Japanese/Korean Linguistics

Volume 5

Edited by
Noriko Akatsuka,
Shoichi Iwasaki,
& Susan Strauss

Published for the
Stanford Linguistics Association by
CSLI Publications
Center for the Study of Language and Information
Stanford, California

Printed in the United States
00 99 98 97 96 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Conference on Japanese/Korean Linguistics (1st : 1989 : University of Southern California)
Japanese/Korean Linguistics/edited by Hajime Hoji
p. cm.
Includes bibliographical references and index.
ISBN 1-57586-045-7
ISBN 1-57586-044-9 (pbk.)
1. Japanese language–Congresses. 2. Korean language–Congresses.
3. Japanese language–Grammar, Comparative–Korean–Congresses.
4. Korean language–Grammar, Comparative–Japanese–Congresses. 5.
Linguistics–Congresses. I. Hoji, Hajime. II. Stanford Linguistics Association. III.
Center for Study of Language and Information (U.S.) IV. Title.
PL503.C6 1989 90-2550
495.6–dc20 CIP

♾The acid-free paper used in this book meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Barriers for A-Adjunction*

MAMORU SAITO
University of Connecticut

1. Introduction

It was widely held throughout the 1980's that a CP/IP pair constitutes a barrier for movement and government. (See Chomsky (1981, 1986a) for detailed discussion.) The Wh-island effect illustrates the barrierhood of this pair for movement. As shown in (1), neither a CP nor an IP blocks Wh-movement, but these nodes together form an island.

(1)a. Who_i does John think $[_{CP}\ \underline{t}_i'\ [_{IP}$Mary gave the book to $\underline{t}_i]]$
b.?*Who_i does John wonder $[_{CP}\text{what}_j\ [_{IP}$Mary gave $\underline{t}_j$ to $\underline{t}_i]]$

---

* The material in this paper was presented in the Spring 1993 syntax seminar at the University of Connecticut, in colloquia at the University of Rochester and Hokkaido University, and at the 1994 Japanese/Korean Linguistics Conference. It was presented also in a more preliminary form at the Spring 1992 general meeting of the English Literary Society of Japan. I would like to thank the audience, in particular, Željko Bošković, Hideki Maki, Roger Martin, and Daiko Takahashi for comments and suggestions. I also benefited from discussions with Naoki Fukui, Howard Lasnik, Keiko Murasugi, Naoko Nemoto, Keun-Won Sohn, Hiroaki Tada, and Asako Uchibori, among others.

The barrierhood of a CP/IP pair for government, on the other hand, enables us to account for contrasts such as the following:

(2)a. Mary decided [$_{CP}$[$_{IP}$PRO to go to college]]
b. *Mary decided [$_{CP}$[$_{IP}$him to go to college]]

(3)a. *John believes [$_{IP}$PRO to be a genius]
b. John believes [$_{IP}$her to be a genius]

In (2b), the verb *decided* does not govern, and hence Case-mark, the embedded subject *him* because of the intervening CP/IP pair. The example is thus excluded by the Case Filter. In (3b), on the other hand, *believes* successfully Case-marks *her* because the CP node is absent. But (3a) is excluded exactly for this reason. The matrix verb governs PRO, and the example is in conflict with the PRO theorem.

The following asymmetry in raising is explained in a similar way:

(4)a. *He$_i$ was decided [$_{CP}$[$_{IP}$t$_i$ to go to college]]
b. She$_i$ is believed [$_{IP}$t$_i$ to be a genius]

(4a) is in violation of the ECP, which requires that a trace be governed by its antecedent or by a lexical category. Because of the intervening CP/IP pair, the matrix subject *he* and the matrix verb *decided* both fail to govern the trace. Lasnik and Saito (1984) propose that a trace of subject-to-subject raising must be licensed by its antecedent, and extend the analysis to examples of "super-raising" such as (5).

(5) *She$_i$ seems [$_{CP}$that [$_{IP}$it is believed [$_{IP}$t$_i$ to be a genius]]]

She fails to license the trace, again, because of the intervening CP/IP pair.[1]

However, the efforts to sharpen and develop these analyses in the 1980's have led us to more abstract hypotheses, and consequently, to abandon the notion 'government', and together with it the special role of CP/IP pairs. For example, it is argued in Lasnik and Saito (1984) that the locality required for "government by an antecedent" has nothing to do with 'government', and is more similar to 'subjacency'. Chomsky and Lasnik

[1] This analysis is extended further and made more precise in Baker (1988) and Chomsky (1986a).

(1993), extending Rizzi (1990), propose an account for the Wh-island effect and super-raising that does not refer to CP/IP pairs. And Chomsky (1991) develops a theory of Case where structural Case is licensed only in SPEC/head configuration.

The purpose of this paper is to discuss some remaining cases where CP/IP pairs seem to block movement. These cases all involve adjunction, and hence, as will be explained below, cannot be straightforwardly explained by Chomsky and Lasnik's hypothesis on Wh-island/super-raising. In the following section, I will briefly go over the more recent analysis of (1)-(5) above that successfully eliminates the special role of CP/IP pairs. There, I will discuss in particular the important proposals by Martin (1992) and Bošković (1993) on the distribution of PRO and raising. Then, in Section 3, I will examine the locality of A-scrambling and heavy NP shift, and argue that they are blocked by CP/IP pairs. Section 4 is concerned with the licensing of the Japanese adjunct Wh-phrase *naze* 'why'. I will present evidence that it can be licensed in LF by adjoining to an argument Wh-phrase, and further, that this adjunction is blocked by CP/IP pairs. Finally, in Section 5, I will suggest a direction toward the analysis of the locality of these adjunction operations without reference to CP/IP pairs.

2. The Barrierhood of CP/IP Pairs: Explanation of the Standard Cases

2.1. Wh-island Effects

As noted above, a Wh-island is formed by a CP/IP pair. The relevant example (1b) is repeated in (6) below.

(6)?*Who$_i$ does John wonder [$_{CP}$what$_j$ [$_{IP}$Mary gave t$_j$ to t$_i$]]

This phenomenon was one of the core cases for the classical Subjacency (Chomsky (1973, 1977)), which prohibits movement across more than one bounding node. But as a more principled theory is developed by Cattell (1976), Kayne (1981) and Huang (1982), among others, the phenomenon has become rather problematic.

Let us take Huang's Condition on Extraction Domain (CED), which is stated in a somewhat simplified form in (7).

(7) An XP is a barrier for movement if it is a non-complement.

This condition accounts for the islandhood of subjects, adjuncts, and relative

clauses, as illustrated below.

(8)a.?*Who$_i$ did [[$_{NP}$a picture of t_i] please you]
(cf. Who$_i$ did [you see [$_{NP}$a picture of t_i]])
b.?*Who$_i$ did you leave the room [$_{PP}$after [you saw t_i]]
(cf. Who$_i$ do you think [$_{CP}$that [John saw t_i]])
c.?*What$_i$ did you see [the man [$_{CP}$who bought t_i]]

The subject NP in (8a) and the adjunct PP in (8b) are non-complements, and hence, are barriers for movement. The condition subsumes also the complex NP case in (8c), another core case for the classical Subjacency. Since the relative clause CP is an adjunct, the example is accounted for in exactly the same way as (8b). However, it is clear that the condition, as stated in (7), has nothing to say about the Wh-island example in (6). The embedded CP and IP are both complements, and hence, neither is a barrier for movement.[2]

Chomsky (1986a) proposes to extend Huang's CED so that it covers Wh-islands as well. The barrierhood of a CP/IP pair can be restated as follows:

(9) The CP is a barrier for movement originating within the IP.

Thus, if the IP can somehow be defined as a "non-complement," and the CED is reformulated so that not the non-complement XPs, but instead the first maximal projections dominating them, are the barriers, then the Wh-island effect is accounted for. A simplified version of Chomsky's (1986a) definition is given in (10).

(10)a. An XP is a *BC* for α $=_{df}$ (a) it dominates α, and (b) it is not a complement of a lexical head.
b. An XP is a *barrier* for α $=_{df}$ it is the first maximal projection dominating γ, where γ is a BC for α.

According to (10), the embedded IP in (6) is a BC for the Wh-phrase *who*, since it dominates the Wh-phrase (prior to movement) and it is a complement but only of a non-lexical head C. Thus, the embedded CP is a barrier for the Wh-phrase. Given (10), the subject NP, the adjunct PP,

[2] See Kayne (1981) on this point. Huang (1982) maintains the classical Subjacency, along with the CED, to accommodate the Wh-island examples.

and the relative clause CP in (6) are not barriers but BCs, and the barriers are the first maximal projections dominating them.

Although Chomsky's (1986a) proposal is quite attractive, it is clearly more desirable if the Wh-island effect can be explained on independent grounds and the complication in the definition of barriers can be dispensed with.[3] And an independent account for Wh-islands is in fact proposed in Chomsky and Lasnik (1993). They extend an idea of Rizzi (1990), and propose to account for the Wh-island effect by their Minimal Link Condition. Here, I adopt the "derivational version" of the condition, and assume that it requires movement to go through every possible landing site. (See Takahashi (1993, 1994a) for much relevant discussion.) Let us first consider the example in (11) to see the effect of the condition.

(11) [$_{CP}$What$_i$ [$_{C'}$do [$_{IP}$you [$_{VP}$think [$_{CP}$ [$_{C'}$that [$_{IP}$John [$_{VP}$bought t_i]]]]]]]]

The movement of the Wh-phrase *what* from the embedded object position to the matrix CP SPEC proceeds as follows: it first adjoins to the embedded VP, then to the embedded IP, and moves into the embedded CP SPEC. From there, it adjoins to the matrix VP, and to the matrix IP before it finally moves into the matrix CP SPEC. Note here that the embedded CP SPEC is a possible landing site for the Wh-phrase, and the movement must proceed through this position. But in the case of Wh-island examples, this step is blocked because of the presence of another Wh-phrase in this position. Hence, a movement out of a Wh-island necessarily results in violation of the Minimal Link Condition.

This account for Wh-islands extends to the super-raising example in (5), repeated below in (12a).

(12)a. *She$_i$ seems [$_{CP}$that [$_{IP}$it is believed [$_{IP}t_i$ to be a genius]]]
b. She$_i$ seems [$_{IP}t_i$' to be believed [$_{IP}t_i$ to be a genius]]

As shown in (12b), the middle subject position clearly is a possible landing site for raising. But the position is occupied by *it* in (12a), and the movement of *she* must skip this position. Hence, the example violates the

[3] In addition, as discussed in detail in Chomsky (1986a), the definition of BC in (10a) makes the status of VP problematic, since it is a complement of I and not of a lexical head.

Minimal Link Condition exactly as in the Wh-island case.[4]

The account for the Wh-island effect illustrated above enables us to go back to the simpler definition of barriers for movement in (7). And more importantly for our purpose, it eliminates the need to assume that CP/IP pairs form barriers for movement. It is now just a coincidence that (6) and (12a) both involve movement across a CP/IP pair. What blocks the movement is the filled CP SPEC in the case of (6), and the filled IP SPEC in the case of (12a).[5]

2.2. The Distribution of PRO

One of the main purposes of Chomsky (1986a) was to propose a unified definition of barriers for movement and government. Thus, the definition in (10) successfully applies to the examples in (2)-(3), repeated below in (13)-(14).

(13)a. Mary decided [$_{CP}$[$_{IP}$PRO to go to college]]
b.*Mary decided [$_{CP}$[$_{IP}$him to go to college]]

(14)a. *John believes [$_{IP}$PRO to be a genius]
b. John believes [$_{IP}$her to be a genius]

In (13), the embedded IP is a BC for the embedded subject. Hence, the embedded CP is a barrier, and it blocks the government of the embedded subject by the matrix verb. Consequently, PRO in (13a) is allowed since it is ungoverned, and *him* in (13b) is excluded as it fails to receive Case. On the other hand, the embedded IP in (14), being the complement of the matrix verb, is not a BC. Thus, the embedded subject position is governed by the matrix verb.

However, as noted above, Chomsky (1991) proposes a theory of Case,

[4] (12a) is considerably worse than (6). See Chomsky (1991), and Chomsky and Lasnik (1993) for an account for this difference in grammatical status.

[5] Chomsky and Lasnik's (1993) analysis is extended further in Takahashi (1993, 1994a). He proposes to derive all island effects from the Minimal Link Condition, and hence, to completely eliminate the notion 'barrier' from the theory of bounding.

known as the AGR-based Case theory, where structural Case is licensed only in SPEC/head configuration. Given this theory, government no longer plays any role in Case licensing, and in particular in distinguishing (13b) and (14b).[6] Then, the only remaining motivation for the notion 'government' is in the account for the distribution of PRO as in (13a) and (14a). This state of affairs raises serious doubt for the PRO theorem as a significant generalization, and for the notion 'government' itself. An alternative explanation clearly should be sought for the distribution of PRO.

A hypothesis is proposed in Chomsky and Lasnik (1993) that raises further questions on the status of the PRO theorem. As discussed in detail in Chomsky (1981, 1986b), the Case Filter, which requires every overt NP to have Case, faces some empirical problems. First, there are cases where overt NPs need not have Case. Typical examples are left-dislocated NPs (*John* in (15a)) and predicate nominals (*my best friend* in (15b)).

(15)a. John, I like him very much
b. I consider [John [$_{NP}$my best friend]]

Secondly, traces of A'-movement must have Case, despite the fact that they are empty, as shown in (16).

(16) *Who$_i$ is it believed [t_i to be a genius]

It seems then that what requires Case are not overt NPs but rather those NPs that are assigned θ-roles. This idea is incorporated in Joseph Aoun's Visibility Condition, which can be stated roughly as in (17).

(17) Only Case marked NPs can be assigned θ-roles at LF.

But if we abandon the Case Filter in favor of (17), PRO becomes a serious problem. If it is ungoverned, then it cannot be assigned Case. Yet, it clearly needs to be θ-marked. Thus, as Chomsky (1981, 1986b) notes, we are led to a strange disjunction as in (18).[7]

[6] I will come back to this contrast in Section 2.3.

[7] Expletives raise another problem. They are not assigned θ-roles, but must have Case as shown in (i).
(i) *It is believed [$_{IP}$there to be a man in the corner]
See Chomsky (1986b) for his solution (LF expletive replacement) to this

(18) Only Case marked NPs and PRO can be assigned θ-roles at LF.

Facing this problem, Chomsky and Lasnik (1993) suggest that PRO is in fact assigned Case. More specifically, they hypothesize that non-finite INFL assigns "null Case" to its SPEC, and that PRO must have this Case to be licensed. This hypothesis makes it virtually impossible to distinguish the subject positions of finite and non-finite clauses in terms of government.

Martin (1992) extends the null Case hypothesis, and argues that it accounts for the general distribution of PRO. He first presents independent evidence for the null Case hypothesis itself. For example, as noted in Lobeck (1990), and Saito and Murasugi (1990), ellipses are allowed in the complement position of a functional category only when the functional head agrees with its SPEC (in the sense of Fukui and Speas (1986)). The following examples illustrate this generalization:

(19)a. I read John's book, and now, I want to read $[_{DP}$Mary's $[_{NP}e]]$
b.*I read about that person, and now, I want to see $[_{DP}$the $[_{NP}e]]$

(20)a. I know that Mary bought something, but I don't know $[_{CP}$what $[_{IP}e]]$
b.*Mary said that she was going to LA, but I don't know $[_{CP}$whether $[_{IP}e]]$

The examples in (19) show that N'(NP)-deletion is possible only when there is a genitive NP in DP SPEC agreeing with the D head, and those in (20) indicate that sluicing is allowed only if there is a Wh-phrase in CP SPEC agreeing with the [+wh] C. But an apparent exception to this generalization is found with VP-deletion. Both finite and non-finite IPs seem to allow VP-deletion as shown in (21).

(21)a. I believe that Mary is smart, and I believe that $[_{IP}$John is $[_{VP}e]]$, too
b. John wants me to attend the meeting, but I don't want $[_{IP}$PRO to $[_{VP}e]]$

If PRO is not Case-marked and does not agree with the I head, it is a mystery why (21b) should be allowed.

Martin (1992) takes (21b) as direct evidence for the null Case

---

problem.

hypothesis. (See also Takahashi (1994b).) If PRO agrees with a non-finite I, and is assigned null Case, (21b) directly falls under the generalization on ellipses mentioned above. He shows further that this analysis receives support from examples of the ECM construction such as (22).

(22) *I believe Mary to be smart, and I believe $[_{IP}$John to $[_{VP}e]]$, too

In this example, *John* is not assigned null Case by I, but its Case is licensed by the higher verb *believe*. Hence, there is no SPEC/head agreement in the IP, and consequently, VP-deletion is not possible.

Having presented independent evidence for the null Case hypothesis, Martin goes on to argue that it by itself explains the distribution of PRO, and hence that the account in terms of government can be dispensed with. More specifically, building on Stowell's (1982) proposal that non-finite I carries tense, i.e., "unrealized future tense," in some cases but not in others, he argues that only non-finite I with tense assigns null Case. The infinitival complement of (13a), repeated in (23), has future tense.

(23) Mary decided $[_{CP}[_{IP}$PRO to go to college]]

Thus, PRO is assigned null Case as required. On the other hand, the embedded I in (14a), repeated in (24), lacks tense.

(24) *John believes $[_{IP}$PRO to be a genius]

Hence, PRO is not licensed in the embedded subject position.

Martin's account for the distribution of PRO basically eliminates the role of 'government' from the theory of syntax, and consequently, that of CP/IP pairs as barriers for government. This leaves us with the contrast in (4), repeated in (25), as the only remaining standard case where a CP/IP pair seems to function as a barrier.

(25)a. *He$_i$ was decided $[_{CP}[_{IP}t_i$ to go to college]]
b. She$_i$ is believed $[_{IP}t_i$ to be a genius]

As noted above, it was argued in Lasnik and Saito (1984) that a trace of subject-to-subject raising must be licensed by its antecedent, and that a CP/IP pair blocks this licensing. Bošković (1993) presents an independent explanation for this contrast, as shown in the next subsection.

2.3. The Locality of Raising

One of the outstanding characteristics of NP-movement is its last resort property. Thus, NP-movement can never originate in a Case position, as shown in (26).

(26)a. It seems to him that Mary is a genius
b.*He_i seems to $\underline{t}_i$ that Mary is a genius

This property is accounted for by Chomsky's (1986b) Last Resort Principle, which implies that NP-movement is possible only when it is necessary to place an NP in a Case position. The matrix subject *he* in (26b) was already in a Case position prior to movement, and thus, the movement is blocked.

Bošković (1993) points out that given the null Case hypothesis, this account for (26b) automatically extends to (25a). The embedded subject position of (25a) must be assigned null Case, since PRO can appear in this position as shown in (23). Then, the movement of *he* originates in a Case position exactly as in (26b). Hence, (25a) is excluded by the Last Resort Principle, independently of the presence of the CP/IP pair. The embedded subject position of (25b), on the other hand, is not assigned null Case. According to Martin's analysis, this is exactly why (24) is excluded. Thus, the raising in (25b) is correctly allowed.

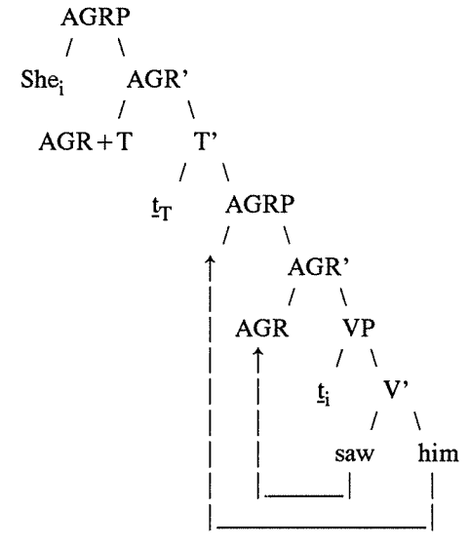
Bošković (1993) also shows that the Last Resort analysis of (26b) extends further to the contrast between (2b) and (3b), repeated below as (27a-b).

(27)a. *Mary decided [$_{CP}$[$_{IP}$him to go to college]]
b. John believes [$_{IP}$her to be a genius]

As noted above, Chomsky (1991) proposes that all structural Cases are licensed in SPEC/head configuration. According to this hypothesis, an NP with objective Case moves in LF to the SPEC position of an AGR head, to have its Case licensed, as illustrated in (28b).

(28)a. She saw him

b.



Thus, *him/her* in (27) must move covertly to the SPEC of AGR associated with the matrix verb. And this movement is blocked by the Last Resort Principle in (27a), since the NP is already in a position where null Case is assigned. Hence, the NP fails to have its objective Case licensed. (27b), on the other hand, does not face this problem.

As the Martin-Bošković analysis of (23)-(24), (25), and (27) is totally independent of the presence/absence of CP/IP pairs, a possibility arises that control complements need not be CPs but can be bare IPs, like ECM and raising complements. Bošković (1993) in fact argues for this conclusion. (See also Ormazabal (1995).) One of his arguments is based on the distribution of empty complementizers. As discussed in detail in Stowell (1981), a CP can be headed by an empty C only when it appears in a complement position. (See also Kayne (1981) and Pesetsky (1994).) Thus, (29a-b) contrast sharply.

(29)a. John believes [$_{CP}\underline{e}$ [$_{IP}$Mary is a genius]]
b.*[$_{CP}\underline{e}$ [$_{IP}$Mary is a genius]] is what John believes

But it has been known that infinitival CPs with PRO subject present exceptions to this generalization. The subject CP in (30), for example, seems to have an empty C despite the fact that it is not a complement.

(30) [$_{CP}\underline{e}$ [$_{IP}$PRO to go to college]] is what Mary decided

Examples of this kind cease to be exceptions if infinitivals with PRO subject can be bare IPs. Then, there need not be a C-projection, and hence an empty C, in (30), and we expect the example to be grammatical. Bošković, thus, concludes that clauses with PRO subject need not be CPs but can be IPs.[8]

3. A-Adjunction in Syntax

As shown above, the special role of CP/IP pairs has been successfully dispensed with in the analysis of (1)-(5). In particular, the Wh-island effect and the illicitness of super-raising are explained by the Minimal Link Condition without any reference to CP/IP pairs. The relevant examples in those cases, however, all involve substitution either to CP SPEC or IP SPEC. And it has been noted in the literature that CP/IP pairs block certain kinds of adjunction operations as well. One of the representative works is Müller and Sternefeld (1993), which examines the clause-boundedness of German scrambling in detail. In this section, I will discuss some cases in Japanese and English where CP/IP pairs seem to constitute barriers for adjunction in overt syntax. As shown in Section 5, those cases cannot be analyzed in terms of the Minimal Link Condition, and hence, they pose interesting theoretical problems.

3.1. A-scrambling in Japanese

As is well known, Japanese allows long-distance scrambling quite freely. Thus, not only (31b) but also (31c) is perfectly grammatical.[9]

(31)a. [$_{IP}$John-ga Bill-ni [$_{CP}$[$_{IP}$Mary-ga sono hon -o motteiru] to]
-nom -to -nom that book-acc have that
itta] (koto)
said fact
([John said to Bill [that [Mary has that book]]])

[8] See also Murasugi (1991), where an argument of the same form is presented to show that prenominal sentential modifiers are bare IPs in Japanese.

[9] For most of the Japanese examples, I will not provide their English translations, but instead will give their rough structures in parentheses.

b. [$_{IP}$John-ga Bill-ni [$_{CP}$[$_{IP}$sono hon -o$_i$ [$_{IP}$Mary-ga t_i motteiru]] to]
-nom -to that book-acc -nom have that
itta] (koto)
said fact
([John said to Bill [that [that book$_i$ [Mary has t_i]]]])

c. [$_{IP}$Sono hon -o$_i$ [$_{IP}$John-ga Bill-ni [$_{CP}$[$_{IP}$Mary-ga t_i motteiru] to]
that book-acc -nom -to -nom have that
itta]] (koto)
said fact
([That book$_i$ [John said to Bill [that [Mary has t_i]]]])

In (31b), the embedded object is moved to the initial position of the embedded clause by clause-internal scrambling. In (31c), it is moved to the initial position of the matrix clause by long-distance scrambling.

But as was first pointed out by Mahajan (1989) for Hindi, phrases moved by clause-internal scrambling and long-distance scrambling have different binding properties. This distinction holds in Japanese also, as discussed by Tada (1990), among others. Thus, a phrase preposed by clause-internal scrambling can serve as the antecedent of a lexical anaphor, as shown in (32).

(32)a.?*[$_{IP}$[Otagai -no sensei]-ga [$_{VP}$karera-o hihansita]] (koto)
each other-gen teacher-nom they -acc criticized fact
([[Each other's teachers] [criticized them]])

b. ?[$_{IP}$Karera-o$_i$ [$_{IP}$[otagai -no sensei] -ga [$_{VP}$$t_i$ hihansita]]] (koto)
they -acc each other-gen teacher-nom criticized fact
([Them$_i$ [[each other's teachers] [criticized t_i]]])

(32a) is out because the lexical anaphor *otagai* 'each other' does not have a c-commanding antecedent. On the other hand, (32b), where the intended antecedent *karera* 'they' is preposed, is far better. Since the lexical anaphor requires a c-commanding antecedent in A-position, this example suggests that clause-internal scrambling can be A-movement.

The examples in (33) show that long-distance scrambling, on the other hand, does not have this property.

(33)a. *[$_{IP}$[Otagai -no sensei]-ga [$_{VP}$[$_{CP}$[$_{IP}$Hanako-ga karera-o
each other-gen teacher-nom -nom they -acc
hihansita] to] itta]] (koto)
criticized that said fact
([[Each other's teachers] [said [that Hanako criticized them]]])

b. *[$_{IP}$Karera-o$_i$ [$_{IP}$[otagai -no sensei]-ga [$_{VP}$[$_{CP}$[$_{IP}$Hanako-ga t_i
they -acc each other-gen teacher-nom -nom
hihansita to] itta]]] (koto)
criticized that said fact
([Them$_i$ [[each other's teachers] [said [that Hanako criticized t_i]]]])

(33a) is out since *karera* 'they' does not c-command *otagai* 'each other'. And (33b) shows that there is no improvement even when *karera* is preposed by long-distance scrambling to a position c-commanding *otagai*. Hence, (32)-(33) indicate that only clause-internal scrambling can be A-movement.

Mahajan points out further that in more strict terms, "clause-boundedness" is not the precise description for the locality of A-scrambling. In Hindi, long-distance scrambling out of control complements can be A-movement. Nemoto (1991) confirms this generalization with Japanese examples such as the following:[10]

(34)a. *[$_{IP}$[Otagai -no sensei]-ga [$_{VP}$Hanako-ni [$_{IP}$PRO karera-o
each other-gen teacher-nom -to they -acc
hihansuru yooni] itta]] (koto)
criticize to said fact
([[Each other's teachers] [told Hanako [PRO to criticize them]]])

b. ?[$_{IP}$Karera-o$_i$ [$_{IP}$[otagai -no sensei]-ga [$_{VP}$Hanako-ni [$_{IP}$PRO t_i
they -acc each other-gen teacher-nom -to
hihansuru yooni] itta]] (koto)
criticize to said fact
([Them$_i$ [[each other's teachers] [told Hanako [PRO to criticize
t_i]]]])

[10] See also Lee (1992) and Sohn (1994a) for Korean data, and much relevant discussion. As discussed in Grewendorf and Sabel (1994), German scrambling in general seems to show the same kind of locality. See this work and the references cited there for more detailed discussion.

Then, what is the precise characterization of the locality of A-scrambling? Here, Bošković's (1993) conclusion on the categorial status of control complements suggests a plausible answer. If control complements can be bare IPs, then the scrambling in (33b) must, but those in (32b) and (34b) need not, involve movement across a CP/IP pair. Then, descriptively, a CP/IP pair seems to constitute a barrier for A-scrambling.[11]

The same kind of locality is observed with VP-adjunction scrambling. So far, we only observed cases of scrambling that involve adjunction to IP. But as shown in (35), VP is also a possible adjunction site.

(35)a. [Mary-ga [$_{VP}$John-ni hon -o watasita]] (koto)
-nom -to book-acc handed fact
([Mary [handed a book to John]])

b. [Mary-ga [$_{VP}$hon -o$_i$ [$_{VP}$John-ni t_i watasita]]] (koto)
-nom book-acc -to handed fact
([Mary [a book$_i$ [handed t_i to John]]])

And, as noted in Saito (1985), VP-adjunction scrambling seems more restricted than its IP-adjunction counterpart. For example, long-distance VP-adjunction scrambling out of a finite clause results in marginality, as shown in (36c) and (37c).

(36)a. [$_{IP}$John-ga [$_{VP}$Bill-ni [$_{CP}$[$_{IP}$Mary-ga sono hon -o motteiru]
-nom -to -nom that book-acc have
to] itta]] (koto) (= (31a))
that said fact
([John [said to Bill [that [Mary has that book]]]])

b. [$_{IP}$Sono hon -o$_i$ [$_{IP}$John-ga [$_{VP}$Bill-ni [$_{CP}$[$_{IP}$Mary-ga t_i motteiru]
that book-acc -nom -to -nom have
to] itta]]] (koto) (= (31c))
that said fact
([That book$_i$ [John [said to Bill [that [Mary has t_i]]]]])

[11] See Nemoto (1993) for more detailed discussion on this point.

c.??[$_{IP}$John-ga [$_{VP}$sono hon -o$_i$ [$_{VP}$Bill-ni [$_{CP}$[$_{IP}$Mary-ga t_i motteiru]
-nom that book-acc -to -nom have
to] itta]]] (koto)
that said fact
([John [that book$_i$ [said to Bill [that [Mary has t_i]]]]])

(37)a. [$_{IP}$John-ga [$_{VP}$Bill-ni [$_{CP}$[$_{IP}$Mary-ga sono mati-ni sundeiru]
-nom -to -nom that town-in reside
to] itta]] (koto)
that said fact
([John [said to Bill [that [Mary lives in that town]]]])

b. [$_{IP}$Sono mati-ni$_i$ [$_{IP}$John-ga [$_{VP}$Bill-ni [$_{CP}$[$_{IP}$Mary-ga t_i sundeiru]
that town-in -nom -to -nom reside
to] itta]]] (koto)
that said fact
([In that town$_i$ [John [said to Bill [that [Mary lives t_i]]]]])

c.??[$_{IP}$John-ga [$_{VP}$sono mati-ni$_i$ [$_{VP}$Bill-ni [$_{CP}$[$_{IP}$Mary-ga t_i sundeiru]
-nom that town-in -to -nom reside
to] itta]]] (koto)
that said fact
([John [in that town$_i$ [said to Bill [that [Mary lives t_i]]]]])

Exactly as in the case of A-scrambling, the relevant locality here is not strict clause-boundedness: long-distance VP-adjunction scrambling is allowed out of a control complement. (38) is perfect, and contrasts clearly with (36c) and (37c).

(38) [$_{IP}$John-ga [$_{VP}$sono hon -o$_i$ [$_{VP}$Bill-ni [$_{IP}$PRO t_i mottekuru yooni]
-nom that book-acc -to bring to
itta]]] (koto)
said fact
([John [that book$_i$ [said to Bill [PRO to bring t_i]]]])

It seems then that VP-adjunction scrambling, exactly like A-scrambling, is blocked by CP/IP pairs.

As VP-adjunction scrambling has the locality of A-scrambling, it is hypothesized in Tada and Saito (1991) that the VP-adjoined position (as a final landing site of movement) is necessarily an A-position. The anaphor binding facts are consistent with this hypothesis: a phrase preposed by long

VP-adjunction scrambling can serve as the antecedent of a lexical anaphor, as shown in (39).

(39)a. *[$_{IP}$John-ga [$_{VP}$[otagai$_i$ -no sensei]-ni [$_{IP}$PRO karera$_i$-o
-nom each other-gen teacher-to they -acc
homeru yooni] tanonda] (koto)
praise to asked fact
([John [asked [each other's teachers] [PRO to praise them]]])

b. ?[$_{IP}$John-ga [$_{VP}$karera-o$_i$ [$_{VP}$[otagai$_i$ -no sensei]-ni [$_{IP}$PRO t_i
-nom they -acc each other-gen teacher-to
homeru yooni] tanonda]] (koto)
praise to asked fact
([John [them$_i$ [asked [each other's teachers] [PRO to praise t_i]]]])

VP-adjunction scrambling, then, is movement to an A-position, and its locality is reduced to that of A-scrambling.[12]

3.2. English Heavy NP Shift

Interestingly, English heavy NP shift seems to exhibit the same locality as VP-adjunction scrambling. As discussed in Ross (1967), rightward movement in English is basically clause-bound (the right-roof constraint effect). Thus, long-distance heavy NP shift out of a finite clause results in total ungrammaticality, as the example (40) from Postal (1974) shows.

(40) *I have expected [$_{CP}$that [$_{IP}$I would find t_i]] since 1939 [$_{NP}$the treasure said to have been buried on that island]$_i$

However, Postal (1974) points out that long-distance rightward movement

---

[12] The locality of VP-adjunction scrambling is discussed in more detail in Tada (1993), Murasugi and Saito (1994), and Saito (1994a). See also Tada (1990) and Nemoto (1991, 1993) for much relevant discussion. At this point, it is not quite clear to me why (33b) is much worse than (36c) and (37c). I speculate that the latter are in fact out as examples of VP-adjunction scrambling, but can marginally be construed as involving IP-adjunction scrambling with the following structure:
(i) [$_{IP}$NP$_j$-ga [$_{IP}$XP$_i$ [$_{IP}$pro$_j$ [$_{VP}$... t_i ...]]]]
The nominative NP in (i) is in the major subject position (in the sense of Kuroda (1986)), binding a pro in the regular subject position.

is possible out of control complements. His examples are shown in (41).

(41)a. I have expected [$_{IP}$PRO to find t_i] since 1939 [$_{NP}$the treasure said to have been buried on that island]$_i$

b. I have wanted [$_{IP}$PRO to know t_i] for many years [$_{CP}$exactly what happened to Rosa Luxemburg]$_i$

Given the discussion above on scrambling, this similarity between VP-adjunction scrambling and heavy NP shift is not surprising. The latter is assumed to be a paradigm case of adjunction operation in English, and further, it has been argued that it involves VP-adjunction. I will here briefly go over two recent arguments for the latter claim. First, Frampton (1991) presents the following paradigm to show that a clause cannot be the adjunction site for heavy NP shift:

(42)a. I saw [[some of our bravest students] killed] with my own eyes
b. I saw [t_i killed] with my own eyes [some of our bravest students]$_i$
c. *I saw [t_i killed [some of our bravest students]$_i$] with my own eyes

These examples contain an embedded small clause with a heavy subject NP. In (42c), the heavy NP is adjoined to the small clause. As the heavy NP in (42b) can be adjoined to the matrix VP, the contrast between (42b) and (42c) follows if a VP, but not a clause, is a possible adjunction site for heavy NP shift. Frampton points out further that this hypothesis automatically explains examples like (43).

(43) * t_i are intelligent [all the students who can solve this problem]$_i$

If VP is the only possible adjunction site, then the movement in (43) necessarily involves lowering and produces an unbound trace.

Another argument is presented in Saito (1991). Let us first consider the following examples from Reinhart (1976):

(44)a. *He$_i$ put cigars [$_{PP}$in Ben's$_i$ box]
b. *[$_{PP}$In Ben's$_i$ box]$_j$, he$_i$ put cigars t_j

(45)a. *He$_i$ put cigars [$_{PP}$in the box that Ben$_i$ brought from China]
b. [$_{PP}$In the box that Ben$_i$ brought from China]$_j$, he$_i$ put cigars t_j

In (44a) and (45a), the pronoun *he* c-commands its antecedent. The examples are thus ruled out by Condition (C) of the Binding theory. In (44b) and (45b), the c-command relation no longer holds because of PP-preposing. The contrast between the two examples indicates that in this case, coreference is allowed when and only when the antecedent is "deeply embedded" within the moved phrase.

With this background, the following examples, also discussed by Reinhart (1976), have a direct implication for the landing site of rightward movement:

(46)a. *After days of search, they finally found him$_i$ [$_{PP}$in Dr. Levin's$_i$ hotel room]
b. After days of search, they finally found him$_i$ [$_{PP}$in a sleazy hotel room that Dr. Levin$_i$ had rented under a false name]
c. *After days of search, he$_i$ was finally found [$_{PP}$in a sleazy hotel room that Dr. Levin$_i$ had rented under a false name]

The contrast between (46a-b) is exactly like that between (44b) and (45b), i.e., coreference is possible only when the antecedent of the pronoun is "deeply embedded" within the PP. And this parallelism is expected since nothing prevents the (string-vacuous) rightward movement of the PP in (46). The configuration of (46a-b), then, will be as in (47).

(47) ... [$_{VP}$ found him$_i$ t_j] ... [$_{PP}$ ... Dr. Levin$_i$ ...]$_j$ ...

As *him* no longer c-command *Dr. Levin* after the movement, the coreference should be allowed as long as the latter is "deeply embedded" within the moved PP.

In (46c), the pronoun is in the subject position, as opposed to the object position. In this case, the coreference is impossible even when *Dr. Levin* is "deeply embedded" within the PP. This is unexpected if the PP can move out of the c-command domain of the pronoun *he*. Then, we expect the coreference to be possible in (46c), exactly as in (45b) and (46b). The example, thus, shows that a complement XP can escape the c-command domain of the object, but not that of the subject, by rightward movement. That is, VP is, but IP is not, a possible adjunction site for the rightward movement of a VP complement. Here, the generalization may appear to be that the adjunction site can in principle be a VP or an IP, but must be the smallest maximal projection dominating the moved phrase. This will exclude IP-adjunction in favor of VP-adjunction in (46). But as Postal's (41a-b) and Frampton's (42b) show, rightward adjunction is not restricted

in this way. Hence, it seems that VP as a category is a possible adjunction site for rightward movement in English, but IP is not.[13]

It was hypothesized above that the VP-adjoined position (as a final landing site) is an A-position. Hence, once it is established that heavy NP shift involves VP-adjunction, we expect it to show the same locality as other A-adjunction operations, i.e., A-scrambling and VP-adjunction scrambling. That is, we expect it to be blocked by CP/IP pairs. And this is exactly what the contrast between (40) and (41) shows.[14] Note that this analysis implies that a phrase moved by heavy NP shift can serve as the antecedent of a lexical anaphor, as it is in A-position. Although it is difficult to construct a relevant example, Daiko Takahashi (p.c., 1993) points out that the prediction is borne out by examples such as the following:

(48)a. *Mary wanted [PRO to meet [$_{NP}$the men who had been accused of the crime]] until each other's trials
b. ?Mary wanted [PRO to meet $\underline{t}_i$] until each other's trials [$_{NP}$the men who had been accused of the crime]$_i$

(48b) is not perfect, but is considerably better than (48a).

The analysis of (40)-(41) sketched above automatically extends to the contrast in (49), discussed by Rizzi (1990), and Lasnik and Saito (1992), among others:

(49)a. We believe [$_{IP}$ $\underline{t}_i$ to have good judgement] [$_{NP}$everyone who took the time to analyze the phenomenon]$_i$
b. *We believe [$_{CP}$(that) [$_{IP}$ $\underline{t}_i$ has good judgement]] [$_{NP}$everyone who took the time to analyze the phenomenon]$_i$

---

[13] See Saito (1985), Fukui and Speas (1986), and Fukui and Saito (1992) for an explanation of this generalization in terms of the X'-theory.

[14] Nishikawa (1990) also argues on the basis of locality that heavy NP shift is A-movement, and proposes further that it is movement to AGR_O SPEC. I will not adopt the latter hypothesis here since the rightward movement of PPs is subject to the same locality, as shown in (i).
(i)a. I wanted [PRO to meet t_i] very much [$_{PP}$with the person who was visiting from Philadelphia]$_i$
b. *I believed [that John met $\underline{t}_i$] sincerely [$_{PP}$with the person who was visiting from Philadelphia]$_i$

If heavy NP shift necessarily involves VP-adjunction, then the embedded subject must be adjoined to the matrix VP in both (49a-b). But the movement crosses a CP/IP pair in the case of (49b). Hence, the example is ruled out in the same way as (40).[15]

4. A-Adjunction in LF

It was shown in the preceding section that overt adjunction operations to A-positions (A-adjunction) are blocked by CP/IP pairs. In this section, I will briefly discuss a case where a covert adjunction operation in LF shows the same locality.

4.1. Wh-licensing in Japanese

As is well known, Japanese lacks syntactic Wh-movement, and examples such as (50a-b) are perfectly fine.

(50)a. John-ga nani-o katta no
-nom what-acc bought Q
(Q [John bought what])

b. John-ga naze sono hon -o katta no
-nom why that book-acc bought Q
(Q [John bought that book why])

But Huang (1982) points out that even in languages without syntactic Wh-movement, adjunct Wh-phrases, as opposed to argument Wh-phrases, exhibit island effects. Thus, argument Wh-phrases can freely appear within islands, but adjunct Wh-phrases cannot, as shown in (51)-(52).

(51)a. Mary-wa [$_{PP}$[$_{IP}$John-ga nani-o katta] kara] okotteru no
-top -nom what-acc bought since angry Q
(Q [Mary is angry [because John bought what]])

---

[15] Rizzi (1990) presents the contrast in (49) as evidence for the head-government requirement on empty categories. (See Lasnik and Saito (1992) for a similar proposal.) Thus, the alternative analysis suggested in the text, if correct, eliminates one kind of evidence for this condition.

b. *Mary-wa [$_{PP}$[$_{IP}$John-ga naze sono hon -o katta] kara]
-top -nom why that book-acc bought since
okotteru no
angry Q
(Q [Mary is angry [because John bought that book why]])

(52)a. John-wa [$_{NP}$[$_{IP}$nani-o katta] hito] -o sagasiteru no
-top what-acc bought person-acc looking-for Q
(Q [John is looking for [the person that bought what]])

b. *John-wa [$_{NP}$[$_{IP}$naze sono hon -o katta] hito] -o
-top why that book-acc bought person-acc
sagasiteru no
looking-for Q
(Q [John is looking for [the person that bought that book why]])

The Wh-phrases *nani* 'what' and *naze* 'why' are contained within an adjunct in (51) and within a complex NP in (52).

Given this generalization, Tsai (1994), for example, hypothesizes that argument and adjunct Wh-phrases are licensed in different ways. The former, being in A-position, can be licensed in situ by virtue of "unselective binding" by Q in COMP. (See Heim (1982), Pesetsky (1987), Nishigauchi (1986), and Chomsky (1992) for relevant discussion.) On the other hand, adjunct Wh-phrases, as they are in A'-position, must undergo covert movement in LF to the SPEC of the CP headed by Q to be licensed. As adjunct Wh-phrases, but not argument Wh-phrases, must move in LF, island effects are observed only with the former.

There are further restrictions on adjunct Wh-phrases in Japanese. As discussed by Watanabe (1992), among others, *naze* 'why' can follow, but cannot precede, another Wh-phrase. This is shown in (53).

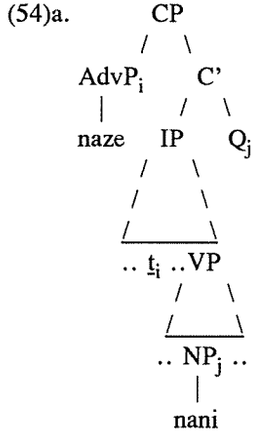
(53)a. *[$_{IP}$John-ga [$_{VP}$naze nani-o katta]] no
-nom why what-acc bought Q
(Q [John bought what why])

b. [$_{IP}$John-ga [$_{VP}$nani -o$_i$ [$_{VP}$naze t_i katta]]] no
-nom what-acc why bought Q

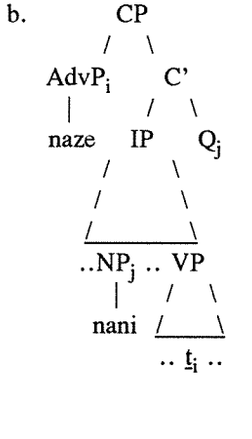
In (53a), *naze* precedes *nani*, and the sentence is out. But as shown in (53b), the sentence becomes grammatical if *nani* is scrambled to a position

preceding *naze*. This contrast shows that descriptively, an adjunct Wh cannot intervene between an argument Wh-phrase and its unselective binder. Or if we refer to the LF configuration, the trace of a non-unselectively bound Wh blocks the unselective binding of an argument Wh by Q. After *naze* moves to CP SPEC to be licensed, (53a-b) have the configurations in (54a-b) respectively.

(54)a.



b.



The trace of *naze*, which is not unselectively bound itself, intervenes between Q and *nani* in the ill-formed (54a), but not in the well-formed (54b).[16]

4.2. Additional-Wh Effects

In the preceding subsection, we observed the following generalizations:

(55)a. An adjunct Wh-phrase cannot be contained within an island.
b. An adjunct Wh-phrase cannot be followed by another Wh-phrase.

However, there are notable exceptions to these generalizations. That is, if an extra Wh-phrase is added in a position c-commanding the adjunct Wh, the examples improve considerably. The examples in (56) from Watanabe (1992) illustrate this "additional-wh" effect.

---

[16] See Watanabe (1992) and Maki (1994) for more detailed discussion of this phenomenon.

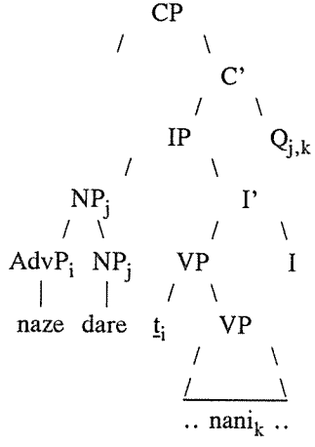
(56)a. *[$_{IP}$John-ga [$_{VP}$naze nani-o katta]] no (= (53a))
-nom why what-acc bought Q
(Q [John bought what why])

b. [$_{IP}$Dare-ga [$_{VP}$naze nani-o katta]] no
who-nom why what-acc bought Q
(Q [Who bought what why])

Both (56a) and (56b) are subsumed under (55b). Yet, the latter is perfect, and contrasts sharply with the former. The only difference between the two examples is that there is an additional Wh-phrase *dare* 'who' in the latter.

Given this "saving effect," it is proposed in Saito (1994b) that *naze* need not necessarily move to a CP SPEC in LF, but can be licensed also by adjoining to another, independently licensed, Wh-phrase. More specifically, being in A'-position, it cannot be unselectively bound by Q in situ, but it can adjoin to a Wh-phrase in A-position and be unselectively bound together with the argument Wh.[17] In the case of (56a), *naze* cannot adjoin to *nani*, since this requires lowering. Hence, it must move to CP SPEC and create the configuration in (54a). But in (56b), the adjunct Wh can adjoin to *dare*. The resulting configuration is shown in (57).

(57)



[17] The precise mechanism of the licensing of *naze* by adjunction differs in Saito (1994b), as it is assumed there that all Wh-phrases undergo LF movement to CP SPEC.

Here, as *naze* is adjoined to *dare*, it can, by hypothesis, be licensed together with this argument Wh through unselective binding by Q. The trace of *naze* still intervenes between Q and the object Wh *nani*. But as the adjunct Wh is now licensed by unselective binding, it does not block the relation between the Q and the object Wh. This trace has no effect exactly like the trace of an (unselectively bound) argument Wh in (58b) and (59b).

(58)a. [$_{IP}$John-ga [$_{VP}$dare-ni nani-o watasita]] no
-nom who-to what-acc handed Q
(Q [John handed what to whom])

b. [$_{IP}$Dare-ni$_i$ [$_{IP}$John-ga [$_{VP}$$t_i$ nani-o watasita]]] no
who-to -nom what-acc handed Q
(Q [To whom$_i$ [John handed what t_i]])

(59)a. [$_{IP}$John-ga dare-ni [$_{CP}$[$_{IP}$Mary-ga nani -o katta] to] itta] no
-nom who-to -nom what-acc bought that said Q
(Q [John said to whom [that Mary bought what]])

b. [$_{IP}$Dare-ni$_i$ [$_{IP}$John-ga t_i [$_{CP}$[$_{IP}$Mary-ga nani -o katta] to]
who-to -nom -nom what-acc bought that
itta]] no
said Q
(Q [To whom$_i$ [John said t_i [that Mary bought what]]])

The hypothesis that *naze* can be licensed by adjoining to an argument Wh receives further support from the exceptions to the generalization in (55a). As noted above, *naze* cannot be contained within an island. The relevant examples (51b) and (52b) are repeated in (60a-b).

(60)a. *Mary-wa [$_{PP}$[$_{IP}$John-ga naze sono hon -o katta] kara]
-top -nom why that book-acc bought since
okotteru no
angry Q
(Q [Mary is angry [because John bought that book why]])

b. *John-wa [$_{NP}$[$_{IP}$naze sono hon -o katta] hito] -o
-top why that book-acc bought person-acc
sagasiteru no
looking-for Q
(Q [John is looking for [the person that bought that book why]])

And an added extra Wh affects the grammaticality of these examples in exactly the way predicted by the hypothesis. First, (60a) becomes virtually perfect when we substitute *dare* 'who' for the embedded subject *John*, as shown in (61).

(61) ?Mary-wa [$_{PP}$[$_{IP}$dare-ga naze sono hon -o katta] kara] okotteru no
-top who-nom why that book-acc bought since angry Q
(Q [Mary is angry [because who bought that book why]])

This is expected since then *naze* need not move out of the adjunct island to be licensed. It can simply adjoin to *dare* and be unselectively bound with this Wh.

Secondly, (62a-b) show that there is no improvement even if we substitute *nani* 'what' for *sono hon* 'that book' in (60a-b).

(62)a. *Mary-wa [$_{PP}$[$_{IP}$John-ga naze nani-o katta] kara] okotteru no
-top -nom why what-acc bought since angry Q
(Q [Mary is angry [because John bought what why]])

b. *John-wa [$_{NP}$[$_{IP}$naze nani -o katta] hito] -o sagasiteru no
-top why what-acc bought person-acc looking-for Q
(Q [John is looking for [the person that bought what why]])

But as shown in (63a-b), when the added Wh *nani* is preposed to a position higher than *naze*, the examples improve significantly.

(63)a. ?Mary-wa [$_{PP}$[$_{IP}$John-ga nani -o$_i$ naze t$_i$ katta] kara] okotteru no
-top -nom what-acc why bought since angry Q
(Q [Mary is angry [because John bought what why]])

b.??John-wa [$_{NP}$[$_{IP}$nani-o$_i$ naze t$_i$ katta] hito] -o sagasiteru no
-top what-acc why bought person-acc looking-for Q
(Q [John is looking for [the person that bought what why]])

This is, again, predicted by the hypothesis. In (62), *naze* cannot adjoin to the added *nani*, since it must lower to do so. Hence, it must move out of the island to the matrix CP SPEC exactly as in the case of (60). But in (63), it can adjoin to the preposed *nani* and be licensed. Thus, it need not move out of an island in this case.

It was shown above that the addition of a higher argument Wh can save

an adjunct Wh *naze* in an otherwise illicit configuration. And interestingly, there is a locality requirement between the saving Wh and *naze*. Let us consider the examples in (64).

(64)a. *Mary-ga [$_{CP}$ [$_{IP}$John-ga naze nani-o katta] to] omotteru no
-nom -nom why what-acc bought that think Q
(Q [Mary thinks [that John bought what why]])

b. Mary-ga [$_{CP}$ [$_{IP}$dare-ga naze nani-o katta] to] omotteru no
-nom who-nom why what-acc bought that think Q
(Q [Mary thinks [that who bought what why]])

c.?*Dare-ga [$_{CP}$ [$_{IP}$John-ga naze nani-o katta] to] omotteru no
who -nom -nom why what-acc bought that think Q
(Q [Who thinks [that John bought what why]])

(64a) contains (53a) as the embedded clause, and is out for the same reason: *naze* precedes another Wh-phrase. The embedded subject is changed to a Wh-phrase *dare* 'who' in (64b), and the example becomes grammatical. This is so because *naze* can adjoin to *dare*, and be unselectively bound together with this Wh, exactly as in the case of (53b). (64c) is obtained from (64a) by substituting *dare* for the matrix subject, instead of the embedded subject. In this case, there is virtually no improvement in sharp contrast with (64b). If *naze* can adjoin to the matrix subject and be unselectively bound, the example should be as good as (64b), contrary to the fact. Hence, it seems that *naze* can adjoin to a higher clausemate Wh, but not to a Wh in a higher clause.

Note that the required locality between the saving Wh and *naze* is structural, and is not string adjacency. Thus, (65b) has the same status as (64c), despite the fact that the matrix subject *dare* is adjacent to *naze*.

(65)a. *Mary-ga [$_{CP}$ [$_{IP}$naze dare-ga sore-o katta] to] omotteru no
-nom why who-nom it -acc bought that think Q
(Q [Mary thinks [that who bought it why]])

b.?*Dare-ga [$_{CP}$ [$_{IP}$naze dare-ga sore-o katta] to] omotteru no
who -nom why who-nom it -acc bought that think Q
(Q [Who thinks [that who bought it why]])

In (65b), as in (64c), the higher Wh *dare* and *naze* are not clausemates. The latter, therefore, cannot adjoin to the former.

Given this locality, it is hypothesized in Saito (1994b) that adjunction to a phrase in A-position is A-movement. Then, the adjunction of *naze* to a higher argument Wh is a subcase of A-adjunction, and it should be subject to the same locality as A-scrambling and heavy NP shift. If this analysis is correct, then it is the CP/IP pair that blocks the adjunction of *naze* to *dare* in (64c) and (65b).[18]

5. The Locality of A-Adjunction

It was shown in the preceding two sections that CP/IP pairs block not only Wh-movement and NP-movement, but also some adjunction operations. It was argued in Section 3 that A-adjunction in overt syntax (A-scrambling and heavy NP shift) is blocked by CP/IP pairs. In particular, A-adjunction is possible out of control complements, but not out of finite embedded clauses. The relevant configurations are shown in (66).

(66)a. ... XP_i ... $[_{IP}$PRO ... t_i ...] ... XP_i ...
b.*... XP_i ... $[_{CP}\,[_{IP}$ ... t_i ...]] ... XP_i ...

And interstingly, the explanation for the blocking effect of CP/IP pairs elsewhere does not seem applicable to this case.

First, it was the Minimal Link Condition that explained the examples in (1b) and (5), repeated below in (67a-b). (Chomsky and Lasnik (1993))

(67)a.?*Who_i does John wonder $[_{CP}what_j\ [_{IP}$Mary gave t_j to t_i]]
b. *She_i seems $[_{CP}$that $[_{IP}$it is believed $[_{IP}t_i$ to be a genius]]]

The filled embedded CP SPEC blocks the Wh-movement of *who* to the matrix CP SPEC in (67a), and the filled middle IP SPEC blocks the raising of *she* to the matrix IP SPEC in (67b). But in the case of (66b), there is no filled CP SPEC that intervenes between the initial position and the final landing site of XP. Further, the movement cannot be blocked by a filled IP SPEC, since that will incorrectly rule out (66a) as well. The absence of the blocking effect by the IP SPEC in (66a) is in fact not surprising. Although A-scrambling and heavy NP shift are A-movement, they are at the same time adjunction operations. Hence, it is reasonable to suppose that an

[18] See Saito (1994b), Nemoto (1993), and especially Sohn (1993, 1994b) for more detailed discussion on the locality imposed on the saving Wh and *naze*.

IP SPEC is not a possible landing site for them. As neither a filled CP SPEC nor a filled IP SPEC seems to play any role in (66), the effect of the CP/IP pair there does not seem to follow directly from the Minimal Link Condition.

Secondly, the example in (4a), repeated in (68), was explained by the Last Resort Principle, which prohibits movement from a Case position to IP SPEC. (Bošković (1993))

(68) *He_i was decided $[_{CP}[_{IP}t_i$ to go to college]]

But it seems impossible to extend this condition to adjunction operations such as scrambling and heavy NP shift. Both clearly allow movement from Case positions. And further, it does not seem possible to distinguish (66a) and (66b) on the basis of the landing site: in the case of heavy NP shift, for example, the XP is adjoined to VP in both cases.[19] The effect of the CP/IP pair in (66), then, must be explained on independent grounds.

The same problem arises with the covert adjunction of *naze* 'why' to a higher Wh-phrase, discussed in Section 4. The relevant configuration is shown in (69).

(69) ... Argument Wh ... $[_{CP}[_{IP}$ ... *naze* ...]] ...
↑_____ X _____|

As in the case of (66b), the embedded CP SPEC is not filled in the relevant examples. Further, (65b) shows that the adjunction is blocked even when the embedded IP SPEC does not intervene between the two Wh-phrases. Hence, the Minimal Link Condition is not a plausible candidate for the explanation of (69). And the Last Resort Principle seems also irrelevant in this case. If the adjunction is prohibited by a condition of this kind, it should be impossible even when the locality is satisfied. Again, an independent explanation is in order.

The blocking effect of CP/IP pairs for adjunction, then, remains an interesting theoretical problem. Unfortunately, I do not have a principled solution to offer at this point. But before I conclude this paper, I would like to mention one possible direction for further research on this issue.

[19] See Tada (1990, 1993), Fukui (1993a), and Fukui and Saito (1992) for detailed discussion on the "optionality" of adjunction operations.

In Saito (1994b), two derivations are considered in the discussion of the configuration in (69). The first is the case where *naze* directly adjoins to the higher Wh. This case is assumed to be blocked by the intervening CP/IP pair. In the other derivation, *naze* first moves to the embedded CP SPEC, and then adjoins to the higher Wh. This derivation, it is argued, creates an improper chain of the form A-A'-A', as the final landing site of *naze*, by assumption, is an A-position. In other words, the proposal is that this derivation is ruled out in the same way as the classical examples of improper movement in (70), discussed in detail by May (1981).

(70)a. *John$_i$ is believed [$_{CP}$t$_i$' [$_{IP}$t$_i$ saw Mary]]
b. *John$_i$ is believed [$_{CP}$t$_i$' [$_{IP}$Mary saw t$_i$]]

In these examples, *John* moves through embedded CP SPEC on the way to the matrix IP SPEC, and hence, this NP heads an improper chain of the form A-A'-A. Similarly, Müller and Sternefeld (1993), discussing the blocking effect of CP/IP pairs for German scrambling, consider the two derivations, one direct and the other through embedded CP SPEC. They also propose that the second is ruled as as an instance of improper movement.[20]

In Murasugi and Saito (1994) and Saito (1994a), this type of analysis is carried one step further, extending the proposals of Takahashi (1993, 1994a). The main idea is that if the movement in (69), for example, can somehow be forced to go through the CP SPEC position or some other A'-position, then it can be ruled out simply as an instance of improper movement, without specific reference to CP/IP pairs as barriers. Takahashi entertains the strong derivational view on the Minimal Link Condition, illustrated in Section 2 above, and presents arguments that movement must proceed through all possible landing sites, including the adjunction positions. Given this, it is quite plausible that an adjunction operation must involve successive-cyclic adjunction to all maximal projections that intervene between the initial position and the final adjunction site of the moved phrase. Then, the contrast in (66), repeated in (71), can be attributed to whether XP moves through a CP-adjoined position or not.

(71)a. ... XP$_i$ ... [$_{IP}$PRO ... t$_i$ ...] ... XP$_i$ ...
b.*... XP$_i$ ... [$_{CP}$ [$_{IP}$ ... t$_i$ ...]] ... XP$_i$ ...

[20] See also Takahashi (1992, 1993) and Fukui (1993b) for much relevant discussion on CP/IP pairs and improper movement.

Here, by assumption, the final adjunction site of XP is an A-position. Hence, if a CP-adjoined position is necessarily an A'-position, as seems plausible, then the Minimal Link Condition forces the movement in (71b) to create an improper chain of the form A-A'-A.[21] Similarly, the movement in (69) must proceed through adjunction to the embedded CP, and create an improper A-A'-A' chain.

If this approach is on the right track, then we can eliminate the special role of CP/IP pairs also in the account for the locality of A-adjunction. And this approach seems plausible on independent grounds. The classical cases of improper movement now receive a straightforward independent account. The movement in (70), for example, involves movement from a Case position to IP SPEC, and hence, is ruled out by the Last Resort Principle. However, many new cases of improper movement have been discussed in the recent literature. Li (1990), for example, discusses the absence of the head movement of a V to a C and on to higher a V, and argues that this is an X^0 counterpart of the A-A'-A improper movement. Bošković (1994) is concerned with contrasts in NP movement such as the following:

(72) [John$_i$ can [$_{VP}$ t$_i$ swim]]

(73)a. *It [$_{VP}$believes [$_{IP}$John to seem that Peter likes Mary]]
b. *It [$_{VP}$John$_i$ believes [$_{IP}$t$_i$ to seem that Peter likes Mary]]
c. *It$_i$ is likely [$_{IP}$t$_i$ to believe [$_{IP}$John to seem that he likes Mary]]

Assuming the VP-internal subject hypothesis, he argues that the possibility of the root reading of *can* in (72) (*John is able to swim* as opposed to *It is possible that John will swim*) indicates that movement from a θ-position to another θ-position is possible. But (73a), for example, shows that an NP (*John*) cannot move from a non-θ-position to a θ-position (the matrix VP SPEC position) in Logical Form to pick up a θ-role. As shown in (73b), this type of movement cannot take place even overtly. Bošković then concludes that movement from a non-θ-position to a θ-position is ruled out

[21] It is hypothesized in Murasugi and Saito (1994) that the IP- and VP-adjoined positions as intermediate adjunction sites are ambiguous between A- and A'-positions. As noted in Section 3, the VP-adjoined position as the final landing site has the properties of an A-position. This dual nature of the VP-adjoined position is one of the main concerns of Murasugi and Saito (1994) and Saito (1994a).

as a kind of improper movement.[22]

Finally, as noted above, Müller and Sternefeld (1993) discuss the clause-boundedness of German scrambling. Relevant examples are shown in (74).

(74)a. *... daß [$_{IP}$niemand [$_{VP}$Pudding$_i$ [$_{VP}$sagt [$_{CP}$daß [$_{IP}$sie t_i mag]]]]]
that nobody pudding says that she likes
(... that nobody says that she likes pudding)

b. *... daß [$_{IP}$Pudding$_i$ [$_{IP}$niemand [$_{VP}$sagt [$_{CP}$daß [$_{IP}$sie t_i mag]]]]]
that pudding nobody says that she likes

They contrast scrambling with Wh-movement, as in (75), and conclude that the latter but not the former can use CP SPEC as an "escape hatch."

(75) Was$_i$ sagt niemand [$_{CP}$daß [$_{IP}$sie t_i mag]]
what says nobody that she likes
(What does nobody say that she likes)

Then, they argue that scrambling through CP SPEC is ruled out as it creates an improper chain. As they assume that scrambling is A'-movement, this analysis implies that A'-positions are not uniform: an adjoined position and a CP SPEC are both A'-positions but are still distinguished in chains.

The works cited above, as well as Takahashi (1992), discuss the general issue of improper movement, and present interesting hypotheses.[23] The proposed analyses are still not quite satisfactory on both conceptual and empirical grounds, as the authors themselves note. But these works clearly show that the phenomenon of improper movement is extensive, and constitutes an interesting theoretical problem. It is, thus, quite reasonable to subsume the problem of CP/IP pairs discussed in this paper under this general problem, and in fact, to take the facts of A-scrambling, heavy NP shift and LF *naze* adjunction as core cases for the investigation of improper

[22] As *John* in (73) has to move finally to the matrix AGR$_O$ SPEC to have its Case licensed, the illicit movement is more precisely from non-θ to θ to non-θ.

[23] See also Saito (1993, 1994a), Fukui (1993b), and Sakai (1994) for relevant discussion.

movement. As Japanese and Korean seem quite rich in A-adjunction, this seems to be one place where research on these languages can lead directly to a contribution to syntactic theory.

References

Baker, M. (1988) *Incorporation: A Theory of Grammatical Function Changing*, University of Chicago Press, Chicago.

Bošković, Ž. (1993) "Clausal Selection, Minimality, and Subjacency," ms. University of Connecticut.

Bošković, Ž. (1994) "D-structure, Theta-Criterion, and Movement into Theta-Positions," *Linguistic Analysis* 24.

Cattell, R. (1976) "Constraints on Movement Rules," *Language* 52.

Chomsky, N. (1973) "Conditions on Transformations," in S. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York.

Chomsky, N. (1977) "On Wh-movement," in P. Culicover, et al., eds., *Formal Syntax*, Academic Press, New York.

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris Publications, Dordrecht.

Chomsky, N. (1986a) *Barriers*, MIT Press, Cambridge.

Chomsky, N. (1986b) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.

Chomsky, N. (1991) "Some Notes on Economy of Derivation and Representation," in R. Freidin, ed., *Principles and Parameters in Comparative Grammar*, MIT Press, Cambridge.

Chomsky, N. (1992) *A Minimalist Program for Linguistic Theory*, MIT Occasional Papers in Linguistics 1.

Chomsky, N. and H. Lasnik (1993) "The Theory of Principles and Parameters," in J. Jacobs, et al., eds., *Syntax: An International Handbook of Contemporary Research*, Walter de Gruyter, Berlin.

Frampton, J. (1991) "Relativized Minimality, A Review," *The Linguistic Review* 8.

Fukui, N. (1993a) "Parameters and Optionality," *Linguistic Inquiry* 24.

Fukui, N. (1993b) "A Note on Improper Movement," *The Linguistic Review* 10.

Fukui, N. and M. Saito (1992) "SPEC/Head Agreement, X'-Compatibility, and Optionality," Colloquium, MIT.

Fukui, N. and M. Speas (1986) "Specifiers and Projection," *MIT Working Papers in Linguistics* 8.

Grewendorf, G. and J. Sabel (1994) "Long Scrambling and Incorporation," *Linguistic Inquiry* 25.
Heim, I. (1982) *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. dissertation, University of Massachusetts.
Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, Ph.D. dissertation, MIT.
Kayne, R. (1981) "ECP Extensions," *Linguistic Inquiry* 12.
Kuroda, S.-Y. (1986) "Movement of Noun Phrases in Japanese," in T. Imai and M. Saito, eds., *Issues in Japanese Linguistics*, Foris Publications, Dordrecht.
Lasnik, H. and M. Saito (1984) "On the Nature of Proper Government," *Linguistic Inquiry* 15.
Lasnik, H. and M. Saito (1992) *Move-α: Conditions on its Application and Output*, MIT Press, Cambridge.
Lee, R. (1992) "Negative Polarity Items Licensing and the Nature of Neg Projection," ms. University of Connecticut.
Li, Y. (1990) "X-zero Binding and Verb Incorporation," *Linguistic Inquiry* 21.
Lobeck, A. (1990) "Functional Heads as Proper Governors," *NELS* 20.
Mahajan, A. (1989) "On the A/A' Distinction: Scrambling and Weak Crossover in Hindi, ms. MIT.
Maki, H. (1994) "Anti-Anti-Superiority," ms. University of Connecticut.
Martin, R. (1992) "On the Distribution and Case Features of PRO," ms. University of Connecticut.
May, R. (1981) "Movement and Binding," *Linguistic Inquiry* 12.
Müller, G. and W. Sternefeld (1993) "Improper Movement and Unambiguous Binding," *Linguistic Inquiry* 24.
Murasugi, K. (1991) *Noun Phrases in Japanese and English: A Study in Syntax, Learnability and Acquisition*," Ph.D. dissertation, University of Connecticut.
Murasugi, K. and M. Saito (1994) "Adjunction and Cyclicity," *WCCFL* 13.
Nemoto, N. (1991) "Scrambling and Conditions on A-Movement," *WCCFL* 10.
Nemoto, N. (1993) *Chains and Case Positions: A Study from Scrambling in Japanese*, Ph.D. dissertation, University of Connecticut.
Nishigauchi, T. (1986) *Quantification in Syntax*, Ph.D. dissertation, University of Massachusetts.
Nishikawa, Y. (1990) "Evidence for the Existence of AgrP: English Heavy NP Shift," *English Linguistics* 7.
Ormazabal, J. (1995) *The Syntax of Complementation: On the Connection between Syntactic Structure and Selection*, Ph.D. dissertation, University of Connecticut.

Pesetsky, D. (1987) "Wh-in-Situ: Movement and Unselective Binding," in E. Reuland and A. ter Meulen, eds., *The Representation of (In)definiteness*, MIT Press, Cambridge.
Pesetsky, D. (1994) *Zero Syntax: Experiencers and Cascades*, MIT Press, Cambridge.
Postal, P. (1974) *On Raising*, MIT Press, Cambridge.
Reinhart, T. (1976) *The Syntactic Domain of Anaphora*, Ph.D. dissertation, MIT.
Rizzi, L. (1990) *Relativized Minimality*, MIT Press, Cambridge.
Ross, J. R. (1967) *Constraints on Variables in Syntax*, Ph.D. dissertation, MIT.
Saito, M. (1985) *Some Asymmetries in Japanese and their Theoretical Implications*, Ph.D. dissertation, MIT.
Saito, M. (1991) "Extraposition and Parasitic Gaps," in C. Georgopoulos and R. Ishihara, eds., *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*, Kluwer Academic Publishers, Dordrecht.
Saito, M. (1993) "Improper Adjunction: Some Notes on Res(ECP)," Colloquium, University of Rochester.
Saito, M. (1994a) "Improper Adjunction," *MIT Working Papers in Linguistics* 24.
Saito, M. (1994b) "Additional-Wh Effects and the Adjunction Site Theory," *Journal of East Asian Linguistics* 3.
Saito, M. and K. Murasugi (1990) "N'-Deletion in Japanese: A Preliminary Study," *Japanese/Korean Linguistics* 1, CSLI, Stanford.
Sakai, H. (1994) "The Uniformity Measure for Derivations," ms. University of California, Irvine.
Sohn, K.-W. (1993) "Additional Argument Effects, Scrambling, and the ECP," ms. University of Connecticut.
Sohn, K.-W. (1994a) "Negative Polarity Items and Rigidity of Scope," ms. University of Connecticut.
Sohn, K.-W. (1994b) "Adjunction to Argument, Free Ride, and the Minimalist Program," *MIT Working Papers in Linguistics* 24.
Stowell, T. (1981) *Origins of Phrase Structure*, Ph.D. dissertation, MIT.
Stowell, T. (1982) "The Tense of Infinitives," *Linguistics Inquiry* 13.
Tada, H. (1990) "Scrambling(s)," The Ohio State Workshop on Japanese Syntax.
Tada, H. (1993) *A/A' Partition in Derivation*, Ph.D. dissertation, MIT.
Tada, H. and M. Saito (1991) "VP-Internal Scrambling," Colloquium, University of Massachusetts.
Takahashi, D. (1992) "Improper Movements and Chain Formation," ms. University of Connecticut.

Takahashi, D. (1993) "Minimize Chain Links," The 11th Annual Meeting of the English Linguistics Society of Japan.

Takahashi, D. (1994a) *Minimality of Movement*, Ph.D. dissertation, University of Connecticut.

Takahashi, D. (1994b) "Sluicing in Japanese," *Journal of East Asian Linguistics* 3.

Takano, Y. (1994) "Scrambling, Relativized Minimality, and Economy of Derivation," *WCCFL* 13.

Tsai, W.-T. D. (1994) *On Economizing the Theory of A-bar Dependencies*, Ph.D. dissertation, MIT.

Watanabe, A. (1992) "Subjacency and S-Structure Movement of WH-in-situ," *Journal of East Asian Linguistics* 1.