

# 琉球大学学術リポジトリ

## 極小主義以前の変形生成文法理論における制約の研究史

メタデータ	言語: 出版者: 琉球大学法文学部国際言語文化学科欧米系 公開日: 2008-01-15 キーワード (Ja): キーワード (En): 作成者: Yoshimoto, Yasushi, 吉本, 靖 メールアドレス: 所属:
URL	<a href="http://hdl.handle.net/20.500.12000/2955">http://hdl.handle.net/20.500.12000/2955</a>

## A Review of the History of Constraints in the Pre-Minimalist Transformational Generative Grammar

Yasushi Yoshimoto

### 1. Introduction

The purpose of this research note is to give a concise review of the partial historical developments of syntactic constraints on movement in the tradition of Chomskyan generative grammar. My review will be limited, covering only the pre-Minimalist conceptions of such constraints. Specifically, it begins with the A-over-A Principle suggested by Chomsky (1964) and ends with the status of constraints on movement put forward in Chomsky (1986).

After Chomsky (1986), the field has seen a dramatic change in the conception of theoretical architecture, with the advent of the Minimalist Program (Chomsky 1995, 2000). How to reinterpret the traditional constraints within the Minimalist framework is a challenging task that the contemporary Chomskyan linguists are facing. I hope that this short review will be of some help in clarifying what has been achieved in the past, and what needs to be accounted for in the Minimalist framework.

### 2. The A-over-A principle

Chomsky (1964a: 930-1, 1964b: 44-5) observed an interesting phenomenon concerning the ambiguous sentence (1).

(1) Mary saw the boy walking towards the railroad station.

One reading of (1) corresponds roughly to sentence (2), and another reading, to (3).

(2) Mary saw the boy walk towards the railroad station.

(3) Mary saw the boy who was walking towards the railroad station.

Interestingly enough, once example (1) is relativized or questioned by transformational rules, the ambiguity disappears. For example, the sentence *The railroad station that Mary saw the boy walking towards is about to be demolished*, which is formed by Relativization from (1), is only associated with the meaning that is parallel to (2). Similarly, the interrogative sentence *What did Mary see the boy walking towards?*, which is derived from (1) by the Question Formation transformation, is unambiguous and can only be interpreted as an interrogative of the reading corresponding to (2), not to (3).

Chomsky (1964a: 931) suggested to account for these facts in terms of a general condition on the application of transformational rules.<sup>1</sup> The hypothetical condition that Chomsky proposed is quoted in (4).

(4) If the phrase X of category A is embedded within a larger phrase ZXW which is also of category A, then no rule applying to the category A applies to X (but only to ZXW).

This condition, which was later to be called the A-over-A principle, explains the above facts as follows. The ambiguity of sentence (1) can be expressed in terms of the two distinct structural descriptions, (5) and (6), underlying it:

(5) Mary saw [the boy] [walking towards [<sub>NP</sub> the railroad station]]

(6) Mary saw [<sub>NP</sub> [the boy] [walking towards [<sub>NP</sub> the railroad station]]]

In (5), the phrase *walking towards the railroad station* is a complement of the matrix verb *saw*, while in (6), it is a relative clause forming a single NP with *the boy*. The structure (5) corresponds to the interpretation analogous to (2), and the structure (6), to (3). In order to derive from (5) and (6) the Wh-question *What did Mary see the boy*

*walking towards?*, the NP *the railroad station* has to move to the front of the sentence. However, the A-over-A principle in (4) prohibits this movement in (6) because the NP *the railroad station* is contained within a larger NP. In such a case the principle dictates that the constituent that could be affected by a transformational rule is the larger, dominating constituent. On the other hand, this movement is allowed in (5), since the NP *the railroad station* is not dominated by another NP. This would explain why the sentence *What did Mary see the boy walking towards?* has the unique interpretation expressed in the structural description (5), excluding the potential interpretation associated with the structural description (6). The similar argument also explains why the relativized phrase *the railroad station that Mary saw the boy walking towards* is associated with the interpretation analogous to (5), but not to (6).

The A-over-A principle demonstrated a possibility that abstract structural constraints on transformational rules may be uncovered by carefully examining general properties shared by different transformational rules. Since its very inception, however, the A-over-A principle was known to be too strong in that it incorrectly excluded some grammatical sentences. Thus, Chomsky (1964b: 46, f.n.10) notes that such sentences as *Who would you approve of my seeing?*, *What are you uncertain about giving to John?*, etc. are wrongly predicted to be ungrammatical by the principle. Such inadequacies of the A-over-A principle inspired J. R. Ross to explore further the issue of constraining transformations (i.e., movement rules).

### 3. Ross's Constraints

In his very influential Ph.D. dissertation, Ross (1967) carefully examined Chomsky's A-over-A principle. Chomsky's (1964b) own

counterexamples to the principle had a structure in which an NP has been moved out of another NP which dominated the former NP before the movement. Showing the examples in (7), Ross (1967: § 2.1) pointed out that in fact an NP embedded in more than one NP could be moved out of them.

- (7) a. The reports, [the height of the lettering on the covers of which] the government prescribes.  
b. The reports, [the lettering on the covers of which] the government prescribes the height of.  
c. The reports, [the covers of which] the government prescribes the height of the lettering on.  
d. The reports [which] the government prescribes the height of the lettering on the covers of.

The relative clauses of these examples derive from the same underlying structure, roughly as in (8), according to the analysis of Relativization that Ross (1967) was assuming.<sup>2</sup>

- (8) [<sub>S</sub> the government prescribes [<sub>NP1</sub> the height of [<sub>NP2</sub> the lettering on [<sub>NP3</sub> the covers of [<sub>NP4</sub> the reports]]]]]

Roughly speaking, example (7a) is derived by movement of NP1 in (8) to the sentence-initial position. Similarly, examples (7b), (7c), and (7d) are derived by movement of NP2, NP3, and NP4, respectively, to the front of S. Of these examples, only (7a) obeys the A-over-A principle. In all the other examples, an NP is moved out of one or more dominating NPs. Nevertheless, all the examples in (7) are acceptable.

In order to explain these kinds of examples, Ross abandoned the A-over-A principle in favor of proposing several independent constraints. These constraints were designed in such a way that they would avoid the difficulties associated with the A-over-A principle, but still explain all the cases that the A-over-A principle successfully accounted for. The

four most important of his constraints are cited below from Ross (1967):

(9) *The Complex NP Constraint*

No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.

(10) *The Coordinate Structure Constraint*

In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

(11) *The Left Branch Condition*

No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule.

(12) *The Sentential Subject Constraint*

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.

Let us see some representative examples that the above constraints can account for. First, the Complex NP Constraint explains such examples as (14) that derives from the underlying structure (13).

(13) I chased [<sub>NP</sub> the boy who threw [<sub>NP</sub> a snowball] at our teacher]

(14) \*Here is the snowball **which** [<sub>S1</sub> I chased [<sub>NP</sub> the boy [<sub>S2</sub> who threw \_\_\_ at our teacher]]]

The then-current transformational rule of Relativization moves the NP *a snowball* in (13) to the clause-initial position and replaces it with the relative pronoun *which* as in (14). The A-over-A principle prohibits this movement because the NP *a snowball* is contained in a larger NP in (13). Ross's Complex NP Constraint, on the other hand, explains the same fact as follows. To derive (14), the NP *a snowball* has to be

extracted out of S2, which is dominated by an NP with a lexical head noun (=boy). Thus (14) violates the Complex NP Constraint.

Consider next examples in (15)-(17). Neither of the conjoined NPs in example (15) may be questioned, as demonstrated by the ungrammaticality of (16) and (17).

(15) He will put the chair between [<sub>NP</sub> [<sub>NP</sub> some table] and [<sub>NP</sub> some sofa]]

(16) \***What sofa** will he put the chair between [<sub>NP</sub> some table and \_\_ ]?

(17) \***What table** will he put the chair between [<sub>NP</sub> \_\_ and some sofa]?

The Coordinate Structure Constraint is specifically constructed in such a way that examples like (16) and (17) would be excluded. Note that the A-over-A principle straightforwardly explains this fact: since the conjunct NPs are constituents of a larger NP, no transformational rule can move the conjuncts out of the larger NP.

Examples (18)-(20) below are relevant for the Left Branch Condition. Both (19) and (20) are derived from (18), but only (19) is well-formed.

(18) I found [<sub>NP</sub> [<sub>NP</sub> someone's] book]]

(19) **Whose book** did you find \_\_ ?

(20) \***Whose** did you find [<sub>NP</sub> \_\_ book]?

Example (20) is ungrammatical because the leftmost NP (*someone's*, to be later replaced by *whose*) has moved out of the dominating NP, violating the Left Branch Condition. The derivation of example (19) is allowed because it is the entire NP (*someone's book*) that has been moved. Again, the facts observed here are also successfully accounted for by Chomsky's A-over-A principle.

Finally, let us consider an example relevant to the Sentential Subject Constraint. This constraint accounts for the ungrammaticality of (21).

(21) \*<sub>[NP Who]</sub> is <sub>[NP [s that Mary will meet \_\_ ]]</sub> obvious?

The NP *who* has been extracted in (21) out of the sentential subject,

which is what the Sentential Subject Constraint prohibits. Here too the A-over-A principle correctly disallows this derivation, since the sentential subject is dominated by NP, the same category as the moved element *who*.

As far as the above examples are concerned, the A-over-A principle is still certainly more attractive than Ross's constraints because it captures generalization. But it was a generalization made possible at the cost of serious empirical problems. Ross's constraints successfully eliminated such problems. To give just one example, the following grammatical sentence (22) is in violation of the A-over-A principle, as Chomsky himself noted, but it violates none of Ross's constraints.

(22) [<sub>NP</sub> Who] would you approve of [<sub>NP</sub> my seeing \_\_ ]?

In addition, Ross's constraints could explain examples that the A-over-A principle could not. For instance, the A-over-A principle has nothing to say about the ill-formedness of example (23), in which the adjectival phrase has been moved to the sentence-initial position.

(23) \* [<sub>AP</sub> How intelligent] will you work for [<sub>NP</sub> a boss [<sub>S</sub> who is \_\_ ]]? But the Complex NP Constraint successfully accounts for the ungrammaticality of this example. Similarly, the Coordinate Structure Constraint predicts the ill-formedness of example (24) because an element (*who*) has been extracted out of a conjunct (*tired of* \_\_ ).

(24) \* [<sub>NP</sub> Who] is Bill [<sub>AP</sub> proud of [<sub>NP</sub> his father]] and [<sub>AP</sub> tired of \_\_ ]? The A-over-A principle on the other hand fails to exclude (24) since the moved element *who* and the conjunct out of which it moved are not of the same category. There are numerous examples such as these that show the superiority of Ross's constraints over the A-over-A principle.

Examples like (23) and (24) indicate that the impossibility of certain extractions is not due to movement of a category out of the same category — as the A-over-A principle suggested — but rather due to the

kinds of structure out of which an element is extracted. This is what Ross realized. He named such special kinds of structure "islands," which soon proved to be one of the most important concepts in syntactic theory.

#### 4. Subjacency

The appearance of Ross (1967) stimulated many syntacticians to investigate further the notion of "islands." While a number of linguists proposed to refine and reanalyze Ross's constraints in subsequent years, the next seminal work on the issue of constraining transformations is perhaps Chomsky (1973). As we have seen in the previous section, although Ross's constraints were very successful in overcoming the deficiency of the A-over-A principle, Ross had to pay the price of introducing several structure-specific constraints instead of a single principle. By proposing the Subjacency Condition, Chomsky (1973) opened the way to investigate a more abstract principle that could unify some of Ross's constraints. His initial formulation of the Subjacency Condition is cited in (25) together with the definition of the term "subjacent" in (26).

(25) *The Subjacency Condition*

No rule can involve X, Y, X superior to Y, if Y is not subjacent to X.

- (26) If X is superior<sup>3</sup> to Y in a phrase marker P, then Y is "subjacent" to X if there is at most one cyclic category C ≠ Y such that C contains Y and C does not contain X.

Among other things, the Subjacency Condition prohibits movement across more than one cyclic node, which were initially assumed to be S' and NP in Chomsky (1973)<sup>4</sup>. Later, Chomsky (1977) considers including S as one of the cyclic nodes in English. Once we assume that

S is a bounding node (=cyclic node) in English as well as S' and NP are, the four constraints of Ross's that we discussed in the previous section all fall under the Subjacency Condition. To confirm this point, let us consider once again the following examples that violate Ross's constraints.

(27) *Complex NP Constraint*

\*Here is the snowball **which**<sub>1</sub> I chased [<sub>NP</sub> the boy [<sub>S</sub> who [<sub>S</sub> threw *t*<sub>1</sub> at our teacher]]]

(28) *Coordinate Structure Constraint*

\***What sofa**<sub>1</sub> will [<sub>S</sub> he put the chair between [<sub>NP</sub> some table and *t*<sub>1</sub>]]?

(29) *Left Branch Condition*

\*Whose<sub>1</sub> did [<sub>S</sub> you find [<sub>NP</sub> *t*<sub>1</sub> book]]?

(30) *Sentential Subject Condition*

\***Who**<sub>1</sub> is [<sub>NP</sub> [<sub>S</sub> that [<sub>S</sub> Mary will meet *t*<sub>1</sub>]]] obvious?

In all these examples, a *wh*-phrase has crossed more than one bounding node, violating Subjacency. Note that if S is not taken to be a bounding node, then examples (28) and (29) will not be accounted for by Subjacency. In other words, the Coordinate Structure Constraint and the Left Branch Condition would be subsumed under Subjacency only when S counts as a bounding node. On the other hand, the Complex NP Constraint and the Sentential Subject Condition would be subsumed under Subjacency regardless of whether or not S counts as a bounding node. In this connection, it is worth noting that Rizzi (1978) offered evidence that for Italian, S' but not S is a bounding node for Subjacency. This work by Rizzi led to a general agreement that the choice of bounding nodes is parameterized in UG.

We have just seen that the Subjacency Condition nicely unifies Ross's major constraints, given that the bounding nodes in English are

S, S', and NP. However, Chomsky was aware of a number of problems associated with Subjacency, including the choice of S, S', and NP as the bounding nodes in English. Let us take a look at two of such problems. First, when S was not considered a bounding node, Subjacency correctly allowed examples such as (31).

(31) Who<sub>i</sub> did [<sub>S</sub> you think [<sub>S</sub> t'<sub>i</sub> [<sub>S</sub> Bill saw t<sub>i</sub>]]] ?

Under the assumption that the bounding nodes are S' and NP, but not S, successive-cyclic movement of *who* in (31) complies with the Subjacency Condition: the first movement to the COMP in the embedded clause crosses no bounding node, and the second movement to the COMP of the matrix clause crosses only one bounding node, S', which is allowed by Subjacency. However, under the assumption that S is also a bounding node, the movement from the embedded-clause COMP to the matrix-clause COMP crosses two bounding nodes, S' and S. Thus, (31) would be incorrectly ruled ungrammatical by the Subjacency Condition. In order to allow examples like (31), Chomsky (1980) suggested, following Koster (1978a,b), that when the matrix verb is a bridge verb such as *think*, *believe*, etc., this S' does not count as a bounding node for Subjacency. Given this stipulation, (31) will be correctly allowed, while examples such as (32), which contains a nonbridge verb, will be correctly disallowed.

(32) \*Who<sub>i</sub> did [<sub>S</sub> you whisper [<sub>S</sub> t'<sub>i</sub> [<sub>S</sub> Bill saw t<sub>i</sub> ]]] ?

A much more serious problem concerning the choice of S as one of the bounding nodes in English, it seems to me, is how to explain the grammaticality of (33).

(33) Who<sub>i</sub> would [<sub>S</sub> you approve of [<sub>NP</sub> my seeing t<sub>i</sub>]] ?

Given that S is also a bounding node, (33) should be ill-formed since the *wh*-phrase has crossed two bounding nodes, S and NP. But actually (33) is grammatical. As far as I know, (33) remained problematic for the

Subjacency Condition that took S, as well as S' and NP, as a bounding node.

In spite of the problems just mentioned, the Subjacency Condition is appealing because it can explain a wide range of phenomena that were treated disparately before. For instance, in addition to the island-violation cases that we saw above, Subjacency explains the *wh*-island phenomenon observed in examples such as (34).

(34) \*Who<sub>1</sub> do you know [<sub>S</sub> when<sub>2</sub> [<sub>S</sub> John met *t*<sub>1</sub> *t*<sub>2</sub>]]

The ill-formedness of (34) follows from the fact that *who* had to move across more than one bounding node due to the presence of *when* in the embedded-clause COMP. Before the appearance of Subjacency, Chomsky (1964: 43) had accounted for the *wh*-island phenomenon in terms of an independent condition which stated that *Wh*-Movement cannot apply twice within a clause. With Subjacency at hand, this condition became unnecessary.<sup>5</sup>

Another locality phenomenon that Subjacency successfully handles is found in PP-extrapolation sentences. In the following examples, both (35) and (36) are derived from the same underlying structure by extraposing PP (indicated by the bold face) out of NP, but only (35) is acceptable:

(35) The fact that [a critical review \_\_ ]<sub>NP</sub> has just appeared **of his latest book** is very worrying.

(36) \*The fact that [[ a critical review \_\_ ]<sub>NP</sub> has just appeared]<sub>S</sub> is very worrying **of his latest book**.

This fact straightforwardly follows from Subjacency. In contrast, Ross (1967) accounted for the same kind of fact by imposing a constraint on the rule of PP-extrapolation using Langacker's (1969) notion of "command."<sup>6</sup> This constraint of Ross's applied not only to PP-extrapolation but also to many other syntactic rules. Yet, it was a

constraint independent of his other major constraints that we saw in section 3. Subjacency, on the other hand, uniformly explained those facts that Ross's constraints separately explained.

## 5. Subjacency and the ECP

In addition to the Subjacency Condition, Chomsky (1973) also proposed the Tensed S Condition (TSC) and the Specified Subject Condition (SSC) that are also constraints on transformations. As the TSC and the SSC developed through successive revisions (Chomsky (1976, 1980)) into the Binding Theory of Chomsky (1981), a new principle emerged concerning empty categories. Chomsky (1981) named it the Empty Category Principle (ECP):

(37) *The Empty Category Principle*<sup>7</sup>

An empty category [ <sub>$\beta$</sub>  e] must be "properly governed," where  $\alpha$  properly governs  $\beta$  if and only if  $\alpha$  governs  $\beta$  and

- a.  $\alpha = [\pm N, \pm V]$  or
- b.  $\alpha$  is coindexed with  $\beta$ .

Among other things, the ECP explains a number of subject/object asymmetries.<sup>8</sup> What I would like to discuss here, however, is the relationship between the ECP and Subjacency. More specifically, can Subjacency be subsumed under the ECP? Kayne (1981) argued that a reformulation of the ECP in (37) into the one in (38) would permit the ECP to treat an important part of what was considered to fall under Subjacency.

(38) *The Empty Category Principle* (Kayne's (1981) formulation)

An empty category  $\beta$  must have an antecedent  $\alpha$  such that (1)  $\alpha$  governs  $\beta$  or (2)  $\alpha$  c-commands  $\beta$  and there exists a lexical category X such that X governs  $\beta$  and  $\alpha$  is contained in some percolation projection of X.

(39) *Percolation Projection*

A is a percolation projection of B if A is a projection of B, or A is a projection of C, where C bears the same superscript as B and governs a projection of B, or a percolation projection of B.

Let us consider how the following example is accounted for by the different versions of the ECP given above.

(40) [<sub>S'</sub> Who<sub>i</sub> do you [<sub>VP</sub> think' [<sub>S'</sub> that Mary [<sub>VP</sub> saw' *t*<sub>i</sub>]]]] ?

The trace in (40) satisfies Chomsky's ECP (37) since it is governed by a lexical category, *saw*. On the other hand, lexical government alone does not suffice for Kayne's ECP (38) to be satisfied: a trace must either be governed by an antecedent or have a c-commanding antecedent in a percolation projection of a governor of the trace. Since the trace in (40) is not governed by its antecedent *who*, it should be the case that this antecedent is contained in some percolation projection of the embedded verb *saw* that governs the trace. Kayne (1981) crucially assumes that when V governs S', V assigns its superscript to S' and the superscript percolates down to the head of S', namely the embedded V. (S' is assumed to be a maximal projection of V.) Given this mechanism, the superscripting indicated in (40) results, and the matrix S' becomes a percolation projection of the embedded verb *saw*. In this percolation projection, the trace has its c-commanding antecedent, *who*, satisfying the ECP in (38). Thus, either version of the ECP, (37) or (38), correctly predicts the grammaticality of example (40).

It is in examples such as (41) that the two versions of the ECP differ crucially. Chomsky's ECP has nothing to say about the ungrammaticality of (41), but Kayne's ECP is formulated in such a way that it will be relevant exactly in such cases.

(41) \*Who<sub>i</sub> don't you believe [<sub>NP</sub> the hypothesis [<sub>S'</sub> that John loves *t*<sub>i</sub>]]?

According to Kayne (1981), the noun *hypothesis* in (41) cannot be co-superscripted with the S' indicated. More generally, he assumes that N can never assign a superscript, in contrast to V.<sup>9</sup> Given this, the maximal percolation projection of the trace in (41) is the embedded S'. There is no antecedent for the trace in this embedded S'. Thus, for Kayne (1981), the ungrammaticality of (41) is accounted for as an ECP violation. For Chomsky (1981), on the other hand, (41) is fine as far as the ECP is concerned, since the trace there is governed by V, a lexical category.

Of course, (41) is a typical example of a Subjacency violation (and of a Complex NP Constraint violation). So Chomsky would argue that (41) should be handled independently of the ECP. However, if Kayne's (1981) ECP can explain all the phenomena that have been considered to fall under Subjacency, then we would no longer need Subjacency. Attractive though Kayne's proposal was, a number of empirical problems in subsuming Subjacency under the ECP were quickly pointed out by Aoun (1981) and Huang (1982).

Aoun (1981) demonstrated that even if one adopts Kayne's (1981) definition of the ECP, there are constructions in Standard Arabic and Modern Hebrew that cannot be explained by the ECP, but can be explained by Subjacency. Thus he concluded that Subjacency cannot be subsumed under the ECP. Huang (1982) showed that even in English, there are constructions that Subjacency can account for but Kayne's ECP can not. (To be fair to Kayne, I should emphasize that Kayne (1981) did not claim that all of the Subjacency phenomena could be handled by his ECP.) For example, consider example (42), which violates the *Wh*-island Condition.

(42) \*?[<sub>s'</sub> What<sub>i</sub> [<sub>s'</sub> do you wonder<sup>i</sup> [<sub>s'</sub> where [<sub>s'</sub> I bought<sup>i</sup> *t*<sub>i</sub>]]]] ?

Given Kayne's mechanism of superscripting, the representation

indicated in (42) results. Thus, the matrix S' is a percolation projection of the embedded verb *bought*, which governs the trace. Within this percolation projection, the trace finds its antecedent, *what*. Hence, (42) will be allowed, contrary to fact. Note that the superscripting mechanism that yields the representation in (42) is exactly what Kayne needs in order to explain the grammaticality of such sentences as (43):

(43) [<sub>S'</sub> What<sub>i</sub> [<sub>S'</sub> do you think<sub>i</sub> [<sub>S'</sub> that [<sub>S'</sub> I bought<sub>i</sub> t<sub>i</sub> ]]]] ?

While Kayne's ECP fails to account for the contrast between (42) and (43), Subjacency readily accounts for it. In the well-formed example (43), the *wh*-phrase has moved successive cyclically, each movement obeying Subjacency.<sup>10</sup> In contrast, the unacceptable example (42) does not allow successive cyclic movement of *what*, since the embedded COMP position is already occupied by another *wh*-phrase. Thus, *what* in (42) has moved in a single swoop to the matrix COMP position, crossing more than one bounding node, violating Subjacency.

Another example Huang (1982) gave to illustrate that Kayne's ECP cannot replace Subjacency involves super-raising:

(44) \*John<sub>i</sub> seems that [<sub>S</sub> it is certain [<sub>S</sub> t<sub>i</sub> to win]]

Before considering (44), we must take a look at how Kayne (1981) would account for simple raising sentences-like (45):

(45) John<sub>i</sub> is certain [<sub>S</sub> t<sub>i</sub> to win ]

Since (45) is well-formed, the trace must satisfy the ECP. Kayne (1981, f.n.25) conjectures that the trace in constructions such as (45) is governed by the reanalyzed element of *be+A*. For instance, in (45), the trace is governed by a combination of *is* and *certain*, reanalyzed as a single element. Given this analysis, the trace in (45) obeys the ECP, since it finds a c-commanding antecedent within the matrix S', a percolation projection of the governor of the trace. Let us now turn to example (44). According to the analysis of (45) just mentioned, the following

superscripting obtains for (44).

(46) \* $[_s^i \text{ John}_1 \text{ seems}^i [_s^i \text{ that it [is certain]}^i [_s \text{ } t_1 \text{ to win } ]]]$

Then, contrary to fact, Kayne's ECP predicts (44) to be grammatical. Subjacency, on the other hand, again offers an explanation of the ungrammatical status of (44), as Huang (1982) points out. In short, examples such as (42) and (44) demonstrate that even with the reformulated version of the ECP (38), which successfully subsumes a considerable portion of the Subjacency effects, Subjacency is still needed to exclude certain constructions. Then why reformulate the ECP? This question led Huang (1982) to opt for preserving the ECP as it was originally formulated in Chomsky (1981), keeping Subjacency intact.

However, Huang also recognized that Kayne's concept of the ECP contained some important insights. For one thing, Kayne's ECP, with a proper formulation of the superscripting mechanism, can successfully derive the effects of Chomsky's (1973) Subject Condition. The following examples illustrate this point.

(47)  $[_s \text{ Who } [_s \text{ did you see } [_{NP} \text{ pictures } [_{PP} \text{ of } t ]]] ] ?$

(48) \* $[_s \text{ Who } [_s \text{ did } [_{NP} \text{ pictures } [_{PP} \text{ of } t ] ] \text{ please you } ] ] ?$

Example (48) is an instance of a Subject Condition violation. Huang argues that once we assume that nouns assign superscript to PPs that they govern,<sup>11</sup> then Kayne's ECP accounts for the grammaticality contrast between (47) and (48): in (47), S' becomes a percolation projection of *of*, the governor of the trace, allowing the trace to satisfy the ECP; in (48), on the other hand, the percolation projection of *of* does not extend beyond NP since the verb *please* cannot assign its superscript to NP, NP being in subject position; the trace in (48) thus violates the ECP. As before, Chomsky's (1981) formulation of the ECP has nothing to say about the ungrammaticality of (48), and (48) will fall under Subjacency in his system. What caught Huang's attention was that (47) and (48) exhibit

another subject/object asymmetry, although they do not fall under Chomsky's original ECP. Since the major effect of the original ECP is to yield subject/object asymmetry, it makes intuitive sense that Kayne's reformulation ties together the kind of subject/object asymmetry observed in (47) and (48) with the standard subject/object asymmetry explained by the original ECP. This observation led Huang (1982) to explore the possibility of the existence of a condition closely related to the ECP, but nonetheless distinct from it. The next section starts with a discussion of that condition.

## 6. The Condition on Extraction Domains

One of Huang's (1982) major insights was that the subject/object asymmetry could be taken as a special case of an even more general asymmetry between complements on the one hand and non-complements (subjects and adjuncts) on the other. To capture this general asymmetry, Huang (1982: 505) proposed the following condition.

(49) *Condition on Extraction Domains (CED)*

A phrase A may be extracted out of a domain B only if B is properly governed.

The basic effect of the CED is to allow extraction out of object phrases, but to prohibit extraction out of subject and adjunct phrases. The examples below illustrate this.

(50) Who did you see [<sub>NP</sub> pictures of *t* ] ? (=47)

(51) \*Who did [<sub>NP</sub> pictures of *t* ] please you ? (=48)

(52) \*Who did Mary cry [<sub>PP</sub> after John hit *t* ] ?

NP in (50) is properly governed, but NP and PP in (51) and (52), respectively, are not. Hence, (50) obeys the CED, but (51) and (52) violate it, yielding the paradigm.

The CED is assumed to apply only at S-structure, not at LF. This

is a crucial assumption, because if the CED applied at LF, such examples as (53) would be problematic.

(53) Who cried [after John hit who] ?

At S-structure, (53) does not pose any problem. But because every *wh*-phrase must be in COMP at LF in Huang's (1982) theory, the second *who* in (53) must be extracted out of the adjunct phrase at LF. Thus, if the CED applied at LF, (53) would wrongly be excluded. By restricting the application of the CED to S-structure, this problem can be avoided.

The CED resembles the ECP in that the notion of proper government is crucially involved. The basic difference between the two is that the former concerns the position of the construction from which extraction takes place, while the latter concerns the position of the trace itself. In addition, they differ in the level of application. The CED applies at S-structure, as we have just seen, while the ECP must apply at LF, if the following contrast – a “superiority effect” – is to be accounted for by the ECP:

(54) Who bought what?

(55) \*What<sub>2</sub> did who buy t<sub>2</sub>?

At the level of S-structure, the ECP cannot exclude the ungrammatical (55), since the trace is lexically governed by the verb. But the ECP does yield the desired contrast at LF:

(56) [what<sub>2</sub> who<sub>1</sub>]<sub>1</sub> [ t<sub>1</sub> bought t<sub>2</sub> ] ? (=LF of (54))

(57) \*[who<sub>1</sub> what<sub>2</sub>]<sub>2</sub> [ t<sub>1</sub> bought t<sub>2</sub> ] ? (=LF of (55))

In (56), both traces satisfy the ECP: *t*<sub>1</sub> is antecedent governed by COMP and *t*<sub>2</sub> is lexically governed. But in (57), *t*<sub>1</sub> is not antecedent governed because COMP bears index 2. It is not lexically governed, either. Thus, an ECP violation results in (57), accounting for the ill-formedness of (55).

Before closing this section, let us consider briefly the relation between the CED and Subjacency. Subjacency remains to play a crucial

role in Huang's (1982) theory, in order to handle the Complex NP Constraint cases, the *wh*-island phenomena, and so forth, which cannot be taken care of by the CED. On the other hand, there is at least one case that Subjacency cannot handle but the CED can: extraction out of an adjunct phrase. Nevertheless, there is a considerable amount of redundancy to include both the CED and Subjacency in grammatical theory. For example, the Subject Condition violations and the Sentential Subject Condition violations are excluded by both the CED and Subjacency. One would expect that a better theory would eliminate the redundancy detected between Subjacency and the CED, just as Binding Theory eliminated the redundancy of its predecessors.<sup>12</sup> A work by Chomsky (1986), our next topic, can be seen as an attempt to derive the effects of the CED in terms of Subjacency or the ECP, thus eliminating the redundancy in Huang's (1982) system.

## 7. Barriers

In Chomsky (1986), a number of modifications of the already existing conditions have been made. Among them, what concerns us most in the present context is the modification of Subjacency and the revision of the definition of "proper government." In Chomsky (1986), the term "barrier" replaces the term "bounding node" for Subjacency. While "bounding nodes" were absolute categories such as NP and S, "barriers" are determined contextually.<sup>13</sup> Thus, the same category can be a barrier in one context, but not a barrier in another. As for the definition of "proper government," the notion of  $\theta$ -marking is now crucially involved.<sup>14</sup>

Let us consider how the revised Subjacency explains the CED effects. The relevant examples are (58) and (59), which involve extraction out of a subject and out of an adjunct, respectively.

(58) \*the man **who** [<sub>IP</sub> [<sub>NP</sub> pictures of *t* ] are on the table]

(59) \***To whom** did [<sub>IP</sub> they leave [<sub>PP</sub> before speaking *t* ]]?

In (58), the embedded subject NP is a barrier for the trace in it, because it is not L-marked. IP (=S) inherits barrierhood from this NP, so that altogether *who* has crossed two barriers in this example. The basic idea of Subjacency in Chomsky (1986) remains the same as its initial formulation in Chomsky (1973): movement may not cross more than one special kind of category. Only now the “special kind of category” is understood as a “barrier” which is determined contextually. Thus, (58) violates Subjacency. Turning to (59), PP is not L-marked, as the verb *leave* does not  $\theta$ -mark it. Thus, PP is a barrier for the trace, and IP inherits barrierhood from PP, forming two barriers for the phrase *to whom* to cross. Hence, the unacceptability of (59). As is clear from these examples, Huang’s (and Kayne’s) insight that a construction that is not properly governed forms an island is incorporated in Chomsky’s (1986) conception of barriers by defining it in terms of the notion of L-marking.

Let us next consider the Complex NP Constraint violations. Chomsky’s (1986) notion of barriers, which is sensitive to the notion of  $\theta$ -marking, opens the way to explain the difference in acceptability in the following examples, which both violate the CNPC. (60) is much worse than (61).

(60) Which book did John meet [<sub>NP</sub> a child [<sub>CP</sub> who read *t* ]]?

(61) Which book did John hear [<sub>NP</sub> a rumor [<sub>CP</sub> that you had read *t* ]]?

Chomsky claims that the relative clause construction (60) yields a Subjacency violation, while the noun-complement construction (61) does not. This is so because the relative clause CP in (60) is not L-marked by the head noun *child*, whereas the complement clause CP in (61) is L-marked by the head noun *rumor*. As a result, between the trace and the

moved *wh*-phrase, there exist two barriers in (60) (NP inherits barrierhood from CP), but there is no barrier in (61). Thus, (61) is more or less acceptable, but (60) is not.<sup>15</sup>

It is worth mentioning at this point that a problem of the old Subjacency that I raised in section 4 concerning example (33) is naturally solved by the Subjacency of Chomsky (1986). Example (33), repeated here as (62), was problematic under the old Subjacency because NP and S (=IP) were absolute barriers for movement.

(62) Who<sub>i</sub> would [<sub>IP</sub> you approve of [<sub>NP</sub> my seeing *t*<sub>i</sub>]] ?

With the relativized notion of barriers at hand, however, NP in (62) does not qualify as a barrier because it is L-marked (presumably by the reanalyzed element *approve-of*). Hence, the movement involved in (62) crosses no barrier, as desired.<sup>16</sup>

An important assumption in Chomsky's *Barriers* theory is that the same notion of barriers is involved for both the theory of movement (including Subjacency) and the theory of government (including the ECP).<sup>17</sup> Let us just take one example of an ECP violation to illustrate how it is accounted for within the framework of *Barriers*:

(63) \*John seems [<sub>CP</sub> that [<sub>IP</sub> it is certain [<sub>IP</sub> *t* to win]]]

We have seen in section 5 that Huang (1982) attributed the ungrammaticality of this example (= (46)) to Subjacency. In Chomsky's (1986) theory, the same fact is accounted for by the ECP. In (63), the trace is governed, but not properly governed, by *certain*, because *certain* does not  $\theta$ -mark the trace. It is not antecedent-governed by *John* either, because CP is a barrier by inheritance from IP'. (The movement involved here is NP-movement; thus successive cyclic movement is not available.) Hence, (63) violates the ECP.

To conclude this section, we have seen how Chomsky's *Barriers* theory may eliminate the need for the CED and how it may

successfully overcome a deficiency of Ross's CNPC by virtue of the relativized notion of barriers. We have also seen very briefly that the ECP and Subjacency are now susceptible to the same notion of barriers. As is always the case with any new proposals, however, there are a number of problematic examples for the approach that Chomsky (1986) has taken. I will not discuss them here as it will take us beyond the bounds of this paper. Despite the problems, it is clear that there are some conceptual and empirical advantages of the *Barriers* approach over the previous approaches.

## 8. Concluding Remarks

I have summarized the history of constraints on syntactic rules from Chomsky's (1964) A-over-A principle to Chomsky's (1986) *Barriers*. The A-over-A principle was replaced by Ross's (1967) constraints, in the course of which the notion "island" was introduced. Chomsky's (1973) Subjacency successfully unified some of Ross's constraints, and Subjacency has proved to be a very resilient condition.

As the ECP was introduced in Chomsky (1981), it was pointed out by Kayne (1981) that a portion of the Subjacency phenomena could be subsumed under the ECP, if the ECP were reformulated in such a way that the existence of a "local" antecedent is a necessary condition in its definition. The works of Aoun (1981) and Huang (1982) showed some problems associated with an attempt to reduce Subjacency to the ECP. Furthermore, Huang (1982) argued that the subject/object asymmetry which is typically handled by the ECP is in fact a subcase of the more general complement/non-complement asymmetry. To capture this asymmetry, Huang proposed the CED, which in effect prohibits extraction out of non-complement clauses.

Chomsky (1986) then showed that if the notion of barriers is

defined contextually, crucially using the notion of  $\theta$ -government, then the CED effects could be obtained by Subjacency. Thus, Chomsky's (1986) system of constraints on movement contains Subjacency and the ECP, both of which involve the same notion of barriers. In addition, movements are constrained by a number of subtheories in this framework, including Binding Theory, Case theory, and  $\theta$ -theory. All of them, together with Subjacency and the ECP, "conspire" to filter out ill-formed sentences. This is roughly how the transformational rule of Move- $\alpha$ <sup>18</sup> was constrained at the final stage of the pre-Minimalist Chomskyan syntactic theory.<sup>19</sup>

### Notes

1. Interestingly, this suggestion is made in the main text in Chomsky (1964a), but is hidden in footnote 10 in Chomsky (1964b: 46), which is a revised and extended version of Chomsky (1964a). Chomsky (1964b) opts to explain the unambiguous status of examples (2) and (3) by proposing a condition that "[the *Wh*-movement transformation] is limited to a single application to each underlying terminal string" (p.43).
2. Strictly speaking, according to Ross's (1967) analysis, there are three more NPs within NP1 in (8), since he assumed that a phrase that consists of a preposition and a noun phrase is of the category type NP, not PP.
3. The category A is "superior" to the category B in the phrase marker if every major category (i.e., N, V, A and the categories that dominate them) dominating A dominates B as well but not conversely.
4. A note on terminology is in order. When one reads Chomsky (1973), one notices that Chomsky actually states that he assumes the cyclic

nodes to be S and NP, rather than S' and NP. However, this is due to the fact that Chomsky (1973) employed rather unusual usage of the symbols S and S'. That is, he used the notation in which a PS rule has the form of [S → Comp S'] rather than the standard notation in which it has the form of [S' → Comp S]. Thus, although Chomsky (1973) states that he assumes the cyclic nodes to be S and NP, what he meant was S' and NP in the standard terminology. In Chomsky (1977) and his subsequent works, the standard notation is used.

5. Incidentally, the “*wh*-island constraint,” which examples like (34) are said to violate, is often mistakenly attributed to Ross (1967) in the literature. In fact, Ross (1967 : § 2.4.1) specifically argued against such a constraint, claiming that sentences such as (i) are more or less acceptable:

(i) He told me about a book which<sub>1</sub> [<sub>s</sub> I can't figure out [<sub>s</sub> how<sub>2</sub>  
       [<sub>s</sub> to read *t*<sub>1</sub> *t*<sub>2</sub>]]

The Subjacency Condition as formulated in (25) fails to explain why (i) is acceptable.

6. See Ross (1967: § 5.2.1.2).
7. This definition is the paraphrase by Kayne (1981) of Chomsky's (1981) original formulation of the ECP. Note that the pronominal anaphor, PRO, must be excluded from the class of empty categories that are subject to the ECP, since the PRO theorem requires that PRO be ungoverned.
8. For example, the subject/object asymmetry observed in the following pair of sentences is handled by the ECP.

- (i) Who do you think that John saw *t* ?  
 (ii) \*Who do you think that *t* saw Bill ?

The trace in (i), which is in object position, is properly governed by

the verb *saw*. But the trace in (ii), which is in subject position, is neither lexically governed nor antecedent governed, and therefore violates the ECP. The ungrammaticality of examples like (ii) was formerly explained by such devices as the *that*-trace Filter (Chomsky and Lasnik (1977)), but now follows from a general principle, the ECP.

9. See Kayne (1981 : § 3.2.3) for arguments that motivate this assumption.
10. Recall that S' after a bridge verb like *think* does not count as a bounding node.
11. This assumption of Huang, however, undermines Kayne's analysis of nominal constructions such as (i), which contrasts to its clausal counterpart (ii) in grammaticality:
  - (i) \*Mary's (recent) reference to *t* (in the newspaper)
  - (ii) Mary was recently referred to *t* in the newspaper.

Kayne (1981) argues that the ungrammaticality of (i) can be explained by his ECP under the assumption that N can never assign a superscript to any category, including PP. Once this assumption is denied, as Huang did, (i) will not fall under the ECP.

12. Binding Theory finds its origin in the Tensed S Condition and the Specified Subject Condition of Chomsky (1973). To reduce the redundancy between these two conditions, Chomsky (1980) proposed the Opacity Condition and the Nominative Island condition. The Binding Theory of Chomsky (1981) eliminated the redundancy between the Opacity Condition and the Nominative Island Condition.
13. The definition of "barrier" is complex. To understand its definition in (iv), one has to start with the definition of  $\theta$ -government in (i). (See Chomsky (1986: 14-5)):

- (i)  $\alpha$   $\theta$ -governs  $\beta$  iff  $\alpha$  is a zero-level category that  $\theta$ -marks  $\beta$ , and  $\alpha$ ,  $\beta$  are sisters.
  - (ii)  $\alpha$  L-marks  $\beta$  iff  $\alpha$  is a lexical category that  $\theta$ -governs  $\beta$ .
  - (iii)  $\gamma$  is a BC for  $\beta$  iff  $\gamma$  is not L-marked and  $\gamma$  dominates  $\beta$ .
  - (iv)  $\gamma$  is a barrier for  $\beta$  iff (a) or (b).
    - a.  $\gamma$  immediately dominates  $\delta$ ,  $\delta$  a BC for  $\beta$ ;
    - b.  $\gamma$  is a BC for  $\beta$ ,  $\gamma \neq \text{IP}$ .
14. Chomsky's (1986: 17) definition is as follows:  
 $\alpha$  properly governs  $\beta$  iff  $\alpha$   $\theta$ -governs  $\beta$  or antecedent-governs  $\beta$ .
15. Regarding the "intermediate status" of (61), Chomsky conjectures that although the definition of "barriers" he has given does not create any barrier for the movement in (61), there is nevertheless one barrier. See Chomsky (1986: 35-36) for discussion.
16. IP in (62) is a BC, but not a barrier. IP is special in that it cannot be a barrier just by being a BC. See footnote 13.
17. In addition, the theory of government is assumed to involve another kind of barrier called "Minimality barrier."
18. The rule Move- $\alpha$  states "move any category  $\alpha$  anywhere."
19. I take Chomsky (1986) to be the last major work by Chomsky within the pre-Minimalist framework. Chomsky (1991) appears to employ the *Barriers* framework, but it certainly explores a Minimalist approach, as Chomsky (1995: 9) himself mentions.

### References

- Aoun, J. (1981) "ECP, Move  $\alpha$ , and Subjacency," *Linguistic Inquiry* 12, 637-645.
- Chomsky, N. (1964a) "The Logical Basis of Linguistic Theory," in H. Lunt, ed., *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., 1962*, Mouton, The Hague.

- Chomsky, N. (1964b) *Current Issues in Linguistic Theory*. Mouton, The Hague.
- Chomsky, N. (1973) "Conditions on Transformations," in S. R. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York.
- Chomsky, N. (1976) "Conditions on Rules of Grammar," *Linguistic Analysis* 2, 303-351.
- Chomsky, N. (1977) "On Wh-Movement," in P.W. Culicover, T. Wasaw, and A. Akmajian, eds., *Formal Syntax*, Academic Press, New York.
- Chomsky, N. (1980) "On Binding," *Linguistic Inquiry* 11, 1-46.
- Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.
- Chomsky, N. (1986) *Barriers*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (1991) "Some Notes on Economy of Derivation and Representation," in R. Freidin, ed., *Principles and Parameters in Comparative Grammar*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (1995) *The Minimalist Program*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. (2000) "Minimalist Inquiries," in R. Martin, D. Michaels, and J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, MIT Press, Cambridge, Massachusetts.
- Chomsky, N. and H. Lasnik (1977) "Filters and Control," *Linguistic Inquiry* 8, 425-504.
- Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT.
- Kayne, R. (1981) "ECP Extensions," *Linguistic Inquiry* 12, 93-133
- Koster, J. (1978a) *Locality Principles in Syntax*, Foris Publications,

Dordrecht, The Netherlands.

Koster, J. (1978b) "Conditions, Empty Nodes, and Markedness,"  
*Linguistic Inquiry* 9, 551-593.

Langacker, R. W. (1969) "Pronominalization and the Chain of Command,"  
in D. Reibel and S. Schane, eds., *Modern Studies in English*,  
Prentice-Hall, Englewood Cliffs, New Jersey.

Rizzi, L. (1978) "Violations of the *Wh*-Island Constraint in Italian and  
the Subjacency Condition," in C. Dubuisson, D. Lightfoot, and Y.  
C. Morin, eds., *Montreal Working Papers in Linguistics* 11.

Ross, J. R. (1967) *Constraints on Variables in Syntax*, Doctoral  
dissertation, MIT.

## 要 旨

### 極小主義以前の変形生成文法理論における制約の研究史

吉本 靖

チョムスキーの極小主義プログラムが世に問われて以来、変形生成文法の研究の傾向にも大きな変動があった。極小主義以前盛んに研究されていた島の制約に代表されるような統語的制約については、極小理論では目覚ましい展開がこれまでなかったと言っていいであろう。そこで、極小主義の枠組みでそれらの制約を考える際の参考に供するため、本研究ノートでは極小主義以前の理論で統語的制約がどのように研究され発展してきたのかを概観してみた。

要約すると以下のとおりである。Chomsky (1964) は変形操作を制約する原理の存在を主張し、それは後に A-over-A 原理と呼ばれるようになった。A-over-A 原理に触発され Ross (1967) はその後の研究に重大な影響を与える制約群（ロスの制約）を提唱した。その成果を受け Chomsky (1973) は下接の条件を提案し、ロスの制約のいくつかが統合可能であることを示した。その後 Chomsky (1981) によって空範疇原理 (ECP) が提案され、移動の結果生じる痕跡に関わる重要な制約として認識されるに至った。Kayne (1981) は ECP を修正し下接の条件の大部分を ECP に包含することを提案したが、Huang (1982) らの批判を受け、下接の条件は存続することになった。Huang (1982) は抽出領域条件 (CED) を提案したが、CED と下接の条件の間に見られる重複性の問題が課題として残った。その重複性を克服するために Chomsky (1986) は、下接の条件における境界節点の定義を環境に依存するものに変更し、障壁という概念が誕生した。これにより下接の条件が CED を包含することになった。その結果、極小主義移行直前のチョムスキー派の変形生成文法理論においては、下接の条件や ECP が移動にかかる制約として残され、それらが束縛原理や  $\theta$ -基準等と「共謀」して移動操作を拘束するという状況を呈していた。