In this paper, I argue that a certain case of scrambling of Wh-phrases counts as syntactic Wh-movement in Japanese: Specifically, I argue that long-distance scrambling of a Wh-phrase to a clause headed by a [+WH] COMP behaves exactly like Wh-movement. I defend this hypothesis by showing (i) that such movement lacks LF 'undoing' effects, a quite surprising fact considering that scrambling can be freely undone in LF, and (ii) that it exhibits the Superiority effects. This observation provides support for Mahajan's (1990) proposal that long-distance scrambling, unlike local scrambling, is uniformly A'-movement. Further, I show, on the basis of LF undoing effects, that multiple Wh-movement is possible in Japanese, and argue that this fact supports Kuroda's (1988) 'multiple SPEC' analysis for the phrase structure of Japanese.

0. INTRODUCTION

The purpose of this paper is to argue for the hypothesis that a certain case of scrambling of Wh-phrases counts as syntactic Wh-movement in Japanese. Following Mahajan's (1990) proposal that while local scrambling may be A- or A'-movement, long-distance scrambling is uniformly A'-movement, I will show that long-distance scrambling of a Wh-phrase to a clause headed by a [+WH] COMP behaves exactly like Wh-movement with respect to LF undoing effects and the Superiority effects. Kuroda (1988) suggests that his AGREEMENT parameter, from which he attempts to derive certain differences between English and Japanese, predicts (optional) syntactic Wh-movement to be present in Japanese. The material in this paper, therefore, can be considered to be an extension of his idea.

In Section 1, based on Saito (1989), I will show that while scrambling can be undone in LF, Wh-phrases scrambled to the initial position of a clause headed by a [+WH] COMP cannot undergo further LF movement,

* I would like to thank Howard Lasnik and Mamoru Saito for their valuable comments and suggestions. I am also very grateful to two anonymous NLLT reviewers for helpful comments and criticisms. Comments from and discussions with Jun Abe, Sam Epstein, Akira Kikuchi, Laurel LaPorte-Grimes, Alec Marantz, and Masaru Nakamura were also helpful in improving this paper. All the remaining errors are, of course, my own.

1 As I was working on this research, an important work appeared which also claims the existence of syntactic Wh-movement in Japanese, based on a different kind of data: Watanabe (1991). I regret that I am unable to consider the connection between his theory and the material in this paper, and leave it for future study.
and argue that this discrepancy can be accounted for by assuming that syntactic Wh-movement does take place in such a case. In Section 2, I will consider Superiority effects in Japanese as further evidence for syntactic Wh-movement. In Section 3, I will point out that more than one Wh-phrase can undergo syntactic Wh-movement in Japanese, and provide an analysis of these facts. Section 4 will be a summary.

1. LF Undoing Effects

1.1. Wh-movement and Scrambling

It is a well-known fact that once a Wh-phrase undergoes Wh-movement in overt syntax, it cannot move any further in LF. Consider the following example:

(1) ?What; do you wonder who; bought where?

Here the embedded subject and object have undergone Wh-movement to the embedded CP-SPEC and the matrix CP-SPEC, respectively. As first pointed out by Baker (1970), the unmoved Wh-phrase where can take either matrix or embedded scope. This indicates that where can move in LF either to the matrix or to the embedded CP-SPEC. The moved Wh-phrases, in contrast, cannot move in LF, which is shown by the fact that they must have scope in their surface positions. Thus, what cannot have embedded scope and must have matrix scope (see Lasnik and Saito (1992) and Epstein (1992) for some discussion).

In contrast with Wh-movement, scrambling can be freely undone in LF, as pointed out by Saito (1989). Consider the following examples:

   NOM NOM what ACC bought Q knows

John knows what Mary bought.

   lit. What, John knows Mary bought.

Example (2a) is a sentence with an embedded question. In (2b), the embedded Wh-objects is scrambled to the matrix clause, but the example has the same interpretation as (2a). Note that since the Wh-phrase takes only embedded scope in (2b), it must be in the embedded CP-SPEC at

---

2 In Japanese, questions are indicated by a Q(uestion)-marker such as ka or no. Since the Q-marker appears only in the embedded clause in (2), it is a declarative sentence with an embedded question.
MOVEMENT OF WH-PHRASES IN JAPANESE

657

LF; note also that since the verb sitteiru ‘knows’ selects [+WH] COMP here, the specifier of the COMP must contain a Wh-element. In order to meet these requirements, the Wh-phrase must be lowered to the embedded clause. Since (2b) is grammatical, this fact differentiates scrambling from Wh-movement (see Saito (1989) for more detailed discussion).

Now we have a discrepancy between Wh-movement in English and scrambling in Japanese; Wh-moved phrases are struck in their surface position, whereas scrambled phrases can move back in LF:

(3)a. A Wh-phrase which has undergone syntactic Wh-movement cannot move any further in LF.

b. A scrambled phrase can move (back and forth) in LF.

Whatever principle(s) may explain (3), we may safely assume that (3) is a correct descriptive generalization.

1.2. Absence of Undoing Effects

The generalization (3) can be used as a diagnostic test for determining whether a movement in question is Wh-movement or scrambling. First let us consider the following examples:

(4)a. John-wa [CP Mary-ga nani-o tabeta ka] siritagatteiru

\(TOP \quad NOM \quad what\ Acc \quad ate \quad Q \quad wants\ to\ know\)

no?

Q

Does John want to know what Mary ate? or

What does John want to know whether Mary ate?

b. Nani-r-o John-wa [CP Mary-ga ti tabeta ka] siritagatteiru no?

What does John want to know whether Mary ate?

Notice that the examples in (4) differ from those in (2) in that they have a Q-marker also in the matrix clause. Hence, (4a–b) are interrogative sentences with an embedded question. Note also that \(ka\) is ambiguous between a scope-marker for a Wh-phrase (a lexical realization of [+WH] COMP) and a complementizer corresponding to \(whether\) in English. Sentence (4a) is ambiguous with respect to the scope of the Wh-phrase \(nani-o\), as indicated in the translations. In (4b), where the embedded Wh-object

\[3\] If we replace \(ka\) in (4a) with \(kadooka\), which unambiguously corresponds to \(whether\), then the resulting sentence has just one interpretation, in which the Wh-phrase takes matrix scope:
is preposed to the matrix clause, one of the two readings associated with (4a) disappears; the Wh-phrase can have only matrix scope in (4b). Under the assumption that scrambling can be freely undone, this fact could not be accounted for. If a scrambled phrase could move back freely in LF, then why is it impossible for the Wh-phrase in (4b) to take embedded scope?

There seem to be two possibilities to account for this fact. One is that when a phrase with some quantificational force is scrambled, it cannot move any further in LF. Thus, since what is scrambled in (4b) is a Wh-phrase, which obviously has a quantificational force, it cannot go back to the embedded clause; hence the absence of the embedded scope reading there. But this account encounters difficulty in (2b), repeated below as (5):

    lit. What, John knows Mary bought.

To assign a proper interpretation to (5), it is necessary for the scrambled Wh-phrase to be lowered to the embedded CP-SPEC. Since the matrix COMP is not [+WH], the Wh-phrase cannot have matrix scope in (5). But once we state that scrambled quantificational phrases cannot move any further in LF, (5) would have no interpretation, which is not the case. The first possibility, therefore, does not work.

The alternative approach is to take the Wh-phrase in (4b) to have undergone Wh-movement, rather than scrambling. A question that arises now is what determines whether a moved Wh-phrase undergoes Wh-movement or scrambling. Let us tentatively assume that when a Wh-phrase is moved to the initial position of a clause headed by [+WH] COMP, that movement counts as Wh-movement. Then since the Wh-phrase in (5) is moved to the initial position of the matrix clause, headed by [−WH] COMP, it counts as scrambling, rather than as Wh-movement. Hence it can return to the embedded CP-SPEC in LF. In (4b), on the other hand, the matrix CP as well as the embedded CP has [+WH] COMP (i.e., *ka* or *no*), and thus the movement of the embedded Wh-object to the matrix clause is Wh-movement, prohibiting the Wh-phrase from being

(i) John-wa [Mary-ga nani-o katta kadooka] siritagatteiru no?
    \[TOP \textit{NOM} \textit{whatACC bought whether wants to know} Q\]

What does John want to know whether Mary ate

The lack of the embedded interpretation of the Wh-phrase here may be due to the (universal) fact that *whether* (or its counterpart in various languages) cannot undergo absorption with Wh-phrases.
lowered to the embedded CP in LF; hence the absence of ambiguity in (4b). Now we have the following descriptive generalization:


This is further supported by the following examples:

4 A reviewer pointed out that for him/her, (7b) is unambiguous and has only one interpretation – that in which the Wh-phrase takes matrix scope – and that this as well as the disambiguity of (4b) follows from (i), which he/she argues should be derived from principles of economy:

(i) Whenever there is an option between raising and lowering, raising must be chosen as a less costly operation.

At present I have no explanation of the dialectal (or idiolectal) variance between the reviewer and me in our judgements about (7b), and simply provide below examples similar to (7), to indicate that my judgement is not special to (7):


Which book does John think that Mary knows whether Bill is interested in?

b. John-wa [dono hon-ni Mary-ga [Bill-ga t kyoomi-o motteiru ka] sitteiru to] omotteiru no?

Which book does John think that Mary knows which book Bill is interested in?

c. Dono hon-ni John-wa [Mary-ga [Bill-ga t kyoomi-o motteiru ka] sitteiru to] omotteiru no?

Which book does John think that Mary knows whether Bill is interested in?

For me, the examples in (ii) have the same interpretations as the corresponding ones in (7): (iia) is ambiguous as indicated; (iib) is ambiguous in the same way as (iia); and (iic) unambiguously has the reading where the Wh-phrase takes matrix scope.

Akira Kikuchi, with whom I checked my judgements of (7) and (ii), pointed out to me that it is much easier for him to get the 'lowest scope' reading of the Wh-phrases in (7b) and (iib) than in (7c) and (iic), respectively. Although he himself refrained from judging (7c) and (iic) to be completely unambiguous, this at least shows that there are people whose judgement is not explained by the reviewer's alternative suggestion, given in (i).
(7a) Ki-mi-wa [CP John-ga [CP Mary-ga nani-o tabeta ka] sitteiru to] omotteiru no?

Do you think that John knows what Mary ate?

What do you think that John knows whether Mary ate?

b. Ki-mi-wa [CP nani-o John-ga [CP Mary-ga ti tabeta ka] sitteiru to] omotteiru no?

lit. Do you think what, John knows (Q/whether) Mary ate?

c. Nani-o ki-mi-wa [CP John-ga [CP Mary-ga ti tabeta ka] sitteiru to] omotteiru no?

What do you think that John knows whether Mary ate?

Sentence (7a) is ambiguous with respect to the scope of the Wh-phrase in the most embedded clause; it can take scope over the lowest clause (then the whole sentence is a yes/no-question), or it can have the matrix clause as its scope (then the entire sentence is a Wh-question with such an interpretation as 'for which x, x a thing, you think that John knows whether Mary ate x'). Sentence (7c), on the other hand, has just one reading; i.e., it is a Wh-question, but not a yes/no-question. This is because the Wh-phrase moves to the matrix clause, which is headed by [+WH] COMP (that is, no), so that it cannot move any further in LF. Of interest is case (7b), where the Wh-phrase moves to the intermediate clause, headed by [−WH] COMP. According to our assumption, the movement counts as scrambling, rather than as Wh-movement. So it is expected that the Wh-phrase is able to undergo LF-movement and have either of the two interpretations associated with (7a). This prediction is borne out; (7b) is ambiguous between a Wh-question and a yes/no-question.5

5 If the intermediate and the matrix verbs in (7b) are replaced with a verb selecting [−WH] C and one selecting [+WH] C, respectively, then the Wh-phrase has only the intermediate clause as its scope:

(i) Ki-mi-wa [nani-o John-ga [Mary-ga ti tabeta to] omotteiru ka] kikimasita ka?

Did you ask what John thought that Mary ate?

Here the intermediate verb is omotteiru 'think', and the matrix verb is kikimasita 'asked'.
Before moving on to the next section, let us note that Mahajan's (1990) hypothesis that local scrambling can be A-movement makes a prediction concerning what appears to be local Wh-movement. Let us consider the following examples:

(8)a. John-wa [CP Mary-ga nani-o tabeta ka] TOP NOM what ACC ate Q siritagatteiru no? Does John want to know what Mary ate? or What does John want to know whether Mary ate?

b. John-wa [CP nani-o Mary-ga ti tabeta ka] siritagatteiru no? Does John want to know what Mary ate? or What does John want to know whether Mary ate?

Sentence (8a), which is identical with (4a), is ambiguous with respect to the scope of the Wh-phrase, as indicated above. In (8b), the Wh-phrase moves to the initial position of the embedded clause. Although the Wh-phrase is in the initial position of the clause headed by a [+WH] COMP in (8b), the sentence is still ambiguous, as shown above. That (8b) has the matrix Wh-question interpretation means that the Wh-phrase can move to the matrix clause in LF. This is apparently a counterexample to the generalization in (6). Now recall that local scrambling, unlike long-distance scrambling, can be A-movement. I argue that the Wh-phrase in (8b) can occupy an A-position, specifically the embedded IP-SPEC (cf. Kuroda 1988):

(9) [CP [IP John-wa [CP [IP nani-o [V [VP Mary-ga ti tabeta]]]]] [C ka]] siritagatteiru] [C no]]?

The ambiguity of (8b) is then accounted for, given the following generalization, instead of (6):


According to (10), the movement of the Wh-phrase in (9) does not count

which selects an interrogative complement. Unlike (7b), (i) is unambiguous; the Wh-phrase takes scope in its surface position. This indicates that the Wh-phrase moved to a clause headed by [+WH] C fails to undergo not only lowering but also raising in LF, as Mamoru Saito (personal communication) suggested to me.
as Wh-movement, because it is A-movement. Since the Wh-phrase is scrambled, it can undergo further LF-movement: hence, the ambiguity of (8b). Notice that in the cases which I have so far argued involve Wh-movement (that is, (4b) and (7c)), the Wh-phrase undergoes long-distance movement; thus, there is no possibility of A-movement. To the extent that (10) is a correct statement, the discussion here provides another piece of evidence for Mahajan's hypothesis that local scrambling can be A-movement.

To summarize, I have argued in this section that some distinction is necessary between scrambling and Wh-movement in order to account for the possibility of LF-movement. The tentative assumption is that when a Wh-phrase (A'-) moves to a clause headed by [+WH] C in the syntax, it counts as syntactic Wh-movement, rather than as scrambling.

2. SUPERIORITY EFFECTS IN JAPANESE

2.1. Local Scrambling and Superiority

Among the phenomena related to syntactic Wh-movement is the superiority effect (cf. Chomsky 1973; Hendrick and Rochemont 1982; Pesetsky 1982; and Lasnik and Saito 1992). Let us consider the following examples:

(11)a. Who did see what?
   b. *What did who see?

(12)a. Who did you persuade to read what?
   b. *What did you persuade who to read?

In the good examples, the trace of the moved Wh-phrase c-commands the other Wh-phrase in situ, whereas in the ungrammatical cases, the hierarchically lower Wh-phrase undergoes movement. As a descriptive generalization, let us assume the following:

(13) *[CP WH, [IP . . . WH, . . . ti . . .]],

where a Wh-phrase WH, is in the specifier position of a [+WH] COMP and another Wh-phrase WH, is in situ and asymmetrically c-commands the variable ti of WH,.

From whatever principle(s) (13) may be derived, it correctly rules out (11b) and (12b), while allowing (11a) and (12a).

Now let us turn to a discussion of Wh-movement in Japanese. Nishigau-uchi (1990) gives the absence of the Superiority effects as one of his two
arguments against syntactic Wh-movement in Japanese. His argument is based on the following cases:

(14)a. Dare-ga nani-o tabeta no?
   \textit{who \textit{NOM} what\textit{ACC} ate \textit{Q}}
   Who ate what?

b. Nani-o dare-ga tabeta no?
   lit. \textit{What, who ate?}

This pair appears to have the same configuration as the pair in (11); the subject Wh-phrase precedes the object Wh-phrase in (14a), whereas the latter is preposed over the former in (14b). Unlike (11b), however, (14b) is as perfectly good as (14a). Nishigauchi (1990) argues that if Japanese had syntactic Wh-movement, then (14b) would exhibit the Superiority effect just as (11b) does, and that the lack of this effect indicates the absence of syntactic Wh-movement in Japanese.

As a careful reader might notice, however, it is somewhat hasty to conclude the lack of Superiority effects in Japanese based on such an example as (14b): since (14b) is a simple sentence, the preposed Wh-phrase \textit{nani-o} may undergo A-movement, rather than A’-movement. Recall that local scrambling can be A-movement, given Mahajan’s (1990) hypothesis. Then the moved Wh-phrase can be in an A-position in (14b), as indicated below:

(15) $[CP [IP Nani-o [IP [VP dare-ga \textit{t_i} \textit{tabeta}]]] \textit{no}]$?

This is somewhat similar to the following examples:

(16)a. $[CP [IP [VP \textit{strikes who} [IP \textit{who as being crazy}]]]]$

b. $[CP \textit{Who_o} [IP t_i' \textit{strikes who} [IP t_i as being crazy]]]]$

We do not say that since the embedded subject Wh-phrase is lower than the other Wh-phrase in (16a), (16b) should be a Superiority violation. In this case, A-movement of the lower Wh-phrase creates another trace which is hierarchically higher than the Wh-in-situ, and serves as a variable, saving (16b) from violation of (13).

2.2. \textit{A’}-movement of Wh-phrases and Superiority

When discussing Superiority effects in Japanese, we must consider the examples where a Wh-phrase is moved across a clausal boundary over
another Wh-phrase, since long-distance movement is uniformly A'-movement.

Thus such cases as (17) are relevant:

(17)a. John-ga dare-ni [Mary-ga nani-o tabeta to]

\[
\text{NOM who DAT NOM what ACC ate COMP}
\]

itta no?

\text{said Q}

Who did John tell that Mary ate what?

b. ??Nani-o John-ga dare-ni [Mary-ga ti tabeta to] itta no?

\text{lit. What did John tell who that Mary ate?}

In (17a), which is a grammatical multiple Wh-question, the matrix and the embedded objects are Wh-phrases, and they are in situ. But once the lower Wh-phrase is moved to the initial position, the result is marginal, as (17b) shows. Note that since this movement is long-distance scrambling, it is unambiguously A'-movement. That this is due to the movement of the lower Wh-phrase is shown by replacing it with a lexical NP:

(18) Pizzai- o John-ga dare-ni [Mary-ga ti tabeta to]

\[
\text{pizza ACC NOM who DAT NOM ate COMP}
\]

itta no?

\text{said Q}

\text{lit. Pizza, did John tell who that Mary ate?}

Here, instead of the Wh-phrase, the non-Wh-NP is scrambled over the Wh-in-situ, but this does not produce a decrease in acceptability.

Further consider the following example:

(19) John-ga [Bill-ga dare-ni [Mary-ga nani-o tabeta to] itta to] omotteiru no?

\text{COMP said COMP think Q}

\text{lit. Does John think that Bill told whom that Mary ate what? Whom does John think that Bill told that Mary ate what?}

This is also a multiple Wh-question with two Wh-phrases in situ. Then let us try moving each Wh-phrase to the sentence-initial position; note that each movement is long-distance scrambling:

(20)a. Dare,-ni John-ga [Bill-ga ti [Mary-ga nani-o tabeta to] itta to]
omotteiru no?

lit. Whom, does John think that Bill told that Mary ate what?

b. ??Nani-o John-ga [Bill-ga dare-ni [Mary-ga t tabeta to] itta to] omotteiru no?

lit. What, does John think that Bill told whom that Mary ate?

When the higher Wh-phrase is preposed, the result is as good as (19), as (20a) shows. When the lower Wh-phrase is moved, on the other hand, the result gets degraded, as indicated in (20b). 6

Now let us replace the Wh-phrase in (19) and (20a–b) by non-Wh-NPs and examine whether the replacement affects acceptability:


COMP said COMP think

John thinks that Bill told Jennifer that Mary ate pizza.


lit. Jennifer, John thinks that Bill told that Mary ate pizza.


lit. Pizza, John thinks that Bill told Jennifer that Mary ate.

All the cases here are good, and in particular there is substantial differ-

---

6 If the Wh-in-situ in (17b) and (20b) receives heavy stress, then the examples sound acceptable. In that case, however, the stressed Wh-phrase may be discourse-oriented, so that it could escape the Superiority effect for some reason. Nishigauchi (1990) argues that although his dialect shows the Wh-island Constraint effect on LF Wh-movement, a Wh-phrase which carries extra stress can violate the condition and be 'extracted' out of a Wh-island. Also, Pesetsky (1987) argues that a D-linked Wh-phrase behaves differently from a non-D-linked one with regard to Superiority.
ence in acceptability between (20b) and (21c). This shows that the degraded status of (20b) is caused by movement of the Wh-phrase. 7

Let us also consider the following examples:

(22)a. John-wa [Mary-ga naze [Jennifer-ga nani-o tabeta
   TOP NOM why NOM what ACC ate
to] itta to] omotteru no?
COMP said COMP think Q
lit. Does John think that why Mary said that Jennifer ate what?
Why does John think that Mary said that Jennifer ate what?

b. Naze; John-wa [Mary-ga ti [Jennifer-ga nani-o tabeta to] itta
   to] omotteru no?.

c. *Naniro John-wa [Mary-ga naze [Jennifer-ga ti tabeta to] itta to]
   omotteru no?

Here one of the two Wh-phrases is an adjunct naze ‘why’. Sentence (22a) is a grammatical multiple Wh-question, which has such an interpretation as ‘for which x, x a reason, and for which y, y a thing, John thinks that for x Mary said that Jennifer ate y’. 8 There, the adjunct Wh-phrase is in a hierarchically higher position than the argument Wh-phrase. When the adjunct is moved as in (22b), the grammaticality does not change. But preposing of the lower Wh-phrase over the adjunct does cause a decrease in acceptability. That this is due to the movement of the Wh-phrase is indicated by the following example:

(23) pizza-ro John-wa [Mary-ga naze [Jennifer- ga ti tabeta
   pizza ACC TOP NOM why NOM ate
to] itta to] omotteru no?
COMP said COMP think Q

7 If the Wh-in-situ in (20b), dare-ni, is preposed to the initial position of the embedded clause, then the sentence improves:

(i) Naniro John-ga [dare-ni Bill-ga ti [Mary-ga ti tabeta to] itta to] omotteru no?

A Wh-phrase moved to an A’-position, rather than in situ, is exempt from Superiority, as stated in (13). It is not clear whether we can obtain a relevant English example to confirm this point.

8 Mamoru Saito pointed out that a multiple Wh-question where the adjunct naze precedes another Wh-phrase is uniformly ungrammatical in his dialect. The argument based on (22), therefore, holds only for those who do not have that constraint. See Nishigauchi (1990), Watanabe (1991), and Saito (1992b) for relevant discussions.
Lit. Pizza, does John think that why Mary said that Jennifer ate? Why does John think that Mary said that Jennifer ate pizza?

This is perfect and thus far better than (22c), from which it seems to follow that (22c) is a case of Superiority. Note also that (22c) seems to be worse than (20b). This conforms to the following contrast in English:

(24)a. *?What did you persuade who to read t? (cf. Hendrick and Rochemont, 1982)

b. *What did you buy t why?

As (24) shows, the adjunct Wh-in-situ yields a more severe violation than the object Wh-in-situ, because (24b) also violates the ECP. The trace created by LF-movement of who is properly governed in (24a) (by the matrix verb), while the trace of why is not in (24b) (cf. Lasnik and Saito (1984), among others). If the contrast between (20b) and (22c) is a substantial one, then it provides further support for the proposal that Wh-movement is involved in examples like (20b) and (22c).

These considerations suggest that in the environment where A-movement of Wh-phrases is excluded, Japanese also exhibits Superiority effects. This in turn suggests the existence of syntactic Wh-movement in Japanese.

Further note that when the lower Wh-phrase in (19) is moved to the intermediate clause, rather than to the matrix clause as in (20b), the result is more acceptable than (20b):

9 A reviewer pointed out the following example:

(i) Nani-o John-ga [dare-ga t, tabeta to] itta no?
what ACC NOM who NOM ate COMP said Q
lit. What did John say that who ate?

Unlike (17b), this example does not display the Superiority effect, although the Wh-phrase nani-o undergoes long-distance movement. As noted by the reviewer, this is consistent with our claim, since the Wh-phrase can first undergo local A-scrambling over the embedded Wh-subject and then A'-move to the matrix clause, an option which is impossible in (17b).

10 A reviewer pointed out that to reach this conclusion, I must show that the Superiority effect obtains only in Wh-movement. If, for example, the Superiority effect is a property of A'-movement in general, then the argument in the text only shows that the movement of Wh-phrases in (20b) and (22c) is A'-movement, rather than Wh-movement.

It is, however, quite difficult to construct 'Superiority' examples with other A'-movement operations like null operator movement. Howard Lasnik (personal communication) pointed out that the following examples may show that topicalization does not exhibit the Superiority effect:

(i)

a. Baseball, JOHN like t.

(cf. *What does who like?)
John-ga [nani, o Bill-ga dare-ni [Mary-ga t, tabeta to] itta to] omotteiru no?

lit. Does John think that what, Bill told whom that Mary ate?

The contrast between (20b) and (25) suggests that movement of a Wh-phrase to the clause headed by [-WH] COMP (that is, to) does not yield a Superiority effect. In this respect, the data concerning Superiority give additional support for the generalization in (10).

Note also that in order for (17b), (20b), and (22c) to be excluded by the Superiority condition in (13), the Wh-phrases moved there must be in the matrix CP-SPEC. Then what excludes the possibility that they are adjoined to IP, rather than moved to CP-SPEC? Since Japanese is a strictly head-final language, it is almost impossible to test whether a given phrase is in CP-SPEC or is IP-adjoined. I will discuss this question in detail in the next section.

3. MULTIPLE SPECIFIERS

3.1. Multiple Wh-fronting

Before going into an analysis of the facts mentioned above, let us note another fact concerning Wh-movement in Japanese.

Aoun, Hornstein, and Sportiche (1981) observe that matrix Wh-movement is optional while embedded Wh-movement is obligatory in French, as indicated below:

(26)a. Jean a vu qui?

has seen who

b. Qui Jean a-t-il vu?

(27)a. *Je me demande [CP (que) Jean a vu qui].

I wonder that has seen who

b. Je me demande [CP qui Jean a vu].

Thus, syntactic Wh-movement is partially optional in French. In Japanese,

b. Baseball, I told JOHN I like t.

(cf. ?*What did you tell who you like?)

Here topicalization moves baseball over the stressed phrase. If a stressed phrase undergoes A'-movement in LF (Chomsky 1977), we might expect (i) to exhibit a Superiority effect, which is not the case. Although it has not been proven that Superiority obtains only in Wh-movement, we may assume at present that it is a property of Wh-movement.
on the other hand, syntactic Wh-movement is fully optional; embedded Wh-movement as well as matrix Wh-movement is optional, as shown in (28) and (29):

(28)a. Kimi-wa [Mary-ga nani-o tabeta to] itta no?
you TOP NOM what ACC ate COMP said Q
What did you say that Mary ate?

b. Nani-o kimi-wa [Mary-ga ti tabeta to] itta no?

(29)a. Boku-wa [John-ga [Mary-ga nani-o tabeta to]
I TOP NOM NOM what ACC ate COMP
itta ka] sitteiru.
said Q know
I know what John said that Mary ate.


Since the Wh-phrase has moved to the initial position of a clause headed by [+WH] C in (28b) and (29b), it counts as Wh-movement. Further, the Wh-phrase remains in situ, as in (28a) and (29a). Thus, overt Wh-movement is fully optional in Japanese.

Further, Lasnik and Saito (1984) observe that languages like Polish have multiple syntactic Wh-movement, as indicated below:

(30) (= (3) in Lasnik and Saito, 1984)

a. Zastanawiam sie kto co przyniesie.
(I) wonder who what will bring
I wonder who will bring what.

b. *Zastanawiam sie kto przyniesie co.

When a Wh-phrase remains in situ at S-structure as in (30b), it results in ungrammaticality.

Now let us consider the following Japanese data:

(31)a. John-wa [Mary-ga dare-ni nani-o ageta ka]
TOP NOM who DAT what ACC gave Q
siritagatteru no?
want to know Q
lit. To whom does John want to know whether Mary gave what?

b. Darej-ni nanij-o John-wa [Mary-ga ti ti ageta ka] siritagatteiru no?

lit. To whom, what, does John want to know whether Mary gave?

Sentence (31a) is ambiguous in at least two ways: both of the two Wh-phrases may take embedded scope or matrix scope. Sentence (31b), where the Wh-phrases are both moved to the matrix clause, is not ambiguous; it is just a matrix multiple Wh-question asking the identity of the person x and the thing y such that John wants to know whether Mary gave y to x. The lack of ambiguity is expected, since the Wh-phrases are moved to the clause headed by [+WH] COMP, so that they cannot be undone in LF. This demonstrates that multiple syntactic Wh-movement is permitted in Japanese, although it is fully optional in this case, too.

3.2. An Analysis

Thus far I have argued that the generalization (10), repeated as (32), is a correct statement.


Then the question that naturally arises is what explains (32). In particular, it is necessary to give a precise account of the distinction between scrambling and Wh-movement. An intuitively appealing way is to state that while (A'-) scrambling is adjunction to IP, Wh-movement is substitution to CP-SPEC. However, since Japanese, a strictly head-final language, does not have such visible operations as Subject/Aux Inversion, it seems almost impossible to determine whether a preposed phrase is adjoined to IP or is in CP-SPEC.

My analysis, to be presented below, is based on Kuroda (1988). Kuroda argues that the agreement-forcing nature of English requires one and only one element to stand in SPEC-Head relation with an agreement-inducing head; hence, English has obligatory but not multiple Wh-movement, and a single subject per clause, for example. On the other hand, agreement is not forced in Japanese, so that the specifier position of an agreeing head may or may not be occupied, and further, more than one element
may enter into SPEC-Head relation with the head; hence, Japanese has multiple scrambling and multiple subjects, for instance (see Kuroda (1988) for more precise and detailed discussion). According to Kuroda's hypothesis, Japanese is expected to have optional multiple Wh-movement, an assumption we have defended.

Departing slightly from Kuroda's original idea, let us assume that A-scrambling is movement to IP-SPEC whereas A'-scrambling is to CP-SPEC. Thus, local scrambling may be either to IP-SPEC or to CP-SPEC, while long-distance scrambling is uniformly to CP-SPEC.\(^{11}\)

\[
\begin{align*}
(33)a. & \quad [\text{CP} \ [\text{IP} \ X \ \overline{\text{[r} \ [\text{VP} \ \ldots \ ]]}]]] \quad \text{(local A-scrambling)} \\
(33)b. & \quad [\text{CP} \ X \ [\text{CP} \ C \ [\text{IP} \ [\text{VP} \ \ldots \ ]]]]]] \quad \text{(local A'-scrambling)} \\
(33)c. & \quad [\text{CP} \ X \ [\text{CP} \ C \ [\text{IP} \ \ldots \ [\text{CP} \ \ldots \ ]]]]]] \quad \text{(long-distance scrambling)}
\end{align*}
\]

Then, following Lasnik and Saito (1984, 1992), let us assume the following LF-filters, which are slightly different from Lasnik and Saito's original formulation but are the same as theirs in essential respects:

\[
\begin{align*}
(34)a. & \quad \text{A [+WH] COMP must have a Wh-phrase in its SPEC.} \\
(34)b. & \quad \text{A Wh-phrase must be in the SPEC of a [+WH] COMP.}
\end{align*}
\]

Under (34), the inability of overtly moved Wh-phrases to move in LF is accounted for by the theory of economy, as argued by Epstein (1992). I assume with Epstein (1992) the following ECONOMY PRINCIPLE (cf. Chomsky 1991):

\[
(35) \quad \text{Satisfy filters by using the fewest possible applications of } \text{Affect-} \alpha.
\]

Epstein argues that the Economy Principle accounts for the following example:

\[
(36) \quad \text{Who} \ i \ \text{wonders [CP where [IP Mary bought what]]?}
\]

Without (35), the derivation in which \textit{where} and \textit{what} move in LF to the matrix and the embedded CP-SPECs, respectively, would be possible; then, \textit{who} and \textit{where} would be in the matrix CP-SPEC and \textit{what} would

\(^{11}\) The status of adjunction remains unclear under this analysis. I tentatively assume that an adjoined position is reanalyzed as a specifier position in this system. This may follow from Kuroda's hypothesis that agreement can be many-to-one in Japanese, so that an adjoined element may also enter into an agreement relation with a head. In English, an adjoined position is distinguished from a SPEC, since SPEC-Head agreement is a one-to-one relation.
be in the embedded CP-SPEC. Thus, (34a–b) would both be satisfied. The fact is that although the Wh-phrase in situ can move to either of the two positions, where is stuck at the S-structure position. Epstein (1992) argues that since the moved Wh-phrase already satisfies the filters in (34), economy prohibits it from moving further; LF movement of where is unnecessary and thus prohibited (see Epstein (1992) for details).

The Japanese example (4b), repeated below as (37), can be accounted for in the same fashion:

(37) Nani-o John-wa [Mary-ga t, tabeta ka] siritagatteiru no?

What does John want to know whether Mary ate?

Here the Wh-phrase nani-o is moved from the embedded clause to the sentence initial position. Notice that the matrix COMP is occupied by a [+WH] COMP, and further that as we observed in Section 1, the Wh-phrase cannot take embedded scope. Since the Wh-phrase is moved long distance, it is A’-moved and is in the matrix CP-SPEC:

(38) \[cP WH \[c' liP-- \[cP.-- t... \] [+WH]\]

Representation (38) meets filters (34a–b) as it is. Economy, therefore, requires the Wh-phrase to remain in the surface position at LF.

Let us also consider (7b), whose structure is schematized below:

(39) \[cP [IP... [cP WH [IP... [cP [IP... t... \] [+WH]]

\[-WH]] [+WH]]

In (39), the Wh-phrase is moved to the SPEC of [–WH] COMP. To satisfy filter (34b), it must move either to the matrix or to the most embedded CP-SPEC. (Note that the Q-marker in Japanese is ambiguous between [+WH] COMP and whether.) Given that LF movement is unbounded as long as the ECP is met, and that lexical government is enough to satisfy the ECP, the Wh-phrase in the intermediate CP-SPEC can move in either direction; hence the ambiguity of (7b).12

Concerning Superiority, (13) applies in Japanese directly, since Wh-movement is to CP-SPEC in Japanese. Let us consider an example of

12 Strictly speaking, economy requires the Wh-phrase to be in situ in (7b): if it is in situ, then the subsequent LF derivation will need just one application of Move-a (note that since the Wh-phrase is the object of the lowest verb, its trace can meet the ECP through lexical government by the verb). On the other hand, since (7b) already involves syntactic movement, the total derivation needs at least two applications of Move-a to reach an appropriate LF-representation (see Epstein (1992) for details). To avoid this problem, I tentatively assume with Tada (1990) and Epstein (1992) that scrambling is not counted by economy. Of course, this needs an explanation, which I leave for future study.
Superiority violation – (17b) – which is repeated below as (40) with relevant structure:

(40) ??[CP Nani-o [IP John-ga dare-ni [CP Mary-ga ti
what ACC NOM who DAT NOM
ate COMP said Q
lit. What did John tell who that Mary ate?

Here the Wh-phrase in situ asymmetrically c-commands the variable of the moved Wh-phrase, which is in the specifier position of a [+WH] COMP. Hence, this is in violation of (13).

As for multiple Wh-fronting sentences like (28b), (29b), and (31b), I give the following structure:

(41) [CP WH_i [CP WH_j [c, [~v WH_i [~v... C]] +WH]]]

Here the first Wh-phrase WH_i adjoins to CP and the second one WH_j is in CP-SPEC. Since Japanese allows agreement to be many to one, WH_i may also enter the agreement relation with [+WH] C. This ‘additional’ agreement enables the CP-adjoined position to be reanalyzed as another specifier position of the COMP (see note 11 and Saito (1992a) for relevant discussions).

Although this analysis provides a fairly straightforward account of the facts about Wh-movement in Japanese, one might argue that IP-adjunction, rather than reanalysis of a CP-adjoined position as another SPEC, is involved in cases of multiple Wh-fronting.13 Under this approach, a ‘multiple Wh-fronting’ sentence would have the following structure, instead of (41):

(42) [CP WH_i [c, [IP WH_j [IP ... ] C]]]

Here the first Wh-phrase occupies CP-SPEC, while the second one adjoins to IP. Perhaps the IP-adjoined Wh-phrase is licensed by government by [+WH] COMP (Nishigauchi 1990).

There is, however, an empirical argument against this ‘IP-adjunction’ approach. The argument comes from consideration of what Ross (1969) calls SLUICING. Lobeck (1990), and Saito and Murasugi (1990) provide examples such as the following:

13 The idea that IP-adjunction serves as Wh-movement is entertained in Nishigauchi (1990) and Mahajan (1990), among others.
(43)a. They say John loves someone, but I don’t know who (he loves).
b. They say John will go there, but I don’t know whether *(he will go).
c. They say John loves Mary, but I don’t believe that *(he loves her). *(that as a COMP)

Observing that IP-ellipsis is possible only when a Wh-phrase occurs in the specifier position of the CP immediately dominating the empty IP, they argue that the empty IP is licensed by an agreeing COMP:

(44)a. \[
\ldots [\text{CP WH} [C' [+WH] [IP e]]]\]

(association)

b. \[
\ldots [\text{CP } [C' \text{ that/whether } [IP e]]]\]

(no association)

In (43a), the Wh-phrase agrees with [+WH] COMP, as shown in (44a), so that the embedded IP can be empty. In (43b-c), on the other hand, the missing IP is not licensed, since the dominating COMP is not an agreeing head, as shown in (44b).

Baltin (1982) and Lasnik and Saito (1992) argue that embedded topicalization involves IP-adjunction. With this in mind, let us consider the following example:

(45) They say that this diamond ring, John gave to some girl. But I don’t know to whom, this diamond ring, *(he gave).

(cf. \ldots I don’t know to whom)

Example (45) shows that when adjunction creates two (or more) IP-segments, the lower segment cannot be ‘sluiced’ even if a Wh-phrase is present in the SPEC of the dominating CP. Thus, the following structure is ill-formed:

(46) \[
\ldots [\text{CP WH} [C' [+WH] [IP X [IP e]]]]\]

Now let us turn to Japanese. This language also allows ‘sluicing’, with almost the same restriction observed in (43):
(47)a. Minna-wa John-ga dareka-o aisiteiru to
everyone TOP NOM someone ACC love COMP
itta ga, boku-wa dare-o (kare-ga aisiteiru) ka
said but I TOP who ACC he NOM love Q
wakaranai.
not know
Everyone said John loved someone, but I don’t know who (he
loves).

b. Minna-wa John-ga sokoni iku to itta ga,
everyone TOP NOM there go COMP said but
boku-wa *(kare-ga iku) kadooka wakaranai.
I TOP he NOM go whether not know
Everyone said John would go there, but I don’t know whether *
(he will go).

c. Minna-wa John-ga Mary-o aisiteiru to itta
everyone TOP NOM ACC love COMP said
ga, boku-wa *(kare-ga kanojo-o aisiteiru) to
but I TOP he NOM she ACC love COMP
omowanai.
not think
Everyone said that John loved Mary, but I don’t think that *(he
loves her).

Given that the empty IP is licensed by the agreeing [+WH] COMP also
in Japanese, (47a) provides support for our proposal that syntactic Wh-
movement (optionally) takes place in Japanese.14

Further, the following example, where two Wh-phrases are fronted,
permits sluicing:

14 Saito and Murasugi (1990) argue that N'-Deletion (reanalyzed as NP-Deletion under the
DP-Hypothesis) is possible in Japanese as well, and that the empty NP is licensed by an
agreeing DET. Their analysis as well as ours provides an argument for the existence of
agreeing functional categories in Japanese. Furthermore, note that [+WH] COMP ka re-
mains after sluicing in (47a), indicating that what is ‘sluiced’ is IP. This does not obtain
directly in English, since [+WH] COMP is covert in English. See Takahashi (1992) for a
detailed analysis of sluicing in Japanese.
(48) Minna-wa John-ga dareka-ni nanika-o okutta to iwu ga, boku-wa dare-ni nani-o sent COMP say but I TOP who Dat what ACC
(kare-ga okutta) ka wakaranai.
he NOM sent Q not know

Everyone says John sent something to someone, but I don’t know what he sent to whom/*what, to whom, (he sent).

According to the ‘multiple CP-SPEC’ and the ‘CP-SPEC/IP-adjunction’ analyses, the embedded CP of the second conjunct in (48) has the structures schematized in (49a–b), respectively:

(49)a. \[\ldots [\text{CP WH}_i \ [\text{CP WH}_j \ [c, [\text{IP e}] \ ka]]] \]

b. \[\ldots [\text{CP WH}_i [c^* \ [\text{IP WH}_j \ [\text{IP e}] \ ka]]] \]

The two Wh-phrases are in CP-SPEC in (49a), whereas the first Wh-phrase is in CP-SPEC and the second one is adjoined to IP in (49b). Now recall the discussion of (45), which involves ‘sluicing’ with embedded topicalization. We argued that the lower IP-segment created by IP-adjunction cannot be ‘sluiced’, as shown in (46). If (46) holds also in Japanese, (49b) should be excluded, since it has the same structure as the ungrammatical example (45). The grammaticality of (48), therefore, leads us to choose (49a) over (49b). Although my discussion of ‘sluicing’ in Japanese needs careful examination, it seems to provide an empirical argument for Kuroda’s (1988) ‘multiple SPEC’ analysis. Thus I assume that multiple Wh-movement at least can utilize multiple CP-SPECs in Japanese.

4. Summary

I have argued for the hypothesis that long-distance scrambing of a Wh-phrase to a clause headed by \([+WH] \ COMP\) counts as syntactic Wh-movement in Japanese, on the basis of the facts concerning the LF undoing effects and the Superiority effects. The observation here provides support for Mahajan’s (1990) proposal that long-distance scrambling is uniformly A’-movement, while local scrambling may be A-movement. I have also pointed out that multiple Wh-movement is possible in Japanese, and attempted to account for this fact based on Kuroda’s (1988) ‘multiple SPEC’ analysis. Although the conclusion I have reached here is far from decisive and needs further examination, I believe that the discussion here
MOVEMENT OF WH-PHRASES IN JAPANESE

contributes to a further understanding of scrambling and Wh-movement in Japanese.

REFERENCES


Tada, Hiroaki: 1990, 'Scrambling(s)', talk given at OSU.

Received 5 May 1991
Revised 21 September 1992

Department of Linguistics, U-145
University of Connecticut
Storrs, CT 06269
takahash@uconnvm.bitnet