Multiple Wh-Questions and Parasitic Chains

Miyuki NOJI *
(Received October 30, 1995)

ABSTRACT

The present study deals with the superiority effects on multiple wh-questions. I will argue that an in-situ wh-phrase is not subject to 'sbsorption' but forms a kind of operator-variable chain at LF with the help of an already existing operator-variable chain, which is referred to as a parasitic chain in this paper. It is shown that the superiority effect appears if the formation of operator-variable chains makes some difference in cost among competing multiple wh-questions, and that the analysis here can not only explain the presence or absence of superiority between two wh-phrases but also cover the cases with three wh-phrases.

KEY WORDS
Multiple Wh-question Superiority Effect
Wh-movement Economy

1. Introduction

There have been many proposals to the superiority effects on multiple wh-questions like (1) since Chomsky's (1973) superiorit condition (2) [1]. In the latest work such as Chomsky (1992) and Kitahara (1993), the superiority effects are given an explanation under the economy principles. Some of the basic facts, however, remain unexplained. The aim of this paper is to shed light on them too and to show that the superiority effects can be given a principled explanation within the economy framework under a proper interpretation of the LF legitimacy of in-situ wh-phrases.

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

(2) Superiority Condition:
   a. No rule can involve X, Y in the structure

* Division of Languages: Department of Foreign Languages
...X... [α... Z... ·WYV]...

where the rule applies ambiguously to Z and Y, and Z is superior to Y

b. "... the category A is "superior" to the category B in the phrase marker if every major category dominating A dominates B as well but not conversely."

(Chomsky (1973: 246))

In section 2, basic facts to be discussed are presented. In section 3, I claim that every wh-phrase in a multiple wh-question has to form an operator-variable chain to be a legitimate LF object, and that while one of the wh-phrases forms an operator-variable chain by movement, the other wh-phrase(s) forms a 'parasitic chain', a kind of operator-variable chain fromed not by movement but by linking with a [+WH] operator already in [Spec, CP]. The operator-variable chain formation of the wh-phrases may give rise to a difference in cost among competing multiple wh-questions. It is shown that the difference in cost causes superiority effects. The final section is a summary and conclusion.

2. Basic Facts

2.1. Superiority Effects between Arguments

In this section, we will see in what configurations a wh-phrase becomes superior to another with respect to wh-movement. Let us first consider some cases where two arguments, subject and object, are questioned.

(3) a. Mary asked who t read what
    b. *Mary asked what who read t

(4) a. Mary asked which man, t_i read which book
    b. Mary asked which book, which man read t_i

(Pesetsky (1987: 104, 106))

(3) is called the 'pure' superiority effect since the ECP is irrelevant, both the embedded subject and object positions being properly governed if the two wh-phrases moved to the embedded [Spec, CP] position by LF (cf. (1))². (4) shows that an object wh-phrase can move over a subject wh-phrase to the specifier position of CP, that is, there is no superiority effect between them, when the subject consists of a wh-word and some other element(s). It is not important here that the object also consists of more than one element. We therefore have (5) alongside of (4b).

(5) Mary asked what, which man read t_i

Pesetsky (1987) refers to wh-phrases like which man as D(iscourse)-linked in that when we ask which man John saw, for example, we presuppose that the possible answers can be found in the previous discourse, and the set of possible answers are fixed in advance. As Contreras (1993) argues, however, the crucial property differentiating (4) from (3) is not whether the subject is D-linked, but whether the wh-word in the subject c-commands the one in the object at the stage before the wh-movement takes place. (6) is the case in
point.

(6)  a. What did whose mother buy?
    b. What type of book does what type of man read? (Contreras (1993))

Notice that the wh-word in the subject does not c-command the object position from which the wh-phrase moves to the sentence-initial position. The superiority effect is not observed here, although whose mother and what type of man are non-D-linked$^3$.

Let us next consider the following examples, where the direct object and the indirect object are questioned.

(7)  a. What did you give to whom?
    b. *Who did you give what to?

(8)  a. *Who did you introduce who to?
    b. *What did you tell who about?

(9)  a. Who did you introduce which people to?
    b. What did you tell which people about? (Fiengo (1980: 122, 123, 126))

The situation is the same as the subject and object cases above. (7, 8) illustrate that the direct object must be moved overtly when it consists only of a wh-word. When a wh-word is embedded in the direct object and doesn't c-command another wh-word as in (9), the superiority effect disappears$^4$.

We have seen some examples of superiority between arguments. In the next section, some examples of superiority between an argument and an adjunct are presented.

2.2. Superiority Effects between an Argument and an Adjunct

Let us begin with the cases where the subject and the adjunct are wh-phrases.

(10) a. Who gets her groceries where?
    b. Who got his answer when? (Lasnik and Saito (1992: 14))

(11) a. *Where does who get her groceries?
    b. *When did who get his answer?

(12) a. Which woman will live where?
    b. Where will which woman live?

(10–12) suggest that a subject is superior to an adjunct with respect to wh-movement only when the subject consists merely of a wh-word, that is, the subject wh-word c-commands the adjunct wh-word.

These data contrast with (13, 14), which contain an object wh-phrase and an adjunct wh-phrase.

(13) a. What did you buy when?
    b. When did you buy what?

(14) a. Where did you buy what?
    b. What did you buy where? (Kuno and Robinson (1972: 474))

(13, 14) show that both an object and an adjunct can be fronted by wh-movement even if
they are simple wh-phrases. This fact seems to relate closely to the structural position of the adjunct wh-phrase, which will be discussed in section 3.2.

So far, we have seen the superiority effects between two wh-phrases. To sum up briefly, we can say that the data presented above fall under either (15a) or (15b), putting aside the object and adjunct cases (13, 14).

\[
(15) \quad \begin{array}{c}
\text{[CP } [\text{AGRSP } \ldots \text{ wh}_1 \ldots \text{ wh}_2 \ldots ]] \\
\text{(3), (7), (8), (10), (11)}
\end{array}
\]

\[
\begin{array}{c}
\text{[CP } [\text{AGRSP } \ldots \text{ wh}_1 \ldots \text{ wh}_2 \ldots ]] \\
\text{(4), (6), (9), (12)}
\end{array}
\]

A descriptive generalization is that the superiority effect arises if one wh-element ccommands the other as in (15a).

2.3. **Superiority among Three Wh-Phrases**

In this section, we will consider the superiority effects on multiple wh-questions with three wh-phrases. (16) and (17) have one wh-phrase in the matrix clause and two wh-phrases in the embedded clause.

(16) Who wonders what who bought? \quad (\text{Lasnik and Saito (1992: 118)})

(17) Which man knows where which woman will live? \quad (\text{Pesetsky (1987: 106)})

(16) is not ambiguous with the embedded who interpreted only in the matrix clause, whereas (17) is ambiguous with which woman interpreted either in the matrix clause or in the embedded clause. The point is that (16) is acceptable if the embedded who is interpreted as a matrix interrogative. If we ruled out the movement of what over who in (3 b) by a derivational constraint, the acceptability of (16) would be unexplained. Lasnik and Saito (1992) thus conclude that superiority effects stem from an interpretive constraint but not from a derivational one.

It should be noted here that in both (16) and (17), the 'superiority condition' is still active between the given two wh-phrases. In (16), for example, it is due to the superiority effect between the two embedded wh-phrases that the embedded question reading of who is blocked. And it is because the matrix who is in a position superior to (i.e., c-commanding) the embedded who and the former moves first to the matrix [Spec, CP] that the matrix question reading of the embedded who becomes possible. This entails that the schematic generalization (15) holds even if the two wh-phrases are not clausemates.

What happens, then, when there are three wh-phrases in a clause and all have the same clause as their scope? As Bolinger (1978) and Kayne (1984) point out, superiority effects are weakened and essentially any of the three wh-phrases can be fronted:

(18) \begin{align*}
\text{a. Who took what where?} \\
\text{b. What did who take where?}
\end{align*}
c. Where did who take what? (Bolinger (1978: 108))

(18) illustrates that not only the subject wh-phrase but also the object wh-phrase and the adjunct wh-phrase can move to the sentence initial position.

We have seen various kinds of facts above, which seem to be governed by one and the same principle. In the next section, we will try to give them an explanation in a uniform fashion.

3. An Analysis

3.1. Parasitic Chains

We will assume the minimalist framework proposed in Chomsky (1992) and Chomsky and Lasnik (1993) in the following discussion. The background assumptions are the following.

(19) a. LF-movement is cheaper than overt movement. (Procrastinate)

b. Given two convergent derivations D₁ and D₂ with the same LF output, both minimal and containing the same number of steps, D₁ blocks D₂ if its links are shorter. (Shortest Link Requirement)

c. The so-called transformational operation is not Move α but Form-Chain

d. There is no LF wh-movement.

As is well-known, for wh-question sentences to be well-formed in English, one and only one wh-phrase has to move to the specifier position of [+WH] C, forming an operator-variable construction. As for in-situ wh-phrases, it has been considered that they undergo absorption. The concept of absorption, however, seem to be not clear especially when we consider the LF legitimacy of in-situ wh-phrases. So we will argue that an in-situ wh-phrase, as well as a moved wh-phrase, forms a kind of operator-variable construction to be a legitimate LF object:

(20) Every wh-phrase forms an operator-variable construction at LF.

Suppose that a wh-phrase in-situ functions as a variable at LF if c-commanded by an operator already in [Spec, CP] and a kind of operator-variable chain is formed without movement⁶⁰. We will call the chain a parasitic chain in that it is formed with the help of an existing operator-variable chain. There are three ways for wh-phrases to form an operator-variable construction as in (21)⁶⁰.

(21) i) overt movement

```
  c-c
 _/\_
Op variable Chain (Op, variable)   
```

[Diagram: Overt movement of wh-phrase to form a parasitic chain]
ii) covert linking for parasitic chain formation

\[
\text{Op} \quad \text{variable}^1 \quad \text{variable}^2 \quad \text{Chain (Op, variable}^3\text{)}
\]

(21i) is a familiar way of chain formation. (21ii) illustrates two ways of parasitic chain formation. As stated above, we assume that a wh-phrase in-situ functions as a variable at LF if c-commanded by a [+WH] operator in [Spec, CP]. This is true both in (21iia) and (21iib). In (21iia), a link is formed between Op and variable^2, which yields a parasitic chain (Op, variable^2). But when another member of the overtly formed chain, variable^1, also c-commands variable^2 as in (21iib), a more economical operation is chosen to yield a parasitic chain (Op, variable^2): a link is formed between variable^1 and variable^2. In other words, Op makes the best use of the already existing chain of its own to from a parasitic chain with variable^2.

Although the operations in (21) are all counted as one, they give rise to a difference in cost:

\[(22) \quad \text{Cost: (21i) > (21iia) > (21iib)}\]

(21i) is an operation in the overt syntax and hence more costly than the other two by the principle of procrastinate (19a). The operations in (21ii) are, however, apparently equal in cost, both being LF options. To see why (21iia) is regarded as more costly than (21iib), let us consider the following configuration:

\[(21iib')\]

(21iib) and (21iib') start with the same structure. In the latter, however, overt movement takes place from the position of variable^2 to the position of Op. To become a legitimate object, variable^2 has to from a parasitic chain with Op by means of (21iia). Option (21iib) is not available here because variable^2 does not c-command variable^1. The overall length of links in (21iib') is, then, longer than the one in (21iib), although both are derived in two steps, one overt and the other covert. The derivation in (21iib') is blocked by the derivation in (21iib), which is due to the Shortest Link Requirement. We can say from this that option (21iib) allows more economical derivations than (21iia) and is therefore always chosen over (21iia) when it is available.

Here we take a position that in multiple wh-questions, all the operations involved in the interpretation of wh-phrases (i.e., three options in (21) should be taken into account in calculating the cost, which is slightly different from Chomsky (1992) and Kitahara (1993)). They deduce the superiority effects from the difference in length of overt movement among
3.2. Presence or Absence of Superiority

Recall that the superiority effect appears in a configuration like (15a), but not (15b). A large part of the basic facts in section 2 is classified into either of the two groups:

(15) a. \[ \text{c-c} \]
\[ \text{[CP [AGRSP ... wh1 ... wh2 ...]]} \]
\[ \text{[3, (7), (8), (10), (11)]} \]

b. \[ \text{c-c} \]
\[ \text{[CP [AGRSP ... wh1 ... wh2 ...]]} \]
\[ \text{[4, (6), (9), (12)]} \]

Let us begin with (3), a typical pair of superiority. They have (3') as their LF representations.

(3) a. Mary asked who t read what
b. *Mary asked what who read t

(3') a. Mary asked \[ \text{[CP who, [AGRSP t, read[TP[AGRPOP what[VP...]]]]]} \]

b. Mary asked \[ \text{[CP what, [AGRSP who read[TP[AGRPOP t, [VP...]]]]]} \]

Each of them involves one overt movement of a wh-phrase, that is, one application of From Chain. Furthermore in (3'a), where t_i c-commands what, a link is formed between them at LF. In (3'b), where t_i does not c-command who, a link is formed between what, the head of the chain (what, t_i), and who. While both are convergent derivations, (3'a) is chosen over (3'b) because they differ in cost: (3'a) receives a more economical interpretation of the wh-phrases with (21ii) and (21iib). The same story holds in (7, 8, 10, 11).

Let us reconsider (4) next, where two options for wh-movement are both possible. The LF representations of (4) are shown in (4').

(4) a. Mary asked which man, t_i read which book
b. Mary asked which book, which man read t_i

(4') a. Mary asked \[ \text{[CP which, [AGRSP [DP t_i man] read[TP[AGRPOP [DP which book][VP...]]]]]} \]

b. Mary asked \[ \text{[CP which, [AGRSP [DP which man] read[TP[AGRPOP [DP t_i book][VP...]]]]]} \]

Here again, one overt movement is involved in both cases. Different from (3), however, LF linking for in-situ wh-phrases are required between the operator in the embedded [Spec, CP] and the in-situ wh-phrase in both sentences, which entails that they are equally economical, both being derived with (21i) and (21iia). In fact, the derivations in (4') contain the same number of steps and result in the same length of links. This is why the
sentences in (4) are both permitted. The same explanation is applicable to (6, 9, 12). Whether the option (21iib) is available or not is then a decisive factor triggering the superiority effect in our analysis.

With this in mind, let us tackle an exceptional case that apparently does not fall under (15a) nor (15b). (13) is such an example.

(13) a. What did you buy when?
    b. When did you buy what?  (Kuno and Robinson (1972: 474))

First, we have to make the position of the adjunct explicit. If *when* is generated in a position that c-commands the position of the object wh-phrase at LF, [Spec, AGROP], only (13b) should be permissible since the remaining wh-phrase in (13b), *what*, can be legitimate by means of (22iib), whereas the one in (13a), *when*, calls for (22iia). If *when* is generated in a position that is c-commanded by the object wh-phrase, on the other hand, only (13a) should be permissible for the very contrary reason. Whichever position we will take, one of the sentences in (13) is wrongly expected to be unacceptable. This dilemma is easily resolved if VP adjuncts such as *when* and *where* can be generated in at least two distinct positions. This is not untenable, but, in fact, natural if we adopt the following sentence structure, which is independently motivated by Koizumi (1993).

(23) Koizumi's (1993) Split VP Hypothesis:

\[
\begin{align*}
\text{AGRSP} & \quad \text{TP} & \quad \text{VP}_1 & \quad \text{SU} & \quad \text{V} & \quad 1 \quad \text{vp} \quad \text{AGROP} & \quad \text{VP}_2 & \quad \text{V} & \quad 2 \quad \text{OB}
\end{align*}
\]

In (23), AGROP intervenes between the subject and the object. In English, both the subject movement to [Spec, AGRS] and the object movement to [Spec, AGRO] are assumed to take place in the overt syntax. VP-adverbs are assumed to be adjoined either to VP1 or to VP2.

(23) with the assumptions above enables us to explain the distribution of VP-adverbs in the following examples.

(24) a. I have found Bob recently to be morose
    b. I've believed John for a long time now to be a liar

As (24) shows, the matrix adverbial elements can appear after the exceptionally Case-marked subject, which is observed in Postal (1974) and Kayne (1985). Under the split VP hypothesis, (24a, b) are analyzed to have the following structures.

(24') a. [AGRSP I, have [VP1, t, found], [AGROP Bob, [VP2 recently, t, [AGRSP t, to be morose]]]]
    b. [AGRSP I,'ve [VP1, t, believed], [AGROP John, [VP2 for a long time now, [VP2, t, [AGRSP t, to be a liar]]]]

In (24'), the matrix main verb moves from the V2 position to the upper V1 position and the subject of the infinitival clause moves over the adverb to the specifier position of AGRO. The two overt movements result in the order of ... V NP Adv ....

(23) can also provide an account for (25), which is problematic under the Larsonian
shell analysis.

(25) a. Bob put all three books intentionally on the table.  
    b. Aaron gave the ring secretly to her.  

In (25), the VP-adverbs show up between the direct object and the PP complement. The Larsonian shell does not have a position between them in which the VP-adverbs can appear\(^{19}\). Given (23), however, (25) is properly analyzed as in (25')\(^{11}\).

(25') a. \([\text{AGRSP} \text{ Bob}_i[\text{VP}_1 \ t_i \ put_j \ [\text{AGROP} \ all \ three \ books}_s \ [\text{VP}_2 \ t_s \ t_j \ on \ the \ table]]]]\]

b. \([\text{AGRSP} \ Aaron[\text{VP}_1 \ t_i \ gave_j \ [\text{AGROP} \ the \ ring}_s \ [\text{VP}_2 \ t_s \ t_j \ to \ her]]]]\]

In both cases, the verb overtly moves to the V1 position, and the direct object to the specifier position of AGRO. The movements give the order of ... V NP Adv PP.

We are now in a position to deal with (13). (23) contains two distinct VP positions for when:

(26) a. \([\text{AGRSP}[\text{TP} \ [\text{VP}_1 \ SU \ V1 \ [\text{AGROP}[\text{VP}_2[\text{VP}_2 \ V2 \ OB \ when]]]]]]\]

b. \([\text{AGRSP}[\text{TP} \ [\text{VP}_1 \ SU \ V1 \ [\text{AGROP}[\text{VP}_2 \ V2 \ OB]] \ when]]\]

The point is that when is located lower than the specifier position of AGRO in (27a), while it is located higher than the position in (26b). (26a) makes only (13a) possible because when can come from a parasitic chain with the operator via the trace in [Spec, AGROP] at LF\(^{10}\).

(13') a. \([\text{CP} \ \text{what}_i \ \text{did}_i \ [\text{AGRSP} \ \text{you}_j \ [\text{TP} \ [\text{VP}_1 \ t_j \ buy}_s \ [\text{AGROP} \ t_i \ [\text{VP}_2[\text{VP}_2 \ t_s \ t_j \ when]]]]]]\]

b. \([\text{CP} \ \text{when}_i \ \text{did}_i \ [\text{AGRSP} \ \text{you}_j \ [\text{TP} \ [\text{VP}_1 \ t_j \ buy}_s \ [\text{AGROP} \ \text{what}_i \ [\text{VP}_2 \ [\text{VP}_2 \ t_s \ t_j \ t_i]]]]]]\]

(27b), on the other hand, makes only (13b) possible because the object wh-phrase in [Spec, AGROP] can form a parasitic chain with the operator via the trace of when:

(13") a. \([\text{CP} \ \text{what}_i \ \text{did}_i \ [\text{AGRSP} \ \text{you}_j \ [\text{TP} \ [\text{VP}_1 \ t_j \ buy}_s \ [\text{AGROP} \ t'_i \ [\text{VP}_2 \ t_s \ t_i]] \ when]]]]\]

b. \([\text{CP} \ \text{when}_i \ \text{did}_i \ [\text{AGRSP} \ \text{you}_j \ [\text{TP} \ [\text{VP}_1 \ t_j \ buy}_s \ [\text{AGROP} \ \text{what}_i \ [\text{VP}_2 \ [\text{VP}_2 \ t_s \ t_i]]]]]]\]

The analysis here, coupled with the assumptions in (23), is successful in explaining the apparent exception (13) in a consistent way.

3.3. Multiple Wh-Questions with Three Wh-Phrases

So far, we have seen how the analysis with parasitic chain formation is applied to the multiple wh-questions that contain two wh-phrases. The basic facts to be discussed next are the multiple wh-questions which contain three wh-phrases, that is, (16, 17, 18) repeated below.

(16) who wonders what who bought  
(Lasnik and Saito (1992: 118))

(17) which man knows where which woman will live  
(Pesetsky (1987: 106))

(18) a. Who took what where?
b. What did who take where?

c. Where did who take what? (Bolinger (1978: 108))

Let us start with (16). Recall that (16) is acceptable if the embedded who is interpreted in the matrix clause. Its LF representation is (28) under our analysis.

\[ (27) \]
\[
\begin{align*}
\text{a. } & [CP \ \text{who}_i [AGRSP \ t_i \ \text{wonders}_i [CP \ \text{what}_i [AGRSP \ \text{who bought}_i [TP \ [AGRSP \ t_i [VP \ ...]]]]]]
\end{align*}
\]
\[
\begin{align*}
\text{b. } & [CP \ \text{who}_i [AGRSP \ t_i \ \text{wonders}_i [CP \ \text{what}_j [AGRSP \ \text{who bought}_j [TP \ [AGRSP \ t_j [VP \ ...]]]]]]
\end{align*}
\]

(27a,b) represent the embedded question reading and the matrix question reading of the embedded who respectively. Following Kitahara (1993), we assume that they do not compete with each other because they do not count as the ‘derivations with the same LF output’ in (19b): while the scope of the embedded who covers the embedded clause in (27a), while it covers the matrix clause in (27b). What competes with (27a, b) is (28a, b) respectively.

\[ (28) \]
\[
\begin{align*}
\text{a. } & [CP \ \text{who}_i [AGRSP \ t_i \ \text{wonders}_i [CP \ \text{who}_j [AGRSP \ t_j \ \text{bought}_j [TP \ [AGRSP \ \text{what}_j [VP \ ...]]]]]]
\end{align*}
\]
\[
\begin{align*}
\text{b. } & [CP \ \text{who}_i \ \text{does}_i [AGRSP \ \text{who wonder}_i [CP \ \text{what}_j [AGRSP \ t_j [\text{bought}_j [TP \ [AGRSP \ t_j [VP \ ...]]]]]]]
\end{align*}
\]

(27a) is blocked by (27a) for the same reason as (3b). (27b) is, on the other hand, not blocked by (28b), but rather blocks it\(^\text{13}\). Only (27b) remains as the LF of (16). The unambiguity of the sentence, thus, follows.

Let us consider (17) next. As stated before, an ambiguity arises in this case with respect to the interpretation of which woman. (17) has the following LF representations.

\[ (29) \]
\[
\begin{align*}
\text{a. } & [CP \ \text{which}_i [AGRSP \ [DP \ t_i \ \text{man}] \ \text{knows}_i [CP \ \text{which}_j [AGRSP \ [DP \ t_j \ \text{woman}] \ \text{will live}_j [VP \ \text{where}_j]]]]
\end{align*}
\]
\[
\begin{align*}
\text{b. } & [CP \ \text{which}_i [AGRSP \ [DP \ t_i \ \text{man}] \ \text{knows}_i [CP \ \text{where}_j [AGRSP \ [DP \ \text{which woman}] \ \text{will live}_j [t_j]]]]
\end{align*}
\]

(29a, b) correspond to the embedded multiple wh-question reading and the matrix multiple wh-question reading respectively. What we have to show is that both of them are the possible LFs of (17). (30) are the LFs that compete with (29).

\[ (30) \]
\[
\begin{align*}
\text{a. } & [CP \ \text{which}_i [AGRSP \ [DP \ t_i \ \text{man}] \ \text{knows}_i [CP \ \text{where}_j [AGRSP \ \text{which woman}] \ \text{will live}_j [t_j]]]]
\end{align*}
\]
\[
\begin{align*}
\text{b. } & [CP \ \text{which}_i \ \text{does}_i [AGRSP \ \text{which man}] \ \text{knows}_i [CP \ \text{where}_j [AGRSP \ [DP \ t_i \ \text{woman}] \ \text{will live}_j [t_j]]]]
\end{align*}
\]

(29a) and (30a) do not block each other: in both cases, the embedded multiple wh-question is derived in two steps, and the total length of the links involved is the same. Both of them are thus permitted by the same reasoning as (4). Here again, we do not compare just the
length of the overt movements (cf. Chomsky (1992) and Kitahara (1993)). Simply measuring the overt movements and calculating their cost does not lead us to expect that they are both possible.

The same is true of (29b) and (30b): they are equivalent in coat. (30b) is, however, ruled out independently by subadjacency.

Finally, let us see how (18) are represented at LF. To give them appropriate LFs, we need an assumption that the parasitic chain formation by (21iib) is available only once, that is, the chain formed by overt wh-movement can feed only one parasitic chain of the type (21iib). Once a new link is formed between the tail of the existing operator-variable chain and a wh-element in-situ, and, in effect, a kind of ‘extended chain’ is formed, it has not had the same status as the chain formed by overt movement, and thus cannot ‘extend’ further when one more wh-element is added. This is why the chain formed by overt movement cannot feed two parasitic chains of the type (21iib). If there are two wh-elements in-situ, one of the two which is structurally closer to the tail of the existing chain forms a link with it. Let us assume that this is true and try to explain (18). (18) are represented at LF as follows.

(18') a. \[ \text{CP who}_i [\text{AGRSP} \ t_i \ \text{took}_{TP} \ [\text{VP}_1 \ [\text{VP}_1 \ \text{AGROP} \ \text{what} \ [\text{VP}_2 \ldots]] \ \text{where} \]] \]

b. \[ \text{CP what}_i \ \text{did} \ [\text{AGRSP} \ \text{who} \ \text{take}_{TP} \ [\text{VP}_1 \ [\text{VP}_1 \ \text{AGROP} \ t_i \ [\text{VP}_2 \ldots]] \ \text{where} \]] \]

c. \[ \text{CP where}_i \ \text{did} \ [\text{AGRSP} \ \text{who} \ \text{take}_{TP} \ [\text{VP}_1 \ [\text{VP}_1 \ \text{AGROP} \ \text{what} \ [\text{VP}_2 \ldots]] \ t_i \]] \]

(18'') a. \[ \text{CP who}_i [\text{AGRSP} \ t_i \ \text{took}_{TP} \ [\text{VP}_1 \ [\text{VP}_1 \ \text{AGROP} \ \text{what} \ [\text{VP}_2 \ldots] \ \text{where}]]] \]

b. \[ \text{CP what}_i \ \text{did} [\text{AGRSP} \ \text{who} \ \text{take}_{TP} \ [\text{VP}_1 \ [\text{AGROP} \ t_i \ [\text{VP}_2 \ldots] \ \text{where}]]] \]

c. \[ \text{CP where}_i \ \text{did} [\text{AGRSP} \ \text{who} \ \text{take}_{TP} \ [\text{VP}_1 \ [\text{AGROP} \ \text{what} \ [\text{VP}_2 \ldots] \ t_i \]]] \]

As we have already seen in section 3.2., VP-adverbs can be generated either in a VP1-adjointed position or in a VP2-adjointed position. The former case is illustrated in (18') and the latter case in (18'').

Let us being with (18'). (18'a, b, c) compete with each other, the three wh-phrases having the same clause as their scope. (18'a, c) are chosen over (18'b) because the former two are cheaper in cost than the latter to the extent that they make use of (21iib). By the same logic, in (18''), (18''a, b) will be left. After all, (18a, b, c) are analyzed as (18'a, 18''a), (18''b), and (18''c) respectively, and are all permitted correctly. The analysis suggested in this paper, but not other analyses that have been so far proposed in the
literature, is thus successful in answering the question why superiority effects disappear when there are three wh-phrases in one and the same clause.

4. Concluding Remarks

We have seen in this paper that the superiority effect appears in multiple wh-questions if the chain formation of wh-phrases gives rise to some difference in cost among competing multiple wh-questions, under the assumption that every wh-phrase forms an operator-variable construction at LF. In English multiple wh-questions, one of the wh-phrases is subject to overt movement, and an operator-variable chain is formed as in (21i). The rest of the wh-phrases makes use of the chain in one of the two ways in (21ii) to from a parasitic chain with the same wh-operator as its head.

(21) i) overt movement

\[
\begin{array}{c}
\text{Op} \quad \text{variable} \\
\hline
\end{array}
\quad \text{Chain (Op, variable)}
\]

ii) covert linking for parasitic chain formation

a. \[
\begin{array}{c}
\text{Op} \quad \text{variable}^1 \quad \text{variable}^2 \\
\hline
\end{array}
\quad \text{Chain (Op, variable}^2)
\]

b. \[
\begin{array}{c}
\text{Op} \quad \text{variable}^1 \quad \text{variable}^2 \\
\hline
\end{array}
\quad \text{Chain (Op, variable}^2)
\]

This approach enables us to explain the basic facts in section 2 in a principled way without recourse to the ‘superiority condition’. Some of the basic facts are problematic to such an analysis that deduces the superiority phenomena from a derivational constraint of economy that prohibits a wh-phrase from moving to the specifier position of CP skipping another wh-phrase c-commanding it. Multiple wh-questions with three wh-phrases in a clause are typical examples that the analysis cannot handle since any wh-phrase of the three can be fronted. As we have seen in section 3.3, they are straightforwardly explained under our approach.

NOTES

This is a revised version of the paper presented at Tokyo Area Circle of Linguistics (TACL) monthly meeting held at Tokyo University, Jun. 18, 1994. I would like to thank the audience for various comments. I am also grateful to Masayuki Ike-uchi and Shigeo
Tonoike for valuable comments and suggestions.

1) As the name of multiple wh-question shows, we exclude an echo question reading in (1), under which not only (1a) but also (1b) can be grammatical, given an appropriate context and intonation. We continue to restrict our attention to multiple question readings in the following discussion.

2) As mentioned in the introduction, we don't take a position in this paper that in-situ wh-phrases move to the specifier position of CP at LF.

3) One exception is (i), where there is a superiority effect between how many people and whom.

(i) a. I need to know how many people, ti, voted for whom

   b. *I need to know who(m), how many people voted for ti

   (Pesetsky (1987: 107))

In (ib), how does not c-command the trace of who(m), and therefore it should be acceptable too, contrary to fact. It might be relevant here that the subject wh-phrase contains a quantifier. If such a subject as a whole functions as a wh-phrase like who and what at LF, it would not be a problem. But a question still remains about why (ii) is not OK.

(ii) *I need to know {who, which candidate} was [introduced, talked] to how many people

   (Hornstein and Weinberg (1987: 319))

We do not go into the matter anymore, just noticing it.

4) Note that we have three possibilities of wh-movement in a configuration ... V NP1[PP, P NP2], where NP1 and NP2 are wh-phrases: the target can be NP1, NP2, and PP as a whole. We deal only with the former two cases, since the acceptability of a sentence like (i) varies in grammaticality among people.

(i) a. To whom did you give what?

   b. Tell me to whom you gave what

Whereas Chomsky (1973) and Fiengo (1980) mark (i) grammatical, Kuno and Robinson (1972) and Bolinger (1978) do not. This might be due to the structure of double object constructions, which is not our concern here. So let us restrict our attention to the former two cases whose judgments seem to be consistent among people.

5) This assumption does not introduce any special device for the chain formation of in-situ wh-phrases since c-command is a general condition on chain formation. It follows from the assumption that an in-situ wh-phrase cannot be interpreted in the specifier position of [+WH] C which does not c-command it.

6) For the ease of exposition, the options in (21) are shown with a solid line, a wavy line, and a dotted line respectively.

7) It is presupposed that in (21iib), c-command relation is established between Op and variable4 as in (21i). Thus the c-command relation between variable4 and variable5 entails that Op c-commands variable5 by transitivity.

8) We assume here that what moves to [Spec, AGROP] at LF. The movement results in one A-chain, which does not concern us here. What forms a parasitic chain with the operator in [Spec, CP] is what in [Spec, AGROP]. As for the level where the object
moves to [Spec, AGROP], some modification will be needed below. We can, however, ignore the A-movement for the time being anyway.

9) In (23), $\Omega$P is a projection for the checking of the indirect object, which does not concern us here.

10) (25) would not be a problem to the Larsonian shell analysis if we assume Chomsky's (1994) bare phrase structure framework that allows multiple specifiers, which is pointed out by M. Ike-uchi. If an adverb appears in the specifier position of the lower VP and verb movement takes place as in (i), the adverb intervenes between the object and the PP.

(i) Bob$_{[\text{VP1}}$ put$_{[\text{VP2}}$ all three books intentionally $t_o$ on the table]$

11) Needless to say, VP-adverbs can also appear in a preverbal position. (23) provide them with a VP1-joined position.

12) The overt movement of the object from the original position to [Spec, AGROP] is an instance of A-movement. Hence the original position of the object is irrelevant in the present discussion. What behaves as a second variable of the operator in [Spec, CP] and forms an operator-variable chain with it is the trace in [Spec, ARGOP].

13) (28b) is independently blocked for another reason. The overt movement of the embedded who crosses a wh-island, causing a wh-island violation.

14) See note 13.

References


_____. 1994. Bare phrase structure. MIT occasional papers in linguistics 5.


Kitahara, Hisatsugu. 1993. Deducing 'superiority' effects from the shortest chain require-