not an unreasonable conclusion in the framework of the binding theory there proposed [i.e., in (Chomsky 1981a) --- M.S.], it is surely one that requires careful consideration". He does not elaborate on the latter point.

(v) A c-command failure for overt anaphors

In terms of the definition of 'bound' adopted by Chomsky (1981a:184), a binder must c-command the element bound by it. In constructions of the form (203a), with each other in the position of OCC, this requirement is apparently violated. Consider, for example, the sentences in (277).

(277) a. the rumors about each other irritated the men (were annoying to the men)

b. stories about each other disturbed the men

In both (277a) and (277b) each other is bound by the object of the verb, namely the men. However, the men does not c-command each other in (277). Chomsky (1981a:229 fn. 64) suggests that such cases may require "a slight modification of binding theory, relaxing the notion of c-command". No further detail is provided. Chomsky also claims that none of the binding theories received in (Chomsky 1981a) can accommodate these cases.

(vi) Arguments in PPs

There are several problems relating to the status of arguments within PPs. Consider firstly sentences like (278), with each other in the NP position of PP.
Chomsky (1981a:225, fn. 37) claims that the order of the two PPs in (278) is free, with a preference for the to-phrase preceding the about-phrase. Only in (278a) can the NP of the first PP be the antecedent of the anaphor. The unacceptability of (278b) follows from the GB binding theory, since the men, which must bind each other, does not c-command it. This leaves unexplained the acceptability of (278a). Chomsky proposes that (278a) might be the result of a rule of reanalysis applying to speak to, so that the men c-commands each other. He then considers various consequences of this proposal, some of which seem plausible, while others are wrong. He concludes that "it is not clear whether this approach is on the right track".

Consider, secondly, sentences with pronouns in the NP position of PPs. Chomsky (1981a:289) discusses an observation by Jean-Yves Pollock, that in Romance languages pronouns in PP can regularly be coindexed with NP antecedents in the same clause, as in (279).

(279) Jean m'a parler de lui
("John spoke to me about himself")

According to the different versions of the GB binding theory, including GB-SUBJECT, the full clause in (279) is the governing category for the pronoun lui. In (279) lui is bound in its governing category, being coindexed with Jean. Principle (B) of (172) is thus violated. Chomsky (1981a:289) considers three possible approaches to the problem raised by sentences such as (279). A first possibility is that PP serves as a governing category. This possibility is rejected on "theory-internal considerations". In terms of the GB-SUBJECT binding theory, governing categories have accessible SUBJECTS. According to Chomsky, there is "little reason" to assume that PP in Romance languages has a SUBJECT while its analogue in English does not have a SUBJECT.

A second possibility is to relate the binding properties of pronouns in PPs to limitations on the distribution of reflexives. There is (near) complementary distribution between reflexives and proximate pro-
nouns. Chomsky rejects this approach as "unlikely" to be correct, without giving any particulars.

The third possibility, which "seems more plausible" to Chomsky (1981a:289), is to relate the appearance of proximate pronouns in sentences such as (279) to the option of cliticization in the Romance languages. It has been observed that clitics in the Romance languages behave in the manner of the English pronouns with regard to disjoint reference. According to Chomsky, this suggests that it is clitics, rather than full pronouns, which fall under principle (B) of (172) in those languages that allow cliticization. He points out that in general full pronouns do not observe principle (B). A possible explanation is that, in languages with the clitic option, full pronouns should be regarded as "somehow emphatic, thus immune to principle (B) of the binding theory".

Chomsky (1981a:290) provides English examples similar to (279).

\begin{align*}
(280) & \quad \text{a. John always keep his wits about him} \quad [5.2 (2)] \\
& \quad \text{(*himself, *Bill)} \\
& \quad \text{b. the melody has a haunting character to it} \\
& \quad \text{(*itsel, *Bill)} \\
& \quad \text{c. John likes to take his work home with him} \\
& \quad \text{(*himself, *Bill)}
\end{align*}

\begin{align*}
(281) & \quad \text{a. John pushed the book away from him} \quad [5.2 (3)] \\
& \quad \text{b. John drew the book towards him} \\
& \quad \text{c. John turned his friends against him} \\
& \quad \text{d. John saw a snake near him}
\end{align*}

In the examples of (280) a proximate pronoun is obligatory, while in (279) it is optional. Chomsky points out that judgments vary as to whether a proximate pronoun or reciprocal should be used in some of the examples of (281), and "obscure factors enter into preference one way or another ...". Thus, (281) contrasts with (282).
(282) John turned the argument against himself (*him
where the reference is to John)

Sentences such as (280) and (281) pose a problem for the GB binding theory similar to the one posed by (279). In all cases the matrix clause is the governing category for the pronoun. By principle (B) the pronoun must be free in its governing category. In each case, however, the pronoun is bound in the matrix clause, thus violating principle (B).

One possible approach to these cases, according to Chomsky (1981a: 290), is to argue that in cases where reflexives are excluded, PP is a governing category. If this is correct, reciprocals should also be excluded. In some cases, the correlation is "reasonably straightforward". Compare, for instance, (281c) and (282) with (283a, b), respectively.

(283) a. *they turned their friends against each other
   b. they turned the arguments against each other

The relevant interpretation of (283a, b) is with each other bound by the subject. Chomsky notes that often in such cases judgments are "unclear". He states that "a proper theory dealing with these matters should explain the choice of elements and also the haziness of the judgments concerning them in many cases".

Chomsky (1981a:290) observes that it is "tempting" to suppose that such examples as (281) should be treated as analogous to (284), with the proximate interpretation of the pronoun.

(284) a. John considers Mary angry at him (*him-
   self, *her, herself)
   b. John strikes Mary as angry at himself (*him,
   her, *herself).

It is assumed that these sentences have the representations (285) with
embedded clauses at the level of LF (and in fact at every level of representation). 124)

(285) a. John considers [Mary angry at him]  [5.2 (7)]
b. John \(_1\) strikes Mary \([t_1 \text{ as angry at himself}]\)

According to Chomsky, one could argue that, correspondingly, (281) has the representation (286) at LF, where John controls PRO, the subject of the predication near him.

(286) John saw a snake \([S \text{ PRO near him}]\)  [5.2 (8)]

Referring to work by Manzini, Chomsky rejects this approach. Manzini's own proposal is that PP, like other categories, has a kind of PRO subject which functions as an agreement element. PP should then be a governing category, allowing pronouns and anaphors accordingly. It follows from this assumption that (287) should be acceptable in English, with the reciprocal bound by the subject.

(287) they saw snakes near each other  [5.2 (11)]

Chomsky notes that, as throughout this category of examples, judgments "tend to be uncertain". He himself does not make a firm suggestion about (281).

5.7 The elimination of structure building rules in the GB-framework

In §4.6 above I outlined a problem for the OB binding theory noted in (Chomsky 1980b). The problem concerned sentences like (288). 125)

(288) a. They regard me as very much like each other (them). \([(38)a, b]\]
b. I impress them as very much like each other (them).

In both cases each other cannot be bound by they, and them can be coin-
dexed with they, exactly as if these sentences contained a subject that invokes opacity. However, Chomsky (1980b:17) observes that there seems to be no syntactic motivation for assigning "anything beyond the obvious surface structure" to such sentences. In particular, there is no subject in the surface structures of these sentences that could invoke Opacity. Chomsky solves this problem by adopting structure building rules that assign to sentences such as (288) representations such as (289) at the level of LF.126)

(289) a. they regard me as \[S PRO be very much like each other (them)] 
\[(39) a, b\]

b. I impress them as \[S PRO be very much like each other (them)]

In (289a) PRO is controlled by me, and in (289b) by I. At the level of LF sentences such as (288) thus do contain a subject that can invoke Opacity. Since it is assumed in (Chomsky 1980b) that the binding conditions apply at the level of LF, the problem posed for the OB binding theory by sentences such as (288) is thus solved.

Chomsky (1981a:29) adopts the projection principle, presented in (290), as a general principle of UC.

(290) "Representations at each syntactic level (i.e., LF, and D- and S-structure) are projected from the lexicon, in that they observe the subcategorization properties of lexical items."

He (1981a:32) points out that a theory which incorporates structure building rules violates the projection principle. It follows that structure building rules are not allowed within the GB framework. The question then arises how the GB theory makes the correct predictions about sentences such as (288). In essence, the answer is that it is assumed that verbs like regard and impress take clausal as-complements. By the projection principle, sentences such as (288) will thus contain an embedded subject at every level of representation.

Chomsky (1981a:109-110) argues that the structure of sentences with
regard is of the form (291a), while the structure of sentences with
impress is of the form (291b). The analysis (291a) follows from the
projection principle, while (291b) follows from the projection principle
and Case theory.

(291)  a. John regards [Bill as foolish] 
       [2.6(25ii)]
   b. John impressed me [t as intelligent]

Given these assumptions, Chomsky (1981a:198) provides the following
examples that correspond to the sentences in (288).

(292)  a. they regard [Bill as too critical of them]  [3.2.3(37ii)]
       (*themselves, each other)]
   b. Bill regards [them as too critical of
       themselves (*them, each other)]  [3.2.3(37iii)]
   c. I impressed them [t as too critical of
       them (*themselves, *each other)]  [3.2.3(37iv)]
   d. they impressed me [t as too critical of
       themselves (*them, each other)]  [3.2.3(37v)]

The GB binding theory yields the correct results in each case. The
embedded clause is the governing category throughout. In (290a, c)
them is free in its governing category, as required by principle (B),
while themselves and each other are free in their governing category
in violation of principle (A). In (290b, d) themselves and each other
are bound in their governing category, as required by principle (A),
while them is bound in its governing category in violation of principle (B).

Note that Chomsky (1981a:196) argues that binding theory applies at
the level of S-structure, rather than at the level of LF. This revi-
sion has an effect on the applicability of principle (C) only, i.e.,
it has no effect on the binding of anaphors and pronominals. Thus,
the GB binding theory will yield the correct results for (290), irre-
respective of whether it applies at S-structure or at LF. The reason is
that the sentences in question will have essentially the same structure at the level of S-structure and the level of LF, given the projection principle.

5.8 The Empty Category Principle

Chomsky (1981a:chapter 4) discusses the RES(NIC) phenomena in considerable detail.128 He argues that these phenomena follow from the Empty Category Principle/ECP. A number of possible formulations of the ECP are considered. Chomsky also discusses various problematic consequences of these different formulations. Since our main concern is with the GB binding theory, I do not provide a detailed exposition of the ECP. I merely present one of the proposed formulations of the ECP, and illustrate how it explains the unacceptability of sentences such as (160) above, i.e., sentences in which wh-phrases are apparently subject to something like the NIC.

The ECP is formulated as follows by Chomsky (1981a:250).

(293)  "ECP: \( \alpha \) must be properly governed." \([4.4. (11)]\)

The notion 'proper government' that features in (293) is defined as follows.

(294)  "\( \alpha \) properly governs \( \beta \) if and only if \( \alpha \) governs \( \beta \) \([\text{and } \alpha \neq \text{AGR.}]."\)\([4.4 (10)]\)

Chomsky (1981a:250) extends the notion of 'government' defined in \([3.2.1 (11)]\), so that coindexed elements are also governors.

(295)  "Consider the structure (i): \( \ldots \gamma \ldots \alpha \ldots \gamma \ldots \) \n
(i) \( \ldots \gamma \ldots \alpha \ldots \gamma \ldots \) \n
(a) \( \alpha = X^o \) or is coindexed with \( \gamma \) \n
(b) where \( \phi \) is a maximal projection, if \( \phi \) dominates \( \gamma \) then \( \phi \) dominates \( \alpha \) \n
(c) \( \alpha \) c-commands \( \gamma \) \n
In this case, \( \alpha \) governs \( \gamma \)."
Consider again sentence (160), presented here as (296).

(296) the men, who I wonder [which stories t told to your brother], were very troubled

Chomsky (1981a:159-160) argues that the unacceptability of (296) partly follows from what appears to be a violation of the NIC. The trace t is free in the subject position of a tensed clause. Since in other cases wh-traces do not obey the NIC — see e.g. (159) above — an apparent contradiction with respect to the NIC arises. However, the unacceptability of (296) can be explained by the ECP, without any reference to the binding theory. who in (296) does not properly govern t, because the maximal projections S and VP intervene. Government by AGR does not count as proper government, by (294). Apart from the wh-island violation involved in (296), the unacceptability of (296) is thus explained by the ECP. The third conceptual problem of the OB binding theory — see §5.2.2.3 above — is thus solved. wh-traces, as variables, are not subject to the SSC and NIC, since they fall under principle (C) of the GB binding theory — see §5.3.2.4.2 above. Cases in which wh-traces apparently obey the NIC fall under an independent principle, the ECP.

Chomsky claims that the conceptual problem caused by the *[that - t]* filter for the OB theory — see §5.2.2.4 above — is solved by the ECP. He (1981a:244f.) considers various ways of deriving the *[that - t]* filter from the ECP. If this could be done, it would be possible to eliminate the filter.

5.9 Chomsky's view of the development of the GB binding theory

Chomsky (1981a:14) bases his approach to linguistic theory on the following "guiding intuition" about the structure of language.

(295) "... that the theory of core grammar, at least, is based on fundamental principles that are natural and simple, and that our task is to discover them, clearing away the debris that faces us when we explore the varied phenomena of lan-
He concedes that this guiding intuition might be mistaken. The approach characterized by this guiding intuition "can be justified only in terms of its success in unearthing a more 'elegant' system of principles that achieves a measure of explanatory success" (Chomsky 1981a:15).

Chomsky claims that the development of the GB binding theory as an alternative to the OB binding theory illustrates the application of the guiding intuition explicated above. He (1981a:338) claims that the improvement in the binding theory was motivated by "intuitive considerations of simplicity" and the "search for unification of principle". That is, the improvement was motivated by the kinds of considerations "that have proven so fruitful in the natural sciences for the past several hundred years ...". The good results obtained in this case (and others mentioned by Chomsky (1981a:339)) suggest to Chomsky that "it makes a good deal of sense to pursue the working hypothesis that the theory of core grammar, for reasons that are not at all obvious, does have some of the properties of the systems studied in the more fundamental natural sciences, and that for some reason neural structures at least in this domain instantiate a perhaps surprisingly simple and unified system of principles".

5.10 Summary of main points

The main points of the description in §§5.2-5.9 of the fourth stage in the development of the SSC and TSC/PIC --- i.e., the development of the GB binding theory --- are as follows.

(i) The OB binding theory, and the larger OB theory of which it forms part, faces a number of conceptual and empirical problems. These problems, in particular the conceptual problems,
motivate the development of an alternative to the OB binding theory, and the overall OB theory —— see §5.2.

(ii) In (Chomsky 1979b), (Chomsky 1981d), and (Chomsky 1981a) the GB binding theory is proposed as an alternative to the OB binding theory, and the GB theory as an alternative to the OB theory —— see §5.3.

(iii) The OB binding conditions —— SSC/Opacity and PIC/NIC —— follow from the GB binding theory in the case of arguments in clauses —— see §5.3.2.4.2.

(iv) The OB binding conditions do not follow from the GB binding theory in the case of arguments in NPs, i.e., in some cases where the OB binding conditions make the correct predictions, the GB binding theory makes the wrong predictions —— see §5.3.2.4.3.

(v) The GB binding theory manages to solve some of the empirical problems of the OB binding theory —— see §5.3.3.

(vi) The GB binding theory, in conjunction with the GB government theory, the GB indexing theory, and the ECP, manages to solve most of the conceptual problems of the OB theory —— see §§5.3.3 and 5.8 (but see also the problems connected with the GB indexing theory noted in §5.3.3).

(vii) There are certain differences between the 1979 and 1981 formulations of the GB binding theory —— see §5.4.2.

(viii) In the works dealing with the GB binding theory that appeared in 1979, the differences between the predictions of the OB and GB binding theories are interpreted in terms of the notion ‘markedness’. In particular, cases about which the GB binding theory apparently makes the wrong predictions, are taken to be marked, and the mechanism responsible for them is said to fall outside core grammar —— see §5.4.3.
(ix) Chomsky (1981) does not rule out the possibility that the cases referred to in (viii) are marked. An attempt is nevertheless made to accommodate them within a revised version of the GB binding theory --- see §5.5.2.

(x) The revised GB binding theory, GB-SUBJECT, overcomes a conceptual problem of the original GB binding theory --- see §5.5.2.

(xi) Apart from the revisions referred to in (ix) and (x) above, additional modifications to the GB binding theory are proposed in (Chomsky 1981a) --- see §5.5.3.

(xii) The GB-SUBJECT binding theory still faces some empirical problems. For some of these, tentative solutions are discussed in (Chomsky 1981a) --- see §5.6.

(xiii) Cases for which structure-building rules are required within the GB theory, are analyzed differently within the GB framework --- see §5.7.

(xiv) Chomsky sees the development of the GB binding theory as an illustration, and in fact, a justification of the guiding intuition that the theory of core grammar is based on principles that are natural and simple --- see §5.9.
FOOTNOTES

1. Cf. §2.2.1 below for a more detailed exposition of the content of the SSC and TSC.


5. In cases where a work has been published more than once, the reference is to the first year of publication, except where otherwise noted.

6. Chomsky (1973:246) defines the notion 'superior' as follows: a category A is superior to the category B in the phrase marker if every major category dominating A dominates B as well, but not conversely. Chomsky takes N, V, A, and the categories that dominate them, to be the major categories.

7. I adopt the following convention for the use of brackets around numbers: numbers in square brackets represent numbers in the work under discussion. Numbers in this work are always in round brackets.

8. Chomsky (1973:230, note 17) leaves open the possibility that the rule relating NP - each other - in sentences such as (7)-(11) is an interpretive rule, rather than a syntactic transformation.

discussion were formally proposed, a language is defined as a set of structural descriptions of sentences. A grammar is a system of rules that generates this language. Cf. Chomsky 1973:232.

In terms of these definitions, knowledge of grammar is thus equivalent to knowledge of language. Recently --- cf. e.g., Chomsky 1980a:90 --- Chomsky explicitly draws a distinction between knowledge of grammar and knowledge of language. Knowledge of grammar now constitutes only a subcomponent of knowledge of language. Knowledge of language is thus no longer equivalent to knowledge of grammar. Cf. also Chomsky 1981a:4 for the derivative status of the concept 'language'. The basic ideas outlined in §2 are not affected by this change.


11. Cf. Chomsky 1977c:125 for a simple explanation of why the essential goal is to restrict the class of available grammars, rather than the class of possible grammars.

12. The condition of descriptive adequacy requires that UC make available a descriptively adequate grammar for each natural language. A descriptively adequate grammar correctly describes the intrinsic competence of the idealized native speaker. Cf. Chomsky 1965:24. The condition of explanatory adequacy requires that UC succeed in selecting a descriptively adequate grammar on the basis of primary linguistic data. Cf. Chomsky 1965:25. The condition of explanatory adequacy is clearly related to the attempt to solve the problem of language acquisition.

13. Cf. e.g., Chomsky and Lasnik 1977:427 for a similar characterization of the nature of the conflict between explanatory and descriptive adequacy.

14. The A-over-A condition and the principle of recoverability of deletion are early examples of such conditions. Cf. Chomsky 1964:
41, 44 for early discussions of these conditions.


16. Chomsky (1976a:315 makes it clear that the universality of conditions, while it constitutes the ideal case, is not necessary for the reduction of the expressive power of transformations. The notion of parametric variation makes it possible to reduce expressive power even with non-universal conditions. The notion of parametric variation will be discussed in detail in §§2.2.5 and 3.4 below.

17. In the case of some of the examples discussed below, I indicate more structure than Chomsky does.

18. Cf. Chomsky 1976a:310 for a list of various other enrichments of transformational theory proposed in the literature.

19. Apparently the rejection of a solution in terms of a rule-specific condition is based on the assumption that the incorporation of general conditions such as TSC in UG will eventually make it possible to eliminate all rule-specific conditions. The issue is of course more complicated than this. As Chomsky (1976a:315) points out, rule-specific conditions may be regarded as parameters to be fixed for particular rules in language learning. The implications of the possibility of parametric variation are discussed in §§2.2.5 and 3.4 below.

20. With reference to another possible enrichment of the theory of transformations, Chomsky (1973:255, fn. 34) declares that "in the absence of other considerations, the general point that the theory of transformations should not be extended to permit this option is compelling, if not decisive".
21. The sentences numbered [14] in (Chomsky 1973) are the following:

(i) Advantage was taken of Bill.
(ii) Bill was taken advantage of.

22. In fn. 24 Chomsky suggests that there is independent justification for the hypothesis that an element in COMP must move to another COMP position. In particular, this condition must block the improper passivization of (i) to give (ii).

(i) John asked what to read.
(ii) *What was asked by John to read

23. Chomsky (1973:263) observes that provision (43)/[161] cannot be assigned to case [160a] of the SSC --- i.e., where Z is not controlled at all. If it is added, then the SSC can no longer block the derivation of (ii) from (i), with Z = it.

(i) It is pleasant for the rich [S COMP poor immigrants to do the hard work] [165a]
(ii) *The hard work is pleasant for the rich for poor immigrants to do [165b]


27. Informally, the Coordinate Structure Constraint stipulated that no conjunct in a coordinate structure, or any element in a conjunct may be moved. In a generalized form, it prohibits any relation between an element that is part of a coordinate structure and an element outside that structure.
28. Although Chomsky does not mention it, Coreference Assignment also violates the SSC, as in the following example:

   John thought that Bill liked him.

   The rule thus also constitutes a problem for the SSC.

29. Chomsky (1976a:319) explains that the rule of Reciprocal Interpretation assigns an appropriate sense to sentences of the form NP ... each other.


31. In fn. 27 Koster refers to (Williams 1977), in which some properties of discourse grammar are discussed.

32. Cf. e.g., Chomsky 1980a:11, 218 f., and Chomsky 1979a:57 f. on the importance of idealization for Chomsky.

33. The Complex NP Constraint, the Coordinate Structure Constraint and the Sentential Subject Constraint were all proposed by Ross (1967). For an informal statement of the Coordinate Structure Constraint, cf. fn. 27 above. The Complex NP Constraint stipulates that no element may be extracted from a sentence dominated by a noun phrase with a lexical head noun. The Sentential Subject Condition prohibits the movement of any element from the sentential subject of a sentence.

34. As Chomsky (1976a:322) points out, this would not be an "untolerable" consequence. Given the option of regarding rule-specific conditions as parameters to be fixed for rules during language-learning, it might still be possible to restrict the expressive power of transformations. Cf. in this connection the discussion on p. 315 of (Chomsky 1976a).

36. In (82) and (83) $X^n$ stands for the categories NP, VP, AP, and $t$ is the trace of the quantifier.

37. Cf. §2.3.2 above for a discussion of the latter point.

38. Although, in the remarks referred to above, Chomsky only mentions the possible reinterpretation of the SSC, he clearly has in mind the possible reinterpretation of the PIC/TSC as well. On p. 317, where he also refers to the possible reinterpretation, he clearly includes the latter condition.

39. Chomsky (1977c:76) stipulates that two terms in the structural description of a transformation are adjacent only if each is constant, and if any term that intervenes between them is a variable.

40. Cf. also the discussion on p. 14 f. above, on the conflict between descriptive adequacy and explanatory adequacy.

41. Cf. Chomsky 1981a:3, 7 f. for a recent discussion of the exact nature of core grammar, and of its role in overcoming this dilemma.

42. The formulation in (Chomsky 1976a) is similar to that of (Chomsky 1977c) in the relevant respect. Cf. in this connection the discussion in §2.3.5 above. Cf. also Chomsky 1977c:fn. 17 for an explication of the different implications of the 1973 and 1977 formulations of the SSC for COMP-COMP Movement.

43. Cf. fn. 33 above for an informal statement of the Complex Noun Phrase Constraint/CNPC. The wh-island constraint stipulates that no element can be moved out of a clause introduced by a wh-complementizer.

44. Two footnotes, (21) and (22), are omitted from the quotation presented in (121).
45. Chomsky (1980b:10) mentions another condition that restricts the binding of anaphors. This is the Command Condition, which stipulates that an antecedent must c-command its anaphor. The c-command requirement on anaphors is built into the definition of the notion 'bound' adopted by Chomsky. Cf. the discussion immediately below. Consequently, no separate Command Condition is required in Chomsky's theory. That Chomsky sees the binding theory as consisting only of the NIC and the Opacity Condition, is clear from the summary he provides on p. 38, where only these two conditions are mentioned as binding conditions.

46. Cf. in this connection Chomsky 1980b:10, 15, 39. Pronouns are partly like lexical NPs, and partly like anaphors. Cf. in this connection the discussion in the Appendix to (Chomsky 1980b), and in §4.5 below. Except where otherwise noted, the definitions presented below are from p. 10 of (Chomsky 1980b).

47. The sentences in (126) were discussed above, where they were numbered (7a), (13b), and (12b) respectively.

48. Cf. Chomsky 1980b:fn. 15 for a discussion of these cases.

49. Cf. the discussion in §3.3 above.

50. Cf. in this connection the definition of 'involve' presented in Chomsky 1976a:316, fn. 22), and discussed in §2.3.3 above.

51. Cf. §3.4 above for a discussion of the possibility of parametric variation as regards the notion 'subject' that features in the SSC. From Chomsky's reference to English, it is quite clear that the notion 'subject' which features in the Opacity Condition can be subject to parametric variation as well.

52. Chomsky (1980b:25) has the following to say about the notion 'government':
"The notion 'government' will no doubt be related to grammatical relations. In a configurational language such as English, we can specify it in terms of c-command, perhaps as follows:

\[(69) \alpha \text{ is governed by } \beta \text{ if } \alpha \text{ is c-commanded by } \beta \text{ and no major category or major category boundary appears between } \alpha \text{ and } \beta. \]

In fn. 29 Chomsky (i) explains that (69) builds in the "adjacency and c-command condition of the * [NP to VP] filter", (ii) explains that structures such as \( \beta [\gamma \alpha] \) and \( \beta \gamma \alpha \) are excluded, where \( \gamma \) is a major category, (iii) points out that the notion 'government' must be defined at a level of abstraction that excludes from consideration parenthetical elements, interpolated adverbs, etc.


54. Cf. Chomsky 1980b:Appendix for more detail on these indexing conventions.

55. Some of these problems are also briefly and informally mentioned in (Chomsky 1981b). Since this work provides only an informal (and incomplete) account of the ideas contained in the other works, no further reference is made to it.

56. Recall that a similar redundancy between the OB binding theory and the * [NP to VP] filter has led to the elimination of this filter --- see the discussion in §4.7 above.

57. Cf. in this connection the remarks in (Chomsky 1977c:111) on the "naturalness" of the SSC and PIC.

58. These are the violations resulting from taking only \( \overline{S} \), and not \( S \), as a bounding node for Subjacency.

59. Chomsky (1981a) uses the term "SSC" for the pre-1978 formulations of the relevant condition, as well as for its reformulated ver-
sion, the Opacity Condition. I will use the notation "SSC/Opacity Condition" to refer to the relevant condition.

60. Examples are separately numbered in each subsection of (Chomsky 1981a). I will refer to these numbers by indicating the relevant subsection and number in square brackets. Thus, \[3.1 (8)\] refers to number (8) of subsection 3.1 of (Chomsky 1981a).


63. Chomsky (1981a:232) claims that the relevant contradiction is only a near contradiction, since there are certain assumptions available in the OB theory that would make it possible to avoid the problem.


65. Chomsky (1981a:161) mentions a further fact relating to the *that - t] filter that must be explained. This filter does not apply in languages that allow missing subjects, i.e., languages that have the "pro-drop parameter". Apparent violation of the *that - t] filter is one of a clustering of properties related to pro-drop in these languages. Cf. Chomsky 1981a:240. Chomsky (1981a:161) says that "we want to explain this clustering, if possible, in terms of a single parameter, which should be related to RES(NIC)."


67. This requirement follows from the θ-criterion for D-(=deep)structures. Cf. Chomsky 1981a:42f. for discussion.
68. Cf. Chomsky 1981a:§§2.2, 2.6 for an exposition of θ-theory.


70. Chomsky (1981a:47) stipulates that the grammatical functions/GFs determined in A-positions are called "A-GFs". GFs determined in A-positions are called "A-AGFs". Cf. Chomsky 1981a:42 for further discussion of the notion 'grammatical function'.

Chomsky (1981a:185) observes that the definition (176) of the notion 'variable' is inadequate in cases such as [3.2.3 (7)].

(i) the man [to whom I gave the book t]
(ii) the man [whose picture I saw t]
(iii) John, [a picture of whom I saw t yesterday]

In these cases the trace t is not the variable bound by the phrase in COMP. Chomsky leaves this problem open.

71. Cf. also Chomsky 1981a:§4.4 for a further modification.

72. Cf. Chomsky 1981a:52 for further discussion of the properties of INFL and AGR. Cf. also Chomsky 1981a:164 for the possibility of regarding INFL itself as the governor.


75. If the principle (i) is adopted, then the Case filter can be reformulated as the Extended Case Filter (ii).

(i) \([_{NP} e]\) is a variable if and only if it has Case \([3.2.2 (16)]\)
(ii) \([_{NP} \alpha c]\) if \(\alpha c\) has no Case and \(\alpha c\) contains a phonetic matrix or is a variable.

Cf. Chomsky 1981a:175 for discussion of this point.
76. Cf. for example the discussion of \[3.2.3 (9)\] by Chomsky (1981a: 186).

77. Chomsky (1981a:§S.1) presents evidence that this simple indexing theory is in fact inadequate. Cf. the discussion in §5.3.3 below.

78. In some of the examples discussed in §§5.3.2.4.2 and 5.3.3.4.3 I indicate more structure than Chomsky does. This is done in order to make certain points clearer. Since the "additional" structure indicated by me does not conflict with the structure assumed by Chomsky, I do not comment on it in individual cases.

79. Cf. Chomsky 1981a:154 for this example, and also for (191), (192b), (193b).

80. Cf. e.g., Chomsky 1981a:66f. for a discussion of exceptional Case-marking in such instances.

81. If the principle \[3.2.2 (16)\] holds (cf. §5.3.2.2.3 for a formulation of this principle), the Case-less status of NP-trace would follow. Cf. also Chomsky 1981a:334, 345, fn.5 for further discussion of this issue.

82. Cf. the discussion of the ECP in chapter 4 of (Chomsky 1981a) in this connection.

83. NP-trace is excluded from position \(\alpha_3\) by conditions on preposition stranding. Cf. Chomsky 1981a:292f. for a discussion of the latter phenomenon. NP-trace is excluded from position \(\alpha_4\) because it is an ungoverned position. Cf. Chomsky 1981a chapter 4 for a discussion of the ECP, from which it follows that NP-trace must be properly governed.

84. Not all the examples presented below are discussed in §3.2.3 of (Chomsky 1981a). In cases where there are "gaps" in the data discussed in the latter section, I have tried to fill them with examples discussed in other subsections of (Chomsky 1981a).
85. Note that in (194) and (195a) the governor of NP-trace is a Case-assigner, which, strictly speaking, is not allowed.

86. Cf. Chomsky 1981a:330 for formal definitions from which it follows that PRO is both a pronominal and an anaphor.

87. Cf. Chomsky 1981a:§2.4 for a discussion of the ungoverned status of PRO.

88. This sentence is the same as [2.4.2 (7iii)], but with PRO in place of you. Cf. also Chomsky 1979b:22 on the impossibility of PRO after for.

89. This sentence is the same as [2.4.2 (7ii)], with PRO in place of him.

90. The requirement of government for trace follows from the ECP. Cf. the reference in fn. 81 above.

91. Cf. Chomsky 1981a:156 for these conditions.

92. Cf. Chomsky 1981d:140ff. for more examples. These are discussed in §5.4.3 below.

93. Cf. Chomsky 1981a:154 for this example. There is clearly an error in [3.2.3 (57ii)]. As printed, each other is bound by their in NP*, while the intention clearly is that each other must be free in NP*.

94. Definition [3.2.1 (6)] reads as follows:

"[β ... γ ... α ... γ ...], where
(i) \( α = x^0 \)
(ii) where \( φ \) is a maximal projection, \( φ \) dominates \( α \) if and only if \( φ \) dominates \( γ \)."

\( α \) then governs \( γ \).
95. Definition [3.2.1 (4)] reads as follows:

"[ β ... γ ... α ... γ ... ], where

(i) α = x^0

(ii) where φ is a maximal projection, if φ dominates γ then φ dominates α

(iii) α is an immediate constituent of β."

α then governs γ.

96. Cf. Chomsky 1981a:165 for a discussion of these possibilities.


98. Cf. Chomsky 1981a:210 for an explanation of why the reference in (218) is to "subject-of-AGR" rather than "subject-of-Tense".

99. Chomsky (1981a:314, fn. 1) thanks Howard Lasnik for the observations discussed in his §5.1.

100. The sentences (i) and (2) referred to in (221) are the following:

(i) *we lost my way

(ii) John_told Bill_j that he_k should leave

Cf. Chomsky 1981a:285 for an explication of the problems raised by these sentences for the indexing theory.


103. Cf. e.g. Chomsky 1981d:133 for an explication of this point.

104. Cf. Chomsky 1981a:275, 334 for a discussion of these points.
105. Cf. Chomsky 1979b:26 for a brief discussion of similar cases with pronouns, e.g.

   John saw [a picture of him]

106. Cf. Chomsky 1973:239, fn. 19 for a discussion of such examples. Cf. also §2.2.4.3 above.

107. Chomsky also refers to (232)-(240) as "structures". Cf. e.g., Chomsky 1981d.


110. Chomsky (1981a:25) adopts the base rule 

    \[2.1 (25)] --- S \rightarrow NP INFL VP --- from which it follows that S will always contain a subject.

111. Cf. the references cited in fn. 109 above.

112. Cf. e.g., Chomsky 1981a:52 for a discussion of the latter point. Note also that AGR is not a binder with respect to the binding theory. Cf. Chomsky 1981a:211-212.

113. Chomsky (1981a:229, fn. 63) discusses the possibility that (247), as it stands, is too strong, and that an "unless"-condition should be added to it.

114. Cf. e.g., Chomsky 1973:239, fn. 19, and Koster 1978b:166f. Cf. also the discussion in §2.2.4.3 above.

115. Chomsky (1981a:230, fn. 67) says that the discussion of these examples is based on suggestions by Tim Stowell and Dominique Sportiche.

117. Emonds (1976:2) defines 'root sentence' as follows:

"Root Sentence: A root S ('sentence') is an S that is not dominated by a node other than S"

118. Cf. Chomsky 1980b:15 for a discussion of the problematic status of reflexives, and of the possibility of parametric variation with regard to what counts as anaphors for the binding theory. Cf. also the discussion in §4.2 above.


120. Cf. the discussion in §2.3.3 above on the applicability of the SSC to Quantifier Movement.

121. Cf. Chomsky 1981a:166 for an explanation of why there is no c-command in this case.

122. The cases discussed directly above in fact also involve arguments within PPs.


125. These examples are presented as (145a, b) in §4.6 above.

126. These structures are presented as (147a, b) in §4.6 above.


128. Cf. also the discussion in (Chomsky 1979b:33f.).
129. Chomsky (1981a:250) explains that pro-drop languages exclude the condition "\( \Box \neq AGR \)" from their notion of proper government.

130. The unacceptability of (296) is partly the result of a wh-island violation. Cf. the discussion in §5.2.2.3 above.

131. Cf. e.g., Chomsky 1980a:8f. for a more detailed discussion of the ideas mentioned in (295).
REFERENCES

Anderson, Stephen and Paul Kiparsky (eds.)

Botha, Rudolf P.

Belletti, A., L. Brandi, G. Nencioni, L. Rizzi (eds.)

Chomsky, Noam


1979b "Lectures on core grammar". Delivered at the workshop of the GLOW meeting, Pisa, April 1979.


1981d "Markedness and core grammar". In Belletti et al. (eds.) 1981:123-146.

Chomsky, Noam and Howard Lasnik
1977 "Filters and control". Linguistic Inquiry 8:425-504.

Culicover, P., A. Akmajian and T. Wasow (eds.)

Emonds, J.
Fiengo, R. and H. Lasnik  
1976  

Freidin, R. and H. Lasnik  
1981  

Harnad, Stevan R., Horst D. Steklis and Jane Lancaster (eds.)  
1976  

Kayne, R.  
1975  

Koster, J.  
1978a  

1978b  

Lasnik, H.  
1976  

Lightfoot, David  
1976  
"Trace theory and twice moved NP's". *Linguistic Inquiry* 7: 559-582.

Newmeyer, Frederick J.  
1980  

Postal, P.  
1976  
Ross, J.R.

Williams, E.
1977  "Discourse and logical form". Linguistic Inquiry 8:101-139.