ICOT Technical Memorandum: TM-0813

TM-0813

Islands in Syntax

by

K. Hashida, T. Gunji, H. Nakajima,
M. Ikeuchi, A. Ishikawa & Y. Harada

October, 1989

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

Islands in Syntax

Proceedings of the Symposium on
Island Constraints

edited by
Gunji, Takao
Osaka University

preface by
Heizo Nakajima
Tokyo Metropolitan University

contributions by

Masayuki Ike-uchi
Joetsu University of Education

Akira Ishikawa
Sophia University

Harada, Yasunari
Waseda University

Hasida, Kôiti
Institute for New Generation Computer Technology

Table of Contents

Preface

Heizo Nakajima
Tokyo Metropolitan University

I am pleased at being conceded the privilege of writing the preface of this special issue, as the organizer of the Unbounded Dependency Workshop held at Tokyo Metropolitan University, May 31, 1987. This issue involves the papers which were presented in the afternoon symposium session under the title of "On Island Constraints." (The proceedings of the papers read in the morning colloquium sessions are available as *Metropolitan Linguistics*, Vol. 7 (Tokyo Metropolitan University).)

The symposium chose, as its topic, the problem of so-called island constraints. The study of island constraints has a long history in the generative research of language, and has constantly stimulated wide and deep interests among many people. The study in this area has, furthermore, served to promote the development of various current linguistic theories, such as the transformational grammar, Lexical Functional Grammar (LFG), or Generalized Phrase Structure Grammar (GPSG). In these theories, the necessity of island constraints has been motivated almost in the same way: the rule or process which deals with the unbounded dependency must be formulated in such a general way that it guarantees, on one hand, that the distance between a displaced element and its gap may be potentially unlimited, and on the other, that any sort of major constituent in a sentence can become a displaced element; but such a general way of formulation gives rise to the problem of overgeneration; then, some kinds of "island constraints" are needed to rule out the overgenerated ill-formed sentences. These lines of reasoning are essentially the same among the distinct linguistic theories. Then, each theory has energetically attempted to incorporate island constraints somehow, and has proposed numerous ideas to do so. Thus, the island problem may be regarded as one of the most intriguing "interfaces" where the concerns of the different theories converge, and many people with different background may actively participate in the discussion.

The symposium was organized to discuss the island problem from the four distinctive views — GB-theory (i.e., the recent version of the transformational grammar), LFG, GPSG, and Cognitive Model. The panelists, Masayuki Ike-uchi, Akira Ishikawa, Yasunari Harada, and Kôiti Hasida, each of whom took charge of one of these theories, had been assigned, in advance, the several common tasks, that is, explanations of the ways, for instance, how to create the gap-binder relation, how to ensure the one-to-one correspondence between a gap and its binder, or how to guarantee the unboundedness of the relation. However, the panelists had been requested to use the common examples sentences when they would make the explanation of the island effects. All these were efforts to make it easier to compare the features of the theories and to clarify the differences and similarities among them. I hope these efforts have succeeded to some extent, and have brought some gains to the participants.

The accomplishment of the symposium is due to the cooperation of many people. I would like to thank Takao Gunji, who chaired the symposium and negotiated the publication of this volume with ICOT, to whom I am grateful for offering the opportunity for the publication. My thanks also go to the four panelists mentioned above. I owe acknowledgment to the more than 100 people who earnestly participated in such an ad hoc conference, and to the students at Tokyo Metropolitan University who extended much energy in doing the dirty work. The workshop was made possible by the Grant-in-Aid for Special Project Research from the Education Ministry of Japan under Grant No. 62210014.

Introduction

Gunji, Takao
Osaka University

The workshop held at Tokyo Metropolitan University on May 31, 1987 was unique in the sense that the focus was on a single issue: *unbounded dependency*. The existence of the phenomena of unbounded dependencies has been perhaps the strongest motivation for the need of a grammatical theory with the descriptive power more than that of the classical phrase structural grammar. As it turned out, the movement transformation has been, and *is*, one of the most important devices in the grammatical apparatuses in transformational grammar. Recent nontransformational grammars must also have systematic ways of treating unbounded dependencies, even though they take the form quite different from moving a constituent around.

In fact, one of the earliest papers written in the framework of nontransformational grammar was entitled "Unbounded dependencies and coordinate structure" (Gazdar 1981). Even though the mechanisms proposed in this paper, now known as *Generalized Phrase Structure Grammar*, assumed no transformation, the fact that unbounded dependency was one of the key phenomena to be treated in the new framework exemplifies the important status of the phenomenon.

In the afternoon symposium of the workshop, on Island Constraints, the speakers presented several possible ways of treating unbounded dependencies and focused on the discussion of how to avoid overgeneration, *i.e.*, how to constrain the behavior of general mechanism of establishing unbounded dependency. As the term suggests, this kind of dependency is *unbounded* in the sense that it can in principle relate any two constituents separated by a large number of sentence boundaries. For example, in the following topicalized sentence or relative clause, the sentence-initial topic or the relative pronoun is related to the gap in an embedded sentence:

(1) a. Who do you think [$_S$ that John said [$_S$ he saw _]]?

b. the person who Mary believes [$_S$ that John loves _]

However, there have been known to exist several "islands" at which the dependency is somehow blocked. Thus, a theory of syntax must not only assure the unboundedness of the dependency but also block the dependency at certain designated points. This is where syntactic theories start to get complicated; in fact, the nature of such constraints is one of the key issues in many recent grammatical theories.

Concerning unbounded dependencies, each speaker was asked to explain the following facts:

(2) a. how to generate the displaced element (filler) and the corresponding gap

b. how to assure the one-to-one correspondence between the filler and the gap

c. how to assure the unboundedness of the dependency

d. how to explain the island constraints

In the case of topicalized sentences, the filler is the sentence-initial topic, while it is the relative pronoun in the case of relative clauses. Transformational grammar usually generate the filler in the d-structure at the place where a gap (namely, a *trace*) will appear in the s-structure after moving the filler to the displaced position. On the other hand, in nontransformational theories, the filler and the gap are generated *in situ*.

In this respect, establishing one-to-one correspondence is a relatively straightforward task in the transformational theory, since a gap is only generated after moving a constituent. Nontransformational theories can in principle allow overgeneration of gaps and displaced elements and thus somehow block ill-formed sentences. The papers by Ishikawa and Harada describe the respective mechanisms for properly propagating the information on gaps in a constituent structure.

The unbounded nature of the dependency is achieved in both transformational and nontransformational theories by placing virtually no constraints on the mechanism to establish one-to-one correspondence between the filler and the gap. This, however, leads to a problem because of the existence of the effects of *island constraints*. For example, the following sentences show the effect of island constraints and the speakers were asked to explain them in their respective frameworks:

(3) a. *(I wonder) which book John met a child who read _.

b. *Which book did you wonder who bought _?

c. *Who did a story about _ surprise you?

The first is known as the *Complex NP constraint*. In short, you cannot have a gap in a clause that modifies a noun phrase (a relative clause, an appositive clause, etc.). The second is called the *Wh-island condition*; you cannot have a gap in a clause led by a *wh*-phrase. The third constraint, the *Subject condition*, prohibits you to have a gap in a sentential subject. These constraints have been known for more than twenty years, since the appearance of Ross (1967). In traditional generative grammar, however, they have long been stipulated as independent constraints which must be added to the grammar. The recent trend in generative grammar, however, is to assume as few stipulations as possible. Instead, grammatical phenomena are explained by interactions between several general principles assumed in the grammar, on the one hand, and the nature of the particular lexical items involved in the construction, on the other.

In the symposium, four such approaches were selected and their respective attempts were compared with one another: the so-called *Government-Binding theory* (or the *theory of principles and parameters*, as it is recently called), *Lexical-Functional Grammar*, *phrase structure grammar*, and a theory of grammar based on *cognitive model*.

In the framework of the recent transformational theory, Ike-uchi's explanation of the effects of the island constrains is rather straightforward: the movement transformation ("Move α") cannot cross more than a single *barrier*. This general statement is not specific to a particular construction, and in fact covers all the above cases, since all of the above sentences can be shown to have at least two barriers between the gap and the filler. Of course, the effectiveness of the argument owes much to how you define a *barrier* in the grammar and this in fact is not a simple task in this grammar. The paper included in the current proceedings elaborate on this issue and suggests a modification to the presently prevalent definitions.

In the framework of *Lexical-Functional Grammar*, the unbounded (or long-distance) dependency is described by a chain of grammatical functions—known as a *functional annotation*—attached to a node in the phrase structure (constituent structure). This in effect expresses the path in the functional structure between the gap and the filler. Thus, in Ishikawa's approach, the island constraints are straightforwardly expressed as constraints on such chains. Whether the above three constraints could be merged to a more general constraint on functional annotations seems to be an interesting topic of future research in LFG.

The version of phrase structure grammar presented by Harada assumes a certain number of local constraints for particular types of local (minimally branching) tree. Unlike transformational grammar, the constraints are associated with particular types of local constructions (complementation, specification, etc.), rather than with somewhat global operations (movement, etc.). In a sense, apparently independent island constraints are incorporated as well-formedness conditions on possible branching of phrase structure trees. Thus, one advantage of this approach is that one needs to see only a limited local domain even if the phenomenon to be explained is apparently unbounded.

One of the limitations, however, of the above approach is that it cannot eliminate a certain type of ill-formed sentence, one type of *Complex NP* construction involving *that* such as the following:

(4) a. *(I wonder) who John believes the rumor that Mary loved _.

b. *Who do you believe the claim that she loves _?

Even though whether this is a defect of the theory or not is an interesting topic of its own right, Hasida presents a completely different approach—based on *cognitive model*—rather than proposing a revision of the syntactic theory. In this model, certain grammatical constructions and the "rules" (in a loose sense) responsible for them are deemed as inherently difficult for a child to acquire because of the burden on processing. Thus, independently of the purely syntactic constraints for the *Complex NP constraint*, certain constructions of this form are never learned by a child due to cognitive difficulty. Even though Hasida's argument is intended to complement Harada's, it may open a way to explain away other constraints which are considered purely syntactic in nature.

At the symposium, the presentations of the speakers were followed by a discussion session, which contained many interesting questions and comments from the floor. Regrettably, none of those discussions are included in the current proceedings, although, hopefully, the current versions of the papers included in the proceedings reflect many of the suggestions given at the occasion.

References

Gazdar, G. 1981. "Unbounded dependencies and coordinate structure." *Linguistic Inquiry*, **12**, 155–184.

Ross, J.R. 1967. *Constraints on Variables in Syntax*. Ph.D. dissertation, Massachusetts Institute of Technology, published as *Infinite Syntax!* by Ablex, Norwood, N.J., 1986.

New Barriers and Islands[*]

Masayuki Ike-uchi
Joetsu University of Education

In his lectures given in Tokyo and Kyoto in 1987, Noam Chomsky proposed a new theory of barriers which attempts to unify the theory of government and the theory of bounding. The main aim of this paper is to examine his new theory, to clarify the issues, and to point out some problems. Before entering into the main discussion, let us briefly summarize the proposed definitions of various notions and principles which are relevant to it.[1]

Let us begin with the notion of barrier. There are two types of barriers. One is the inherent barrier and the other is the barrier involving the Minimality Condition which is called the minimality barrier here. The inherent barrier is defined as follows:

(1) α is an inherent barrier for β iff α is not H-marked or L-marked.

(2) a. α H-marks β iff β is the complement of α.

b. α L-marks β iff α is lexical and β is an immediate constituent of γ, γ is the complement of α.

Notice that we have dropped the notion of inheritance of barrierhood.

The minimality barrier for government is defined as:

(3) α is a barrier for β iff α (improperly) includes γ, δ, where $\gamma = \gamma^{max}$, $\delta = \delta^{0}$ and β is included in γ and the complement of δ.[2]

The minimality barrier for bounding is defined as:

---

[*] I wish to thank Heizo Nakajima and Takeru Suzuki for reading an earlier version of this paper. I am also grateful to Jeffrey Jones, who corrected stylistic errors. This work was supported in part by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Grant No. 61710262 and No. 62510249.

A slightly different version of this paper appeared in *The Bulletin of Joetsu University of Education*, Vol. 7 (1988) under the title of "Notes on New Barriers."

[1] For a more informal exposition of the new theory of barriers, see Ike-uchi (1987). See also Nakajima (1987).

[2] In order to block the extraction of the complement of N from NP (shown in (i)), it is required that β be "improperly" included in the complement of δ:

(i) $\ldots [_{NP} \ldots [_{N'}\ N\ t]]$

However, β must not be "improperly" included both in γ and in the complement of δ, since the extraction of the complement of N over N′ would be ruled out if it were permitted. See (ii). This is due to Yasuaki Abe.

(ii) $\ldots [_{NP} \ldots [_{N'}\ N\ t]]$

(4) $[\alpha, \gamma]$ is a barrier for β iff α (improperly) includes γ, δ, where $\gamma = \gamma^{\max}$, $\delta = \delta^0$ and β is included in γ and the complement of δ.

It should be stressed here that both the minimality barrier for government and that for bounding are defined referring to one and the same configuration.

Using these barriers, we can define the notion of government (5), the ECP (8), and the Subjacency (9):

(5) government
α governs β iff (i) α c-commands β and
(ii) there is no γ, γ a barrier for β, γ excludes α.

(6) c-command
α c-commands β iff (i) α does not cover β and
(ii) for all γ, if γ covers $\alpha^{\max}$, then γ covers/includes β.

(7) cover
α covers β iff β is properly included in a segment of α.

The barrier γ in (5ii) includes both the inherent barrier (1) and the minimality barrier (3).

We now have the ECP defined as (8):

(8) ECP
Chain links must satisfy antecedent government.

Notice that the ECP is now a condition on chains and that proper government is reduced to antecedent government, dropping θ-government.

(9) Subjacency
β is n-subjacent to σ iff there are n barriers for β that c-commands σ.[3]

Now the Subjacency Condition stipulates that β must be less than 2-subjacent to σ. This means essentially that movement cannot cross more than a single barrier. The relevant barrier (in principle) includes both the inherent barrier (1)[4] and the minimality barrier (4).

We have one more government condition for traces:

(10) L-government
Traces must be L-governed.

In the following sections I will sometimes touch on other relevant assumptions where necessary.

[3] $[\alpha, \gamma]$ c-commands β iff α c-commands β.

[4] Assuming slight modification of (9), since there need be no c-command requirement for the inherent barrier.

1 Some Problems with the Minimality Barrier $[\alpha, \gamma]$ System

In this section I will examine the minimality barrier $[\alpha, \gamma]$ system and will point out some problems (and suggest a slight modification).

1.1 The minimality barrier $[\alpha, \gamma]$ system, as it stands, incorrectly rules out completely acceptable sentences like (11):[5]

(11) a. Who$_i$ do you $[_{VP_1}$ t_i^2 $[_{VP_2}$ think $[_{CP}$ t_i^1 that John saw $t_i]]]$

b. Who$_i$ did you $[_{VP_1}$ t_i^2 $[_{VP_2}$ see $[_{NP}$ t_i^1 a picture $t_i]]]$

(11a) is an example where extraction is from an embedded complement sentence and (11b) is an instance where it is from an object NP phrase. According to the $[\alpha, \gamma]$ system, there are two barriers [VP_2, VP_2] and [VP_2, CP] in (11a) and [VP_2, VP_2] and [VP_2, NP] in (11b). Thus, both (11a) and (11b) involve subjacency violations. This is not a desirable result.

There will be several possible modifications which could avoid this result. I will tentatively suggest that the "improper" inclusion in the definition (4) be replaced by the "proper" inclusion. The "proper" inclusion version of the minimality barrier $[\alpha, \gamma]$ system allows only one barrier ([VP_2, CP] and [VP_2, NP], respectively) in both cases and thus correctly predicts their grammaticality. In the subsequent discussion, I will adopt this version. See §5.[6]

1.2 The $[\alpha, \gamma]$ system correctly predicts the status of *wh*-island violation like (12):

(12) What did you wonder to whom John gave?

There exist two offending barriers [VP_2, CP] and [VP_2, IP]:

(13) What$_i$ did you $[_{VP_1}$ t_i^2 $[_{VP_2}$ wonder $[_{CP}$ to whom$_j$ $[_{C'}$ ϕ $[_{IP}$ John t_i^1 gave t_i $t_j]]]]]$

This system, however, cannot explain the difference in grammaticality between (12) and (14):

(14) Which book did the students forget who wrote?

Grimshaw 1986: 365

It is well-known that *wh*-island violations are weakest when extraction occurs from indirect questions where the subject is *wh*-questioned. According to the system in question, *wh*-movement crosses two barriers in (14) and thus it should be no different from (12) above.[7]

---

[5] Cf. Nakajima (1987).

[6] Accordingly, an alternative treatment must be given to the *wh*-island violation in Topicalization like (i):

(i) *This book, I wondered to whom I gave.

[7] See also a case of extraction from the relative clause in which the subject is relativized:

1.3 Chomsky assumes that indefinite determiners such as *a*, ϕ, and *some* do not occupy the specifier position in NP and thus that position is vacant in such NP's. Hence, *wh*-phrases moved by Move-α can stop by there and then move upward. On the contrary, definite determiners such as *the*, *that*, and genitive NP's do occupy the NP specifier position and thus *wh*-phrases cannot move upward via that position. These assumptions predict that there will be differences in grammaticality between (15a) and (15b) and between (16a) and (16b):

(15) a. Which book did John hear a rumor that you had read?

b. Which race did John believe the/Tom's claim that Mary had won?

(16) a. Which book did John meet a child who read?

b. Which book did John meet that child who read?

The structures to be examined are the following:

(17) a. Which book$_i$ did John t_i^4 hear $[_{NP}\ t_i^3\ [_{N'}\ \text{a}\ [_{N}\ \text{rumor}][_{CP}\ t_i^2\ [_{C'}\ \text{that}\ [_{IP}\ \text{you had}\ t_i^1\ \text{read}\ t_i]]]]]]$

b. Which race$_i$ did John $[_{VP_1}\ t_i^3\ [_{VP_2}\ \text{believe}\ [_{NP}\ \text{the/Tom's}\ [_{N'}\ \text{claim}\ [_{CP}\ t_i^2\ [_{C'}\ \text{that}\ [_{IP}\ \text{Mary had}\ t_i^1\ \text{won}\ t_i]]]]]]]$

(18) a. Which book$_i$ did John t_i^3 meet $[_{NP}\ t_i^2\ [_{N'}\ [_{N'}\ \text{a child}][_{CP}\ \text{who}_j\ [_{C'}\ \phi\ [_{IP}\ t_j\ [_{I'}\ \text{I}\ [_{VP}\ t_i^1\ [_{VP}\ \text{read}\ t_i]]]]]]]]$

b. Which book$_i$ did John $[_{VP_1}\ t_i^2\ [_{VP_2}\ \text{meet}\ [_{NP}\ \text{that}\ [_{N'}\ [_{N'}\ \text{child}]\ [_{CP}\ \text{who}_j\ [_{C'}\ \phi\ [_{IP}\ t_j\ [_{I'}\ \text{I}\ [_{VP}\ t_i^1\ [_{VP}\ \text{read}\ t_i]]]]]]]]]]]$

According to the minimality barrier $[\alpha, \gamma]$ system, there are two barriers $[VP_2, NP]$ and $[VP_2, CP]$ in (17b), but there is only one barrier $[N', CP]$ in (17a).[8] Similarly, in (18b) three barriers $[VP_2, NP]$, $[VP_2, CP]$ and $[VP_2, IP]$ are crossed, but in (18a) only two barriers $[N', CP]$ and $[N', IP]$ are. This will mean that in both (15) and (16) b. is less acceptable (i.e., involves a more serious subjacency violation) than a. Then the question is whether this is factually correct.[9]

2 On Inherent Barriers

2.1 Redundancy

Let us first point out the redundant nature of inherent barriers in certain cases.

Consider the extraction of *wh*-phrases from the subject NP or the adverbial clause:

(i) That's one trick that I've known a lot of people who've been taken in by.

Chung and McCloskey 1983: 708

[8] Note that it is in itself a problem that Move-*wh* crosses just one barrier in (17a), since (15a) is not perfectly grammatical. In other words, it is not as grammatical as (11a), where there is only one offending barrier.

[9] This problem was also (partly) pointed out by M. Saito, if I remember correctly.

See, for example, Ross 1967 and Chomsky 1975: 90, 1981: 49 and 1986: 34–36.

(19) *The man who pictures of were on the table came in.

(20) *To whom did John leave before you spoke?

They have the following structures, respectively (irrelevant details omitted):

(21) The man [$_{CP}$ who$_i$ [$_{C'}$ ϕ [$_{IP}$ [$_{NP}$ t_i^1 [$_{N'}$ pictures t_i]] were on the table]]] came in.

(22) To whom$_i$ did John [$_{VP_1}$ t_i^3 [$_{VP_2}$ [$_{V'}$ leave][$_{PP}$ [$_{P}$ before][$_{CP}$ t_i^2 [$_{C'}$ ϕ [$_{IP}$ you t_i^1 spoke t_i]]]]]]

In both cases *wh*-movement crosses two barriers ([C′, IP], [C′, NP] in (21) and [VP_2, PP], [VP_2, CP] in (22)) according to the [α, γ] system. Notice that there is one more relevant barrier. Thus NP in (21) and PP in (22) are both inherent barriers as well. Consequently, *wh*-movement crosses three barriers in all in these cases. Crossing three barriers should by definition yield a severer subjacency violation than crossing two barriers. The problem, then, is whether this is factually correct in these examples. If it is not, then it follows that inherent barriers are redundant here.

Furthermore, consider the extraction of *how* from the relative clause or the adverbial clause which involves the ECP:

(23) *How did John meet a man who fixed the car?

(24) *How did you leave before he fixed the car?

Examine their respective structures:

(25) How$_i$ did John t_i^3 meet [$_{NP}$ t_i^2 [$_{N'_1}$ [$_{N'}$ a man][$_{CP}$ who$_j$ [$_{C'}$ ϕ [$_{IP}$ t_j [$_{I'}$ I [$_{VP}$ t_i^1 [$_{VP}$ fix the car t_i]]]]]]]]]

(26) How$_i$ did you [$_{VP}$ t_i^2 [$_{VP}$ [$_{V'}$ leave][$_{PP}$ [$_{P}$ before][$_{CP}$ t_i^1 [$_{C'}$ ϕ [$_{IP}$ he t_i fixed the car t_i]]]]]]

The ECP correctly excludes both (23) and (24), since N'_1 in (25) and PP in (26) can be the minimality barriers which block the antecedent government of t_i^1 by t_i^2.[10] Now notice that there also exists an inherent barrier between t_i^2 and t_i^1, CP in (25) and PP in (26). In the ECP violation cases, different from the subjacency violations above, these inherent barriers are truly redundant, since a single barrier, whatever it may be, is enough for the ECP to function. The existence of another barrier is by definition meaningless.[11]

In sum, inherent barriers are redundant and unnecessary at least in certain cases.

---

[10] At the final stage of derivation which I will ignore here, (25) becomes ungrammatical because of the improper chain. See §6.

[11] See, however, §6 and note 24.

2.2 Necessity

Let us next consider the contrary arguments which evidence the necessity of inherent barriers.

First, consider (27):

(27) *How is it time for John to fix the car?

In order to explain the ungrammaticality of (27) by the ECP, there must be at least one barrier which blocks antecedent government. Examine the structure of (27) in (28):[12]

(28) How$_i$ it [$_{VP}$ t$_i^3$ [$_{VP}$ is time [$_{CP}$ t$_i^2$ [$_{C'}$ for [$_{IP}$ John to t$_i^1$ fix the car t$_i$]]]]]

Notice that there is no minimality barrier intervening t$_i^3$ and t$_i^2$. CP, however, constitutes a barrier blocking antecedent government as an inherent barrier. Here, then, the inherent barrier is not redundant and is necessary as an independent barrier.

Next consider (29):

(29) a. ?What did you figure out how to finance?

b. *What did you figure it out how to finance?

Culicover and Wilkins 1984: 161.

If Culicover and Wilkins' judgement is correct, then (29) can be regarded as motivating the necessity of inherent barriers. As predicted by the minimality barrier [α, γ] system, in (29a) *wh*-movement crosses two barriers [VP$_2$, CP] and [VP$_2$, IP] and thus causes a subjacency violation.[13] See (30a):

(30) a. What$_i$ did you [$_{VP_1}$ t$_i^2$ [$_{VP_2}$ [$_V$ figure out][$_{CP}$ how$_j$ [$_{C'}$ ϕ [$_{IP}$ PRO [$_{I'}$ to [$_{VP}$ t$_i^1$ [$_{VP}$ [$_{V'}$ finance t$_i$] t$_j$]]]]]]]]

b. What$_i$ did you [$_{VP_1}$ t$_i^2$ [$_{VP_2}$ [$_{VP}$ figure it out][$_{CP}$ how$_j$ [$_{C'}$ ϕ [$_{IP}$ PRO [$_{I'}$ to [$_{VP}$ t$_i^1$ [$_{VP}$ [$_{V'}$ finance t$_i$] t$_j$]]]]]]]]

Notice that (29b) is far worse than (29a) although *what* crosses the same number of barriers (i.e., [VP$_2$, CP] and [VP$_2$, IP]) as shown in (30b). We can attribute this difference in grammaticality to the existence (or absence) of an inherent barrier. Thus, in b. CP is in an adjunct position and hence serves as an inherent barrier as

[12] I believe that the structure given in (28) is essentially correct for the *it is time for NP to VP* construction. If it has a structure like (i) (mentioned by Chomsky) where CP is the complement of *time*, then there arises a problem in the case of the movement of the argument *wh*-phrases:

(i) It [$_{VP}$ is [$_{NP}$ [$_{N'}$ time [$_{CP}$ for John to fix the car]]]]

Notice that a *wh*-phrase in the position of *the car* crosses two barriers, [VP, NP] and [VP, CP], yielding a subjacency violation contrary to fact:

(ii) What is it time for John to fix?

But see §2.3 for possible inheritance of barrierhood by a lower segment in (28).

[13] But for some not exactly known, but familiar reasons, (29a) is not so bad. We will not enter into the details here.

well. In a., however, the CP in question is a complement of the verb and hence does not constitute an inherent barrier.[14]

It should be noticed that the system which has only the minimality barrier available cannot capture the difference between complements and non-complements with respect to barrierhood. As is already exemplified above, it takes any maximal projection that fits the defined configuration as γ in the definition (4) whether it is a complement or a non-complement. Thus it cannot in principle distinguish, for example, objects and complement sentences from subjects and adjuncts. It is, fundamentally, for this reason that we also need the notion of inherent barrier which can by definition capture the (non-)complementhood.[15]

2.3 On the Necessity of Inheritance of Barrierhood

Let us now turn to the phenomena which suggest that a barrier by inheritance in the sense of (31) is still necessary contrary to Chomsky's proposal to abolish it:

(31) α is an inherited barrier for β iff α immediately dominates an inherent barrier for β.

First, consider (32):

(32) ??This is the company of which every new idea frightens the president.

Following Belletti and Rizzi 1986, we can assign a structure like (33) to (32):[16]

(33) This is the company [$_{CP}$ of which$_i$ [$_{C'}$ ϕ [$_{IP}$ every new idea$_j$ [$_{I'}$ I [$_{VP_1}$ t_i^2 [$_{VP_2}$ [$_{V'}$ frighten t_j][$_{NP}$ t_i^1 the president t_i]]]]]]]

Notice that there are no minimality [α, γ] barriers here and that there is only one inherent barrier, i.e., NP. In order to explain the unacceptability of (32) by the Subjacency Condition, we need at least two barriers intervening between t_i^2 and t_i^1. Thus we will have to assume that VP_2 inherits barrierhood from the inherent barrier NP, constituting a second barrier. This assumption yields a desired result.

Next consider (34):

(34) *Who did John see in the gallery a picture of?

Nakajima 1984: 46.

This is assigned a structure like (35a):

(35) a. Who$_i$ did John [$_{VP_1}$ t_i^2 [$_{VP_2}$ [$_{VP}$ see t_j in the gallery][$_{NP_j}$ t_i^1 a picture t_i]]]

Once again there are no minimality barriers. In order to make VP_2 a second barrier which is necessary for the Subjacency Condition, we must again resort to the inheritance of barrierhood from the inherent barrier NP_j, assuming that VP_2 is a lower segment of the maximal projection.

[14] In fact, the situation is a bit more complicated. We will return to the problem in §2.3

[15] For Incorporation phenomena, see Baker 1985.

[16] We tentatively assume that *of which* moves upward via the NP's specifier position, avoiding the ECP violation.

This argument crucially depends on the ordering of the application of rules: i.e., Complex NP Shift–*wh*-movement, whether the Subjacency Condition is a condition on movement or on the resulting representation.

Suppose, alternatively, that *wh*-movement is applied prior to Complex NP Shift and that the Subjacency is a condition on the representation. Then the structure to be examined is not (35a) but (35b) below:

(35) b. Who$_i$ did John $[_{VP}\ [_{VP_1}\ t_i{}^2\ [_{VP_2}$ see t_j in the gallery$]][_{NP_j}\ t_i{}^1$ a picture $t_i]]$

This time the inheritance of barrierhood from NP_j by VP_2 is impossible. And in fact there can be no effective inheritance of barrierhood, although NP_j is again an inherent barrier, occupying an adjunct position. Thus there is only one offending barrier. This yields an incorrect result.

Suppose, now, that *wh*-movement is applied before Complex NP Shift and that the Subjacency is a condition on the application of rules. Then there exist no offending barriers when *wh*-movement is applied. Thus (34) is again predicted to be grammatical.

It can thus be concluded that the ordering must be: Complex NP Shift–*wh*-movement in order to explain the ungrammaticality of (34) in terms of the present system.[17]

3 On the Notion of the "Best" Derivation in Syntax

Chomsky has presented one crucial assumption that "in syntax you always have to pick up the best derivation." Suppose that this assumption means that we have to pick up the derivation where there are as few syntactic (e.g., subjacency) violations as possible, or no syntactic violations, if possible.

---

[17]A similar remark holds in the case of (30b) above, where the rules involved are extraposition and *wh*-movement.

It might be argued that one possible solution is to assume that VP also constitutes a cycle.

Notice that the argument here presupposes that H-marking is done not at the level of D-structure, but at a later level, e.g., S-structure. Suppose now that H-marking is done at D-structure and that the category will be marked, say, [+H] when it is H-marked at that level. And suppose also that the [+H] marked category does not constitute an inherent barrier. These assumptions lead to an undesirable result, since neither interpretation of the Subjacency Condition gives any necessary offending inherent barriers under either ordering of the rules in question.

There is another class of possibly relevant phenomena which involve psych-verbs (and, for that matter, passive constructions). Consider, for example, a sentence like (i):

(i) This is the plan $[_{CP}$ on which$_i$ $[_{C'}\ \phi\ [_{IP}\ [_{NP_j}\ t_i$ a newly proposed idea $t_i][_{I'}$ I $[_{VP}\ [_{V'}$ frightened $t_j][_{NP}$ the president of the company$]]]]]]$

If NP_j were H-marked as [+H] at D-structure and thus did not constitute an inherent barrier, then *wh*-movement would cross only two minimality barriers [C′, IP] and [C′, NP_j] in (i). But recall that extraction from the subject NP like (21) above crosses (at least) three barriers including one inherent NP barrier. Then the question is whether there is a difference in acceptability between (i) and (21). I myself doubt that there is.

To sum up, the assumption is wrong that H-marking is carried out at D-structure.

In this connection, see also Chomsky 1986: 26.

I have benefited from the comments of Takeru Suzuki.

If IP adjunction is permitted, then we must assume inheritance of barrierhood to exclude the extraction of a *wh*-phrase from the subject NP (as is suggested by Chomsky) and Topicalization from it (as is pointed out by Belletti and Rizzi 1986: 45).

Assuming that his is correct, let us consider two possible problems which it brings about. First, observe (36):

(36) *What did John arrive yesterday $[_{AP}$ sad about$]$?

Huang 1982: 487.

Note that AP hangs from IP. If *what* takes a usual derivation process, moving upward as shown in (37), then it crosses two minimality barriers, [C′, IP] and [C′, AP] and an inherent barrier, AP. Thus this correctly yields a subjacency violation.

(37) $[_{CP}$ what$_i$ $[_{C'}$ did $[_{IP}$ John $[_{VP}$ arrive yesterday$][_{AP}$ sad $[_{PP}$ t_i^1 $[_{PP}$ about $t_i]]]]]]$

Alternatively, suppose that we choose a different syntactic derivation where *what* first adjoins to PP, then adjoins to VP, and finally moves to the specifier position of the matrix clause:

(38) $[_{CP}$ what$_i$ $[_{C'}$ did $[_{IP}$ John $[_{VP_1}$ t_i^2 $[_{VP_2}$ arrive yesterday$]][_{AP}$ sad $[_{PP}$ t_i^1 $[_{PP}$ about $t_i]]]]]]$

Notice that the second *wh*-movement is a kind of lowering process. It should be noted that this second *wh*-movement crosses one inherent barrier, AP, and the final movement, one minimality barrier, [C′ IP]. Thus this derivation yields no subjacency violation. If it is a possible syntactic derivation and thus is the best derivation in syntax with no syntactic violation, then (36) should be grammatical. How can we exclude a derivation like (38)? One possible solution is to check the resultant chain (what$_i$, t_i^2, t_i^1, t_i). In this chain t_i^2 does not c-command t_i^1. We could therefore claim that the chain in question is excluded for this reason and thus a derivation like (38) is not possible, and hence (36) is ungrammatical. It should be emphasized, however, that we must check a chain at the level of S-structure, not at the level of LF, since it is assumed that both t_i^2 and t_i^1 as intermediate $\bar{A}$-traces of an argument are deleted at LF to get operators to have operator-variable constructions. It seems that there are no other phenomena which require that a chain-checking like the above be done at S-structure. If this perspective is correct, then the solution under consideration requiring a chain-checking at S-structure is dubious. If so, there will be no way to block a derivation like (38).

Next consider NP-movement. It is usually assumed that NP-movement in (39) is blocked by the ECP, since t_i is not antecedent-governed:

(39) *John$_i$ seems that it appears t_i to be intelligent.

Notice, however, that the NP-movement in question violates the Subjacency Condition as well. Thus if we must pick up the best derivation in syntax and it is defined as above, then a derivation like (39) is not permitted in syntax. Hence the possible derivation which can be regarded as the best derivation must involve a process of VP-adjunction and movement to the specifier position in CP, and thus will yield a derivation like (40):

(40) $[_{IP}$ John$_i$ $[_{I'}$ I $[_{VP}$ t_i^3 $[_{VP}$ seem $[_{CP}$ t_i^2 $[_{C'}$ that $[_{IP}$ it $[_{I'}$ I $[_{VP}$ t_i^1 $[_{VP}$ appear $[_{IP}$ t_i to be intelligent$]]]]]]]]]]]$

All the four NP-traces satisfy antecedent government requirement. Thus, we must attribute the ungrammaticality of (39) not to the ECP, but to the violation of Condition C of the binding theory by t_i, assuming that the binding theory applies at LF (as is also assumed in Chomsky 1986).[18] Here again the notion of the best derivation is crucially involved.

4 L-Government

Chomsky has proposed that the sentence in (41) is ungrammatical due to the L-government condition on traces:

(41) *How fix the car John can?

This is illustrated by the following derivation where VP fronting is applied before *wh*-movement of *how*:

(42) a. $[_{CP}\ [_{C'}\ \phi\ [_{IP}$ John can $[_{VP}$ fix the car how]]]]

b. $[_{CP}\ [_{C'}\ \phi\ [_{IP}\ [_{VP}$ fix the car how]$[_{IP}$ John can t_{VP}]]]]

c. $[_{CP}$ how$_i$ $[_{C'}\ \phi\ [_{IP}\ [_{VP}\ \underset{[-\lambda]}{t_i{}^1}\ [_{VP}$ fix the car t_i]]$[_{IP}$ John can t_{VP}]]]]

Notice that $t_i{}^1$, which is $[-\lambda]$ marked, does not satisfy the L-government requirement. This yields a correct result.

It should be noted, however, that an alternative derivation is possible which ends in avoiding the above result. Suppose that *wh*-movement is applied before VP fronting:

(43) a. $[_{CP}\ [_{C'}\ \phi\ [_{IP}$ John can $[_{VP}$ fix the car how]]]]

b. $[_{CP}$ how$_i$ $[_{C'}\ \phi\ [_{IP}$ John can $[_{VP}\ \underset{[+\lambda]}{t_i{}^1}\ [_{VP}$ fix the car t_i]]]]]

c. $[_{CP}$ how$_i$ $[_{C'}\ \phi\ [_{IP}\ [_{VP}\ \underset{[+\lambda]}{t_i{}^1}\ [_{VP}$ fix the car t_i]]$[_{IP}$ John can t_{VP}]]]]

Notice that $t_i{}^1$ is L-governed and thus $[+\lambda]$ marked in b. after *wh*-movement is applied.[19] A later application of VP fronting does not affect the L-government of $t_i{}^1$. Hence the proposed system predicts that (41) should be grammatical.[20]

---

[18] Note that if the binding theory applies at LF, then Condition C cannot explain the status of (39) in the *Barriers* system—contrary to the claim made in Chomsky 1986: 22 and n. 20—, since $t_i{}^1$ may be deleted at LF after having licensed t_i at S-structure in that system.

[19] We assume that $t_i{}^1$ is L-governed in a configuration like (43b), following Chomsky.

[20] It seems that we could not cope with this problem by resorting to the notion of strict cycle, since it is usually assumed that only CP, not IP, constitutes an independent cycle in the case at issue.

5 On the Minimality Barrier for Government and Lower Segments in Adjunction Structure

Chomsky has suggested the possibility that lower segments in adjunction structure serve as minimality barriers for government and block antecedent government.

Consider, for instance, (44) which is an illustration of the effect of the Left Branch Condition:

(44) *How many$_i$ did you [$_{VP_1}$ t_i^1 [$_{VP_2}$ read [$_{NP}$ t_i books]]]

If VP_2 is a minimality barrier which blocks government,[21] then (44) is ruled out by the ECP, since t_i is not antecedent-governed by t_i^1. Similarly, (45):

(45) *How$_i$ John [$_{VP_1}$ t_i^1 [$_{VP_2}$ is [$_{AP}$ t_i tall]]] (= *How is John tall?)

Furthermore, (46) and (47b) can also be treated in a similar manner:[22]

(46) *From which city$_i$ did you [$_{VP_1}$ t_i^1 [$_{VP_2}$ meet [$_{NP}$ the man t_i]]]

Chomsky 1986: 80.

(47) a. John predicted [$_{NP}$ a crisis on account of widespread unemployment]

b. *On account of what$_i$/Why$_i$ did John [$_{VP_1}$ t_i^2 [$_{VP_2}$ predict [$_{NP}$ t_i^1 [$_{N'}$ a crisis] t_i]]]

Although these examples are in favor of the assumption that lower segments are minimality barriers, the following are against this assumption:

(48) How well$_i$ the meat [$_{VP_1}$ t_i^1 [$_{VP_2}$ is [$_{AP}$ cooked t_i]]]

Chomsky 1986: 79.

(49) How$_i$ do you [$_{VP_1}$ t_i^3 [$_{VP_2}$ want [$_{CP}$ t_i^2 [$_{IP}$ to t_i^1 fix the car t_i]]]]

In these cases, VP_2 must not be counted as a barrier blocking antecedent government.[23]

---

[21] Though VP_2 does not exclude t_i^1.

[22] If V′ exists in the structures at issue, then V′ instead of VP_2 could be taken as a barrier. See Chomsky 1986: 4 and 47.

[23] V′ must not appear here. See note 22.

See also (35a) as an instance of rightward movement. If the structure derived by Complex NP Shift is (roughly speaking) (i), then a lower segment VP* must not be taken as a minimality barrier which blocks antecedent government:

(i) ... [$_{VP}$ [$_{VP^*}$ [$_{V'}$ V t_j] PP] NP_j]

Cf. note 2.

6 On the Inadequacy of LF-VP Fronting

Chomsky assumes that LF-VP fronting preposes VP (VP_1, VP_2) and adjoins it to the matrix IP in (50) and (51), finally yielding an improper chain:

(50) *[$_{CP}$ How$_i$ [$_{C'}$ did [$_{IP}$ John [$_{I'}$ I [$_{VP}$ t_i^3 [$_{VP}$ meet [$_{NP}$ t_i^2 [$_{N'}$ [$_{N'}$ a man][$_{CP}$ who$_j$ [$_{C'}$ ϕ [$_{IP}$ t_j [$_{I'}$ I [$_{VP_1}$ [t_i^1 [$_{VP_2}$ fixed the car t_i]]]]]]]]]]]]]]]

(51) *[$_{CP}$ How$_i$ [$_{C'}$ do [$_{IP}$ you [$_{I'}$ I [$_{VP}$ t_i^2 [$_{VP}$ wonder [$_{CP}$ who$_j$ [$_{C'}$ ϕ [$_{IP}$ t_j [$_{I'}$ will [$_{VP_1}$ t_i^1 [$_{VP_2}$ fix the car t_i]]]]]]]]]]]]

However, if the VP in question is adjoined to the embedded IP—this possibility is allowed in the system, then neither an ECP violation is yielded nor an improper chain is produced. Examine (52) and (53) in which VP is adjoined to the embedded IP:

(52) [$_{CP}$ how$_i$ [$_{C'}$ did [$_{IP}$ John [$_{I'}$ I [$_{VP}$ t_i^3 [$_{VP}$ meet [$_{NP}$ t_i^2 [$_{N'}$ [$_{N'}$ a man][$_{CP}$ who$_j$ [$_{C'}$ ϕ [$_{IP}$ [$_{VP_1}$ t_i^1 [$_{VP_2}$ fixed the car t_i]][$_{IP}$ t_j [$_{I'}$ I [$_{VP}$ *e*]]]]]]]]]]]]]]

(53) [$_{CP}$ how$_i$ [$_{C'}$ do [$_{IP}$ you [$_{I'}$ I [$_{VP}$ t_i^2 [$_{VP}$ wonder [$_{CP}$ who$_j$ [$_{C'}$ ϕ [$_{IP}$ [$_{VP_1}$ t_i^1 [$_{VP_2}$ fix the car t_i]][$_{IP}$ t_j [$_{I'}$ will [$_{VP}$ *e*]]]]]]]]]]]]

Neither CP (or C′) nor N′ is now a minimality barrier, since t_i^1 (β) is not included in the complement of ϕ (δ). Thus there is no ECP violation.[24] Notice that chains are proper in both of these cases. Hence we have no explanation for the ungrammaticality of (50) and (51).[25] [26]

References

Baker, M. 1985. *Incorporation: A Theory of Grammatical Function Changing.* Ph.D. dissertation, Massachusetts Institute of Technology.

Belletti, A and L. Rizzi. 1986. "Psych-Verbs and Th-theory." Unpublished paper. MIT-Scuola Normale Superiore di Pisa and MIT-Université de Genève.

Chomsky, N. 1975. *Reflections on Language.* New York: Pantheon.

________. 1981. "Principles and Parameters in Syntactic Theory." *Explanation in Linguistics*, ed. by N. Hornstein and D. Lightfoot, pp. 32–75. London: Longman.

________. 1986. *Barriers.* Cambridge, Mass.: The MIT Press.

[24] But CP as a relative clause in (52) can be an inherent barrier. See §2.1.

[25] I was able to think of this argument thanks to T. Suzuki.

[26] Notice that LF-VP fronting refers to a full category VP which consists of the two segments VP_1 and VP_2, preposing the whole VP. (See also §4.) Full categories are also referred to in the definition of cover and c-command. Then why can't we refer to a full category consisting of two segments in the case of θ-role assignment? Chomsky assumes that an argument category must not have an adjunction structure, i.e., must not be covered by a bigger category, because if it does (or is), then it is invisible at LF and thus cannot be assigned a θ-role. As far as I can see, there is no independent reason why a full category cannot be referred to and assigned a θ-role at LF.

Chung, S. and J. McCloskey. 1983. "On the Interpretation of Certain Island facts in GPSG." *Linguistic Inquiry*, **14**.4, 704–13.

Grimshaw, J. 1986. "Subjacency and the S/S' Parameter." *Linguistic Inquiry*, **17**.2, 364–69.

Huang, C.-T.J. 1982. *Logical Relations in Chinese and the Theory of Grammar.* Ph.D. dissertation, Massachusetts Institute of Technology.

Ike-uchi, M. 1987. "Chomsky Riron no Genzai (1), (2), (3) (On Chomsky's Current Theory (1), (2), (3))." *The English Teachers' Magazine*, **36**, nos. 7, 9, 10. Tokyo: Taishukan.

Nakajima, H. 1984. *Eigo no Ido Gensho Kenkyu (A Study of Movement Phenomena in English).* Tokyo: Kenkyusha.

_________. 1987. "Barrier riron no Shintenbo (A New Perspective on the Theory of Barriers—A summary of Chomsky's Lectures in Japan)." *The Rising Generation*, **133**, nos. 2, 3. Tokyo: Kenkyusha.

Ross, J. 1967. *Constraints on Variables in Syntax.* Ph.D. dissertation, Massachusetts Institute of Technology.

Received December 25, 1987

Long-Distance Dependencies and Their Functional Constraints

Akira Ishikawa
Sophia University

1 Introduction

Long-distance dependencies are known to follow various constraints such as those proposed by Ross (1967). It has been one of the major goals of generative syntax to reduce such constraints to more fundamental principles. In this paper, I will discuss one possible approach to this problem.

The approach in this paper comes from the framework of Lexical Functional Grammar (LFG). I will first introduce a device recently proposed in Kaplan and Zaenen (1987) to handle the phenomenon of long-distance dependencies in LFG. This device marks a clear departure from the former LFG account of this phenomenon. The former account made crucial use of c-structure information, such as the syntactic categories of gapped elements. Kaplan and Zaenen have revealed that c-structure information is unnecessary if the expressive power of functional annotations is enhanced. This idea is in line with the recent developments of LFG, which have been termed extensions of the interpretation of functional equations. The main purpose of this paper is to show how the new extension should be modified to capture some of the well-known constraints. It will be shown that this functional approach reveals a universal regulating principle closely associated with a deterministic parsing mechanism.

2 Control equations for long-distance dependencies

Kaplan and Zaenen (1987) proposed an extension of the interpretation of functional equations in order to capture long-distance dependencies without using double-arrow meta-variables ($\Uparrow, \Downarrow$). Double-arrow meta-variables were regarded as indispensable in Kaplan and Bresnan (1982) in order to explain the matching of the syntactic category of the displaced element with that of its gap. The following pairs were cited as an example calling for the information of the syntactic categories of correlated elements.

(1) The girl wondered how tall she would grow.
*The girl wondered how tall she would reach.

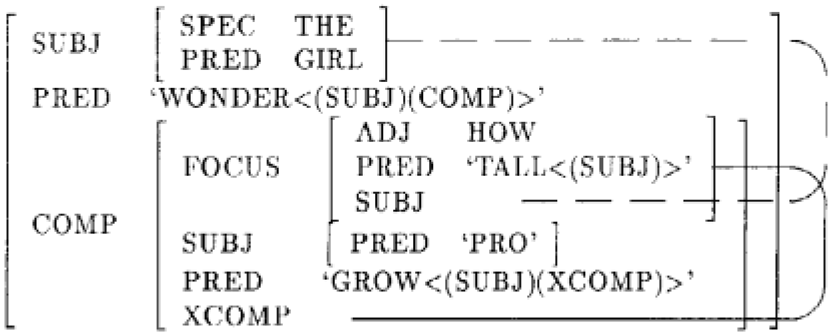
(2) *The girl wondered what height she would grow.
The girl wondered what height she would reach. (p.241)

The dependencies between *how tall* and *grow*, on the one hand, and *what height* and *reach*, on the other, as shown in (1) and (2) can be described by associating the following functional equations with *how tall* and *what height*, respectively.

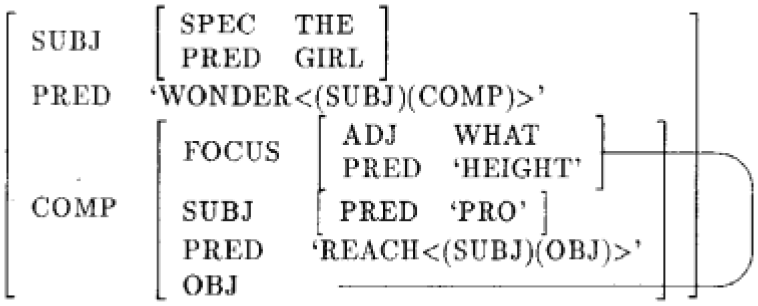
(3) a. (↑ FOCUS) = (↑ XCOMP)
 b. (↑ FOCUS) = (↑ OBJ)

These equations represent the control relations involved as the following f-structures show.

(4) a.



b.



The solid lines indicate the control relations induced by the equations in (3). The dotted line in (4a) comes from the functional-control equation (↑ SUBJ) = (↑ XCOMP SUBJ), which becomes equivalent to (↑ SUBJ) = (↑ FOCUS SUBJ) through the equation (↑ FOCUS) = (↑ XCOMP) (cf. Bresnan (1982)).

The use of double-arrow meta-variables becomes necessary when there are more intervening elements.

(5) The girl wondered how tall John had said that Bill would predict that she would grow.

In (5), the control relation between *how tall* and the gapped element would need the following equation, which violates the principle of functional locality prohibiting more than two symbols (excluding the leftmost arrow) from occurring on either side of a functional equation, thus limiting the expressive power of functional equations.

(6) (↑ FOCUS) = (↑ COMP COMP COMP XCOMP)

The principle of functional locality embodies the idea that the overall structure of a sentence is derived from the integration of the locally consistent sub-structures of its constituents. This restriction of the expressive power of functional equations made it necessary to transmit such a long-distance control relation as in (6) in piecemeal fashion, i.e. from one node to another. Such a transmission was effected in the c-structure using equations involving double-arrow meta-variables subscripted with the syntactic categories of correlated elements (cf.Kaplan and Bresnan (1982)).

Taking the other tack, Kaplan and Zaenen (1987) extended the expressive power of functional equations without infringing the functional locality principle. They augmented the expressive power by allowing regular language notation within functional equations. In the case of long-distance dependencies, the notations for closure and disjunction play a crucial role. The closure notation collapses an indefinite number of occurrences of the same symbol in one complex symbol.

(7) a. Mary John telephoned yesterday.
(↑ TOPIC) = (↑ OBJ)

b. Mary John said that Bill telephoned yesterday.
(↑ TOPIC) = (↑ COMP OBJ)

c. Mary John claimed that Bill said that Mike telephoned yesterday.
(↑ TOPIC) = (↑ COMP COMP OBJ)

The three sentences in (7) all involve a long-distance dependency between TOPIC and OBJ. But the path linking the two grammatical functions can get indefinitely stretched through the intervention of COMP. By using the closure notation, we can specify this indefinite number of occurrences of COMP on the path.

(8) (↑ TOPIC) = (↑ COMP* OBJ)

When there are more than one type of grammatical function on the path, the alternative is indicated by disjunction.

(9) a. Mary John made it clear that Bill said that Mike telephoned.

b. (↑ TOPIC) = (↑ XCOMP COMP COMP OBJ)

c. (↑ TOPIC) = (↑ {XCOMP, COMP}* OBJ)

(10) a. Mary John claimed that Bill made it clear that Mike telephoned.

b. (↑ TOPIC) = (↑ COMP XCOMP COMP OBJ)

c. (↑ TOPIC) = (↑ {XCOMP, COMP}* OBJ)

In (9) and (10), the paths contain both COMP and XCOMP in different combinations indicated in (9b) and (10b). The use of disjunction together with closure can collapse these two paths into one as indicated in (9c) and (10c).

A parallel extension is carried out on the side of the interpretation of functional equations.

(11) $(f\ a) = v$ holds iff $((f\ s)\ \mathit{Suff}(s,a)) = v$ for some symbol s, where $\mathit{Suff}(s,a)$ is the set of suffix strings y such that $sy \in a$.

In (11), the argument a denotes a set of strings in a regular language.

On the basis of these extensions, Kaplan and Zaenen represent the general schema for constructions involving long-distance dependencies.

$$
\begin{array}{lll}
(12)\ S' \rightarrow & \Omega & \Sigma \\
 & (\uparrow \text{DF}) = \downarrow & \\
 & (\uparrow \text{DF}) = (\downarrow \textit{ body bottom}) &
\end{array}
$$

DF indicates a discourse function such as TOPIC and FOCUS (cf. Bresnan and Mchombo (1987)). The second equation containing *body* and *bottom* acts as the control equation for long-distance dependencies. *Body* corresponds to the path linking DF and the controlled grammatical function, which is what *bottom* corresponds to. An instance of this schema is the rule for English.

$$
\begin{array}{ll}
(13)\ S' \rightarrow & \text{XP or S} \\
 & (\uparrow \text{DF}) = \downarrow \\
 & (\uparrow \text{DF}) = (\downarrow \{\text{XCOMP, COMP}\}^{*}\ (\text{GF-COMP}))
\end{array}
$$

The body for English specifies that the path must be a chain made up of an indefinite number of XCOMP or COMP. On the other hand, the bottom can be any grammatical function except for COMP.

Kaplan and Zaenen used the rule in (13) to explain, among other things, the contrast of grammaticality in (14).

(14) a. Kevin persuaded Roger that these hamburgers were worth buying.

b. *That these hamburgers were worth buying, Kevin persuaded Roger.

c. Louise told me that Denny was mean to her.

d. That Deny was mean to her Louise told me (already).

The difference in the grammaticality of the topicalized versions fall quite naturally out of the interaction of rule (13) and the lexical forms of *persuade* and *tell*. Besides the lexical form involving COMP as in (14c), *tell* has another lexical form taking OBJ instead of COMP, whereas *persuade* does not have this alternative.

(15) a. *Kevin persuaded Roger the news.

b. Louise told me the story.

It should be noted that the grammaticality of (14d) could not be accounted for without the separation of the level of syntactic categories from that of grammatical functions.

3 How to Implement the Constraints on Long-distance Dependencies

It is apparent that the schema in (13) does not cover all the possibilities of long-distance dependencies in English. Moreover, the task still remains of motivating the choice of the grammatical functions that can occur in the body or the bottom in the schema.

Let us first consider the problem of extending the coverage of the schema. It can easily be noticed that the present formulation does not cover the case of preposition stranding in English.

(16) a. *Which country* did John schedule *a visit to*?
FOCUS OBL

b. *What subject* did Mary show *some interest in*?
FOCUS OBL

(17) (↑ FOCUS) = (↑ OBJ OBL OBJ)

The two sentences in (16) require the control equation in (17) which is not covered by the schema (↑ DF) = (↑ {XCOMP, COMP}* (GF-COMP)). This is due to the fact that the current schema takes only the subcategorization of verbs into account. Rappaport (1983) pointed out that derived nominals do not allow for semantically unrestricted GFs as their arguments. In other words, SUBJ, OBJ, and OBJ2 cannot be arguments for derived nominals. Another well-known characteristic of derived nominals is the fact that their arguments are optional. It is not unreasonable to suppose that these characteristics may lead to different behavior of long-distance dependencies. As it stands, the schema cannot deal even with non-stranding cases.

(18) a. *To which country* did John schedule *a visit*?
TOPIC OBJ

b. *In what subject* did Mary show *some interest*?
TOPIC OBJ

(19) (↑ TOPIC) = (↑ OBJ OBL)

These sentences might be taken to suggest that *body* should also be instantiated by GFs other than XCOMP and COMP in English. Indeed, besides OBJ there are a number of GFs which appear on the path, or the chain of GFs leading up to the bottom.

(20) a. *What* was everyone surprised *by the news about*?
FOCUS OBL_{ag}

b. (↑ FOCUS) = (↑ OBL_{ag} OBL_{about} OBJ)

(21) a. *Who* did John show Mary *a picture of*?
FOCUS OBJ2

b. (↑ FOCUS) = (↑ OBJ2 OBL_{of} OBJ)

(22) a. *What subject* did They go *to a lecture on*?
FOCUS OBL_{go}

b. (↑ FOCUS) = (↑ OBL_{go} OBL_{on} OBJ)

Not only subcategorizable GFs, but also ADJ my appear in the position in question.

(23) a. *What* did John open the bottle *with*?
FOCUS ADJ

b. (↑ FOCUS) = (↑ ADJ OBJ)

However, it is also clear that we should not indiscriminately extend the range of GFs allowed for in this position. One of the generalizations made by Kaplan and Zaenen is the impossibility of extraction out of clauses bearing the ADJ function. In our present terms, ADJ cannot occur on a path when the constituent bearing it is clausal.

(24) *Which picture did they all blush *when they saw*?
ADJ

This generalization is attractive since the complex NP constraint becomes its particular case. This is because in LFG every non-argument constituent including a modifier of a noun bears the ADJ function.

(25) a. *Which picture did they deny the fact *that they saw*?
ADJ

b. *Which picture did they count the boys *who saw*?
ADJ

If the range of GFs for *body* is simply expanded to include ADJ, this generalization cannot be retained. The difference between ADJ in (23) and that in (24) and (25) seems to come from the nature of the constituent bearing the function.

There is another phenomenon which suggests that COMP and XCOMP are the only GFs appearing on the path which can correspond to clausal constituents. The phenomenon is known as the Sentential Subject Condition, which prohibits extraction out of a sentential subject.

(26) *What would *for me to give up* be a pity?
SUBJ

When the subject is not sentential, extraction is not totally impossible.

(27) a. *To which cities* were *visits* most popular?
FOCUS SUBJ

b. He is the person *of whom pictures* are on the table.
TOPIC SUBJ

(Chomsky 1986, p.32)

In (27), the displaced elements can be taken to control the OBL arguments of the underlined SUBJs.

In view of these facts, we can distinguish between two types of paths. One type of paths involve clausal constituents, whereas the other type of paths do not. In LFG terms, clausal constituents correspond to those nuclei that contain SUBJ. The above facts indicate that SUBJ and ADJ can appear on the path only when their f-structures do not contain SUBJ. In other words, GFs corresponding to clausal constituents can appear on the path only when they are COMP or XCOMP in English. For ease of reference, let us call GFs which correspond to clausal constituents *clausal GFs*, and those which do not *non-clausal GFs*.

As we have seen above, both clausal and non-clausal GFs may constitute paths in English, although clausal GFs are restricted to COMP and XCOMP. There seems to be another restriction on the possible GFs forming paths.

(28) a. ?? *Which room* did John keep [the promise [to clean]]?
FOCUS OBJ COMP

b. ??? *Who* did Mary follow [the instruction [to meet]]?
FOCUS OBJ COMP

c. (↑ FOCUS) = (↑ OBJ COMP OBJ)

The control relation involved in (28a,b) is shown in (28c). Compared to examples (16,20 to 22), (28a)and (28b) are marginal although the length of the paths in terms of the number of GFs is the same. The only difference is the occurrence of COMP in (28c). We can explain the marginality of (28a,b) if we hypothesize that *clausal GFs must precede non-clausal GFs on the path.*

The two restrictions we have considered suggest that the schema in (13) be replaced by the one in (29).

(29) (↑ DF) = (↑ {XCOMP, COMP}* (GF-COMP)$^+$)

The only apparent alteration is the addition of the + closure symbol to the bottom of the schema. The interpretation by means of suffix strings in (11) remains the same. But this alteration holds some implications as we will see below.

First, a path no longer has the two-way distinction between *body* and *bottom* as in (12). Rather, it can be broken into two sub-paths, the first of which is a chain of clausal GFs and the second one of non-clausal GFs.

(30) (↑ DF) = (↑ *clausal-path non-clausal-path*)

Second, the exclusion of SUBJ and ADJ from the set of clausal GFs suggest the possibility of introducing a universal principle regulating the path. What the schema in (30) represents is a situation in which a greater variety of GFs are allowed to form paths than in (12). With an adequate general characterization of GFs, we might even be able to drop any restriction on such GFs. On the other hand, the impossibility of having SUBJ and ADJ as path-forming clausal GFs in English and their corresponding restriction to COMP and XCOMP seem to indicate a more general principle for path-formation.

Let us consider what this principle will be like so that the schema in (29) might capture the constraints on long-distance dependencies correctly. First of all, we have to take heed of the fact that the GF of a constituent in a sentence is often undecided or uncertain (as Kaplan and Zaenen aptly termed) before the whole sentence is parsed. Long-distance dependencies are a typical case of functional uncertainty. In our present formulation, this uncertainty is captured by the newly introduced complex symbols representing all the possible combinations of GFs for forming paths. The actual combination for a particular path is progressively determined as the sentence is parsed from left to right. We do not go into the details of this determination process. But it is not unreasonable to suppose that such successive stages of determination of the GFs forming the path is *synchronized* with the identification of the constituents of the sentence. For example, consider (7c).

(7c) Mary John claimed that Bill said that Mike telephoned.
x

When the parsing process has advanced to the point marked by x, we have already identified the main-clause SUBJ (John) and detected the presence of COMP, of which *that Bill* is the initial part. By *synchronization*, we mean that the control schema instantiating (29) is activated at this point to register the newly detected GF as part of the path. If this is the case, we will have a partially instantiated equation of the form as in (31).

(31) (↑ TOPIC) = (↑ COMP ...)

The dots indicate the still undecided part of the equation. Under this hypothesis, the *partial determinations* of the path-forming GFs occur at certain synchronization points. Then, what would happen if we did not have the Sentential Subject Condition? If sentential subjects can contain gapped elements, we would have to expect a partial determination to occur when the presence of a sentential subject is detected just as in the case of (7c). But the situation would be quite different. For now partial determinations sometimes fail because there are sentences which contain a sentential subject (or clausal SUBJ) and a COMP.

(32) *That John did not show up at the party* meant *that he had quarreled with*
SUBJ COMP
Mary again.

By contrast, there are no lexical forms which take COMP and XCOMP as arguments at the same time. Moreover, the syntactic position of these two GFs guarantees that a partial determination involving either of them never occurs until all the other sister GFs have been tested for the gapped element. As the result, partial determinations involving COMP and XCOMP are free from *undoing*. In the case of clausal SUBJs, their syntactic position does not guarantee this.

(33) *Who* would *that John did not show up* surprise?
FOCUS SUBJ x

In (33), a partial determination would induce the following partial equation.

(34) (↑ FOCUS) = (↑ SUBJ ...)

But this partial determination has to be undone because the gapped element is not contained within the SUBJ.

From these considerations, we can hypothesize that the constraints on long-distance dependencies have to do with making the parsing process as free from the undoing of partial determinations as possible. In other words, they serve to make the parsing process as *deterministic* as possible.

What of non-clausal GFs? They seem to have a different mode of synchronization. In the case of a clausal GF, the detection of its presence is sufficient for assuming the presence of the controlled element (bottom) somewhere within. On the other hand, non-clausal GFs do not allow this move. This is because non-clausal GFs forms a nucleus only when its predicate is argument-taking. Moreover, unlike clausal GFs, non-clausal GFs do not allow their ADJs to be controlled in a long-distance manner.

(35) a. John received [*frequent* visits] from the ghost.
OBJ ADJ

b. *How frequent did John receive visits from the ghost?

(36) a. John remembers [the visit of condolence *after the fire*].
OBJ ADJ

b. *When does John remember the visit of condolence?

So, it seems that the synchronization of a partial determination with the identification of a constituent occurs after the identification of an argument-taking predicate in the case of non-clausal GFs. Even then, we cannot avoid the possibility of undoing partial determinations without some restrictions. But the undoing of partial determinations is far less costly compared with the case of clausal GFs because of the relative number of elements involved. Moreover, we can keep the cost of undoing to a minimum, depending on the mechanism of synchronization. For example, we can imagine that a partial determination occurs not directly after the identification of an argument-taking predicate, but after the identification of its argument. This provision makes partial determinations free from undoing in the cases (16), (18), (20) through (23). It is only when an argument-taking predicate occurs within another argument-taking predicate that we have the possibility of undoing.

(37) Who did John make [the schedule [of a visit [to the city]]] with?
OBJ OBL OBL

Thus, we have speculated on the nature of a universal principle regulating the long-distance schema. As we have seen, the fundamental theme is to avoid the undoing of partial determinations, or to achieve deterministic processing of long-distance dependencies. In the case of clausal GFs, this avoidance is fully complied with, whereas it is only partially, though in fact almost fully, complied with in the case of non-clausal GFs. It may be worthwhile to note here that the *Wh-island Constraint* follows directly from this speculation.

(38) *Who did Mary wonder *what John broke with*?
COMP

A wh-island violation like (38) involves a COMP which introduces a second control equation. Within such a COMP, we could have two controlled elements looking for their controllers. If the two control equations were active inside the same COMP, we would have the possibility of undoing a partial determination as to a clausal GF, a clear violation of our speculated principle.

(39) When did Mary wonder *what John broke frequently*?
COMP

So, we have to conclude that partial determination should not occur after the detection of a COMP which introduces a second control equation. In other words the first control equation is not active within such a COMP. This situation is to be contrasted with a case like (40), where the first control equation is temporarily suspended.

(40) a. I wonder *which violin* *the sonata* is tough for her to play _ on _.

b. *I wonder *which sonata* *the violin* is tough for her to play _ on _.

In (40), the first control equation is suspended within the second one which is introduced by tough. The unacceptability of (40b) can be explained by the unavailability of the first control equation until the second one is fully instantiated. It is clear that (38) should become acceptable if the first control equation is simply suspended instead of not being active.

(38') *Who* did Mary wonder *what* John broke _ with _?

It is also noteworthy that the order between clausal and non-clausal GFs is a direct consequence of the different modes of synchronization. That is, COMP and XCOMP trigger a partial determination as soon as their presence is detected, whereas non-clausal GFs trigger one only after the identification of their argument. When COMP appears as the argument within a non-clausal GF, it requires a partial determination at the earliest occasion of its detection, which obviously conflicts with the suppression of a partial determination until after its full identification.

4 Conclusion

We have seen that long-distance dependencies can be characterized using control equations in LFG. The original idea of Kaplan and Zaenen's has been shown to be further extendable by considering parsing processes in natural language. We have focused on two aspects of their general control schema. One aspect is concerned with the question of what GFs can instantiate the schema, and the other with their relative order.

It has been argued that the schema should be broken into two sub-paths rather than into body and bottom. We have also suggested that any GF can be part of a path. These two points follow from our hypothesis that parsing processes tend to be as deterministic as possible. If this is a universal principle for parsing, it is straightforward to arrive at the well-known constraints on long-distance dependencies. We have looked at the Complex NP Constraint, the Sentential Subject Condition, and the WH-island Constraint. We have established that these constraints are direct consequences of our hypothesis, i.e. the universal tendency to make parsing as deterministic as possible, and that they are captured by the control schema for long-distance dependencies.

References

Bresnan, J. 1982. "Control and Complementation." *The Mental Representation of*

Grammatical Relations, ed. by J. Bresnan. Cambridge, Mass. The MIT Press.

Bresnan, J. and S. Mchombo. 1987. "Topic, Pronoun, and Agreement in Chichewa." Center for the Study of Language and Information, Stanford University.

Chomsky, N. 1986. *Barriers.* Cambridge, Mass. The MIT Press.

Kaplan, R. and A. Zaenen. (to appear). "Long-distance Dependencies, Constituent Structure, and Functional Uncertainty." *Alternative Conceptions of Phrase Structure*, ed. by M. Baltin and A. Kroch. Chicago: Chicago University Press.

Rappaport, M. 1983. "On the Nature of Derived Nominals." *Papers in Lexical Functional Grammar*, ed. by L. Levin, M. Rappaport, and A. Zaenen. Bloomington: Indiana University Linguistics Club.

Ross, J. R. 1967. *Constraints on Variables in Syntax.* Ph.D. dissertation, Massachusetts Institute of Technology.

Received August 30, 1988

On Island Constraints—A Phrase Structure Grammar Perspective

HARADA Yasunari
Waseda University

In this article we will consider how the so-called "island constraints" are to be incorporated into a phrase structure grammar description of (a fragment of) English Grammar.

"Island Constraints" have long been the central issue of generative transformational studies of English syntax. Although researchers working within unification-based grammars do not necessarily share this interest, we have to consider what a possible 'solution' to this problem within a phrase structure grammar approach to syntax might look like and how it might affect our understanding of the 'adequacy' of our research objective. Thus we will define a very small fragment of English, outline how unbounded dependencies are to be dealt with in that fragment and then go on to discuss how 'island constraints' are to be guaranteed.

0 Constraint-Unification-Based Phrase Structure Grammar

In a constraint-unification-based phrase structure grammar, syntactic regularities are to be stated in terms of local phrase structures and constraints obtaining among feature specifications involved in a given local phrase structure. Here a description of a fragment of a natural language such as English or Japanese is determined if and only if we (1) define what objects constitute grammatical categories in that language, (2) state what phrase structure rules are to be utilized in that language and (3) specify what lexical items belong to what syntactic category or categories. As with other recent grammatical theories, we conceive of grammatical categories as bundles of feature specifications rather than non-decomposable monadic objects. Also, we permit partial specifications of these categories. Thus, phrase structure rules in our fragment will refer to grammatical categories with very few feature specifications, which enables us to efficiently state syntactic regularities.

When embedded into an environment where constraint-unification is executed, our grammar will sanction only the grammatical configurations of (the fragment of) English defined. Partially specified categories that are mentioned in phrase structure rules are 'unified' or matched with more richly specified categories that are assigned to lexical items. In this way, our grammar determines (1) whether a given string of words constitute a grammatical sentence (or some other category) of (our fragment of) English, (2) what are possible strings in the language defined, and (3) what parse trees are to be assigned to these strings.

0.1 Categories and features

As an example of how our description of English might look like, we will give you some typical features together with their intuitive or heuristic 'meanings' and possible ranges of their values. This is of course a very limited subset of features that are needed to adequately describe the grammar of English.

(0.1.1)	pos	part of speech	{n, v, p, a, det}
	pn	person and number	{nil, 1s, 2s, 3s, 1p, 2p, 3p}
	case	case	{nil, nom, poss, acc}
	form	verb form	{nil, base, fin, presp, pastp}
		preposition	{*of, at, on, in, for, to*}
	spec	specifier	list of categories
	subcat	complement	list of categories
	comp	complementizer	list of variables
	sem	semantics	some expression
	gap	syntactic gap	list of categories
	bind	variable	list of variables

In what follows, categories are designated by a left square bracket ("[") followed by an indefinite number of feature specifications separated by commas (",") followed by a right square bracket ("]"). A feature specification is maximally a feature name followed by its value. However, when the value uniquely determines the name, the name can be omitted. Also, when the value is nil or a null-list (< >) the entire feature specification can be omitted. Finally a category of the form [pos *P*, ..., sem *S*] is sometimes designated as *P*[...]:*S*. Some examples follow, together with symbols often used in traditional generative literature. Note that traditional symbols do not necessarily bear as much information as category designations defined here.

(0.1.2) NP [pos n, subcat < >, spec < >, sem **m'**]
n[ ]:**m'**
PP [pos p, subcat < >, spec < >, form *for*]
p[*for*]
VP [pos v, spec <[pos n, subcat nil, sem X]>, form fin, sem **love'**(X,**j'**)]
v[fin, spec <n[ ]:X>]:**love'**(X,**j'**)
P [pos p, subcat <[pos n, subcat < >, case acc, sem **j'**]>, form on]
p[subcat <n[acc]:**j'**>, on]
V [pos v, subcat <[pos n, subcat < >]>, spec<[pos n, subcat < >, pn 3s]>]
v[subcat <n[ ]>, spec<n[3s]>]

0.2 Phrase Structure Rules and Feature Inheritances

Partial specification of grammatical categories enables us to efficiently encapsulate a great part of English syntax into the following two phrase structure rules.

(0.2.1) a. specification

Mr → Sp Hd
where spec @ Hd = <Sp|spec @ Mr>,
subcat @ Mr = subcat @ Hd = nil,
sem @ Mr = sem @ Hd,
head @ Mr = head @ Hd

b. complementation

Mr → Hd Ct
where subcat @ Hd = <Ct|subcat @ Mr>,
spec @ Mr = spec @ Hd,
sem @ Mr = sem @ Hd,
head @ Mr = head @ Hd

Typically these two phrase structure rules sanction the following local phrase structures, respectively. Note that in English, the length of the value of spec cannot be more than 1, but the same does not hold for the length of the value of subcat. However, since an adequate exposition of how double object constructions and control phenomena are to be handled in the framework we have in mind here would take us too far afield, we will restrict our attention to cases where the length of the value of subcat is less than 2.

(0.2.2) a. P[spec < >, ...]:S

X P[spec <X>, ...]:S

b. P[subcat < >, ...]:S

P[subcat <X>, ...]:S X

The constraints or conditions that follow "where" (0.2.1.a,b) are the clauses in which syntactic regularities are stated. These conditions are called feature inheritances. Here an expression of the form "Feature_name @ Category_designator" refers to the value of the feature designated by Feature_name with respect to the grammatical category referred to by Category_designator. The symbol "=" represents unification. However, unification here employed disregards the values of bind. The expression "head @ C1 = head @ C2" is a short hand way of repeating expressions of the form "F @ C1 = F @ C2" with F ranging over head_features, where head_features = {pos, pn, case, form}. Here and in what follows, constants are designated by names beginning with a lowercase letter and variables and meta-variables are designated by names beginning with an uppercase letter.

Let us take a very simple sentence such as "John loves Mary" and see how these phrase structure rules are involved in admitting that sentences like this are grammatical in English.

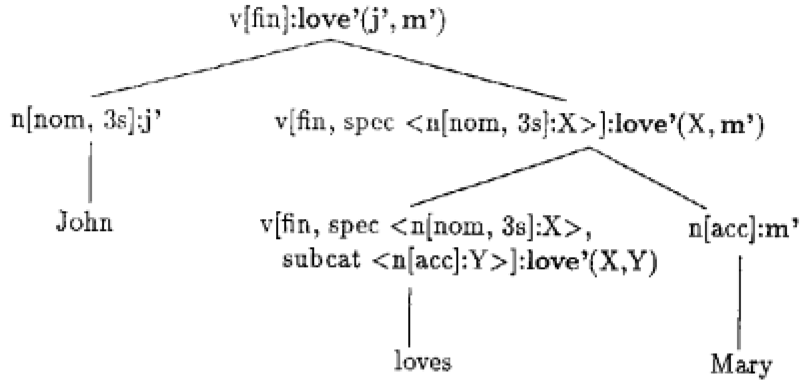
First of all we have to remember that in a phrase structure grammar account of English syntax, all lexical items or words are supposed to be given rich syntactic information in the lexicon, where syntactic and semantic information concerning each and every word is stored. For instance, a transitive verb "love" will be specified in the lexicon as shown in (0.2.3).

(0.2.3) loves ⊨ v[fin, spec <n[nom, 3s]:X>, subcat <n[acc]:Y>]:**love'**(X,Y)

An expression of the form "Spelling ⊨ Category" asserts that a lexical entry whose representation is given as Spelling is assigned feature specifications designated by Category.

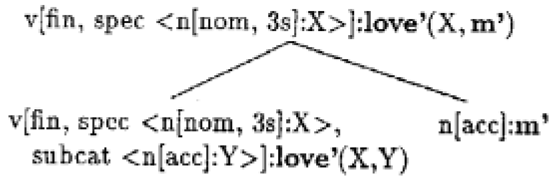
Our two phrase structure rules assigns the following parse tree to the string "John loves Mary".

(0.2.4) John loves Mary.



Let's take a look at the local structure in (0.2.5). The phrase structure rule given in (0.2.1.b), or complementation, sanctions this local structure.

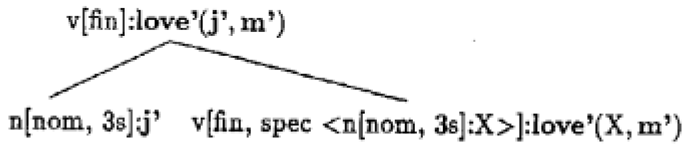
(0.2.5)



If we match the three grammatical categories in (0.2.5) against those in (0.2.1.b), more specifically, v[fin, spec <n[nom, 3s]:X>]:**love'**(X, **m'**) against Mr, v[fin, spec <n[nom, 3s]:X>, subcat <n[acc]:Y>]:**love'**(X,Y) against Hd, and n[acc]:**m'** against Ct, we see that all the conditions that follow "where" are satisfied. Therefore, the three categories in (0.2.5) are said to "unify" with the three categories in (0.2.1.b) and complementation is said to "sanction" this local phrase structure. Incidentally, variable Y in the semantic representation of the transitive verb "love" is bound to or unified with the constant **m'**, which is the semantic representation associated with the proper noun "Mary", as a side-effect of this unification.

Likewise, the 'upper part' of the parse tree in (0.2.4), namely the local structure in (0.2.6) is sanctioned by specification in (0.2.1.a).

(0.2.6)

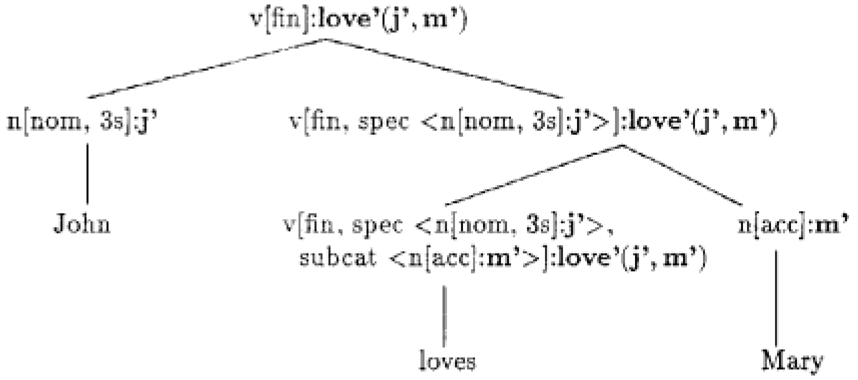


Here, if we match v[fin]:**love'(j', m')** against Mr, n[nom, 3s]:**j'** against Sp, and v[fin, spec <n[nom, 3s]:X>]:**love'**(X, **m'**) against Hd, all constraints are satisfied. As above, variable X in the semantic representation of the verb phrase is bound to **j'**, which is the semantic representation of the subject proper noun. The reason a prepositional phrase cannot be the subject of a transitive verb like "loves" is that its value for pos would be p, thus preventing it from unifying with Sp in the local structure involved. Also, a plural noun phrase cannot be the subject here, because the value of pn with respect to Sp is specified as 3s, rather than 3p.

We have reasons to believe that except for constructions that involve dislocation of elements, only four phrase structure rules, namely complementation, specification, adjunction and coordination are responsible for almost all English sentence constructions. However, we have very little, if any, to say about adjunction or coordination in this article.

Strictly speaking, our constraint-unification executive would not produce parse-trees as shown in (0.2.4) but rather something like (0.2.7).

(0.2.7) John loves Mary.



That is, once unification binds a variable to a constant, all relevant occurrences of the same variable is bound to that constant throughout a given representation. However, if we give parse trees like (0.2.7), the binding process is quite obscured to readers unfamiliar with our approach presupposed here. Therefore, we will talk as if parsing is processed bottom-up, and give parse trees the way (0.2.4) is given, rather than the way (0.2.7) is given.

For brevity, we will omit reference to features such as pn or case in what follows.

1 Unbounded Dependencies in Phrase Structure Grammar

Unbounded dependencies are guaranteed through chains of local constraints in the phrase structure analysis of English we are here considering. The feature gap plays a central role in this, binding syntactically and semantically the dislocated grammatical element and the syntactic gap it binds in an unbounded dependency construction.

Typical examples of unbounded dependencies are sentences such as those in (1.0.1). Here, however, in order not to complicate our grammar to something overly loaded with new lexical items or phrase structure rules that are responsible for Subject Auxiliary Inversion constructions, let us take sentences in (1.0.2). Sentences in (b) and (d) are added because in our grammar, subjects and objects are treated quite differently.

(1.0.2) a. what did you see _

b. what do you think that John saw _

(1.0.3) a. [I wonder] who John loves

b. [I wonder] who loves John

c. [I wonder] who Mary thinks John loves

d. [I wonder] who Mary thinks loves John

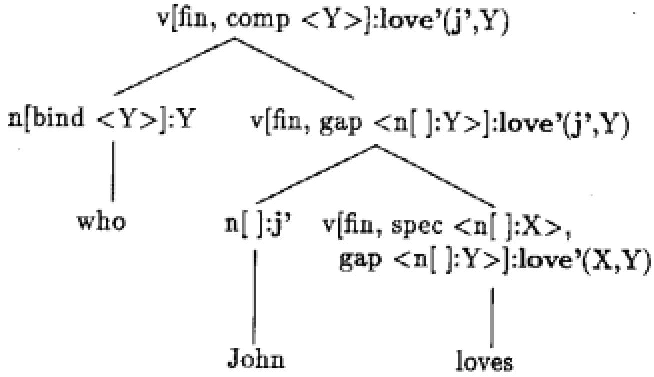
Here, we are presupposing a lexical entry of the following form.

(1.0.4) thinks ⊨ v[fin, spec<n[ ]:X>, subcat <v[fin]:Y>]:**think'**(X,Y)

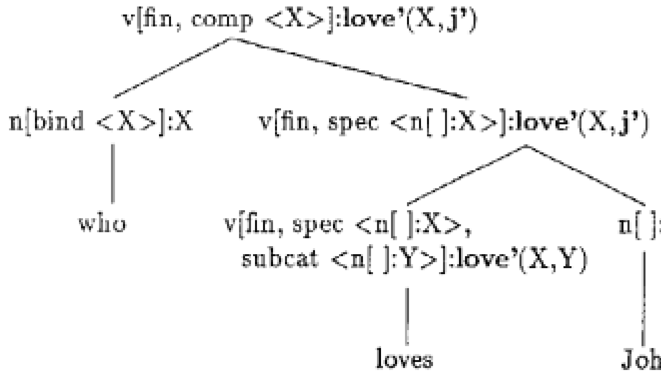
For the (embedded) sentences in (1.0.2) our grammar will assign the following parse trees. Explanations of what these symbols in the diagrams are supposed to mean will follow shortly.

(1.0.4) [I wonder] who John loves

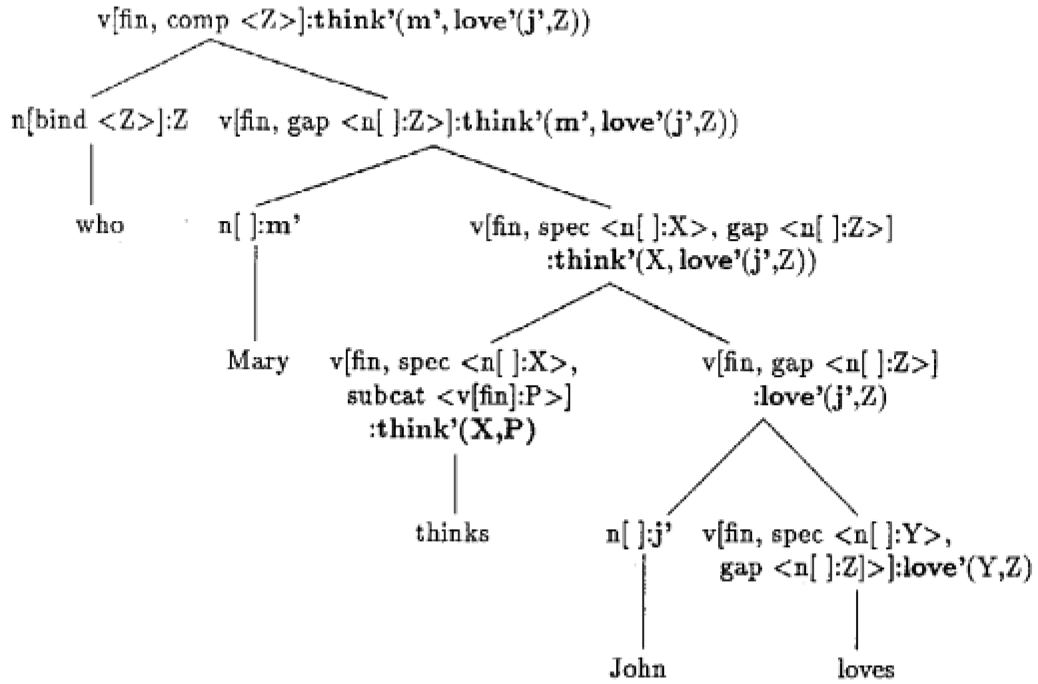
wonder'(speaker,<Y>,love'(j',Y))



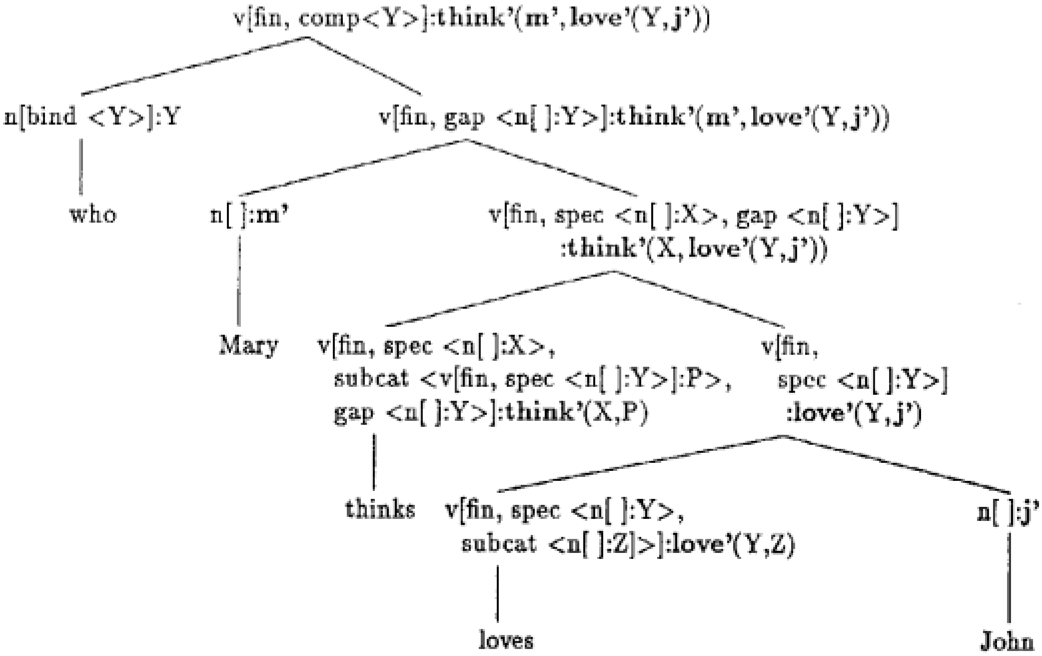
(1.0.5) [I wonder] who loves John
wonder'(speaker,<X>, **love'**(X,**j'**))



(1.0.6) [I wonder] who Mary thinks John loves
wonder'(speaker,<Z>, **think'**(**m'**, **love'**(**j'**,Z)))



(1.0.7) [I wonder] who Mary thinks loves John
wonder'(speaker,<Y>, **think'**(m', **love'**(Y,j')))



1.1 Binding Features

The feature gap takes a list consisting of grammatical categories as its value and propagates information concerning the existence of syntactic gaps through a parse tree. This feature is equivalent, for the most part, to the feature called slash in the previous GPSG and/or HPSG literature. Note incidentally, that binding between semantic representations of dislocated elements and variables that appear in the semantic representations of gaps they bind is also achieved as 'side-effects' of unification chains. The feature bind takes a list of variables as its value and propagates information concerning semantic variables of relative and interrogative constructions. This feature is somewhat similar to the feature wh in previous GPSG studies. These two are called binding features.

(1.1.1) binding_features = {gap, bind}

Binding feature inheritance states local constraints obtaining among values of binding features with respect to grammatical categories involved in a given local phrase structure. In general, the following condition holds in all local structures where binding features are not bound.

(1.1.2) binding feature inheritance

Mr → C1 C2

where F @ Mr = F @ C1 + F @ C2
if F ∈ binding_features

Here "+" designates a list operation not unlike concatenation. For instance, < > = < > + < >, <X> = <X> + < >, <X> = < > + <X> and so on. However, non-distinct elements are merged rather than repeated. Thus, <X> = <X> + <X>, <X|Y> = <X> + <X|Y> and so on. Typically (1.1.2) sanctions local structures of the following forms, among others.

(1.1.3) a.
[gap <X>]
[gap <X>] [gap < >]

b.
[gap <X>]
[gap < >] [gap <X>]

c.
[bind <X>]
[bind <X>] [bind < >]

d.
[bind <X>]
[bind < >] [bind <X>]

Earlier we mentioned that our unification disregards the values of bind. In fact, we have reasons to believe that something like (1.1.4) really captures how bind and gap are inherited with respect to complentation and specification. Consideration of how constructions involving double wh questions and parasitic gaps along with violin-sonata sentences are to be admitted in our framework would lead to the these formulations. However, an adequate explanation of these would justify another article. Thus, we assume simpler inheritance conditions as given in (1.1.2) in the discussions that follow.

(1.1.4) an alternative formulations of phrase structure rules
(Here "⊕" designates concatenation.)

a. complementation
H[subcat L1, gap L2⊕L3, bind L4⊕L5]
H[subcat <C[gap L6]|L1>, gap L2, bind L4] C[gap L6⊕L3, bind L5]

b. specification
H[gap L1⊕L2, bind L4⊕L5, comp L6⊕L7]
C[gap L1, bind L4⊕L6] H[spec <C>, gap L1⊕L2, bind L5⊕L7]

1.2 Lexical Rules

Lexical rules assert regularities that hold among lexical entries. To take a trivial example, given an infinitival form of a verb, we can 'predict' rather safely that a third person singular form of that verb exists, whose spelling is obtained by adding an "s" at the end of the spelling of the infinitival form with some orthographical modifications. Thus, something like (1.2.1) will be part of the lexicon (and grammar) of English.

(1.2.1) third person singular lexical rule
Vs $\models$ [fin, spec <n[nom, 3s]>] if V $\models$ [inf, spec <n[ ]>]
where Vs tpsf V

Here, "if" is a relation obtaining among two lexical entries. An expression of the form given in (1.2.1) states that a lexical entry that appear on the left-hand side of "if" exists in the lexicon if a lexical entry that appear on the right-hand side of "if" exists in it. All feature specifications including spellings and semantic representations that are not explicitly stated in the rule are the same between the two lexical entries. In the example given in (1.2.1) "tpsf" is to be construed as a binary relation that holds between an infinitival form representation of a verb and its third person singular form representation.

Some lexical rules play an important role in our analysis of unbounded dependency constructions. For instance, (1.2.2) is indispensable.

(1.2.2) gap introduction lexical rule
v[subcat < >, gap <n[ ]>] if v[subcat <n[ ]>, gap < >]

In the fragment defined in the previous section, "loves" is specified as a transitive verb, occurring in environments immediately preceding a noun phrase. However, as classical arguments show, they appear without apparent object noun phrases in an unbounded dependency constructions. Thus, only (b) is ungrammatical in (1.2.3).

(1.2.3) a. John loves Mary.
b. *John loves.
c. Mary, John loves.
d. Jane, Mary thinks John loves.

Given a lexical rule of the form in (1.2.2) and a lexical entry of the form in (0.2.3) we obtain the following lexical entry, which is involved in the sentences (1.2.3.c,d).

(0.2.3) loves $\models$ v[fin, spec <n[nom, 3s]:X>, subcat <n[acc]:Y>]:**love'**(X,Y)

(1.2.4) loves $\models$ v[fin, spec <n[nom, 3s]:X>, gap <n[ ]:Y>]:**love'**(X,Y)

Note that the lexical rule in (1.2.2) explicitly refers to subcat. Thus, (1.2.5) is not part of English lexicon.

(1.2.5) (this is not part of English lexicon)
loves $\models$ v[fin, gap <n[ ]:X>, subcat <n[ ]:Y>]:**love'**(X,Y)

Then what about sentences like (1.0.2.d).

(1.0.2.d) [I wonder] who Mary thinks loves John

Here, following the analysis of Gazdar (1981) we assume the following lexical rule.

(1.2.6) subject extraction lexical rule
v[subcat <v[fin, spec <n[ ]:X>]>, gap <n[ ]:X>]
if v[subcat <v[fin, spec < >, gap < >]>]

Since we have a lexical entry of the form in (1.0.3), a lexical entry of the form in (1.2.7) is obtained if we assume that "thinks" is a legitimate input to this rule.

(1.0.3) thinks ⊨ v[fin, spec<n[ ]:X>, subcat <v[fin]:Y>]:**think'**(X,Y)

(1.2.7) thinks ⊨ v[fin, spec<n[ ]:X>, subcat <v[fin, spec <n[ ]:Y>]:Z>,
gap <n[ ]:Y>]:**think'**(X,Z)

1.3 Binding

We first saw how information regarding the existence of gaps are to be propagated, and then introduced lexical rules that introduce gaps. (The bottom- up expressions are intended only as a means of facilitating the understanding of the readers.) What remains to be discussed is how binding is achieved.

We have to take into consideration two kinds of binding. Topicalization is syntactic binding of syntactic gaps, whereas complementization is binding of semantic variables.

1.3.1 Syntactic Binding

Information concerning the existence of syntactic gaps has to be somehow resolved or bound in some local phrase structure. The following phrase structure rule sanctions local structures where this binding takes place.

(1.3.1.1) topicalization
Mr → Bdr Hd
where gap @ Hd = <Bdr|gap @ Mr>,
subcat @ Mr = subcat @ Hd = nil,
spec @ Mr = spec @ Hd = nil,
sem @ Mr = sem @ Hd,
head @ Mr = head @ Hd

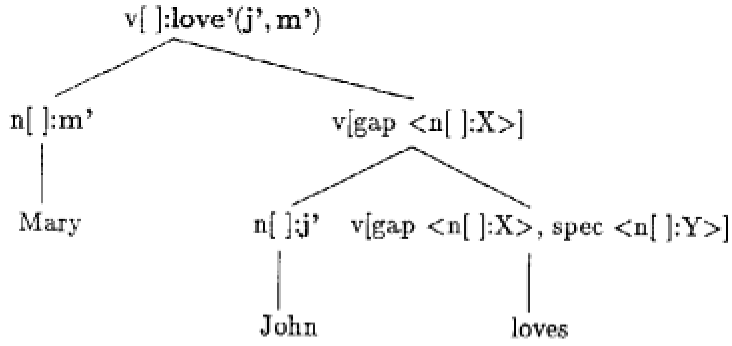
In this local structure, binding occurs with respect to gap @ Bee. Therefore, binding feature inheritance does not hold in this local structure with respect to gap @ Bee and gap @ Mr. The relation among them are explicitly stated as a constraint.

Graphically, the following structure will show what topicalization is all about.

(1.3.1.2) P[gap < >, ...]:S
X P[gap <X>, ...]:S

This will sanction sentences like (1.3.1.3).

(1.3.1.3) Mary, John loves



1.3.2 Semantic Binding

Besides syntactic binding of gaps, we have to take into consideration semantic binding of variables that occur in a relative or interrogative constructions. We call this complementization.

(1.3.2.1) complementization

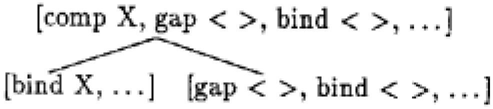
Mr → C1 C2

where comp @ Mr = bind @ C1,
gap @ Mr = nil

Here, since binding occurs with respect to bind @ C1, binding feature inheritance does not hold between bind @ C1 and bind @ Mr.

Graphically, (1.3.2.2) will tell you what this is all about.

(1.3.2.2)



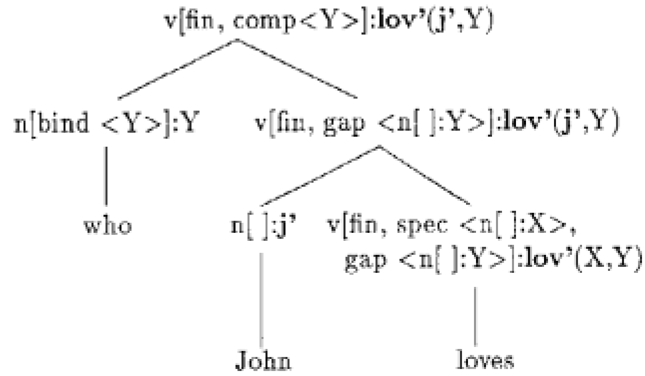
Although we give this as a phrase structure rule, or constraints obtaining among feature specifications involved in a local phrase structure, it can never 'stand alone'. Complementization concurs with topicalization and specification. Or, we can think of complementization as unifying with other phrase structure rules.

(1.3.2.3) complementization unified with topicalization

Mr → Bdr Hd

where gap @ Mr = nil,
gap @ Hd = <Bdr>,
comp @ Mr = bind @ Bdr,
sem @ Mr = sem @ Hd,
head @ Mr = head @ Hd,
subcat @ Mr = subcat @ Hd = nil,
spec @ Mr = spec @ Hd = nil

(1.3.2.4) [I wonder] who John loves

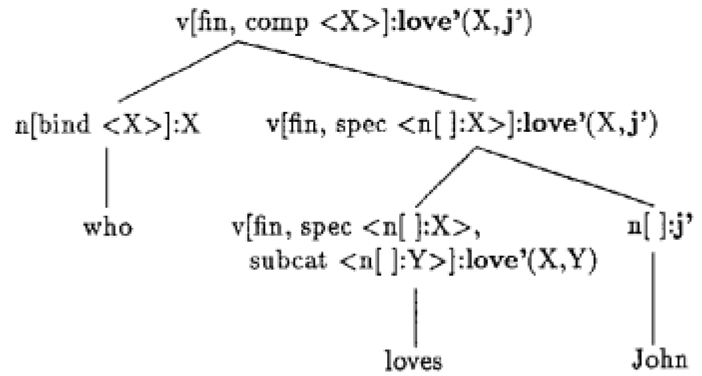


(1.3.2.5) complementization unified with specification

Mr → Sp Hd

where spec @ Hd = <Sp|spec @ Mr>,
comp @ Mr = bind @ Hd,
gap @ Mr = gap @ Hd = nil,
sem @ Mr = sem @ Hd,
head @ Mr = head @ Hd,
subcat @ Mr = subcat @ Hd = nil

(1.3.2.6) [I wonder] who loves John



2 Island Constraints

In this section we will consider how the so-called "island constraints" phenomena are to be guaranteed in a phrase structure grammar description of English.

2.1 Examples

What follows are some typical ungrammatical sentences exemplifying island constraint violations.

(2.1.1) complex NP constraint

a. *[I wonder] which book John met a child who read _

b. *[I wonder] who John believes the claim that Mary loved _

wh-island

c. *which book did you wonder who bought _

subject condition

d. *who did a story about _ surprise you

The ungrammatical examples (2.1.1.a) and (2.1.1.c) both violate the condition that gap @ Mr = nil in local structures where complementization is to take place. We will discuss this point shortly.

It is difficult to account for the ungrammaticality of (2.1.1.b) from our point of view, because with respect to local structures involved, it is quite similar to the sentence in (2.1.2), which is grammatical.

(2.1.2) [I wonder] who John believes that Mary loves _

This disparity could be explained from a cognitive point of view. See Hasida's article in this volume for detail.

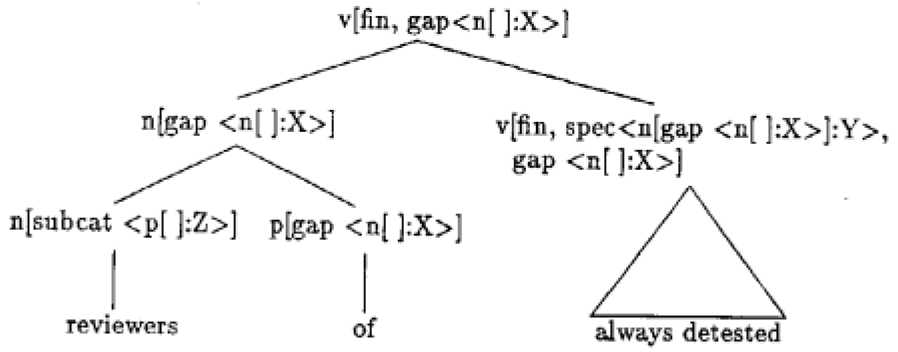
Our simpler grammar with binding feature inheritance as stated in (1.1.2) would admit those strings as given in (2.1.1.d), although sentences like (2.1.3.a) would be disallowed simply because our gap inheritance lexical rule specifically refers to subcat rather than spec in the input category. However, if we assume phrase structure rules as given in (1.1.4), sentences like (2.1.1.d) are disallowed while sentences with parasitic gaps within subjects could be allowed.

(2.1.3) a. *Who did John think that loves Mary?

b. Kim wondered which authors reviewers of always detested.

[GKPS: p.163]

(2.1.4) Kim wondered which authors reviewers of _ always detested _.



A reasonable explanation of how this is achieved will take us too far afield, so let's suffice it to say that in the 'top' local phrase structure in (2.1.4) where specification takes place, the relations in (2.1.5.a) hold among values for gap with respect to the categories involved, which satisfy conditions stipulated in (1.1.4.b), namely those conditions given in (2.1.5.b).

(2.1.5) a. gap @ Sp = gap @ Mr = gap @ Hd = <n[ ]:X>

b. gap @ Sp = L1, gap @ Mr = gap @ Hd = L1⊕L2

2.2 Constraints on Complementization

Thus we have to show how examples (2.1.1.a) and (2.1.1.c) are to be disallowed. For brevity of exposition, however, we will consider (2.2.1.a) and (2.2.1.b) instead.

(2.2.1) a. *[I wonder] which boy John met a girl who loved

b. *[John asked me] which boy Mary wondered who loved

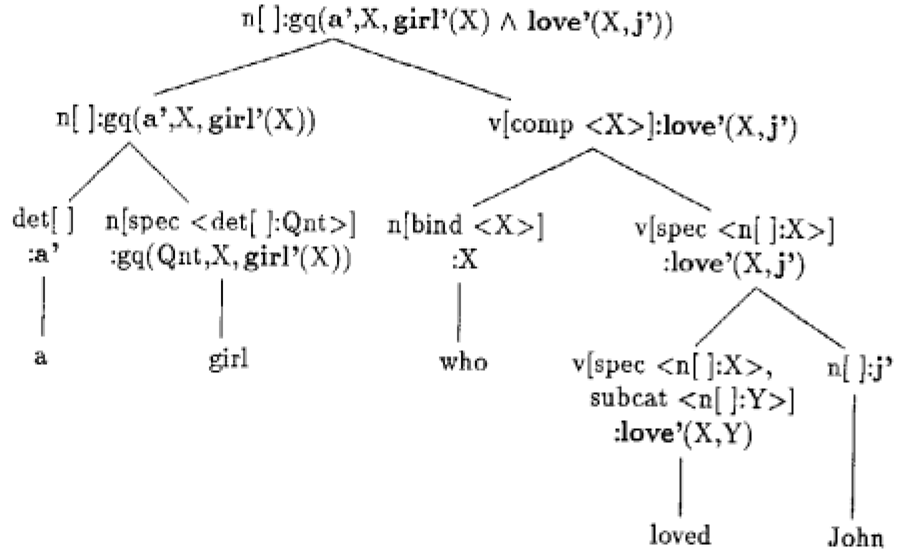
The following treatment regarding syntax and semantics of relative constructions are quite informal, inadequate, and inaccurate. However, since our concern here is to show how island constraints are to be effected within a phrase structure account of English, any attempt to rectify this defect is bound to grow into something grossly out of proportion. Thus, let's simply assume that relative clauses are sanctioned through the following phrase structure rule. Here, all semantic representations are nothing more than notational junks, but if you would try to take the expressions of the form "gq(...)" as generalized quantifiers, you will see what I have in mind here.

(2.2.2) adjunction 1

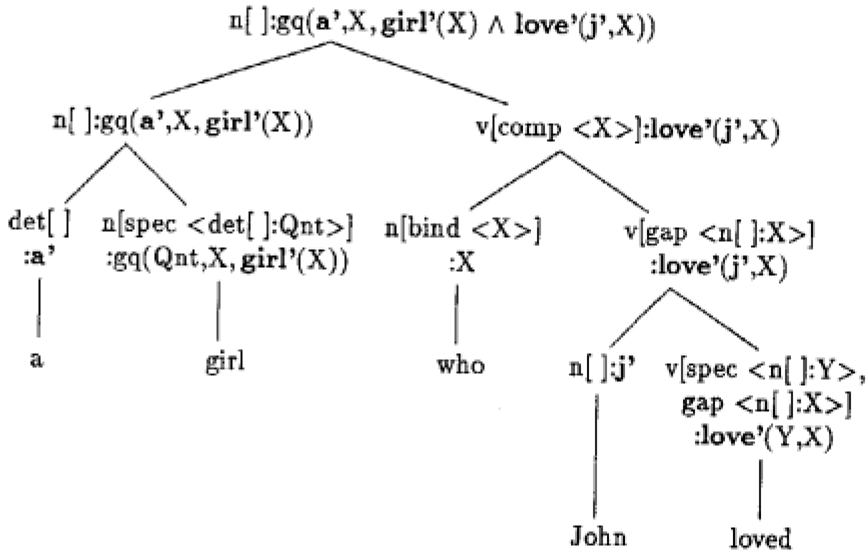
n[ ]:gq(Qnt,X,P1∧P2) → n[ ]:gq(Qnt,X,P1) v[comp <X>]:P2

This phrase structure rule is responsible for the following parse trees, among others.

(2.2.3) a girl who loved John
gq(**a'**,X,**girl'**(X) ∧ **love'**(X,**j'**))

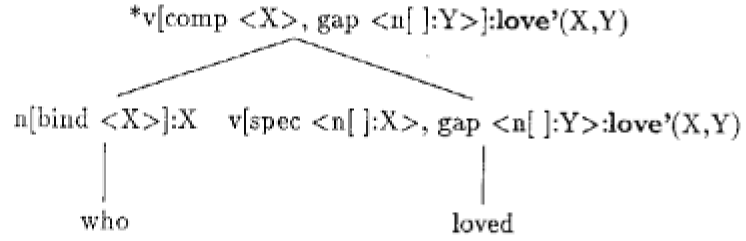


(2.2.4) a girl who John loved
gq(**a'**,X,**girl'**(X) ∧ **love'**(**j'**,X))



In order to see how (2.2.1.a) is disallowed, let's take a look at the following parse tree.

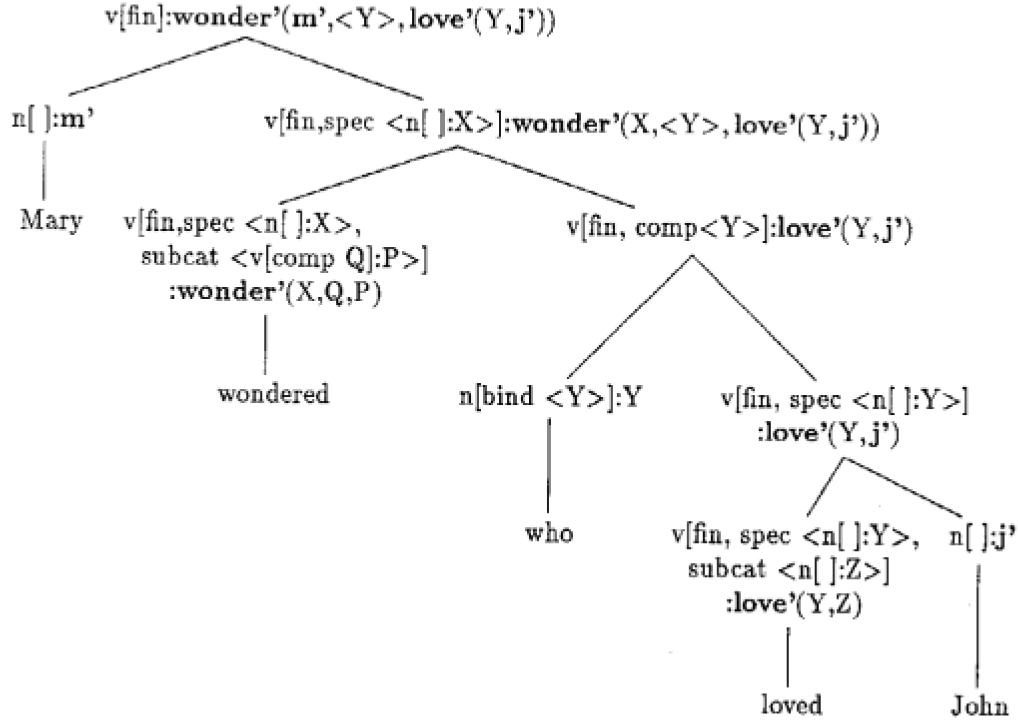
(2.2.5) *[I wonder] which boy John met a girl who loved



If this structure is to be admitted, complementization along with specification must be involved. However, since gap @ Mr is not nil, this is disallowed.

Likewise, embedded questions are assigned parse trees as shown in (2.2.6).

(2.2.6) Mary wondered who loved John
wonder'(**m'**,<X>, **love'**(X,**j'**))

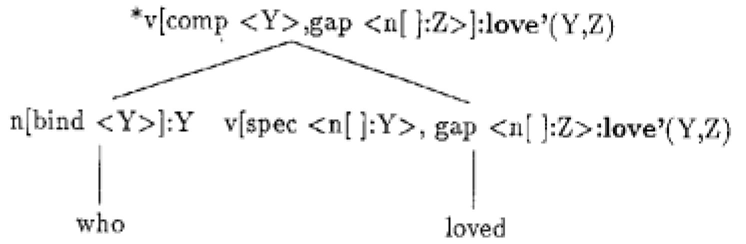


Here we assume the following lexical entry.

(2.2.7) wondered ⊨ v[fin, spec<n[ ]:X>, subcat<v[comp Q]:P>]:**wonder'**(X,Q,P)

The ungrammaticality of (2.2.1.b) can be easily understood if you look at the following local phrase structure.

(2.2.8) *[John asked me] which boy Mary wondered who loved



Here again, we would like to have complementization along with specification, but this is not permitted since gap @ M_T is not nil.

3 Concluding Remarks

Phrase structure grammar descriptions of natural languages as shown above is not developed with giving explanation to the so-called "island constraints" as its central objective. It is a monostratal and unification based theory of grammar, with a particular interest in giving a reasonable account to the fact that human beings rapidly and easily understand and produce utterances of a given natural language. Unlike transformational grammars, a strict dichotomy of competence and performance is not a theoretical prerequisite in constructing theories of grammar in our framework, although we could argue whether a given phenomenon fall within the realm of competence or performance, in the sense that the former is to be thought of as a static characterization of human linguistic capacity, whereas the latter should involve dynamic aspects of this. It is to our advantage that we have real means to consider how these two are related to each other.

Our condition on complementization to the effect that gap @ M_T be nil is in a sense just a notational variant of FCR 20: ([SLASH]&[WH]) in GKPS pp.153-155. This feature cooccurrence restriction states, figuratively translated in everyday prose, that grammatical categories that are commanded by a wh-phrase cannot have gaps in them. This is in a sense what wh-island constraint is all about. Our constraint on complementization is, descriptively speaking, no different from this statement. However, we might try to predict what sort of local structures are more likely to be excluded in a given language by introducing a measure of processing complexity to our theory, while feature cooccurrence restrictions can in principle state any sort of stipulations.

Thus, our fragment of English utilized four category-valued features. Let us informally define the notion "operative in a given local phrase structure" with respect to these four category-valued features.

(3.1) operative in a local structure

In the following local structures (a) and (b), F is operative.

a. [F V1, R]

[F V2, R] X

where length(V1) < length(V2)

b. [F V1, R]

X [F V2, R]

where length(V1) < length(V2)

Typically subcat is operative in complementation, spec in specification,gap in topicalization, bind in complementization. Note that complementization concurs syntactically with specification and topicalization. In these local structures, two category-valued features are operative. It is natural to assume that in performance processing of these structures is heavily loaded, thus dis llowing inheritance of non-null values of gap between the head and the mo her. Although this point has to be further attested through other than mere s eculations, this account is not unlikely to hold.

We notice that the following ordering relation hold among the four category-valued features introduced in the fragment above.

(3.2) subcat < spec < gap < bind

Here the expression "F1 < F2" states that in a local phrase structure in which F2 is operative, F1 @ Mr is to be nil. Note that since lexical rules do not involve phrase structures, this ordering is irrelevant in their application. If we incorporate this ordering into our theory of grammar, phrase structure rules in our grammar will be stated in much simpler forms. Finally, this relation seems quite natural if we take into consideration pre-theoretic significance of these features, that is subcat must be locally processed, or bind is semantically salient and so on. In this way, grammatical constraints might be in part or whole translated into processing complexity of a given configuration.

Acknowledgments

This article is based on a Japanese document I prepared for Unbounded Dependency Workshop Symposium on Island Constraints held at Tokyo Metropolitan University on May, 5, 1987. Although it was entitled "Island Constraints, from HPSG point of view" as the grammatical theory employed in that document resembled those unification-based theories exposed in Pollard (1984) and/or Pollard (1985) [HPSG] rather than those instantiation-based approaches of Gazdar and Pullum (1982) [GPSG-82] and/or Gazdar, Klein, Pullum and Sag (1985) [GKPS], the term "phrase structure grammar" is employed here because it does not really refer to differences among theories of GPSG. A slightly amended version of this Japanese document is available as Harada (1988).

I would like to express my gratitude to those responsible for the Unbounded Dependency Workshop, especially Professor Nakajima Heizo at Tokyo Metropolitan University, among others.

The grammatical description format employed in this article is sort of a by product of discussions with members of Japanese Phrase Structure Grammar Working Group at ICOT, whose chairman is Gunji Takao. I would like to express my thanks to all those involved in the project. Especially Hasida Kôiti and Sirai Hidetosi are at least as responsible for what's written here as I am. Some premature outbursts of ideas adopted here can be found in Harada (1986) and Harada (1987), if you prefer obscure explanations in English to accurate accounts in Japanese. For further details and developments of our approach, see Miyosi et. al. (1986) and Shirai et. al. (1987). Hasida and Shirai(1986) will show you what constraint unification is all about.

References

Gazdar, G. 1981. "Unbounded dependencies and coordinate structure." *Linguistic Inquiry*, **12**.

Gazdar, G. and G.K. Pullum. 1982. *Generalized Phrase Structure Grammar: a theoretical synopsis.* Bloomington: Indiana University Linguistics Club.

Gazdar, G., E. Klein, G.K. Pullum, and I.A. Sag. 1985. *Generalized Phrase Structure Grammar.* Oxford, Basil Blackwell.

Harada, Y. 1986. "A Prolog implementation of SUHG." *Linguistic Research* No. 4, Tokyo University English Linguistics Association.

Harada, Y. 1987. "Toward grammatical descriptions of natural languages within extended unification environments," *Linguistic Research* No. 5, Tokyo University English Linguistics Association.

Harada, Y. 1988. "A phrase structure grammar account of island constraints (in Japanese)." *Humanitas*, No. 26, Tokyo: The Waseda University Law Association.

Hasida, K. and H. Sirai. 1986. "Constraint unification (in Japanese)." *Computer Software*, **3**, No. 4.

Miyosi, H., T. Gunji, H. Sirai, K. Hasida, and Y. Harada. 1986. "JPSG—a phrase structure grammar for Japanese (in Japanese)." *Computer Software*, Vol. 3, No. 4.

Pollard, C.J. 1984. *Generalized Phrase Structure Grammars, Head Grammars and Natural Language.* Ph.D. dissertation, Stanford University.

Pollard, C.J. 1985. "Lectures on HPSG." unpublished manuscript, Stanford University.

Sirai, H., T. Gunji, K. Hasida, and Y. Harada. 1987. "Grammatical descriptions with local constraints (in Japanese)." paper presented at Language Processing

and Communications Research Group, Institute for Electronics, Information, and Communication Engineers of Japan.

Received March 4, 1988

A Cognitive Account of Unbounded Dependency*

Hasida Kôiti
Institute for New Generation Computer Technology

1 Introduction

The purpose of the present article is to account for syntactic constraints on some aspects of **unbounded dependency** (**UBD** hereafter) phenomena, by means of a **computational model**. A computational model is described in terms of formal operations on formal structures. An explanation based on such a model ascribes the phenomenon in question to some computational properties of the model; i.e., such properties as computational complexity with regard to both space and time, accessibility to some parts of memory, and so forth. A major advantage of this sort of computational approach is that it can exploit dynamic aspects of the phenomena, such as temporal order of processing and structural ambiguity arising dynamically in the course of comprehending or producing utterances.

Linguistics, by contrast, has paid little attention to those aspects and has limited itself to investigation of characteristics of language which could be talked about in static terms. To say that S-structure is derived from D-structure, for example, does not imply that the former chronologically precedes the latter. In fact, many linguistic (especially syntactic) phenomena can be understood without referring to ambiguity, processing order, etc. This is partly why linguistics has seen its successes.

Nevertheless, languages have some properties essentially stemming out of dynamic features of language processing (or perhaps of a more general cognitive processer, after Piaget). Our main concern here is with such dynamic aspects of language. In what follows, we shall first touch upon some cognitive viewpoint applied to several types of island phenomena, showing that some significant part of linguistic account is reducible to processing terms. Further shall we go on to demonstrate that a dynamic approach can elucidate some phenomenon, the noun-complement case of the complex noun-phrase constraint, which is unlikely to be explicable in static terms of traditional syntactic theories.

2 Static Account

Consider the (non-) sentences below, which involve UBD constructions; the subscripts i and j indicate coindexations.

(1) Who$_i$ [$_{S_0}$ do [$_{NP}$ you] [$_{VP}$[$_V$ believe] [$_{S'_1}$ that [$_{S_1}$ she loves ϵ_i]]]]?

*A slightly different version of this paper appeared in the Proceedings of COLING'88 under the same title.

(2) *What$_i$ did you see the girl [$_{S'_2}$ who [$_{VP_2}$ ate ϵ_i]]?

(3) *What$_i$ do you wonder who ate ϵ_i?

(4) *Who$_i$ do you wonder what$_j$ I gave ϵ_j to ϵ_i?

(5) *Who$_i$ did a story about ϵ_i surprise you?

The grammatical status of these strings is understood without referring to dynamic terms such as temporal processing order, structural ambiguity, etc. Let us see how.

2.1 Constraints about Dislocated Elements

The syntactic operations of English we will pay attention to in the following discussion are what we might call **complementation**, **specification**, **adjunction**, **binding**, and **passing**, each of which takes place in a branching local tree. Complementation is to associate an object with its head. In (1), for instance, a complementation takes place in the local tree consisting of VP, V, and S'_1; the mother, the head, and the object, respectively. Specification attaches a specifier to the head; e.g., the subject of a sentence to VP (or to IP in the recent transformational theories (Chomsky 1986)). An example of complementation is the local tree expanding S_0 in (1), where the specifier is the subject NP. Adjunction associates an adjunct with the head; e.g., an adverb with VP, and a relative clause to NP. Binding is to bind a **dislocated element** (see next paragraph), associated with a syntactic gap, to its antecedent (e.g., a WH-phrase such as *who* and *on which day*). For instance, the dislocated element associated with ϵ_i gets bound by *Who*$_i$ in the top local tree of (1) above. Passing is to pass a dislocated element between the mother category and some of the daughters. In the local tree expanding VP in (1), the same dislocated element is passed between the mother (VP) and the complement daughter (S'_1).

Thus, binding and passing are both operations on dislocated elements. By a *dislocated element*, we refer to a token in mental representation which syntactically corresponds to several positions in a sentence; Typically, there are two such positions, the filler and the gap, the former being often called the antecedent of the latter. For instance, there is a dislocated element appearing as *Who*$_i$ and as ϵ_i in (1). Different approaches to syntax assume a dislocated element to additionally correspond to different sets of positions between the filler and the gap. In general, transformational grammars tend to postulate fewer such positions than do phrase-structure grammars (PSGs, hereafter, which include, among others, GPSG Gazdar, Klein, Pullum, and Sag 1985) and HPSG (Pollard 1984, 1985) or LFG (Bresnan 1982). Also, these theories assign different status to dislocated elements; Transformational grammarians talk about them as if they *move* through sentence structure, while the others assume they are simply *associated with* grammatical categories, via such means as the SLASH feature.

Such differences among various approaches to syntax, however, is irrelevant to the discussion in the rest of the papaer. We shall exploit no hypothesis specific to any of these syntactic theories, so that our discussion will be neutral across them. We will borrow some useful terminology and metaphors from specific grammar theories, but that is only for explanatory ease, and should not be taken to be any commitment to any of such approaches. The above description of passing, for instance, reads as if a dislocate element were part of grammatical categories, and hence might well remind the reader of PSGs. This by no means implies that we should abandon transformational accounts in favor of PSGs. Although PSGs are neutral with respect to the temporal order of processing, incidentally, we shall sometimes talk about passing as the dislocated element going from the mother to a daughter, reflecting the temporal order of actual sentence processing; cf. transformational grammarians talk about movement as if a dislocated element goes along the opposite direction.

2.2 Explanation

Now let us resume the problem. The distribution of grammaticality over (1) through (5) is accounted for if we assume the following constraints on the syntactic structure of English.

(6) Passing of a dislocated element is permitted only between the mother and the head daughter or between the mother and the complement daughter.

(7) Passing of a dislocated element and binding of another cannot take place simultaneously in one local tree.

In (1), every passing obeys this constraint. For instance, passing the dislocated element bound by *Who*$_i$ into S'_1 and passing it into S_1 are both O.K., because S'_1 is the complement of *know* and S_1 is the complement of *that* (irrespective of whether you might employ a transformational account or such theories as GPSG, HPSG, and LFG.).

However, (2) violates (6) and (7). First, the dislocated element bound by *What*$_i$ is passed into S'_2, which is not a complement but an adjunct of *girl*; i.e., a violation of (6). Second, the same dislocated element is passed into VP_2 where another dislocated element gets bound by *who*, causing a violation of (7). Similarly in (3) and (4), passing and binding co-occur at the local trees introducing *who* and *what*$_j$, respectively, violating (7). (5) is blocked by (7), because of the passing into the subject *a story about* ϵ_8; i.e., the specifier of INFL (in the transformational account) or of VP (in theories like GPSG, etc.).

2.3 Cognitive Aspects of Constraints

The explanations about such phenomena proposed in contemporary linguistic inquiries, especially the accounts in terms of **barriers** (Chomsky 1986), may be roughly regarded as formalizations of the idea sketched above. This line of reasoning is in turn attributed to processing terms, when viewed from the standpoint of cognitive science. The background intuition is that the extent of processing load imposed by a syntactic operation varies from one type of operation to another, and that there is an upperbound on the total processing load for constituting one branching local tree. On account of this, the reason why passing tend to be blocked under the certain sorts of circumstances would be that it is an expensive operation and thus is hard to perform together with other expensive operations such as adjunction or binding.

The processing load for carrying out various syntactic operations could be further reduced to more fundamental aspects of information processing. Consider, for instance, why specification and adjunction should be harder than complementation. The reason seems to be that complementation is lexically licensed and is head-initial (the current discussion is limited to English); i.e., the occurrence of a complement licensed by the lexical entry of its head and thus is predicted from the occurrence of the preceding head. On the other hand, the occurrence of a specifier or of an adjunct is harder to predict, because the former (though lexically licensed) precedes its head and the latter is not lexically licensed by the head. Passing and binding of dislocated elements are also considered to be non-lexical operations, though there are a few exceptions involving, for example, the so-called tough adjectives, as indicated by the following sentence.

(8) [Which violin]$_i$ is [this sonata]$_j$ [$_{AP}$ easy to play ϵ_j on ϵ_i]?

The difference between (4) and (8) is that the binding of the dislocated element carrying index j is lexically licensed in the latter but not in the former. In (8), the binding in effect occurs at the local tree expanding AP, where this binding is sanctioned by the lexical entry of *easy*. This is why (8) is grammatical despite (7); i.e., the lexically licensed binding is not an expensive operation, so that it does not play the same role as ordinary binding would play in regard of (7).

Note that the processing load concerning the examples we have discussed so far is defined within a single sentence structure rather than by taking structural ambiguity into account. This is why the traditional syntactic approaches are able to capture some of such aspects of language; in principle, properties of a single structure can be characterized in static terms.

3 Dynamic Account

However, the above static approach cannot by itself explain some UBD phenomena, especially the so-called Complex NP Constraint (originally termed by Ross (1967)) observed in the examples that follow.

(9) *Who$_i$ do you believe [$_{NP_1}$[$_{NP_2}$the claim] [$_{S'_3}$that [$_{S_3}$she loves ϵ_i]]]?

(10) *What$_i$ did you propose a plan [$_{VP}$ to buy ϵ_i]?

In (9), S'_3 is the complement of *claim*. Hence the dislocated element bound by *Who$_i$* should be permitted to be passed into S'_3 without violating (6). Similarly, VP (or CP, in the transformational account) in (10) is regarded as the complement of *plan*, so that the dislocated element bound by *What$_i$* should be able to pass through, (6) and (7) being respected. Hence the static account in the previous section provides no reason why these examples should be ungrammatical. It is considered because of essentially the same sort of difficulty that Chomsky (1986) leaves unexplained this type of island effect.

3.1 A Model of Language Processing

Now then let us turn to dynamic aspects of language processing, and consider what kind of syntactic structures a human hearer should have built and tentatively maintains when *that* in (9) is encountered. As a basis for this investigation, we adopt the following postulates about human language processing.

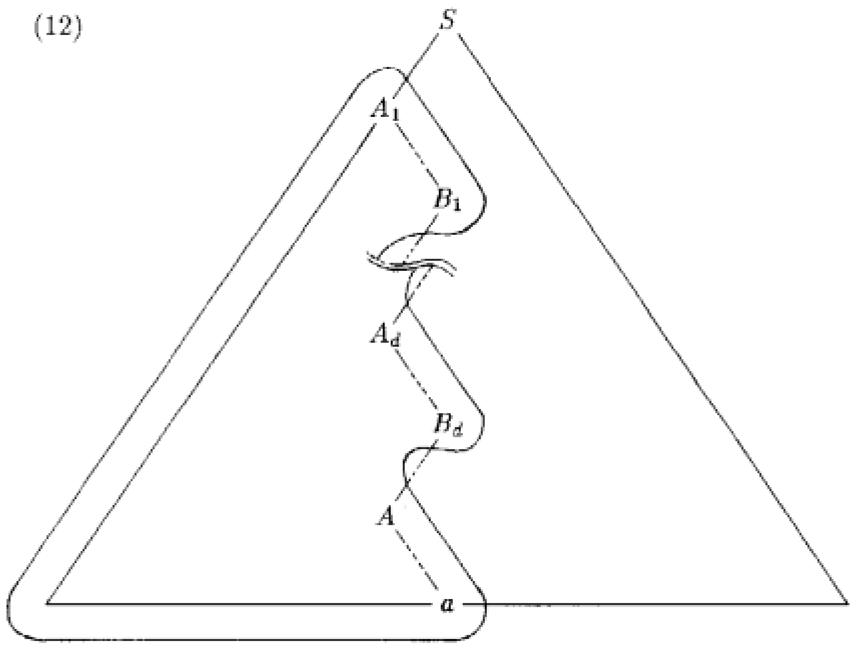
(11) a. When a word is encountered, it is attached to structures previously built, giving rise to new structures. Even when several possible ways of attachment are acknowledged, the processing is not postponed, but as many

new structures corresponding to those ways of attachment are made in parallel.

b. There is a limitation on the size of the memory for storing these structures, and thus it is impossible to retain all the structures potentially sanctioned. Only structures activated strongly enough can survive the competition for seats in the limited memory.

From (11) plus some minor hypotheses, a general processing model follows, which describes both sentence comprehension and generation. This model postulates that just after any word a is encountered, every maximal structure of the sentence currently hypothesized in mind should look approximately like the part enclosed within the curve in (12).

(12)



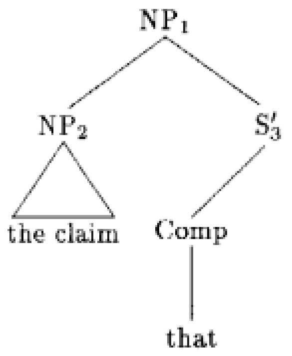
Here every branching local tree is assumed to be binary, without loss of generality. A_1 and S may be identical, and the short-term memory contains the information about A_i, B_i ($1 \leq i \leq d$), and A, plus the informatin about the configuration of these categories relative to each other. Note that, as a whole, enough information is thus retained to control the grammaticality of the way the foregoing context fits the rest of the sentence; Those categories are the points on which the currently hypothesized structure has contacts with the still unknown part of the sentence.

Strictly speaking, the picture shown in (12) should be looked upon as a very rough approximation of the reality. That is, the part of the sentence structure enclosed in the curve might contain some variable parts, rather than being totally definite. Suppose, for instance, that a sentence begins with a noun phrase, say *This man*. The entire tree structure of this NP should be completed as soon as *man* is encountered, but its grammatical case would not be uniquely determined yet, because the sentence as a whole might turn out to be something like *This* *man, I don't know*, rather than *This man is crazy*; The initial NP is accusative in the former sentence, and nominative in the latter. In the following discussion, however, we shall merely exploit very rough properties of the model, so that such an inaccuracy is considered irrelevant. Readers are referred to Hasida (1985) for how this model is obtained and what it predicts, which the limited space of the current article fails to cover.

3.2 Explanation

Let us turn back to (9). According to this model, when *that* is encountered while (9) is being comprehended, the right-branching structure covering the string from Who_i through *claim* has been nearly completed and the most active structure around *that* should look like (13). Here arise two pieces of independent two-way ambiguity, as

(13)



listed below, concerning how this structure might potentially grow.

(14) a. *That* is a relative pronoun.
b. *That* is a conjunction.

(15) a. S'_3 contains a gap bound by $What_i$.
b. S'_3 does not contain a gap bound by $What_i$.

The combination of (14a,b) and (15a,b) gives rise to local structural ambiguity encompassing four hypotheses: (14a&15a), (14a&15b), (14b&15a), and (14b&15b).

Since (16), an instantiation of (14a&15b), is clearly OK, what we have to show is that out of these four hypotheses just (14a&15b) and (14b&15b) enter the grammar to be acquired.

(16) Who_i did you tell ϵ_i the fact $that_j$ he knew ϵ_j?

Hence now let us consider why (14a&15a) and (14b&15a) are rejected. We pay attention to the behavior of dislocated elements, as we did in the static approach. Two dislocated elements are relevant to the grammatical status of (9). The first one, which is bound by *that*, corresponds to the possibility (14a). Let us refer to this element as α. The other dislocated element, the one bound by $What_i$, is present iff (15a) obtains. We shall call it β.

The status of (14a&15a) is parallel to (2). (7) rules out this possibility immediately, because it postulates that the local tree expanding S'_3 accommodates both the binding of α by *that* and the passing of β into S_3. As for (14b&15a), however, (7) as it is fails to work.

We need some preparation before tackling why (14b&15a) is rejected. (7) says that two distinct dislocated elements, one passed and the other bound, cannot simultaneously take part in one local tree. As mentioned above, the cognitive-scientific motivation for the constraints (6) and (7) is that the mental grammar does not admit a rule whose execution accompanies too severe processing load. On account of this motivation, (7) is naturally generalized simply by abandoning the hitherto implicit presupposition that just a single consistent structure of the sentence is talked about at one time. That is, we hold:

(17) The grammar excludes a local tree which involves two distinct dislocated elements in two different manners, binding the one and passing the other, at the same time, irrespective of whether the two dislocated elements appear in the same coherent structure of the sentence.

The essential difference between (7) and (17) is that the latter can simultaneously mention several hypothetical structures of the same sentence. That is, (17) does, but (7) does not, outlaw a local tree construction which subsumes two distinct local hypothetical structures one of which involves a dislocated element being bound and the other of which another dislocated element being passed.

Note that it is not a trivial matter whether a local tree in one hypothetical structure and another local tree in another hypothetical structure are the same local tree to which (17) should apply. The identity of local tree should be defined so that the same local tree should be processed at the same time whichever hypothetical structure it might appear. This is where the model introduced above as in (12) comes in. Let us say that two local trees in two different hypothetical structures are the same iff they appear as exactly the same local tree in a local hypothetical structure constructed according to this model.

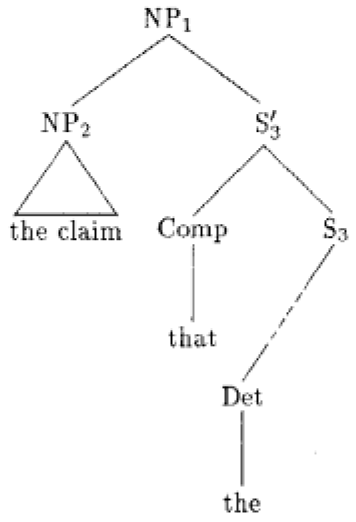
Thus, in (13), the local tree expanding S_3' is the same local tree across the four hypothetical structures corresponding to (14a&15a), (14a&15b), (14b&15a), and (14b&15b). That is, this local tree is built at the same time in all the four possible lines of processing. This is understood by comparing (13) with the next state (18); When you go from (13) to (18), the local tree expanding S_3' is completed.

According to (17), therefore, the rule of syntax in charge of this case must reject the possibility of the existence of either α or β; otherwise these two dislocated elements would be manipulated (bound and passed) simultaneously here. Now note that α is chronologically newer than β. What psychologists call the **recency effect,** consequently, tells us that α (hence (14a&15b)) should survive, defeating β and thus rendering (9) ungrammatical. As a result, incidentally, we are left with only the ambiguity about (14); i.e., whether S_3' is a relative clause or a noun complement clause.

The account of the ungrammaticality of (10) is the same except that the potential binder, which is the counterpart of *that* in (9), is hidden in *plan* and thus is not overt here. This time α is the dislocated element bound by this binder, and β the one bound by *What$_i$*.

Note that this explanation concerns language acquisition by children, rather than language use by adults. It must concern the acquisition stage; otherwise what we have shown would not be the ungrammaticality of (9) and (10) but merely the

(18)



difficulty of processing them. In fact, the above account does apply to language acquisition, because the ambiguity pertaining to (14) and (15) occurs every time a structure like (13) is encountered, so that its disambiguation can be fixated as a part of the grammar of English to be acquired.

3.3 Rules Handling Local Ambiguity

Further discussion are in order here about the generalized constraint (17) and its role in the above explanation. First, the above discussion postulates that the grammar rules are sensitive to structural ambiguity such as (14) and (15) about (9), in the sense that some rules of syntax work on multiple structures, and thus are in charge of disambiguation. Here one might worry which types of ambiguity are handled by the grammar, and which are handled metagrammatically. Not every sort of ambiguity is visible to the grammar, as is demonstrated by the following example, which is grammatical.

(19) Who$_i$ did you tell the man $[_{S'_4}$ that $[_{S_4}$ she loves $\epsilon_i]]$?

The local ambiguity arising here appears similar to that of (9). More precisely, a four-way local ambiguity arises at *that*, as a combination of two pieces of independent two-way ambiguity, one concerning whether *that* is a relative pronoun or a conjunction, and the other whether or not S'_4 contains a gap bound by *Who*$_i$, just as in (9). An outstanding difference between (9) and (19), however, is that the ambiguity in the latter case involves two different hypothetical constituent structures that follow.

(20) $[_{VP}[_{V'_1}$ tell the man$]$ $[_{S'_5}$ that $S_5]]$

(21) $[_{VP}$ tell $[_{NP_2}$the man $[_{S'_5}$ that $S_5]]]$

To children learning UBD constructions, both of these constituent structures should appear ambiguous about whether or not S_5 contains a gap bound by *Who*$_i$.

This ambiguity, unlike the one in (9), is considered invisible to the grammar, presumably because of the above difference. That is, if any single rule were sensitive to this ambiguity, (19) should be rendered ungrammatical for the same reason as why (9) is so, because in (19) a dislocated element would be bound by *that* and another dislocated element would be passed into S_4 simultaneously. To make sure that the binding and the passing should be simultaneous here, notice that the local tree expanding S'_5 is completed simultaneously in the two pairs of hypotheses corresponding respectively to (20) and (21). Hence the binding by *that* and the passing into S_5 must take place simultaneously.

Seemingly the reason why the grammar is not sensitive to this type of local ambiguity is that the four possibilities are not coherent enough, in the sense that they are distributed across the two distinct constituent structures as mentioned above. It appears that the disambiguation of a piece of local structural ambiguity is acquired as a part of the grammar only if the structures (or hypotheses) constituting that ambiguity are coherent enough with each other. Comparing (19) with (9) and (10), one might thus posit the following generalization.

(22) A piece of local structural ambiguity is handled within the grammar only if the parallel structures involved therein share the same constituent structure.

As for (9), for instance, besides (13) there could of course be several other structural possibilities, but they are simply irrelevant to the acquisition process discussed above, rather than systematically abandoned like (14a&15a) and (14b&15a). We would then be able to disregard any interaction across distinct constituent structures when considering the competence grammar.

There are at least two more supports to constraint (22). First, (22) follows from the following more fundamental constraint.

(23) There is a severe limit on the size of the structure which one rule of syntax can refer to at once.

An ambiguity within one constituent structure tends to fall within this limit, because the parallel structures involved therein share most of the storage with each other. An ambiguity across several constituent structures, however, would more often run out of this limit, since the rate of the shared memory is smaller. Note that (23) claims, after all, nothing more than the limitation on the complexity of mentally feasible rules.

Another reason for holding (22) is based on how *stubborn* an ambiguity is. As mentioned earlier, patterns like (9) and (10) constantly accompany the local ambiguity like (14) plus (15). In contrast, patterns like (19) are often less ambiguous, as shown in the example below.

(24) Who_i did you tell him $[_{S'_6}$ that she loves $\epsilon_i]$?

In this sentence, the possibility of *that* being a relative pronoun is very unplausible, the local ambiguity being greatly reduced; we are left with the ambiguity of whether or not the dislocated element bound by $\textit{Who}_i$ is contained in S'_6. In summary, the ambiguity in (9) and (10) is robust, while that in (19) is fragile. An ambiguity within a single constituent structure tends to be robust. In comparison, an ambiguity encompassing several different constituent structures tends to be fragile, because the relationship (as for which is more plausible than which, etc.) between those constituent structures varies from case to case, depending on the internal details of the relevant constituents, context, and so on. The corresponding relation in the former type of ambiguity, on the other hand, is more constant. Robust ambiguity is visible to the grammar, while fragile one is not; Some rules of syntax handle the former, while the latter is treated metagrammatically.

3.4 German Case

The following pair of German examples might fall out of our dynamic account on (9) and (10).

(25) Wen$_i$ glauben Sie, daß er ϵ_i liebt?

(26) *Wen$_i$ glauben Sie die Behauptung, daß er ϵ_i liebt?

(25) and (26) are German counterparts of (1) and (9), respectively. Note that the direct translation between English and German preserves grammaticality across these pairs of examples.

Since *Behauptung* 'claim' is of the feminine gender rather than neuter, *daß* 'that' should not be confused with a relative pronoun, namely *die* or *der*, whose antecedent is *Behauptung*; The relative pronoun pronounced the same as *daß* is *das*, which is of the neuter gender and the nominative or accusative case. So it appears that in the case of (26) children learning German should face no ambiguity like (14). Our current approach, which is essentially based on local ambiguity, hence seems unable to account for the grammatical status of (26).

In reality, however, the above example does not contradict our approach; i.e., constraint (17). Ambiguity parallel to (14) arises in fact, and therefore the ungrammaticality of (26) is predicted by (17). To children learning UBD construction of German, (26) accompanies the same sort of local ambiguity as (9) does, because they acquire the rough framework of UBD construction before the gender system is properly installed into the morphology of relative pronouns. According to Mills (1986), when children begin to use relative clauses around the age of 3, relative pronouns in the relative constructions they make are either simply ommited or lacking information about the gender (and the case, too), as in:

(27) Das ist ein Pilz *mm in Walde ist.
That is a mushroom REL.PRON in wood is
'That is a mushroom which is in wood.'

This means that at the beginning children cannot distinguish a relative pronoun of a gender from another relative pronoun of a different gender; e.g, between *die* and *das*. Mills reports that the use of relative pronouns is equipped with the proper system of inflection only after the age of 4 is reached.

The significance of (25) and (26) is not crystal clear, incidentally, if the nonsentence below is also taken into account.

(28) *Wen glauben Sie, daß er sehen wollte?
Who think you that he see-INF want-PAST
'Who do you think that he wanted to see?'

Thus it might be that the ungrammaticality of (26) is attributed to that of (28), without regard to (17). Otherwise the above pair of examples (25) and (26) should provide a further evidence supporting (17).

4 Final Remarks

We have accounted for some island conditions by means of computational evaluation of relevant syntactic operations; i.e., the evaluation reflected in constraints such as (6), (7), etc. These constraints are regarded as captured by the existing linguistic theories. A generalized constraint (17) applicable to dynamic aspects of language processing, especially local structural ambiguity, has been demonstrated to account for the noun-complement case of the Complex NP Constraint, which seems hard to elucidate in static terms of traditional approaches to syntax.

One important aspect of our approach employed here is the hypothesis that some sort of local ambiguity is visible to and thus handled by the grammar. If this hypothesis finally turns out true, which we have attempted to demonstrate, the static approach pursued so far in the linguistic inquiries must be reconsidered. That is, an explanation on the grammaticality of sentences will have to sometimes take into account several possible structures in parallel.

The explanation of the same sort of island condition by Marcus (1980) is comparable to ours in that it also exploits local ambiguity, postulating rules handling them. Since Marcus pays no attention to what kind of ambiguity is visible and what kind is not, however, his discussion has nothing to say about the contrast between (9) and (19). Besides, a radical difference between the two approaches is that Marcus exploits a stipulation called the **determinism hypothesis**, whereas we employ a more humble working hypothesis of parallel processing plus memory limitation.

A caution would be worthy of noting here. A success of computational explanation does not necessarily support either innatism (à la N. Chomsky, J. A. Fodor, D. Marr, etc.) or constructivism (à la J. Piaget, etc.). If any part of human intelligence could be understood to be the outcome of a simple optimization for some computation, it should subject to two different interpretations: that this part should be a domain-specific innate endowment because such a simple optimization may well be preprogrammed in the course of evolution, or, contrariwise, that it should be generated after birth by the work of the domain-independent general intelligence because such a simple optimization could be carried out through maturation and internal experiences. Further scrutiny would thus be simply needed in order to steer our way either to innatist or constructionist disposition.

Along the line of the present argument, perhaps the first point where we could face the choice between these two doctrines is the problem of how much processing load we should ascribe to various syntactic operations. The evaluation of computational load as we have exploited here should vary across languages, depending on the relative statuses of syntactic operations. For instance, the situation must be drastically different between dominantly head-initial languages like English and Spanish and head-final languages such as Japanese and Korean (and maybe German, too). Also open to further scrutiny is whether the variation is explained by the parameter setting approach of transformational theories, or by more general computational considerations.

References

Bresnan, J. ed. 1982. *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: The MIT Press.

Chomsky, N. 1986. *Barriers.* Cambridge, Mass.: The MIT Press.

Gazdar, G., E. Klein, G.K. Pullum, and I.A. Sag. 1985. *Generalized Phrase Structure Grammar.* Oxford: Basil Blackwell.

Hasida, K. 1985. *Bounded Parallelism: A Theory of Linguistic Performance.* Doctoral dissertation, University of Tôkyô.

Marcus, M.P. 1980. *A Theory of Syntactic Recognition for Natural Language.* Cambridge, Mass.: The MIT Press.

Mills, A.E. 1986. *The Acquisition of Gender: A Study of English and German* (Springer Series in Language and Communication). Berlin & Heiderberg: Springer-Verlag.

Pollard, C. J. 1984. *Generalized Phrase-Structure Grammar, Head Grammars, and Natural Languages.* Ph.D. dissertation, Stanford University.

Pollard, C. J. 1985. "Lectures on HPSG." unpublished manuscript, Center for the Study of Language and Information, Stanford University.

Ross, J. R. 1967. *Constraints on Variables in Syntax.* Ph.D. dissertation, Massachusetts Institute of Technology.

Received May 10, 1988