Proceedings of the 4th GLOW in Asia 2003

GENERATIVE GRAMMAR IN A BROADER PERSPECTIVE

Edited by
Hang-Jin Yoon

Published for
The Korean Generative Grammar Circle
&
Institute for Cognitive Science, Seoul National University

Hankook

Some Remarks on Superiority and Crossing*

Mamoru Saito
Nanzan University

1. Introduction

This paper concerns the superiority and crossing phenomena, illustrated in (1) and (2) respectively.

(1) a. Who_i $\underline{t}_i$ bought what
b.*What_i did who buy $\underline{t}_i$

(2) a.?Which book_i do you know who_j to persuade $\underline{t}_j$ to read $\underline{t}_i$
b.*Who_j do you know which book_i to persuade $\underline{t}_j$ to read $\underline{t}_i$

An elegant explanation for these phenomena has been proposed recently in terms of the operation 'Attract'. (See, for example, Boskovic 1997, Kitahara 1997 and Richards 2001.) By definition, a [+Q] C attracts the closest Wh-phrase to its Spec and hence, the subject Wh is moved to this position in (1). Similarly, in (2), the embedded [+Q] C attracts the closest Wh *who* as the structure is constructed in a bottom-up fashion, leaving *which book* for the matrix [+Q] C. Thus, Attract does not even allow (1b) and (2b) to be generated. The main purpose of this paper is to explore this explanation further.

In the following section, I will discuss examples from Lasnik and Saito (1992) that indicate that superiority is relative to C, and based on them,

---

* This is a shortened version of the paper presented at the 4th Asian GLOW meeting held at Seoul National University in August, 2003. I would like to thank the participants, especially Hisa Kitahara and Akira Watanabe, as well as Jinjin In, Roger Martin, Keiko Murasugi, Koji Sanji, and Keun-Won Sohn for helpful comments. The research reported here was supported in part by the Nanzan University Pache Research Grant IA.

Proceedings of the 4th Asian GLOW in Seoul
571-595

suggest a revi-sion of the account for superiority. I will argue there that C (Q)-Wh binding takes place prior to movement, along the lines suggested in Tsai (1994(. Then, I will extend the discussion to the crossing/nesting phenomenon in English and Japanese in Section 3. The proposal there is that when C has multiple features to be checked, the order in which those features are checked is parametrized accord-ing to the language. Finally, in Section 4, I will discuss some language-specific problems that arise with the proposals in Sections 3. The first problem concerns the scrambling of Wh-phrases observed in Japanese. And the second has to do with the property of English that one and only one Wh-phrase moves overtly to CP Spec in a multiple-Wh question. I will take Richards (2001) as the starting point of the discussion in this paper, and hence, assume LF Wh-movement and also that movement to multiple Specs involves tucking in.

2. The Scope Sensitivity of Superiority Effects

The superiority condition was originally formulated as in (3) in Chomsky 1973.

(3) No rule can involve X, Y in the structure
... X ... [$_{\alpha}$ ... Z ... - W Y V ...] ...
where the rule applies ambiguously to Z and Y and Z is superior to Y.

This formulation implies that the effect is observed when two Wh-phrases take scope at the same CP. For example, in (1), *who* and *what* take the same scope and the superior Wh *who* must move to the CP Spec.

This implication is confirmed by the following examples discussed in Lasnik and Saito (1992):

(4) a. Who$_i$ t_i wonders what$_j$ who bought t_j
b. Who$_i$ t_i wonders what$_j$ you told who to read t_j

These examples do not have the ungrammatical status of (1b) despite the fact

that *what* moves across *who*. And even more interestingly, the Wh *who* in situ cannot have embedded scope but must have matrix scope. Thus, (4a) has the interpreta-tion in (5b) but not the one in (5a).

(5) a. [For which x] x wonders [for which z, y] y bought z
b. [For which x, y] x wonders [for which z] y bought z

This is exactly what is predicted by the formulation of the superiority condition in (3). Under the interpretation in (5a), *who* or *what* can move to the embedded CP Spec. Since *what* is moved over the superior *who*, the example is a superiority violation with this interpretation. On the other hand, *what* is the only Wh that can move to the embedded CP Spec under the interpretation in (5b). The matrix sub-ject *who* and the embedded subject *who* compete for the matrix CP Spec, and the superior one is actually moved. Hence, (3) allows (4a) with the interpretation in (5b).

The grammaticality and the interpretive properties of (4a-b) show that superi-ority is relative to C. That is, superiority is relevant only when two Wh-phrases take the same scope and compete for the same CP Spec position.[1] But unfortu-nately, this is not predicted by the simple, elegant explanation of superiority in terms of Attract mentioned above. Since the explanation makes no reference to the interpretation, a [+Q] C just attracts the closest Wh-phrase. Thus, the embedded C in (4a) should attract *who* in the embedded subject position regardless of the interpretation, and consequently, the example cannot even be generated. A refinement is clearly needed.

If we are to maintain the account for superiority in terms of Attract and make it relative to the scope of the Wh-phrases, it seems necessary to make C (Q)-Wh binding a prerequisite for the C to attract the Wh. Suppose that a

1. Given this, the superiority condition is formulated as in (i) in Lasnik and Saito (1992).
(i) a. A wh-phrase X in Spec of CP is Op-disjoint (operator-disjoint) from a wh-phrase Y if the assignment of the index of X to Y would result in the local A'-binding of Y by X. (S-structure)
b. If two wh-phrases X and Y are Op-disjoint, then they cannot undergo absorption.

[+Q] C "binds" all the Wh-phrases that take scope at its CP, and that a C can only attract those Wh-phrases that it binds. Then, in (4a), the embedded C binds both *who* and *what* under the interpretation (5a). In this case, it must attract the superior *who* to its Spec position. On the other hand, the C only binds *what* under the interpretation (5b). Hence, it can attract this Wh to its Spec. Thus, the grammaticality as well as the interpretive properties of (4a) are correctly predicted.

The analysis outlined above is in accord with the general idea of Tsai (1994) that C-wh (unselective) binding is a prerequisite for Wh-movement. Further, it is consistent with Richards' (2001) account for superiority in Bulgarian, a language with overt multiple Wh-movement. A typical superiority contrast in this language is shown in (6).[2]

(6) a. Koj kogo vizda
who whom see
b.*Kogo koj vizda
whom who see

Richards proposes that the C attracts the superior Wh *koj* first as in (7a), and then the other Wh *kogo* is tucked into the lower (closer) CP Spec position as in (7b).

(7) a. [$_{CP}$ Koj$_i$ [$_{TP}$ $\underline{t}_i$ vizda kogo]]
b. [$_{CP}$ Koj$_i$ [$_{CP}$ kogo$_j$ [$_{TP}$ $\underline{t}_i$ vizda $\underline{t}_j$]]

Making C-Wh binding a prerequisite for Attract does not affect this analysis. As the C with the feature Q binds both Wh-phrases prior to the application of Wh-movement, it attracts the superior *koj* first.

On the other hand, the revised account for superiority affects the analysis of crossing. Let us consider once more the examples in (2) repeated below in (8).

2. See Rudin (1988) for detailed discussion of superiority effects in Bulgarian and other Slavic languages.

(8) a.?Which book$_i$ do you know who$_j$ to persuade $\underline{t}_j$ to read $\underline{t}_i$
b.*Who$_j$ do you know which book$_i$ to persuade $\underline{t}_j$ to read $\underline{t}_i$

Recall that the analysis in terms of Attract, originally suggested by Kitahara (1997), is that the embedded C with the Q-feature attracts the closest Wh *who*, and hence, only (8a) can be generated. However, if Attract presupposes C-Wh binding as argued above, this analysis cannot be maintained as such. When the embedded C binds *which book* but not *who*, it should be able to attract the former Wh to its Spec. This movement is in fact parallel to the movement of *what* to the embedded CP Spec in (4a). Then, the matrix C should be able to bind and attract *who*, which yields the ungrammatical (8b). In the following section, I will examine crossing/ nesting effects in English and Japanese, and address this problem.

3. Crossing and Feature-Checking

As discussed in detail in Richards (2001), there is a further complication with the analysis of crossing. In English, crossing dependencies, as opposed to nesting dependencies, are banned as illustrated in (9).

(9) a.?Which book$_i$ do you know who$_j$ to persuade $\underline{t}_j$ to read $\underline{t}_i$
b.*Who$_j$ do you know which book$_i$ to persuade $\underline{t}_j$ to read $\underline{t}_i$

On the other hand, there are languages like Japanese that exhibit the opposite pattern and disallow nesting dependencies.

In this section, I will first discuss the Japanese pattern, and show that it too raises an interesting problem for the Attract account. Then, I will suggest an analysis of the phenomenon, both in Japanese and English, appealing to the mechanism of feature-checking at the embedded CP Spec.

3.1. Nesting Effects in Japanese

As noted above, Japanese exhibits an opposite pattern from English with respect to crossing/nesting effects. Let us first consider the example in (10).

(10) Taroo-wa [$_{CP}$ dare-ga nani -o katta ka] siritagatteru no
-TOP who-NOM what-ACC bought Q want-to-know Q
'[Q Taro wants to know [Q who bought what]]'

The most salient interpretation for this example is the multiple-Wh question reading of the embedded clause. In this case, the matrix clause is a yes/no ques-tion and the sentence is interpreted as in (11a).

(11) a. Does Taro want to know [for which x, y] x bought y
b. ?[For which x, y] Taro wants to know whether x bought y
c. ?[For which x] Taro wants to know [for which y] x bought y
d.*[For which y] Taro wants to know [for which x] x bought y

The example has two other marginally possible readings. First, both Wh-phrases can take matrix scope as in (11b) especially when the Wh-phrases are stressed. In addition, *dare* can take matrix scope and *nani* embedded scope as in (11c) when strong stress is placed on the former Wh. This leaves the last logically possible reading shown in (11d), where *nani* takes matrix scope and *dare* embedded scope. This interpretation is simply impossible. The generalization here is that an LF of the following form is prohibited:

(12) *[$_{CP}$ Wh$_1$ [$_{TP}$ … [$_{CP}$ Wh$_2$ [$_{TP}$ … t_2 … t_1 …] Q] …] Q]

Note that what is relevant here is the surface order (or c-command relation) of Wh$_1$ and Wh$_2$. If we scramble *nani-o* across *dare-ga* in (10) as shown in (13), the interpretation in (11d) is marginally possible while that in (11c) is totally out.

(13) Taroo-wa [$_{CP}$ nani -o$_i$ dare-ga t_i katta ka] siritagatteru no
-TOP what-ACC who-NOM bought Q want-to-know Q

Given the relevance of the surface order of the Wh-phrases, I suggested a "parsing account" along the lines of Baker (1977) in Saito (1988). Baker proposes that the English crossing effect in (9) is accounted for if a "push-down storage" mechanism is employed in the association of Wh-phrases and their traces. Suppose first that Wh-phrases are stored in the order they appear as the sentence is processed from left to right. Then, suppose that when one encounters traces, the Wh-phrases are associated with them in the reverse order starting with the last Wh stored. This will result in the associations shown in (14), allowing only nested dependencies.

(14) [$_{CP}$ Wh$_1$ [$_{TP}$ … [$_{CP}$ Wh$_2$ [$_{TP}$ … t_2 … t_1 …]]]]

The idea suggested in Saito (1988) is that the same mechanism applies to the association of Wh-phrases and Q-morphemes in Japanese. Then, the Wh-Q rela-tions can only be nested as shown in (15).[3]

(15) [$_{CP}$ [$_{TP}$ … [$_{CP}$ [$_{TP}$ … Wh$_1$ … Wh$_2$ …] Q$_2$] …] Q$_1$]

Interestingly, as illustrated in (16), this yields crossing associations between the Wh-phrases and their traces at LF.

(16) [$_{CP}$ Wh$_1$ [$_{TP}$ … [$_{CP}$ Wh$_2$ [$_{TP}$ … t_1 … t_2 …] Q$_2$] …] Q$_1$]

Thus, the interpretive properties of examples such as (10) and (13) are

3. The crossing phenomena of the dependencies between polarity items and their associated heads are investigated in detail in Tanaka (1997) and his subsequent works.

correctly predicted.

Although this "parsing" account seems to work, it is only informal and speculative. Hence, it is clearly desirable if a more precise, formal alternative can be provided.[4] And such an alternative is in fact proposed in Richards (2001). The proposed account there relies only on multiple Specs and tucking in, mechanisms that are independently motivated by the Bulgarian superiority. Let us consider the movements in (16) for a rough illustration of the proposal. One crucial assump-tion is that Japanese allows multiple Specs as landing sites for LF Wh-movement. Then, Wh_1, in addition to Wh_2, moves to the embedded CP Spec on its way to the matrix CP, yielding the structure in (17).

(17) $[_{CP} [_{TP} \ldots [_{CP} Wh_1\ Wh_2 [_{TP} \ldots \underline{t_1} \ldots \underline{t_2} \ldots]\ Q_2] \ldots]\ Q_1]$

Here, Wh_1 occupies the higher CP Spec, since it is attracted by the embedded C first and Wh_2 is later tucked into the lower (closer) CP Spec. Then, when the matrix C attracts a Wh-phrase, it must be Wh_1 since it is the closer Wh. Thus, this account correctly predicts that only crossing dependencies as in (16) are allowed in the case of Japanese.

This analysis is clearly more attractive than the speculative "parsing" account. Yet, it faces the same problem as the Attract account for crossing in English briefly discussed above. Recall that we have reached the conclusion in Section 2 that C-Wh binding is a prerequisite for a C to attract a Wh-phrase. Then, if the matrix C in (17) binds Wh_2 and not Wh_1, it should be able to attract Wh_2 to its Spec. Thus, basically the same problem arises with the analysis of Japanese nesting effect and with the analysis of English crossing effect.

4. The main goal of Saito (1988) was in fact not to propose a principled account for crossing/ nesting, but to show that the general phenomenon does not straightforwardly follow from Pesetsky's (1982) Path Containment Condition, and thereby to offer indirect support for the classical ECP, as developed in Chomsky (1981) and Huang (1982).

3.2. Feature Checking with Multiple CP Specs

Richards' analysis of nesting effects, as noted above, is quite attractive. Let us then consider the structure in (17) again and examine what revision is required to accommodate this example. If the embedded C attracts both Wh-phrases, then the superior Wh_1 must be attracted first and move into the higher CP Spec, as the configuration is overtly attested in Bulgarian multiple-Wh questions. But as noted above, this does not necessarily make the matrix C attract Wh_1. If it binds only Wh_2, it should be able to attract this Wh-phrase instead. Then, what is needed is an independent way to force the matrix C to attract Wh_1.

Here, the details of the feature-checking that takes place in the embedded CP Spec seem helpful. Note that only one of the two Wh-phrases takes scope at this position. This Wh must check the Q-feature of C. On the other hand, the CP Spec position is only an intermediate landing site for the other Wh. What feature does this Wh check at this position? It is suggested in Chomsky 1998 that Wh-movement to an intermediate CP Spec is triggered by the P(eripheral)-feature that is freely assigned to C. Thus, the embedded C attracts the Wh-phrase *what* to its Spec because of its P-feature in (18).

(18) $[_{CP}$ What$_i$ do $[_{TP}$ you think $[_{CP}\ \underline{t_i}'\ [_{TP}$ John bought $\underline{t_i}]]]]$

It seems reasonable, then, to suppose that any Wh that moves to an intermediate CP Spec, including one of the Wh-phrase in (17) checks a P-feature. If this is correct, the embedded C in (17) must have two features to be checked, Q and P.

The nesting effect then follows if the first Wh that moves into the CP Spec must check the P-feature. The embedded C binds and attracts Wh_1 and Wh_2 to its Spec. By definition, it attracts Wh_1 first because it is superior to Wh_2. If the first Wh-phrase that moves into the CP Spec checks the P-feature, then Wh_1 checks P and Wh_2 checks Q. Thus, only Wh_1 can move on to the matrix CP Spec.

Although this account for the nesting effect relies on a stipulation on the order in which P- and Q-features are checked, it provides a means to account for the crossing effect in English as well. Recall that the simple account for the latter effect in terms of Attract can no longer be maintained as such. Let us consider the crossing configuration in (19).

(19) $[_{CP}$ Wh_1 $[_{TP}$ … $[_{CP}$ Wh_2 $[_{TP}$ … $\underline{t}_1$ … $\underline{t}_2$ …]]]]

If the embedded C binds Wh_2 and the matrix C Wh_1, nothing prevents the structure to be generated at this point. But suppose that English allows multiple CP Specs and that a C may have both P- and Q-features exactly as in Japanese.[5] Then, both Wh-phrases move into the embedded CP Spec as shown in (20).

(20) $[_{CP}$ $[_{TP}$ … $[_{CP}$ Wh_1 Wh_2 $[_{TP}$ … $\underline{t}_1$ … $\underline{t}_2$ …]]]]

In the case of English, it must be Wh_2 and not Wh_1 that moves into the matrix CP Spec. Hence, crossing effects, instead of nesting effects, are observed. And this is accounted for if the first Wh into the CP Spec checks the Q-feature leaving the P-feature for the second Wh in this language.

According to this account, the order in which P- and Q-features are checked varies among languages. Informally put, the more "prominent" Q-feature is checked first in English but last in Japanese. The rough idea can be extended also to the antisuperiority effect in Japanese as well. In multiple-Wh questions like (21), the C with a Q-feature binds both *who* and *what*.

(21) a. Who_i $\underline{t}_i$ bought what
b.*$What_i$ did who buy $\underline{t}_i$

This assumption is necessary in order to force C to attract the superior (closer) Wh *who*. In (21a), C attracts *who* overtly and *what* covertly. It seems

5. I suspect that the latter option is only marginally allowed, and that this is responsible for the mild deviance of Wh-island violations in both languages.

then that both Wh-phrases check the Q-feature. We can make the mechanism here more precise by postulating two sub-features of Q, say Q-primary and Q-secondary, that are checked by the Wh-phrases. The primary sub-feature of $Q_{<P,S>}$ is checked by the first Wh in CP Spec and the secondary sub-feature by the second Wh.

Let us next consider the following examples in the light of this mechanism:

(22) a. Why_i did you buy what $\underline{t}_i$
b.*Who_i $\underline{t}_i$ left early why

Neither of these examples violate superiority since the superior (higher) Wh is moved to CP Spec in both. Yet, (22b) is totally ungrammatical as discussed in detail in Huang (1982). The contrast shows that adjunct Wh-phrases like *why* have a peculiar property that requires them to check the primary sub-feature of Q.

Given this property of adjunct Wh-phrases, the following contrast suggests that feature-checking in Japanese indeed takes place in the opposite order from English:[6]

(23) a. [Taroo-ga nani -o naze katta ka] osiete kudasai
-NOM what-ACC why bought Q tell-me please
'Please tell me why Taro bought what'
b.*[Taroo-ga naze nani -o katta ka] osiete kudasai
-NOM why what-ACC bought Q tell-me please

In the derivation of the LF of (23a), the superior Wh *nani-o* must move to the CP Spec first. The adjunct Wh *naze* then moves to the CP Spec only after *nani-o* has checked a sub-feature of Q. Yet, the example is grammatical. If adjunct Wh-phrases must check the primary sub-feature of Q, this implies that unlike in the case of English, the first Wh in CP Spec checks the secondary sub-feature of Q in Japanese.

6. The phenomenon in (23) is called antisuperiority. See Saito (1988), Watanabe (1992), and the references cited there for detailed discussion of this phenomenon.

This hypothesis automatically explains the ungrammaticality of (23b). Since *naze* is superior to *nani-o*, it must move to the CP Spec first and check the secondary sub-feature of Q. Hence, it fails to fulfill the requirement on adjunct Wh-phrases. If the primary sub-feature of Q is more "prominent" than the secondary sub-feature, the comparison between (22) and (23) thus leads to con-firming evidence that the more prominent feature is checked first in English but last in Japanese.

4. The Interaction of Scrambling and LF Wh-Movement

Thus far, I have suggested revisions in the accounts in terms of Attract for the superiority and crossing/nesting phenomena. In this section and the next, I will consider more complex cases that involve the interaction of overt movement and covert movement. I will first discuss the interaction of scrambling and LF Wh-movement in Japanese in this section. Then, in Section 5, I will examine English multiple-Wh questions more closely.

4.1. Radical Reconstruction

The surface positions of the Wh-phrases were crucial in the account for the nesting effect in Japanese. Let us consider again (13) repeated below as (24).

(24) Taroo-wa [$_{CP}$ nani -o$_i$ dare-ga t_i katta ka] siritagatteru no
-TOP what-ACC who-NOM bought Q want-to-know Q

Here, *nani-o* is scrambled over *dare-ga*, and this makes the interpretation in (25a) possible while blocking the interpretation in (25b).

(25) a.?[For which y] Taro wants to know [for which x] x bought y
b.*[For which x] Taro wants to know [for which y] x bought y

This fact apparently contradicts the well-known radical reconstruction proper-ty of scrambling. Let us consider the examples in (26).

(26) a. [$_{TP}$ Taroo-ga [$_{CP}$ [$_{TP}$ Hanako-ga dono hon -o yonda] ka]
-NOM -NOM which book-ACC read Q
siritagatteiru] (koto)
want-to-know fact
'Taro wants to know [Q [Hanako read which book]]'
b. Dono hon -o$_i$ [$_{TP}$ Taroo-ga [$_{CP}$ [$_{TP}$ Hanako-ga t_i yonda] ka]
which book-ACC -NOM -NOM read Q
siritagatteiru] (koto)
want-to-know fact
'[Which book$_i$, Taro wants to know [Q [Hanako read t_i]]]'

In (26a), the Wh-phrase *dono hon-o* moves in LF to the embedded CP Spec, where it takes scope. In (26b), the Wh is scrambled out of the embedded CP and yet the example is only slightly marginal. If we apply LF Wh-movement directly to the example, the Wh-phrase must "lower" to the embedded CP Spec. The example, then, is predicted to be ungrammatical.

Given this fact, I proposed in Saito (1989) that scrambling is "semantically vacuous" and can be "undone" prior to the application of LF Wh-movement.[7] This hypothesis provides a straightforward explanation for the grammaticality of (26b). But if it is correct, *nani-o* in (24) must also be able to move back to the position of its trace before LF Wh-movement applies. It is then incorrectly predicted that (24) allows the interpretation in (25b). In this subsection, I will briefly discuss the precise mechanism for radical reconstruction proposed in Saito (2003). Then, in Section 4.2, I will show that the mechanism makes it possible to dissolve the apparent contradiction between the nesting phenomenon and the radical reconstruction property of scrambling.

The analysis of the semantically vacuous nature of scrambling proposed in Saito (2003) relies on the general mechanism of chain interpretation. It is proposed in Chomsky (1993) that movement involves copy and deletion. Thus, for (27), the Wh-phrase is initially copied in CP Spec as in (28a).

7. This "undoing" property of scrambling was later named its radical reconstruction pro-perty.

(27) Whose brother$_i$ did you see $\underline{t}_i$

(28) a. [$_{CP}$ [Whose brother] did [$_{TP}$ you see [whose brother]]]

b. [$_{CP}$ [Who] did [$_{TP}$ you see [x's brother]]]

Then, deletion applies as in (28b), retaining only the operator in CP Spec and turning the corresponding part of the object NP into a variable.

This mechanism can be restated in terms of features as illustrated in (30) for (29).

(29) What$_i$ did John buy $\underline{t}_i$

(30) a. What did John buy what
[π,O,D] [π,O,D]

b. What did John buy what
[π,O,~~D~~] [~~π~~,~~O~~,D]

The Wh-phrase *what*, for example, should consist of phonetic (π features, the operator (O) feature, and the D feature.[8] According to the copy and deletion analysis, these features should all be copied in CP Spec as in (30a). Then, deletion of features should apply as in (30b). The π features should be retained at the head position of the chain. This is virtually what overt movement means. The O feature should be in the CP Spec so that the Wh-phrase in this position is interpreted as an operator. And finally, it is plausible that the D feature of the Wh yields its interpretation as a variable in the object position. Then, more generally, the deletion of features seems to be governed by (31a-b).[9]

(31) a. Phonetic features are retained at the head of the chain.

b. Other features are retained in positions where they are selected.

8. Here, I assume that Φ-features and the Case feature are to be treated together with the categorial feature D.

9. (31b) is adapted from Lee's (1994) proposal that only feature-checking positions are retained at LF. See Saito (2001) for a more general discussion of feature-deletion in chains.

The feature-based interpretation of chains suggested above immediately cap-tures the semantically vacuous nature of scrambling. Let us consider the simple example in (32) for the illustration of the point.

(32) Karera-o$_i$ Taroo-ga $\underline{t}_i$ hometa (koto)
they -ACC -NOM praised fact
'Taro praised them'

Because the scrambled phrase lacks an operator feature, only its π features and D feature are copied at the sentence-initial position as in (33a).

(33) a. Karera-o Taroo-ga karera-o hometa
[π,D] [π,D]

b. Karera-o Taroo-ga Karera-o hometa
[π,~~D~~] [~~π~~,D]

The π features are retained at the landing site since the movement is overt. And if scrambling is not feature-driven, as argued for in Saito and Fukui (1998), for example, the D feature is selected at the object position but not at the landing site. Hence, it is retained at the object position according to (31b). The resulting chain in (33b) shows that the representation that scrambling creates is indistinguishable from PF movement.

It is argued in Saito (2003) that this mechanism of chain interpretation, when applied cyclically or phase by phase, also explains the binding properties of the scrambled phrase. As is well known since Mahajan's (1990) work on Hindi, phrases preposed by clause-internal scrambling can serve as the antecedent of an anaphor, but not those preposed by long scrambling. The following examples illustrate this point:

(34) a.?[$_{TP}$ Karera-o$_i$ [[otagai -no sensei] -ga $\underline{t}_i$ hihansita]] (koto)
they -ACC each other-GEN teacher-NOM criticized fact
'Them$_i$, [each other's teachers] criticized $\underline{t}_i$'

b.*[$_{TP}$ Karera-o$_i$ [[otagai -no sensei] -ga [$_{CP}$ [$_{TP}$ Tanaka-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 Tanaka criticized t$_i$'

The scrambling chain in (34a) is interpreted as in (35a).

(35) a. [Karera-o … [… otagai … karera-o …]]]
[π,~~D~~] [~~π~~,D]
b. [Karera-o … [… otagai … [$_{CP}$ karera-o [… karera-o …]]]]
[π] [~~π~~,~~D~~] [~~π~~,D]

Here, the D feature of *karera-o* appears in a position c-commanding *otaga*i at one point of the derivation, i.e., before it is deleted at the landing site. Hence, if Condition (A) is an anywhere condition, as argued in Belletti and Rizzi (1988), and Lebeaux (1988), the grammaticality of (34a) is correctly accounted for.

On the other hand, the scrambling in (34b) involves extraction out of a CP and hence must proceed through the edge of this CP. If chain interpretation takes place phase by phase, then the mechanism should apply first at the point the scrambled phrase moves into the embedded CP Spec.[10] Then, the D feature of *karera-o* is deleted in this CP spec, as indicated by '~~D~~' in (35b). In the matrix clause, only the phonetic features of *karera-o* are copied at the sentence-initial position and they are later deleted in the embedded CP Spec as indicated by '~~π~~'. Thus, the ungrammaticality of (34b) is also correctly predicted because the D feature of *karera-o* c-commands *otagai* at no point of the derivation of the exam-ple.

4.2. The [+Wh] Feature and Nesting Effects

The mechanism of chain interpretation illustrated above captures the

10. Here, I will focus on the CP phase. For discussion of the status of the *v*P phase in this context, see Saito (2003).

seman-tically vacuous nature of scrambling and also the binding properties of scrambled phrases. But since I have only discussed those syntactic features that are selected at specific positions, i.e., operator features and categorial features, questions remain with other features such as [+anaphor] and [+wh]. In Saito (2003), I hypothesized that [+anaphor] is not selected at any specific position and hence, that it can be retained at any position of the chain. There, the following examples from Dejima (1999) was presented as evidence:

(36) a. Taroo-ga$_i$ [$_{CP}$ Hanako-ga$_j$ [$_{CP}$ Ziroo-ga$_k$ zibunzisin-o$_{*i,*j,k}$
-NOM -NOM -NOM self -ACC
hihansita to] itta to] omotteiru (koto)
criticized that said that think fact
'Taro thinks that Hanako said that Ziro criticized self'
b. Taroo-ga$_i$ [$_{CP}$ zibunzisin-o$_{i,j,k}$ Hanako-ga$_j$ [$_{CP}$ Ziroo-ga$_k$ t
-NOM self -ACC -NOM -NOM
hihansita to] itta to] omotteiru (koto)
criticized that said that think fact

Whether *zibunzisin* should be classified as an anaphor is controversial.[11] But as (36a) shows, it clearly prefers a local antecedent. And when it is preposed, the range of its possible antecedents increases as shown in (36b). Intuitively, it can take *Ziroo* as its antecedent in the initial position, *Hanako* at the most deeply embedded CP Spec, and *Taroo* at its final landing site. This is in fact expected under the hypothesis that [+anaphor] can be retained at any position of the chain. When *Taroo* is the antecedent in (36b), for example, we may say that the feature is retained at the landing site as in (37).

(37) [Taroo-ga … [zibunzisin-o … [zibunzisin-o … [… zibunzisin-o …]]]]
[π,D] [π,+anaphor] [~~π~~,~~+anaphor~~,~~D~~] [~~π~~,~~+anaphor~~,D]

Another way to state this analysis would be to say that [+anaphor] is

11. See, for example, Nakamura (1996) and the references cited there for relevant discussion.

retained in the highest position of the chain where it is (or can be) locally bound by the antecedent. The rough idea is that the feature is retained at the highest position of the chain until it is locally bound by the antecedent, and once it is bound, it is retained at that position. This will make it possible to unify the conditions for the selected features, O and D, and for [+anaphor]. Suppose that features such as O and D are licensed by selection, and those like [+anaphor] are licensed by local binding. Then, (31) can be restated more generally as in (38).

(38) a. Phonetic features are retained at the head of the chain.
b. Other features are retained at the highest positions where they are or can be licensed.

Further, it can be extended straightforwardly to the feature [+wh] of Wh-phrases. Suppose that [+wh] is licensed by virtue of binding from C, or more pre-cisely at the highest position of the chain where it is (or can be) bound by a [+Q] C. Then, (38b) makes a precise prediction about where this feature is retained in a chain. Consider the example in (39a) for an illustration:

(39) a. Nani-o$_i$ Taroo-ga t_i katta no
what-ACC -NOM bought Q
'What did Taro buy'
b. [$_{CP}$ [$_{TP}$ Nani-o [Taroo-ga nani-o katta]] no]
[π,+wh,~~D~~] [~~π~~,~~+wh~~,D] [Q]

In (39b), the feature [+wh] is bound by C at the landing site. Hence, it is retained there and deleted at the object position.[12]

This analysis of the [+wh] feature provides a straightforward solution to the problem posed by the nesting effect and radical reconstruction. Let us consider first the radical reconstruction example in (26b), repeated in (40).

12. Here, I assume that the binding takes place as soon as the relevant C is introduced into the structure. Since (39b) constitutes an input to LF, the analysis implies that LF Wh-movement should be feature movement. (cf. Chomsky 1994)

(40)?Dono hon -o$_i$ [$_{TP}$ Taroo-ga [$_{CP}$ [$_{TP}$ Hanako-ga t_i yonda] ka]
which book-ACC -NOM -NOM read Q
siritagatteiru] (koto)
want-to-know fact
'[Which book$_i$, Taro wants to know [Q [Hanako read t_i]]]'

As the most deeply embedded CP is generated, the [+wh] feature of *dono hon-o* is bound by a [+Q] C as in (41a).

(41) a. [$_{CP}$ [$_{TP}$ Taroo-ga dono hon-o yonda] ka]
[π,+wh,D] [Q]
b. [$_{CP}$ dono hon-o [$_{TP}$ Taroo-ga dono hon-o yonda] ka]
[π,~~+wh~~,~~D~~] [~~π~~,+wh,D] [Q]

Thus, after the Wh-phrase is copied at the edge of the CP, the feature is deleted in the CP Spec along with the D feature. The subsequent movement does not in-volve the feature [+wh], and hence, is irrelevant for the interpretation of the Wh. There is no lowering in the LF Wh-movement, now construed as the movement of the feature [+wh], because it originates in the object position.

Let us next consider the example of the nesting effect in (24), repeated below in (42).

(42) Taroo-wa [$_{CP}$ nani -o$_i$ dare-ga t_i katta ka] siritagatteru no
-TOP what-ACC who-NOM bought Q want-to-know Q
'[Q Taro wants to know [Q whati who bought ti]]'

The embedded CP is as in (43).

(43) [$_{CP}$ [$_{TP}$ nani-o dare-ga nani-o katta] ka]
[π,+wh,~~D~~] [π,+wh,D] [~~π~~,~~+wh~~,D] [Q]

For the licit interpretation in (44a), the embedded [+Q] C must only bind *dare-ga*.

(44) a.?[For which y] Taro wants to know [for which x] x bought y
b.*[For which x] Taro wants to know [for which y] x bought y

Since *nani-o* is not yet bound, its [+wh] feature is retained at the highest position of the chain. The illicit (44b) would also yield the structure in (43). In this case, the embedded [+Q] C binds the two instances of *nani-o* and the [+wh] feature is retained in the highest position of the chain. Thus, in both cases, the [+wh] feature of the scrambled Wh-phrase is retained at the surface position.

Given this result, the analysis proposed above for the nesting effect can be maintained as such. When LF Wh-movement takes place, the embedded C is as-signed a P-feature and as the result, binds the other Wh-phrase as well. It attracts the closest Wh *nani-o* first, and by hypothesis, this Wh-phrase checks the P-feature. This is consistent with the interpretation in (44a), but not with that in (44b).

5. The Interaction of Overt and Covert Wh-Movement

Since all instance of Wh-movement are covert in Japanese, its analysis is rela-tively straightforward. But it was shown in the preceding section that interesting problems arise when we consider its interaction with scrambling. I suggested a way to solve this problem based on the analysis of scrambling proposed in Saito (2003). The situation in English is potentially more complex because there are both overt and covert instances of Wh-movement in the language. In this last section, I will discuss one complication that arises in the account of superiority.

Let us first consider (4a) again, which is repeated below as (45).

(45) Who$_i$ $\underline{t}_i$ wonders what$_j$ who bought $\underline{t}_j$

As was pointed out in Section 2, this example allows the interpretation in

(46).

(46) [For which x, y] x wonders [for which z] y bought z

This follows from the proposals made so far in this paper. In overt syntax, the embedded [+Q] C binds *what* and attracts the Wh to its Spec. Then, the matrix [+Q] C binds the matrix subject *who* as well as the embedded subject *who*, and attracts the former to its Spec. In LF, the embedded C is assigned a P-feature and because of this, it is able to bind and attract the embedded subject *who*. The Wh eventually moves on to the matrix CP Spec.

A complication starts to arise when we consider examples with further embed-ding as in (47).

(47) ??Who$_i$ $\underline{t}_i$ knows what$_j$ John think that who bought $\underline{t}_j$

As Roger Martin (p.c.) points out, this example is marginal but has the same interpretive properties as (45). That is, the example allows the matrix interpretation of the embedded subject *who*. Here, the important difference between (47) and (45) is that the movement of *what* across *who* is not to its scope position in the case of (47). More specifically, *what* must initially move to the most deeply embedded CP Spec and check the P-feature assigned to the C. In order to allow this movement, it seems necessary to assume that P, like Q, can select the Wh-phrase it binds and attract the bound Wh-phrase over another.

But this assumption leads to further complication with superiority violations like (48).

(48)*What$_i$ does John think that who bought $\underline{t}_i$

If the most deeply embedded C with a P-feature can bind and attract *what* over *who* in (47), nothing at this point prevents the embedded C in (48) from doing the same. And once *what* is in the embedded CP Spec, the matrix C should be able to attract the Wh to its Spec as it is superior at this point to the other Wh *who*.

The difference between (47) and (48) lies in whether *what* and the embedded subject *who* take the same scope. The clear generalization here is that a Wh-phrase cannot move over another Wh-phrase if they are to be interpreted at the same CP Spec, as already stated Chomsky's (1973) formulation of superiority in-troduced at the outset of this paper. I have tried to incorporate this insight by marking the scope of Wh-phrases by C-Wh binding prior to movement and by making this binding a condition for the application of Attract. But the contrast between (47) and (48) indicates that it is necessary to do this for movement of Wh-phrases not only to their scope positions but also to intermediate landing sites.

In Section 3, I suggested that when two Wh-phrases take scope at the same CP Spec, the C has a single Q-feature with two sub-features Q_P and Q_S, and binds both Wh-phrases. In English, Q_P is checked overtly and Q_S covertly at LF. If this idea is extended to movement to intermediate CP Specs, then the assumptions in (49) can be entertained.

(49) Suppose that two Wh-phrases take the same scope. Then, when they move to the same CP Spec, they must check the same Q/P feature. More speci-fically, the C head must bind them simultaneously and one of the Wh-phrases checks the sub-feature Q_P/P_P while the other checks Q_S/P_S.

Technically, (49) correctly distinguishes between (47) and (48). It allows the movement of *what* over the embedded subject *who* in (47) because the two Wh-phrases take distinct scope. In the overt syntax, the embedded C with a P-feature can bind and attract *what*. Then, another P-feature can be assigned to this C in LF, allowing it to bind and attract *who*. On the other hand, the two Wh-phrases in (48) must check the same P-feature at the embedded CP Spec in the case of (48) because they take the same scope. Then, when a P-feature is assigned to C in overt syntax, the C must bind both Wh-phrases and consequently attract the superior *who*.

Although the extension of "scope marking" to the P-feature solves the problem posed by (47)-(48), the analysis is far from principled as it is stated. Since the C-Wh binding with Q-feature is indeed scope marking, it has a semantic basis. Further, if Q-feature is construed as a scope marker, it makes sense for two Wh-phrases to check a single Q-feature when they take the same scope. On the other hand, P-feature, as it is currently understood, is a purely syntactic device to make movement to an intermediate CP Spec possible. Hence, it is not clear why it should be subject to the restrictions stated in (49).

Thus, even if we maintain (49), it is still necessary to investigate why P-feature has virtually the same properties as Q-feature in the relevant respects. I must leave this for future research, hoping that further investigation of the nature of P-feature will provide the relevant insight.

6. Conclusion

In this paper, I have tried to extend the analysis of superiority and crossing based on Attract. In Section 2, I pointed out that superiority is relative to the scope of the Wh-phrases, and incorporated this property of superiority into the analysis by making C-Wh binding a prerequisite for the application of Attract. In Section 3, I discussed the consequences of this proposal for the analysis of crossing/nesting effects. There, I proposed that the order in which Q- and P-features are checked is parametrized and that this is responsible for the variation between crossing and nesting among languages. The proposals in these sections build on Richards (2001), and I believe they constitute a coherent analysis, at least for the limited range of data considered in this paper.

Then, in Sections 4 and 5, I examined the interaction of overt and covert instances of movement. In Section 4, I considered the apparent contradiction between the proposed analysis of nesting effects in Japanese and the semantically vacuous nature of scrambling. I argued that the contradiction may indeed be only apparent based on a closer examination of the radical reconstruction property of scrambling. In Section 5, I pointed out and discussed a remaining problem with the account of superiority effects in

English.

As noted at the outset of this paper, the recently proposed, purely syntactic analysis of superiority and crossing/nesting in terms of Attract is elegant and appealing. On the other hand, the discussion in this paper suggests that "interpre-tation" may be more involved in the phenomena. The pursuit of the Attract account to cover wider range of data should eventually make the relation and the division of labour between syntax and semantics clearer. This paper is intended to contribute to this endeavour. At the same time, it still remains to be seen whether independent evidence can be found for the mechanism of C-Wh binding, which was proposed to serve the role of bridging the Attract account and interpretation.

References

Baker, C. L. 1977. Comments on the paper by Culicover and Wexler. In P. Culicover, et al (eds.), *Formal syntax*. Academic Press, New York.

Belletti, A. and L. Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory* 6:291-352.

Boskovic, Z. 1997. On certain violations of the superiority condition, AgrO, and economy of derivation. *Journal of Linguistics* 33: 227-252.

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. 1981. *Lectures on government and binding*.Dordrecht: Foris Publications.

Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds.), *The view from building 20*. Cambridge, Mass: MIT Press.

Chomsky, N. 1994. Bare phrase structure. MIT Occasional Papers in Linguistics 5.

Chomsky, N. 1998. Minimalist inquiries: The framework. MIT Occasional Papers in Linguistics 15.

Dejima, M. 1999. A study of remnant-scrambled sentences in Japanese and their LF representations. M.A. thesis, Nanzan University.

Huang, C.-T. J. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.

Kitahara, H. 1997. *Elementary operations and optimal derivations*, Cambridge, Mass.: MIT Press.

Lasnik, H. and M. Saito. 1992. *Move-*. Cambridge, Mass.: MIT Press.

Lebeaux D. 1988. Language acquisition and the form of grammar. Doctoral dissertation, University of Massachusetts.

Lee, R. K. 1994. Economy of representation. Doctoral dissertation, University of Connecticut.

Mahajan, A. 1990. The A/A' distinction and movement theory. Doctoral dissertation, MIT.

Nakamura, M. 1986. *Sokubaku kankei [Binding relations]*, Hitsuji Shobo, Tokyo.

Pesetsky, D. 1982. Paths and categories. Doctoral dissertation, MIT.

Richards, N. 2001. *Movement in language*. Oxford: Oxford University Press.

Rudin, C. 1988. On multiple questions and multiple Wh-fronting. *Natural Language and Linguistic Theory* 6:445-501.

Saito, M. 1988. Crossing and Wh-Q binding. The Ohio State Workshop on Japanese Syntax.

Saito, M. 1989. Scrambling as semantically vacuous A'-movement. In M. Baltin and A. Kroch (eds.), *Alternative conceptions of phrase structure*. University of Chicago Press, Chicago.

Saito M. 2001. Two notes in the analysis of NP-movement in the minimalist program. *Proceedings of the Sophia University Linguistic Society* 16:72-98.

Saito, M. 2003. A derivational approach to the interpretation of scrambling chains. *Lingua* 113:481-518.

Saito, M. and N. Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29:439-474.

Tanaka, H. 1997. Invisible movement in *sika-nai* and the linear crossing constraint, *Journal of East Asian Linguistics* 6:143-178.

Tsai, W.-T. D. 1994. On economizing the theory of A' dependencies. Doctoral dissertation, MIT.

Watanabe, A. 1992. Subjacency and S-structure movement of Wh-in-situ. *Journal of East Asia Linguistics* 1:255-291.