Chapter 2
Toward a Theory of Conditions on Transformations

2.1 The A-over-A Principle

Chomsky's *Current Issues in Linguistic Theory* (1964) marks the first important step toward establishing a general theory of conditions on transformations. Perhaps the most appealing constraint proposed in that book is the *A-over-A Principle*. Consider (1a–c):

(1)
a. Mary saw the boy walking toward the railroad station
b. This is the railroad station which Mary saw the boy walking toward □
c. Which railroad station did Mary see the boy walking toward □

"□" indicates the position from which the *wh*-phrase was moved, by Relative Clause Formation in (1b) and by Question Formation in (1c). In the remainder of this chapter we will refrain, however, from indicating this "gap."

(1a) is ambiguous between the readings (2a) and (2b):

(2)
a. Mary saw $[_S$ the boy walk toward $[_{NP}$ the railroad station$]_{NP}]_S$
b. Mary saw $[_{NP}$ the boy $[_S$ who was walking toward $[_{NP}$ the railroad station$]_{NP}]_S]_{NP}$

Both (1b) and (1c) are unambiguous, however: they can only be understood in the sense of (2a). The A-over-A Principle accounts for this fact by requiring that a transformation that could apply to two NPs, one of which is embedded in the other and both of which match its structural description, can only apply to the higher (more inclusive) one. More formally:

(3)
A-over-A Principle
In a structure ... $[_A$ ... $[_A$ ...$]_A$ ...$]_A$ ..., if a structural description refers to A ambiguously, then that structural description can only analyze the higher, more inclusive, node A.

This principle will correctly prevent Relative Clause Formation and Question Formation from applying to (1a) when it has the structure of (2b).

In *Current Issues* this proposal was simply a more or less plausible and elegant-looking solution to the descriptive problem posed by (1a–c), and it was hidden in a footnote. Furthermore, in the same footnote, Chomsky pointed out several potential counterexamples:

(4)
a. Who would you approve of my seeing
b. What are you uncertain about giving to John
c. What would you be surprised by his reading

But what started out as an interesting attempt at solving the descriptive problem associated with processes of *wh*-movement (processes that move *wh*-phrases—phrases consisting of or containing *who, what, which,* etc.) in examples like (1a–c) was to become the germ for the construction of a general theory of constraints on transformations, which has constituted the main research program throughout the 1970s. Though the idea still remained unexplored in *Aspects,* it was taken up again, more or less simultaneously, by Chomsky in *Language and Mind* (Chomsky (1968)) and by John R. Ross in "Constraints on Variables in Syntax" (Ross (1967)). Chomsky stresses the fact that, despite many empirical problems, the A-over-A Principle has "a certain explanatory force." Ross, on the other hand, concentrates on the empirical problems raised by the principle and tries to formulate more adequate alternatives. This chapter presents the most influential of Ross's proposals, their shortcomings, and their significance.

2.2 A-over-A Phenomena

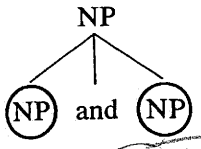
Among the phenomena that are potentially relevant to the A-over-A Principle are the following.

(i) An NP that is a conjunct of a coordinate structure cannot be questioned or relativized:

(5)
a. We have to read some books and some papers
b. *Which books do we have to read and some papers
c. *Those are the papers that we have to read some books and

Since the coordinate structure *some books and some papers* in (5a) has the form (6), the A-over-A Principle prevents transformations from applying to the circled NPs and thereby accounts for the ungrammaticality of (5b) and (5c):

(6)



(ii) An NP that is part of a subject NP cannot be questioned or relativized:

(7)
a. For John to win a medal is easy
b. *What is for John to win easy
c. *It is the gold medal that for John to win is easy

(8)
a. My letter to a friend in Italy got lost
b. *Who did my letter to get lost
c. *Gianni is the friend who my letter to got lost

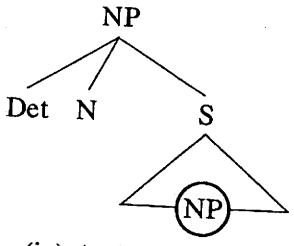
Here the A-over-A Principle straightforwardly prevents transformations from applying to the NPs *a medal* and *a friend,* contained in the subject NPs *for John to win a medal* and *my letter to a friend.*

(iii) An NP that is contained in the sentential complement to a noun cannot be questioned or relativized:

(9)
a. He refuted the proof that you can't square an ellipse
b. *What did he refute the proof that you can't square
c. *The figure that he refuted the proof that you can't square looks a bit like an egg

Again, the grammaticality judgments on these constructions follow directly from the A-over-A Principle, since (9a) is of the following form:

(10)



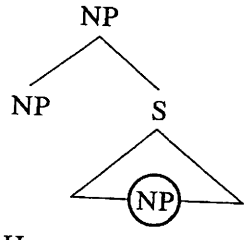
(iv) An NP that is part of a relative clause cannot be questioned or relativized:

(11)

a. Bill found a principle which solves the problem
b. *Which problem did Bill find a principle which solves
c. *The problem that Bill found a principle which solves was very recalcitrant

These cases also have the typical A-over-A structure:

(12)



However, Chomsky (1964) notes that there may be another explanation for the facts in (11). Observe that Relative Clause Formation has applied twice to elements of the most deeply embedded relative clause: the subject and the object. Suppose now that this is ruled out as a matter of principle. This principle will then account for the ungrammaticality of (11c). Furthermore, if the principle is extended to exclude more than one application of either Relative Clause Formation or Question Formation, (11b) is also accounted for. This analysis also explains the following facts:

(13)

a. Who saw what
b. *Who what saw (*What who saw)

This restriction also accounts for a fifth set of facts, which does not fall under the A-over-A Principle in any obvious way.

(v) The *Wh-Island Constraint:* An NP that is part of an indirect question cannot be questioned or relativized:

(14)

a. John wondered who would win a gold medal
b. *What did John wonder who would win
c. *The medal that John wondered who would win was the gold medal

These facts, to which we will return in chapter 4, have come to be known as *wh-island phenomena*.

Since Relative Clause Formation and Question Formation consistently pattern together with respect to these phenomena, they have been identified under the heading *Wh-Movement*. We will therefore often exemplify arguments with instances of one or the other, but not both.

In the next sections we will discuss first some problems with the A-over-A explanation of the facts in (i) through (iv) and then Ross's alternative accounts for them: the Complex NP Constraint for (iii) and (iv), the Sentential Subject Constraint for (ii), and the Coordinate Structure Constraint for (i).

2.3 The Inadequacy of the A-over-A Principle

The A-over-A Principle is both too strong and too weak; that is, it excludes certain grammatical sentences and fails to exclude certain ungrammatical ones. We have already seen examples of the former problem in (4). The latter, however, appears to be much more serious. In fact, it appears that most cases where a phrase of type B is extracted from a phrase of type A are ungrammatical too, and consequently an A-over-B Principle would seem to have at least as much chance to be correct as the intuitively more appealing and plausible A-over-A Principle. To illustrate:

(i) An adjective phrase cannot be extracted from an NP:

(15)

a. You have a very intelligent sister
b. How intelligent a sister do you have
c. *How intelligent do you have a sister

(ii) A PP cannot be extracted from the contexts given in section 2.2:

(16)
a. Bill rejected the suggestion that he should talk to someone
b. *To whom did Bill reject the suggestion that he should talk

(17)
a. He has [an ID which is valid until tomorrow]
b. *Until when does he have an ID which is valid

(iii) In many languages, such as French and German, an NP cannot be extracted from a PP:

(18)
a. Ils ont voté pour de Gaulle
they have voted for de Gaulle
'They voted for de Gaulle'
b. *Qui ont-ils voté pour
who have they voted for

(19)
a. Sie sollte mit ihrem Bruder spielen
she should with her brother play
'She should play with her brother'
b. *Wem sollte sie mit spielen
who should she with play

There are other problems as well, but it is clear even from these examples that the A-over-A Principle is far from adequate. An alternative conclusion is that the constructions listed in (i) through (v) in section 2.2 will not permit the extraction of any kind of phrase. This is what Ross realized. In Ross (1967) he studied large portions of English syntax in this light to discover the properties of these constructions, giving them the name *islands*.

Ross proceeded to impose a number of constraints on transformations in order to account for such island phenomena. Viewing these as constraints on the variables in the structural descriptions of transformations, he assigned them the following general format:

(20)
No rule can involve X and Y (move Y to X, for example) in the structure

$\ldots X \ldots W_1 Y W_2 \ldots$

if ...

What follows *if* is the characterization of the island structures.

2.4 The Complex NP Constraint

The main idea behind Ross's approach to constraints on transformations is that the class of moved items to which they apply must be strongly generalized, in view of the observations in section 2.3, but that some generality with regard to characterization of the island contexts must be sacrificed. In other words, there is no one principle that will cover all the extraction contexts in section 2.2; instead, it is necessary to formulate several constraints, each of which captures only part of the phenomena in section 2.2. However, each one will apply to a wider class of phrases, as required by the data in section 2.3. The first of these, the Complex NP Constraint (CNPC), covers the phenomena of (iii) and (iv), which obviously look quite similar; however, it applies not only to NPs but to any phrase:

(21)
Complex NP Constraint (CNPC)
No element contained in an S dominated by an NP with a lexical head noun may be moved out of that NP by a transformation.
(Ross's (4.20))

The CNPC is not limited to *Wh*-Movement of NPs, as in (9) or (11), but applies to PPs as well, as in (16) and (17). APs also fall under the restriction:

(22)
*How intelligent are you looking for an au pair who is

Other movement processes seem to be subject to the CNPC as well. Consider the processes of VP-preposing and *though*-inversion.

(23)

a. They all argued that Bill would leave the country, and leave the country I must conclude that he did t

b. *They all argued that Bill would leave the country, and leave the country I resent the conclusion that he did t

(24)

a. Intelligent though I agree that your au pair is, I still don't like him

b. *Intelligent though you have an au pair who is, I still would have preferred a girl

2.5 The Sentential Subject Constraint

Ross limits his account of the phenomena in (ii) of section 2.2 to sentential examples of type (7) (for the generalization to all subject NPs, see chapter 4). Again, however, the constraint he proposes applies to any element to be moved out of the island configuration:

(25)

Sentential Subject Constraint (SSC)

No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S. (Ross's (4.254))

The SSC applies not only to *for*-clauses, as in (7), but also to *that*-clauses:

(26)

*Who did that Mary was going out with bother you

And again, the constraint applies to other processes as well:

(27)

*They said that John would win the race, and win the race that he did t was evident to everyone

(28)

*Intelligent though for an au pair to be is unlikely, one shouldn't be prejudiced

It is clear that these facts depend on more general questions regarding the distribution of sentential substructures, which may be governed via base-generation or some process of extraposition (see chapter 3 for discussion). In any event, though we will not pursue the matter here, it

appears that *that*-clauses, and also to some extent *for*-clauses, cannot easily occur in sentence-internal positions, an observation often referred to as the *Internal S Constraint*. This constraint would also account for most of the facts handled by the SSC.

2.6 The Coordinate Structure Constraint

The phenomena in (i) of section 2.2 are also accounted for in Ross's framework by a separate constraint that defines conjuncts of coordinate structures as islands. The main generalization achieved here with respect to the A-over-A Principle is that the new constraint also applies to the extraction of parts of a conjunct, which the A-over-A Principle would not cover in any obvious way:

(29)

Coordinate Structure Constraint (CSC)

In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct. (Ross's (4.84))

The CSC easily accounts not only for (5) but also for sentences like (30a–b):

(30)

a. Bill is $[_{AP}$ proud of $[_{NP}$ his father$]_{NP}]_{AP}$ and $[_{AP}$ tired of $[_{NP}$ his mother$]_{NP}]_{AP}$

b. *Who is Bill $[_{AP}$ proud of $[_{NP}$ ~~his father~~$]_{NP}]_{AP}$ and $[_{AP}$ tired of$]_{AP}$

It also improves on the A-over-A Principle by preventing (31b) from being generated from (31a):

(31)

a. Mary wondered $[_S$ (whether) John would come$]_S$ and $[_S$ who would bring what$]_S$

b. *Mary wondered $[_S$ what John would come$]_S$ and $[_S$ who would bring$]_S$

In (31b) the *wh*-phrase has been moved from the second conjunct to the initial position of the first.

A final observation concerning coordinate structures and the CSC is that *Wh*-Movement can extract parallel *wh*-phrases out of conjuncts when *all* conjuncts of a coordinate structure are affected this way, as in

(32). This property of *Wh*-Movement (and certain other transformations, including Subject-Aux Inversion) is called *across-the-board rule application*.

(32)
I wonder which books Mary hates and Sam likes

These, then, are the beginnings of the theory of constraints. All the principles discussed here have since been modified, generalized, or replaced. The fate of the CSC has been somewhat different, however, because it has not interacted with the other constraints under these revisions. Instead, there have been essentially very few proposals to derive the CSC from specific formalizations of the way transformations apply to syntactic structures that are primarily designed to capture the across-the-board property.

2.7 Pied Piping, the Left Branch Condition, and Upward Boundedness

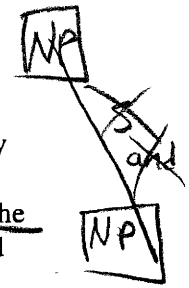
Ross formulated constraints to account for two more important phenomena: cases involving (i) pied piping and the Left Branch Condition and (ii) the upward boundedness of rightward movement rules. Again, these proposals deserve to be mentioned not so much because they offer real explanations for the phenomena in question as because they define standard problems that continue to be highly relevant for present-day research.

The Pied Piping Convention emerged from the observation that the A-over-A Principle is too strong in certain cases. Consider, for example, (33a–b):

(33)
a. This is the book $[_{NP}$ which$]_{NP}$ I have proofread $[_{NP}$ the preface $[_{NP}$ of$]_{NP}]_{NP}$
b. This is the book $[_{NP}$ the preface $[_{NP}$ of $[_{NP}$ which$]_{NP}]_{NP}]_{NP}$ I have proofread

(Note that Ross regards PPs as NPs, hence as structures of the form $[_{NP}$ P $[_{NP}$ *wh*-word$]_{NP}]_{NP}$.) Here the *wh*-word, which is an NP, is contained within a larger NP, a paradigmatic A-over-A situation. Nevertheless, *Wh*-Movement can move the smaller NP as well as the larger one. Ross suggests that in such a situation the more inclusive structure may optionally be carried along in the movement of the *wh*-word. Following a suggestion of Robin Lakoff, he calls this phenomenon *pied piping* (". . . just as the children of Hamlin followed the Pied Piper out of town, so the constituents of larger noun phrases follow the specified noun phrase when it is reordered" (p. 263)). He formulates the Pied Piping Convention as follows:

(34)
Pied Piping Convention
Any transformation that is stated in such a way as to move some specified node NP, where this node is preceded and followed by variables in the structural description of the rule, may optionally apply to this NP or to any noncoordinate NP that dominates it, as long as there are no occurrences of any coordinate node, or of the node S, on the branch connecting the higher node and the specified node. (Paraphrasing Ross's (4.180))



(Adopting the more standard view that PP is distinct from NP would require modifying this formulation.)

The main feature of this principle is its optionality. Consequently, auxiliary principles must be formulated to account for cases where pied piping is either prohibited or obligatory. Preposition stranding (as in (33a)) provides instances of both. On the one hand, when *Wh*-Movement applies to an NP that is the object of a certain type of preposition in Danish, then pied piping cannot occur; in other words, preposition stranding is obligatory. On the other hand, in German, French, Russian, and other languages preposition stranding is impossible. For these languages, then, pied piping of prepositions is obligatory. Finally, in English, where pied piping of prepositions appears to be essentially optional, several additional constraints are nevertheless involved. The only one of these that reaches a fair generality is the Left Branch Condition:

(35)
Left Branch Condition (LBC)
No NP that is the leftmost constituent of a larger NP can be moved out of this NP. (Paraphrasing Ross's (4.181))

This principle will account for facts such as the following:

(36)
a. You saw the president's wife's guard
b. *Whose did you see wife's guard
c. *Whose wife's did you see guard
d. Whose wife's guard did you see

In other words, the LBC imposes obligatory pied piping in such structures. Although some interesting ideas have been raised in the literature in this regard, the LBC remains largely a mystery for which no truly satisfactory explanation has been found.

Ross also formulated the Upward Boundedness Constraint to account for the type of paradigm shown in (38):

(37)
Upward Boundedness Constraint
No element that is moved rightward by a transformation may be moved out of the next higher node S. (Paraphrasing Ross's (5.58))

(38)
a. [$_S$[$_S$ That a serious discussion *of this topic* could arise here]$_S$ was quite unexpected]$_S$
b. [$_S$[$_S$ That a serious discussion could arise here *of this topic*]$_S$ was quite unexpected]$_S$
c. *[$_S$[$_S$ That a serious discussion could arise here]$_S$ was quite unexpected *of this topic*]$_S$

The extraposition of the PP yields an ungrammatical result when it goes out of the first S-node up.

More recent proposals have been made to deal with examples like (38c), but they remain inconclusive in view of the possibility of successive cyclic movement (see section 4.1). Thus, the upward boundedness issue is currently one of the standard unsolved problems.

2.8 The Domain of Transformations

The primary impact of Ross's proposals concerning island constraints was on rules like *Wh*-Movement. Such rules are typically unbounded in their domain of application, and the fact that the variable that covers this unbounded domain of application is subject to the island constraints does not alter this. The question then arises whether all transformations can apply over unbounded domains. Ross partially answered this question: only transformations with a "crucial" variable have this property (that is, transformations with a structural description of the form X – A – Y – B – Z, where A and B are constant terms of the structural description, X, Y, and Z are variable terms, and Y is the "crucial" variable—the variable that separates the two constant terms). Rules like Affix Hopping or Subject-AUX Inversion, which

lack such a crucial variable, cannot apply over unbounded domains. They are local. Subject-AUX Inversion, for example, moves the first auxiliary verb in a sentence to the position immediately to the left of the subject; the structural description of the rule can therefore be written without an essential medial variable, as in (39a):

(39)
a. X NP V_{AUX} Y
b. George can leave → Can George leave?

More or less simultaneously with Ross's work on island constraints, Peter Rosenbaum was exploring the syntax of complement clauses in English (Rosenbaum (1967)). Rosenbaum proposed transformations such as Extraposition and Equi-NP-Deletion (his Identity Erasure). Equi-NP-Deletion accounted for the phenomena of subject control that we will discuss in chapter 8. Curiously, this rule was neither unbounded nor local. In fact, it operated in general exactly across one (cyclic) S-node. In view of this, linguists began to wonder about what we might call the typology of transformations. The following basic typology emerged with respect to domain of application:

(40)
a. Monocyclic transformations: transformations that operate within one clause (S-domain), including the local transformations Passive, Reflexivization, etc.
b. Bicyclic transformations: transformations that operate across exactly one S-boundary, such as Equi-NP-Deletion and Raising-to-Object
c. Unbounded transformations: transformations that operate across arbitrarily many S-boundaries, such as *Wh*-Movement

This typology was quite popular for a while, even though it had two important weaknesses. First, it imposed a heavy burden on the language learner, who was now thought to learn not only the particulars of each transformation (its structural description, its structural change, whether it was obligatory or optional, etc.) but also its domain type. (It might be possible to circumvent this objection by positing a learning strategy based on the assumption that all rules are monoclausal unless evidence is found to the contrary, much as in our discussion of Emonds's typology in the next chapter. In the present case, however, a much more elegant solution is available, which we will present in chapter 4.) The second weakness was that (40) presupposes that a given

transformation belongs to the same type in all its applications. This is not obviously true, given the behavior of Passive, for example:

(41)
a. I believe $[_S$(that) John left$]_S$
b. *John is believed $[_S$(that) left$]_S$ (by me)

(42)
a. I believe $[_S$ John to have left$]_S$
b. John is believed $[_S$ to have left$]_S$ (by me)

If the ungrammaticality of (41b) follows from the fact that Passive is a monocyclic rule, then how can the grammaticality of (42b) be accounted for? Rosenbaum devised an ad hoc solution for this problem: he proposed that there is a bicyclic transformation that raises the subject of the infinitival complement clause into the VP of the matrix clause. This rule was called *Raising-to-Object*, and it provided the intermediate step (42c) for the derivation of (42b), making it possible to maintain Passive as a purely monocyclic rule:

(42)
c. I believe John $[_S$ to have left$]_S$

Much evidence was adduced to motivate Raising-to-Object, but virtually all of it was of the same type: counterexamples to the typology of (40) that could be eliminated if Raising-to-Object was assumed. Apparently nobody wondered about the correctness of (40) itself.

Raising-to-Object was shortly abandoned. Ross's island constraints, on the other hand, are still partly with us as central problems for linguistic research. Very few of the constraints have been regarded as true solutions because they lack sufficient generality. Most of them mirror in a fairly direct way the structural properties of the island configurations. In that sense they are observational generalizations rather than truly general abstract explanatory principles. Their great importance is in having initiated the systematic quest for such an abstract explanatory theory of conditions on transformations.

2.9 Bibliographical Comments

The A-over-A Principle was first introduced in Chomsky's (1964) *Current Issues in Linguistic Theory*. The first part of that book is devoted to largely methodological matters (in a way a precursor of the first chapter of *Aspects*), and the second part deals with phonology. The A-over-A Principle is hidden in footnote 10 on page 46. (Interestingly, the proposal appears in the main text of Chomsky (1962), an earlier version in article form of the same monograph.) The A-over-A proposal first takes real shape in Chomsky (1968) and is simultaneously attacked in Ross (1967), a very important piece of work that regrettably has remained unpublished.

The proposal regarding across-the-board rule application referred to in section 2.6 appears in Williams (1978). More recent proposals concerning the across-the-board property can be found in Gazdar (1981) and Pesetsky (1982b).

The Left Branch Condition has given rise to numerous publications. Of particular interest is the proposal in Bresnan (1976a) to derive the LBC from a reformulation of the A-over-A Principle, her "Relativized A-over-A Principle." Among the many counterarguments to the LBC, the most thoroughly documented is Obenauer (1976). Interesting observations with respect to the Upward Boundedness Constraint appear in Akmajian (1975) and Baltin (1981).

The typology of transformations discussed in section 2.8 was first introduced in Rosenbaum (1967). Its most controversial element, the Raising-to-Object transformation, was defended in Postal (1974). In the aftermath of the battle, which basically had already been won with the appearance of Chomsky (1973), Postal's book was reviewed by Bresnan (1976c) and Lightfoot (1976).