A DERIVATIONAL APPROACH TO SUPERIORITY EFFECTS IN ENGLISH

MEGUMI NISHIMURA
Kyushu University

This paper presents a derivational analysis of the Superiority effect observed in multiple wh-questions and Negative Polarity Items (NPIs) licensing in English, in terms of a general condition on semantic information and a make-up strategy to satisfy this condition. Assuming the definition of derivational c-command provided by Epstein et al. (1998), we argue that the licensing relation is established through c-command, and this c-command relation may not reverse in the derivation in accordance with the general condition. It is shown that our analysis can account for a wide variety of data including ones recalcitrant to previous analyses, and ultimately these phenomena can be dealt with comprehensively and in a uniform fashion.*

Keywords: Superiority effects, wh-in-situ phrases, NPIs, derivational c-command, make-up strategy

1. Introduction

In the framework of the minimalist program advocated by Chomsky (2000, 2001, 2004), the syntactic structure is formed derivationally step by step (phase by phase). Especially in Chomsky (2004), it is argued that the relevant structure is transferred cyclically to the semantic component in addition to the phonological component by the operation Transfer. Following and extending this idea, we argue that a general...

* This is a revised version of a paper read at the 58th General Meeting of the Kyushu branch of the English Literary Society of Japan held at Nagasaki University in 2005, part of which has already appeared in Kyushu University English Review no. 48. I am especially grateful to Nobuaki Nishioka for his valuable comments and suggestions. I would also like to thank the audience at the conference for their useful comments. I am also indebted to two anonymous EL reviewers for useful comments and suggestions. My thanks also go to Mark Volpe for stylistic improvements. Needless to say, responsibility for any errors is my own.
condition on semantic information derivationally accumulated can accommodate the Superiority effect in English multiple wh-questions observed in the contrast in (1), as well as the contrast in (2), where Negative Polarity Items (NPIs) are involved.

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

(2) a. John *(didn’t) eat anything.
    b. *Anyone didn’t eat apples.

Demonstrating the deficiencies of the previous analyses of multiple wh-questions and NPIs, we argue that comprehensive data concerning both the phenomena can be accounted for by the general condition and a make-up strategy to satisfy the condition.

This paper is organized as follows. Section 2 presents some theoretical assumptions and our proposals. In section 3, we examine Kato’s (2002) analysis on NPI licensing, and point out a flaw that makes it untenable. In section 4, we demonstrate how our alternative analysis can successfully deal with comprehensive data including problematic ones for the previous analyses. Concluding remarks are given in section 5.

2. Assumptions and Proposals

In this paper, in conformity with the framework of the minimalist program provided by Chomsky (2000, 2001, 2004), we assume that the syntactic structure is built up derivationally and then the relevant structure is transferred to the semantic component and the phonological component step by step at the phase levels. Moreover, we assume that information transferred cyclically is accumulated, being updated, and semantic information must be consistent to build a well-formed semantic representation. Thus, we propose a general condition on semantic information in (3).

(3) Ban on Contradictory Information (BCI)

Semantic information may not include any contradiction. Specifically, we suppose that a syntactically encoded licensing relation between two elements provides the basis of semantic information in some cases, and therefore, the relation must be consistent to satisfy (3). If a derivation contains reversal of the licensing relation between the same two elements in such cases, it will violate BCI. We propose that a make-up strategy in (4) function in this case to satisfy (3).
The licensing relation is canceled if the relation reverses between the same two elements. Then, in some cases, the derivation involving reversal of the licensing relation between the same two elements results in ungrammaticality, because the licensee, which must be licensed by an appropriate element by definition, loses its licensor.

We assume that the licensing relation in cases we consider is established between a possible licensor (A) and its relevant licensee (B) through ‘c-command’: A licenses B iff A asymmetrically c-commands B. As for the definition of c-command, we will embrace the definition of derivational c-command provided by Epstein et al. (1998).¹

Derivational C-command

\[ X \text{ c-commands all and only the terms of the category } Y \text{ with which } X \text{ was paired/concatenated by Merge or Move in the course of derivation}. \]  

(Epstein et al. (1998: 10))

We also assume the VP-internal subject hypothesis by Koopman and Sportiche (1991: 212, 225), among others.

Since our analysis shares a basic idea with Kato’s (2002) analysis of NPI licensing, we first outline his analysis and make clear the differences between the analyses by pointing out the problem with Kato’s analysis in the following section.


Kato gives an analysis of NPI licensing in terms of the minimalist program of Chomsky (1995, 2000, 2001, 2004) adopting the VP-internal subject hypothesis. Kato’s analysis, in effect, relies on the condition in (6), which is in essence the same as BCI in (3).

A representation may not involve any contradictory information.  

(Kato (2002: 6))

NPIs such as *anything* and *anyone* require negative elements as in (2) (repeated here as (7)). Moreover, they show an asymmetry between the subject and the object. As in (7b), NPIs cannot be in the subject position, unlike in the object position as in (7a). It has been assumed that

¹ Epstein et al. (1998) argue that the derivational definition of c-command can be established by an elementary operation Merge or Move under this derivational definition. It is, therefore, more natural than the representational one.
NPIs must be c-commanded by negative elements in a sentence to account for the differences in (7).

(7) a. John *(didn’t) eat anything.  (= (2a))
   b. *Anyone didn’t eat apples.  (= (2b))

However, the asymmetry disappears when NPIs are structurally embedded, as in (8). In (8), *anything is embedded within the subject NP.

(8) A doctor [who knew anything][about acupuncture] was not available.  (Linebarger (1980: 149))

By adopting the definition of derivational c-command in (5), in terms of which \( \alpha \) comes into a c-command relation with \( \beta \) (and its terms) when \( \alpha \) is merged with \( \beta \) forming a category \( K \) in (9), Kato proposes a derivational condition of (10) on NPI licensing, which requires that an NPI should be c-commanded by a negative element in the course of derivation.

(9) \[ \begin{array}{c} K \\ \alpha \end{array} \begin{array}{c} \beta \end{array} \]

(10) Derivational Licensing Condition on NPIs
An NPI is licensed, if it is c-commanded by a negative element in the course of derivation.  (Kato (2002: 5))

Here Kato presupposes that the concept of licensing should be interpreted by fundamental properties of c-command and feature sharing as in (11), where the relation of “X licenses Y” is designated as an ordered set \(<X,Y>\) at some level of semantic representation.

(11) \(<X,Y>\) holds, iff (i) X c-commands Y, and (ii) X and Y share a neg-feature.  (ibid.: 6)

In addition, Kato claims that the licensing should be restricted by the condition in (12), which is derived from (6).

(12) Asymmetric Licensing Condition
\[ X \text{ and } Y \text{ may not be reversed in derivation with respect to c-command relation where } X \text{ is a licenser (negative), } Y \text{ a licensee (NPI).} \]  (ibid.: 5)

In other words, Kato argues that a semantic representation, which is assembled based on information accessed throughout the course of the syntactic derivation, must not include contradictory pairs such as \(<X,Y>\) and \(<Y,X>\) but remain consistent.

Under these assumptions, Kato explains the cases including NPIs in the following way. Firstly, derivations of (7a, b) are illustrated in
A DERIVATIONAL APPROACH TO SUPERIORITY EFFECTS IN ENGLISH

(13b–d) and (14b–d), respectively. Here, a copy of X is represented as (X) and “X (derivationally) c-commands Y” as “X>Y.”

(13)  a. John didn’t eat anything. (= (7a)) (ibid.: 6)

b. The subject John Merges (and creates a VP)
   \[ VP \text{ John eat anything} \]

\[ not \text{ Merges with VP} \]
   \[ not [ VP \text{ John eat anything} ] \] (not > anything)

c. not Merges with VP
   \[ not [ VP \text{ John eat anything} ] \]

d. The subject Moves to Spec-IP
   \[ IP \text{ John didn’t } [ VP \text{ (John) eat anything} ] \] (not > anything)

(14)  a. *Anyone didn’t eat apples. (= (7b)) (ibid.: 6)

b. The subject anyone Merges (and creates a VP)
   \[ VP \text{ anyone eat apples} \]

\[ not \text{ Merges with VP} \]
   \[ not [ VP \text{ anyone eat apples} ] \] (not > anyone)

c. not Merges with VP
   \[ not [ VP \text{ anyone eat apples} ] \]

d. The subject Moves to Spec-IP
   \[ IP \text{ anyone didn’t } [ VP \text{ (anyone) eat apples} ] \]
   *(anyone > not)*

In (13), the c-command relation between the licenser (negative element) and the licensee (NPI) remains the same, while it is reversed in the course of derivation in (14). The contrast between (13a) and (14a) is thus well captured by (10)/(11) and (12).

The fact that the asymmetry between the subject and the object disappears when the NPI is embedded within the subject, as in (8), is also accounted for on the basis of (10)/(11) and (12). The derivation of (8) (repeated here as (15a)) is illustrated in (15b–d).²

(15)  a. A doctor \[ [ who knew anything] [ about acupuncture] \] was not available. (= (8)) (Linebarger (1980: 149))

b. The complex NP subject Merges (and creates a VP)
   \[ VP \text{ [ NP a doctor [ CP who knew anything about acupuncture] ] was available} \]

\[ not \text{ Merges with the VP} \]
   \[ not [ VP \text{ [ NP a doctor [ CP who knew anything about acupuncture] ] was available} ] \] (not > anything)

² Kato (2002) argues that the representational definition of c-command provided by Reinhart (1976) does not function in example (15), for under this definition anything in (15) is not c-commanded by not, and then this definition would wrongly predict (15) to be ungrammatical.
d. The complex NP subject Moves to Spec-IP
\[\text{IP } [\text{NP a doctor } [\text{CP who knew anything about acupuncture}] \text{ was not } [\text{VP } ([\text{NP a doctor } [\text{who knew anything about acupuncture}] \text{ was available}]) \text{ not > anything}]]\]

Even though the complex NP subject moves to Spec-IP in (15d), the reversal of c-command relation between not and anything does not occur, because anything is embedded within the complex NP subject, and anything does not c-command not under the definition of derivational c-command.

However, it faces a difficulty in explaining such sentences as (16a), as Nishioka (2005) points out. The derivation of (16a) is shown in (16b–e).³

(16) a. Does anyone not attend the party? (Nishioka (2005: 114))
   b. The subject anyone Merges (and creates VP)
      \[\text{VP anyone attend the party}\]
   c. not Merges with VP
      \[\text{not } [\text{VP anyone attend the party}] \text{ (not > anyone)}\]
   d. The subject Moves to Spec-IP
      \[\text{IP anyone does not } [\text{VP (anyone) attend the party}]\]
      *(anyone > not)
   e. \textit{Does} Moves to C⁰
      \[\text{CP \textit{Does} } [\text{IP anyone (does) not } [\text{VP (anyone) attend the party}]]\]

The c-command relation between not and anyone in (16c) reverses at the next stage of (16d). As a result, Kato’s (2002) proposal predicts (16a) to be ungrammatical, contrary to fact. Our analysis crucially differs from Kato’s in that we incorporate the make-up strategy in (4), which accommodates (16) without problem as shown in 4.3.

4. Superiority Effects

In this section, we show that our approach can explain the Superiority

³ Further examples in (i) are given by Nishioka (2005) as counter examples to Kato (2002) besides (16a).

(i) a. Didn’t anyone attend the party?
   b. I don’t think that anyone didn’t attend the party.

   (Nishioka (2005: 114))
effect of multiple *wh*-questions in English as well as NPIs comprehensively and in a uniform fashion.

4.1. Multiple *wh*-questions in English

Let us look at properties of multiple *wh*-questions in English. First, we repeat (1) as (17).

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

It has been argued that English multiple *wh*-questions in (17) show the Superiority effect. Namely, the subject *wh*-phrase must move to Spec-CP as in (17a), unlike the object *wh*-phrase as in (17b).

However, there are some cases in which this effect disappears. First are cases when a *wh*-in-situ phrase is structurally embedded, as shown in (18), where *what* or *whom* is embedded in the subject NP.

(18) a. Who do books about what annoy most?
   b. Who were stories about whom being told to?

(Stroik (1996: 90))

Secondly, multiple *wh*-questions including *which* NPs and *whose* NPs also do not exhibit the Superiority effect, as illustrated in (19) and (20).

(19) a. Which boy bought which book?
   b. Which book did which boy buy? (Hasegawa (2005: 436))

(20) a. Whose mother promised to read what to you?
   b. What did whose mother promise to read to you?

(Stroik (1996: 93))

Third are cases in which more than two *wh*-phrases occur in a sentence shown in (21). The sentences in (21) are grammatical, even though the object *wh*-phrase *what* can move over the *wh*-phrase(s), *who* (and *whom* in (21b)) which c-command(s) it.

(21) a. What did who say to whom? (Stroik (1996: 98))
   b. What did who persuade whom to buy?

(Pesetsky (2000: 50))

Finally, there is a controversial example provided by Lansik and Saito (1992) as in (22).

(22) Who wonders what who bought?

(Lasnik and Saito (1992: 118))

According to Lasnik and Saito (1992), this example is acceptable only when the lower *who* takes the matrix scope. The main point at issue is why the derivation of the embedded clause in (22) is allowed, even though it seems to be a violation of the Superiority effect as much as
4.2. An Analysis

As presented in section 2, we propose the general condition (3) (repeated here as (23)).

(23) Ban on Contradictory Information (BCI)

Semantic information may not include any contradiction.

In addition, we propose the make-up strategy (4) (repeated here as (24)) to satisfy (23). We assume that (24) only applies to the cases which involve semantic sharing between the licenser and the licensee, such as the licensing of wh-in-situ phrases, and NPIs. Therefore, we consider that this is not concerned with a syntactic licensing such as Case, Argument-Predicate relation, movement and others.

(24) The licensing relation is canceled if the relation reverses between the same two elements.

Moreover, we assume the licensing relation is established through the derivational c-command provided in (25)(=(5)), assuming the VP-internal subject hypothesis.

(25) Derivational C-command

\[ X \text{ c-commands all and only the terms of the category } Y \]

with which \( X \) was paired/concatenated by Merge or Move in the course of derivation. (Epstein et al. (1998: 10))

Our main idea is that a wh-in-situ phrase must be licensed by another wh-phrase in multiple wh-questions. The basic reason is that a wh-in-situ phrase needs another wh-phrase for a pair-list reading. Note that multiple wh-questions must involve more than one wh-phrase by definition; one is to move to Spec-CP overtly, the others are to remain in-situ. The relevant case is shown in (26), which reminds us of the fact that NPIs require a negative element as their licensor.

(26) a. *You said what? (non-echo reading)

b. Who said what?

This fact indicates that the wh-in-situ phrase forms a licensing relation with another wh-phrase. Therefore, we suggest that the possible licenser for the wh-in-situ phrase is another wh-phrase.

Here, we propose specific conditions on wh-in-situ phrase licensing, similar to Kato’s (2002) proposals on NPIs in (10)–(12). First, we assume that a wh-in-situ phrase must be c-commanded by another wh-phrase in the course of derivation in order to be licensed appropriately as in (27).
(27) Derivational Licensing Condition on Wh-in-Situ Phrases
A wh-in-situ phrase is licensed, if it is c-commanded by another wh-phrase in the course of derivation.

Next, following Kato (2002), we assume (28): “X licenses Y” is represented as an ordered set <X,Y> at some level of semantic representation. In the case of wh-in-situ phrase, the crucial feature that wh-phrases should share for licensing is assumed to be an interpretable Q-feature.4

(28) <X,Y> holds iff (i) X c-commands Y, and (ii) X and Y share a Q-feature.

We assume here that the Q-feature is assigned to a wh-word in the numeration. Furthermore, it is also supposed that this feature can optionally percolate from the wh-word to the whole phrase including it. Consequently, it can be proposed that the whole phrase can get a wh-status when this percolation occurs.

Moreover, the condition in (29) asserts that the c-command relation holding between wh-phrases may not be reversed. The structurally higher wh-phrase which c-commands another wh-phrase is a licensor, while the wh-phrase c-commanded by the higher wh-phrase is a licensee. Thus (29) is derived from the general condition on semantic information, i.e. BCI in (23).

---

4 In this paper, we adopt Chomsky’s (2000: 128) assumption for feature assignment to a wh-phrase: a wh-phrase which moves to Spec-CP overtly has an uninterpretable wh-feature and an interpretable Q-feature, while a complementizer has an uninterpretable Q-feature and EPP-feature. However, there is a controversy about whether a wh-in-situ phrase has the uninterpretable wh-feature or not. One view is that the wh-in-situ phrase has it. It follows from this view that all wh-phrases have them so that we can treat wh-phrases uniformly. However, given that the wh-in-situ phrase has no locality restriction, in this case, we must assume some particular Agree system which does not work for locality. On the other hand, another possibility is that the wh-in-situ phrase does not have the uninterpretable wh-feature, as argued in Nishioka (2005: 231). This assumption can accommodate the fact that the wh-in-situ phrase may not undergo movement and it is not sensitive to locality. However, the problem with this view is that the wh-in-situ phrase has a different lexical property from the wh-phrase which undergoes overt movement to Spec-CP, even though their lexical forms are the same. Thus, these two views of feature assignment for the wh-in-situ phrase require further research. We just suggest that the wh-in-situ phrase can be licensed by another wh-phrase which shares an interpretable Q-feature.
(29) Asymmetric Licensing Condition

X and Y may not be reversed in derivation with respect to a c-command relation where X is a licensor (wh-phrase), and Y a licensee (wh-phrase).

Now, let us consider the examples of multiple wh-questions in 4.1. The basic contrast of the Superiority effect in (17) (repeated here as (30a) and (31a)) can be accommodated as follows. As for (30a), the c-command relation between who and what remains the same as in the derivation of (30b-d), and (30a) satisfies (27)/(28) and (29). Then, it results in a grammatical sentence.

(30)  
   a. Who saw what? (=(17a))
   b. The subject who Merges (and creates VP)
      \[\text{VP who saw what}\] (who > what)
   c. The subject Moves to Spec-IP
      \[\text{IP who [VP (who) saw what]}\] (who > what)
   d. The subject Moves to Spec-CP through Spec-IP
      \[\text{CP who [IP (who) [VP (who) saw what]]}\] (who > what)

In contrast, however, (31a), which involves the violation of (29) as shown in (31b-d), is ungrammatical.

(31)  
   a. *What did who see? (=(17b))
   b. The subject who Merges (and creates a VP)
      \[\text{VP who saw what}\] (who > what)
   c. The subject Moves to Spec-IP
      \[\text{IP who [VP (who) saw what]}\] (who > what)
   d. The object what Moves to Spec-CP
      \[\text{CP what did [IP who [VP (who) see (what)]]}\] *(what > who)

In the shift from (31c) to (31d), the c-command relation between who and what reverses in violation of the condition in (29). Under our approach, the make-up strategy (24) functions, and then the licensing relation in (31b, c) is canceled. Therefore, this derivation violates (27)/(28), because the wh-in-situ phrase cannot be properly licensed. Consequently, the example results in ungrammaticality. Thus the contrast in the Superiority effect between (30a) and (31a) finds a natural explanation in our analysis.5

5 The example (31) can also be accounted for in terms of Attract Closest (AC)
Next, let us examine an example in which wh-in-situ phrases are embedded within the subject and the Superiority effect disappears such as in (18). The derivation of (18a) (repeated as (32a)) is illustrated in (32b–d).

(32) a. Who do books about what annoy most? (= (18a))
    (Stroik (1996: 90))
b. The complex NP subject Merges (and creates VP)
   [VP [NP books [PP about what]] annoy most who]
c. The complex NP subject Moves to Spec-IP
   [IP [NP books [PP about what]] [VP ([NP books [PP about what]]) annoy most who]]
d. The object who Moves to Spec-CP
   [CP who do [IP [NP books [PP about what]] [VP ([NP books [PP about what]]) annoy most (who)]]] (who > what)

Since what is embedded within the complex NP subject, what does not c-command the object who by the definition in (25). When who moves to Spec-CP at the stage of (32d), who c-commands what, and then what can be licensed by who in terms of (27)/(28). In addition, there is no reversal of the c-command relation in (32), or the condition in (29) is observed, and the example in (32a) results in its grammaticality.

Moreover, we can explain the cases including which NPs and whose NPs, where no Superiority effects are observed, along the same line. We repeat (19) and (20) as (33) and (34), respectively.

(33) a. Which boy bought which book? (= (19a))
b. Which book did which boy buy? (= (19b))
    (Hasegawa (2005: 436))

(34) a. Whose mother promises to read what to you? (= (20a))
b. What did whose mother promise to read to you? (= (20b))

defined as in (i), or the Minimal Link Condition (MLC) as in (ii). In the recent minimalist program, the MLC is part of the definition of the operation Agree (Chomsky (2000: 122)).

(i) Attract Closest (AC)
   \( \alpha \) can raise to target K only if there is no legitimate operation Move \( \beta \) targeting K, where \( \beta \) is closer to K. (Pesetsky (2000: 15))

(ii) Minimal Link Condition (MLC)
   K attracts \( \alpha \) only if there is no \( \beta \), \( \beta \) closer to K than \( \alpha \), such that K attracts \( \beta \). (Chomsky (1995: 311))

However, given the legitimacy of the example (42) in the text, it is suggested that the way the traditional constraints as in (i)/(ii) apply should be reconsidered.
Given that an interpretable Q-feature is assigned to a wh-word in the numeration as assumed in our analysis, it follows that which or whose can have this feature in (33) and (34) as well. In addition, we assume that which NPs and whose NPs have the structure of (35) and (36), for which we adopt the analysis of Radford (1997), respectively. If which or whose are embedded in a DP as in (35) and (36), they do not c-command another wh-phrase just like what embedded in books about what does not in (32a).

(35) \[\text{DP} [\text{D which [+Q]} [\text{N book}]] \quad \text{(Radford (1997: 277))}\]

(36) \[\text{DP} [\text{D who [+Q]} [\text{NP [D's [N mother]]}]] \quad \text{(ibid.: 278)}\]

Moreover, recall that the Q-feature can optionally percolate from the wh-word to the whole phrase including it, and then the whole phrase results in a wh-status, as mentioned here. As a result, in this case, the whole phrase with the Q-feature can undergo movement to Spec-CP, licensing a wh-in-situ phrase in terms of (27)/(28). Given this, in (33a), the Q-feature optionally percolates from which to which boy, and then which boy moves to Spec-CP. Then, it follows that which boy c-commands which book, and consequently, which boy results in a licenser for which book, satisfying (27)/(28) without violation of (29). In the case of (33b), the Q-feature optionally percolates from which to which book, and then which book can undergo movement. On the other hand, the Q-feature allocated to which in which boy does not percolate, and consequently which boy remains in-situ as in (33b). In this case, the Q-feature remains in which of which boy, and then, which keeps a wh-status, qualifying itself as a licensee. As a result, which of which boy can be licensed by which book through c-command relation between them in accordance with (27)/(28). In addition, there is no c-command reversal between them in (33b) without violation of (29), since the Q-feature in which of which boy does not percolate up to the whole DP phrase. Ultimately, its grammaticality is ensured. Also, the same explanation applies to (34). Thus, the fact that the sentence involving which NPs and whose NPs does not exhibit the Superiority

\[\text{Hasegawa (2005) gives a similar proposal for the Superiority effect including which NP in terms of Generalized A-over-A Principle, independently of our analysis here. See Hasegawa (2005) for the details.}\]
effect can be properly explained under our analysis.7

Here, before entering the discussion of more complex data, it would not be amiss to confirm the validity of our licensing condition for wh-in-situ phrases in (27)/(28) and (29), in which no movement of wh-in-situ phrases is assumed. In fact, several analyses which assume covert movement of wh-in-situ phrases at LF have thus far been proposed.8 However, as Simpson (2000) and Hasegawa (2003, 2005) argue, if wh-in-situ phrases were to undergo covert phrasal movement, a certain parallelism in behavior should be anticipated between them and other wh-phrases which move to Spec-CP overtly. Contrary to such expectation there are some cases which do not show the expected parallelism.

Firstly, a wh-phrase which moves to Spec-CP overtly cannot move over syntactic islands, while a wh-in-situ phrase does not show island effects, as illustrated in (37) and (38), respectively (See Fiengo et al. (1988), Reinhart (1995, 1998), and Simpson (2000), among others).9,10

The following examples pointed out by an anonymous reviewer can also be explained in terms of feature percolation.

(i) a. To whom did you give what?
   b. About what did you talk to whom?

In this case, a Q-feature optionally percolates from whom or what to to whom or about what, and consequently the whole category undergoes movement to Spec-CP, licensing a wh-in-situ element as well as (33) and (34). Therefore, examples in (i) result in grammaticality, satisfying (27)/(28) and (29). Specifically, as for (ia), its grammaticality can be explained by assuming that there is no asymmetric c-command between what and to whom as discussed in Oba (1998).

Approaches that a wh-in-situ phrase should undergo covert movement are based on several conditions/principles: the Superiority Condition (Chomsky (1973)), the ECP (Empty Category Principle) analysis (Chomsky (1986), Huang (1982)), the PCC (Path Containment Condition) analysis (May (1985)), and the multiple-specifier requirement analysis based on AC (Pesetsky (2000)) and others.

In (38a), what can be interpreted as a pair-list with both who and where. Our analysis can deal with two possible interpretations, since what can be licensed by either who or where because what has a c-command relation with both who and where.

Pesetsky (2000) argues that the problem of island effects shown in (37) and (38) can be resolved in terms of AC and the PMC (Principle of Minimal Compliance), which is originally proposed by Richards (1997). The PMC claims that the first wh-movement pays an AC tax (Attract Closest tax), which permits following wh-movement to the same position to ignore AC. However, Pesetsky’s analysis still cannot give a convincing account of the examples (39)–(42) shown in the text.

7 The following examples pointed out by an anonymous reviewer can also be explained in terms of feature percolation.

8 Approaches that a wh-in-situ phrase should undergo covert movement are based on several conditions/principles: the Superiority Condition (Chomsky (1973)), the ECP (Empty Category Principle) analysis (Chomsky (1986), Huang (1982)), the PCC (Path Containment Condition) analysis (May (1985)), and the multiple-specifier requirement analysis based on AC (Pesetsky (2000)) and others.

9 In (38a), what can be interpreted as a pair-list with both who and where. Our analysis can deal with two possible interpretations, since what can be licensed by either who or where because what has a c-command relation with both who and where.

10 Pesetsky (2000) argues that the problem of island effects shown in (37) and (38) can be resolved in terms of AC and the PMC (Principle of Minimal Compliance), which is originally proposed by Richards (1997). The PMC claims that the first wh-movement pays an AC tax (Attract Closest tax), which permits following wh-movement to the same position to ignore AC. However, Pesetsky’s analysis still cannot give a convincing account of the examples (39)–(42) shown in the text.
(37)  
a. *Who do you remember where we bought $t$? (wh-island)
b. *Who do you like books that criticize $t$? (complex NP island)
c. *Who do you think that pictures of $t$ are on sale? (subject island)
d. *Who did you get jealous because I spoke $t$? (adjunct island)
e. *What color hair did you meet [students with $t$]? (adjunct in NP island)

(38)  
a. Who remembers where we bought what? (wh-island)
b. Who likes books that criticize who? (complex NP island)
c. Who thinks that pictures of who are on sale? (subject island)
d. Who got jealous because I spoke to who? (adjunct island)
e. Who met [students with what color hair]? (adjunct in NP island)

(Fiengo et al. (1988: 81))

Next, a wh-phrase which moves to Spec-CP overtly shows that-trace effects in (39a). In contrast, a wh-in-situ phrase does not as in (39b).

(39)  
a. *Who do you think that $ti$ left?
b. ?Who $ti$ thinks that who left? (Lasnik and Saito (1992: 116))

In addition, a wh-in-situ phrase containing an anaphor cannot take the matrix subject John as its antecedent when it is in the embedded clause as in (40b), in contrast to (40a), where which picture of himself moves to Spec-CP in the embedded clause overtly.

(40)  
a. John wondered [which pictures of himself Bill liked $t$].
b. *John wondered [when Mary saw [which pictures of himself]]. (Simpson (2000: 30, 31))

Moreover, if a wh-phrase overtly moves from the embedded subject position in the case of wager-class verbs, the sentence is grammatical as in (41a). In contrast, the sentence with wh-in-situ phrase in that position is ungrammatical as in (41b) as well as a lexical object as in (41c).

(41)  
a. Who did John wager $ti$ to be $ti$ crazy?
b. *Who wagered who to be crazy?
c. *John wagered Peter to be crazy. (Bošković (1997: 52))

These differences between overtly moved wh-phrases and wh-in-situ phrases strongly favor no covert movement of wh-in-situ phrases. It might be argued that some factors cause the non-parallelism between
overt and covert movement. Theoretically, however, the null hypothesis predicts the same behaviors between overt and covert movement, and the burden of proof lies on the side of such argument. Without explicit arguments, a non-movement analysis of wh-in-situ phrases advocated here is preferable.  

Thus, our licensing condition in (27)/(28) and (29) which supposes no (covert) movement of wh-in-situ phrases in English is fairly well supported by empirical data.

4.3. Further Considerations

Let us move on to examine how our analysis explains more controversial data with regard to English multiple wh-questions.

First, the Superiority effect disappears in the following case, where acceptability is contingent upon its interpretation. As mentioned in section 4.1, (42) is acceptable only when the lower who takes the matrix scope.

\[(42) \text{ Who wonders what who bought? } (=(22)) \]
(Lasnik and Saito (1992: 118))

\[(43) * \text{What did who see? } (=(31a)) \]

The derivation of (42) proceeds in the following way. Here, who1 represents a wh-phrase in the main clause, and who2 a wh-phrase in the embedded clause.

\[(44) \text{ a. The subject who2 Merges (and creates VP)} \]
\[\text{[VP who2 bought what]} \quad (\text{who2 > what}) \]

\[\text{b. The subject Moves to Spec-IP} \]
\[\text{[IP who2 [VP (who2) bought what]]} \quad (\text{who2 > what}) \]

\[\text{c. The object what Moves to Spec-CP} \]


(i) Principle of Lexical Association (PLA)

An operator like “only” must be associated with a lexical constituent in its c-command domain. (Stroik (1996: 83))

(ii) Who only likes what? (ibid.: 84)

Given that (i) applies at LF, if a wh-in-situ phrase move into Spec-CP at LF, it follows that what cannot be associated with only. However, the fact that what can be associated with only suggests that wh-in-situ phrases must not move at LF.
The verb *wonders* Merges with embedded CP and the subject *who* Merges (and creates VP)

\[ [\text{VP who}_1 \text{ wonders } [\text{CP what who}_2 \text{ bought}]] \]  

\( (\text{who}_1 > \text{who}_2) \)

e. The subject *who*_1 Moves to Spec-CP through Spec-IP

\[ [\text{CP who}_1 [\text{IP (who,) } [\text{VP (who,) wonders } [\text{CP what who}_2 \text{ bought}]]]] \]  

\( (\text{who}_1 > \text{who}_2) \)

The step in (44c) induces a c-command reversal when *what* moves over *who*_2. This reversal violates the condition in (29) as happens in (43/31a), namely our general condition BCI in (23) (repeated here as (45)). Recall that the violation of BCI causes the cancellation of the licensing relation as a result of the make-up strategy to satisfy BCI, as stated in (24) (repeated here as (46)).

(45) Ban on Contradictory Information (BCI)

Semantic information may not include any contradiction.

(46) The licensing relation is canceled if the relation reverses between the same two elements.

However, this does not give rise to difficulty for our proposal, because at the later stage of (44d) *who*_2 can form a c-command relation with a new licenser *who*_1. Consequently, *who*_2 is appropriately licensed by *who*_1 in accordance with (27)/(28) and (29), and (42) results in a legitimate derivation. Namely, in (42), a new licenser can newly establish a proper licensing relation so that the example results in a well-formed sentence. On the other hand, as we argued in 4.2, in (43/31a), a new licenser does not occur after the cancellation of the licensing relation in terms of (46), and then the licensee is not appropriately licensed so that the example results in its ungrammaticality by violating (27)/(28) and (29). Finally, given this reasoning, we can properly account for the fact that (42) is acceptable only when *who*_2 takes the matrix scope. Thus, our analysis equipped with the make-up strategy can successfully resolve the problem of (42).

Next, let us examine the problematic data in multiple *wh*-questions as shown in (21a): the Superiority effect apparently disappears in multiple *wh*-questions involving more than two *wh*-phrases, which is called the additional *wh*-effect. The example of (21a) is repeated as (47).

(47) What did who say to whom? \( (=\) (21a))  \( (\text{Stroik (1996: 98)}) \)

Here, *wh*-in-situ phrases require focal stress, as pointed out by Stroik (1996: 101) and Hasegawa (2005: 440) independently. This is compati-
ble with Clifton et al.’s (2006) observation. Clifton et al.’s experiment based on the acceptability of multiple *wh*-questions, which argues that this additional *wh*-effect does not always take place, leads to the conclusion that this issue should be correlated with some non-syntactic factors such as processing difficulty, stress and prosody. It follows that we need some other solution rather than a complicated syntactic theory.

Translating this into the current approach, it can be argued that the acceptability of (47) can be improved because focal stress assignment should be enforced on *who* and *whom* to establish a pair-list reading and the c-command reversal may be overlooked. The derivation of (47) is shown as in (48).

   (48)  
   a. The subject *who* Merges (and creates VP)  
       [VP *who* said what to whom]  
       (who > what > whom)  
   b. The subject *who* Moves to Spec-IP  
       [IP *who* [VP (who) said what to whom]]  
       (who > what > whom)  
   c. The object *what* Moves to Spec-CP  
       [CP *what* did [IP *who* [VP (who) say (what) to whom]]]  
       (what > who > whom)

The c-command relation of “*who* > *what* > *whom*” turns into another relation of “*what* > *who* > *whom*,” when *what* moves to Spec-CP at the stage of (48c). Then, there is a cancellation of licensing relation in terms of (46), because of c-command reversal between *what* and *who*. However, *who* and *whom* can be appropriately licensed in terms of focal stress assignment. Specifically, following Stroik (1996) based on Chomsky (1981), it can be assumed that a *wh*-in-situ phrase with focal stress can obtain a status as an operator, and license another *wh*-in-situ phrase with focal stress as its variable, since it is assumed that an operator must license a variable. On the other hand, if the *wh*-in-situ phrase with focal stress does not license a variable, it cannot be an operator, and then must be licensed by another *wh*-phrase with focal stress. Then, in (48), even though the licensing relation is cancelled at the stage of (48c) in terms of (46), the succeeding focal assignment permits *who* and *whom* to establish a proper licensing relation as an operator and a variable independently of movement of *what*. As a result, this sentence results in grammaticality in spite of c-command reversal. Thus, under the idea that the non-syntactic factor should be responsible for the acceptability of this example, we can settle this problem by making use of a focal stress assignment. Thereby, this
example is not a counter evidence for our approach.\textsuperscript{12, 13}

Finally, let us see examples involving adverbial \textit{wh}-phrases such as \textit{why} and \textit{how}; \textit{why} and \textit{how} cannot occur in multiple \textit{wh}-questions.

\begin{equation}
(49) \quad a. \text{I wonder what you fixed \{{*why/*how/when/where}\}.}
\end{equation}
\begin{equation}
(49) \quad b. \text{I wonder \{{*why/*how/when/where}\} you fixed what.}
\end{equation}

(Oba (1998: 155))

We can incorporate this matter into our analysis by assuming an interpretable Referential(R)-feature proposed by Oba (1998). Namely, in order to be properly interpreted as multiple \textit{wh}-questions, \textit{wh}-phrases must have a Q-feature involving the R-feature, which can be derived from the fact that \textit{wh}-phrases in multiple \textit{wh}-questions must be referential to make a list for a pair-list interpretation. Then, Oba (1998) proposes that the R-feature is assigned to \textit{wh}-phrases such as \textit{who}, \textit{what}, and \textit{where}, while adverbial \textit{wh}-phrases such as \textit{why} and \textit{how} do not have this feature. Consequently, (28) can be elaborated here based on this idea as follows:

\begin{equation}
(50) \quad \langle X,Y \rangle \text{ holds iff (i) } X \text{ c-commands } Y, \text{ and (ii) } X \text{ and } Y \text{ share a Q-feature and an R-feature.}\textsuperscript{14}
\end{equation}

\textsuperscript{12} Note the difference between (47) and below:

(i) \text{You said what to whom.}

While (47) induces a pair-list reading, (i) is only interpreted as an echo question. Essentially, in \textit{wh}-questions, an uninterpretable Q-feature on C\textsuperscript{0} must be checked. However, C\textsuperscript{0} in (i) can not be assumed to have such a C\textsuperscript{0} so that this sentence cannot be licensed as a \textit{wh}-question. Therefore, in this case, we assert that the point of the matter is not \textit{wh}-in-situ phrases themselves, but independent of them. In short, that (i) is not licensed as a \textit{wh}-question gives rise to an absence of a pair-list reading.

\textsuperscript{13} It might seem that the following example can also be acceptable in terms of the focal stress assignment on a \textit{wh}-in-situ element.

(i) \text{*What did WHO see?}

However, as discussed in the text, we assume that the \textit{wh}-in-situ phrase with focal stress can be an operator, and in this case, it must license another \textit{wh}-in-situ phrase as its variable. Given this, we could account for that the acceptability of (i) cannot improve even with the focal stress, because of the absence of any other \textit{wh}-in-situ phrase which \textit{WHO} can license.

\textsuperscript{14} Similar cases given by an anonymous reviewer can also be explained under our analysis involving an R-feature. In (i), \textit{why} does not have the R-feature, and then these cases cannot be licensed as multiple \textit{wh}-questions in terms of (50).

(i) \quad \text{a. *Who came why?}
\begin{equation}
(50) \quad b. \text{*Why did who come?}
\end{equation}

Our analysis in terms of (50) would be also more favorable than the one in terms of
Under (50), the unavailability of occurrence of *why* and *how* in multiple *wh*-questions can be captured. Note here that this elaboration does not affect our previous arguments.

Up to this point, we have shown that our approach can provide a unified explanation to various cases involving multiple *wh*-questions. By the same token, the current approach can also deal with NPI licensing. Firstly, we can formulate the licensing condition on NPIs in (51) under our approach in terms of (45) and (46), by making modifications to Kato (2002). We suppose that NPIs must be licensed by elements with affective features such as negative elements and Q-elements. In addition, we present the asymmetric licensing condition in (52), which can be derived from BCI in (45). Notice that once a licensing relation violates (52), the make-up strategy in (46) functions in order to satisfy (45).

(51) Derivational Licensing Condition on NPIs
An NPI is licensed, if it is c-commanded by an element with affective features such as a negative element/a Q-element in the course of derivation.15

(52) Asymmetric Licensing Condition
X and Y may not be reversed in derivation with respect to a c-command relation where X is a licenser (element with affective features), and Y a licensee (NPI).

Under our approach, we can account for examples with regard to NPIs the ECP analysis given by Huang (1982) and others. One of the conceptual advantages is that we can deal with the fact that adverbial *wh*-phrases cannot occur in multiple *wh*-questions, without resorting to a complicated idea of government. It could be said that this is a crucial advantage for our analysis, in that government has been abandoned in the framework of the minimalist program. On the other hand, empirically, the current analysis can give an explanation to the following example, which could not be accounted for under the ECP analysis.

(ii) a. *Who fainted when you behaved how?*

Under our analysis, in (iib) *what way* with the R-feature can occur in the multiple *wh*-question, while in (iia) *how* without the feature cannot. As a result, we can easily handle the contrast in (ii).

15 An affective feature is a cover term coined by Klima (1964) for the licenser of NPIs. We here use it purely as a descriptive term to express the fact that NPIs are licensed by negative elements, Q-elements, conditionals and others, as discussed in Klima (1964), Ladusaw (1979), Progovac (1993), Nishioka (1994), among others.
given in section 3, repeated as (53) and (54) respectively here. Namely, along the same lines of Kato’s (2002) explanation, in (53a) and (54), the c-command relation keeps the same in the course of derivation, and then these examples satisfy (51) without violation of (52), resulting in their grammaticality. On the other hand, in (53b), c-command reversal induces a violation of (52), and the licensing relation is canceled in terms of (46). Consequently the example is ungrammatical in violation of (51).

(53)  
   a. John didn’t eat anything. (=13a))
   b. *Anyone didn’t eat apples. (=14a))

(54) A doctor [who knew anything] [about acupuncture] was not available. (=15a))

Furthermore, our strategy can also solve the problem in Kato’s (2002) analysis which is pointed out in section 3. We repeat the example (16a) and its derivation as (55).

(55)  
   a. Does anyone not attend the party? (=16a))
   b. The subject anyone Merges (and creates VP)
      [VP anyone attend the party]
   c. not Merges with VP
      not [VP anyone attend the party] (not > anyone)
   d. The subject Moves to Spec-IP
      [IP anyone does not [VP (anyone) attend the party]]
      *(anyone > not)
   e. Does Moves to C0
      [CP Does [IP anyone (does) not [VP (anyone) attend the party]]]
      (C0 > anyone)

Recall that (55a) was problematic to Kato’s (2002) analysis, in that this example results in grammaticality in spite of the c-command reversal. In contrast, our analysis accounts for (55a) as well as (42) in terms of (45) and (46). Our analysis goes as follows: not loses its status as a licenser for anyone when a contradictory relation between them occurs at the stage of (55d), because this reversal violates (52), and therefore the make-up strategy (46) functions successfully. However, at the later stage of (55e), anyone can get into c-command relation with a new licenser: C0. C0 is assumed to contain a Q-feature, and it can plausibly be argued that the Q-feature can also license NPI, as assumed since Klima (1964). Therefore, anyone is properly licensed by C0, and then the derivation results in legitimacy in accordance with (51) and
(52). Thus, our analysis has explanatory advantage over Kato (2002), since it can account for problematic cases for Kato’s (2002) analysis of NPI licensing.\textsuperscript{16}

To sum up this section, we have shown that our approach can handle some controversial cases involving the Superiority effect as well as NPIs. Consequently, our approach can obtain a theoretical unification for such licensing phenomena.\textsuperscript{17}

5. Concluding Remarks

In this paper, we have argued that the Superiority effect can be dealt with as well as NPI licensing comprehensively and in a uniform manner, in terms of the general condition BCI (56) and the make-up strategy (57) to satisfy the condition.

\begin{itemize}
\item (56) Ban on Contradictory Information (BCI)
\hspace{1cm} Semantic information may not include any contradiction.
\item (57) The licensing relation is canceled if the relation reverses between the same two elements.
\end{itemize}

More specifically, we can present a general licensing condition on \textit{wh}-in-situ phrases and NPIs as in (58), which constrains \textit{wh}-in-situ phrase and NPI licensing.

(58) General Licensing Condition on \textit{Wh}-in-Situ Phrases and NPIs
\begin{itemize}
\item a. A licensee (\textit{wh}-in-situ phrases and NPIs) must be c-commanded by its possible licenser in the course of deriva-
\end{itemize}

\textsuperscript{16} Examples (i) shown in footnote 3 can also be accounted for under the same explanation as the one for (55a). As for (ia), \textit{anyone} can be licensed by \textit{not} in \textit{Didn’t} at the later stage of derivation. On the other hand, in (ib), \textit{anyone} can be licensed by \textit{not} in the matrix clause, after there is a cancellation of a licensing relation between \textit{anyone} and \textit{not} in the embedded clause, in terms of c-command reversal.

\textsuperscript{17} It has been discussed by Rudin (1988), Pesetsky (2000), Richards (2001) among others that some multiple \textit{wh}-fronting languages such as Bulgarian show the Superiority effect. The main idea of this paper is that the possible licenser for a \textit{wh}-in-situ phrase is another \textit{wh}-phrase in English. However, in Bulgarian, it might be supposed that all \textit{wh}-phrases must front to Spec-CP because they have immediate relations to C\textsuperscript{0}. Given this, although some analysis in terms of AC may be valid for multiple \textit{wh}-fronting in Bulgarian, such analysis cannot handle English multiple \textit{wh}-questions (see footnotes 5 and 10).
b. The licensing relation between the licenser and the licensee cannot be reversed in multiple wh-questions and NPI licensing.

The theoretical significance of the proposal advocated here lies in the demonstration that a licensing relation between a licenser and a licensee is a more general one, applying not only to NPIs but also to wh-in-situ phrases in multiple wh-questions.

REFERENCES

A DERIVATIONAL APPROACH TO SUPERIORITY EFFECTS IN ENGLISH 443

444, Kaitakusha, Tokyo.


Tancredi, Christopher (1990) “Not Only *Even*, But Even *Only*,” ms., MIT.

3–9–49–201 Hakozaki, Higashi-ku
Fukuoka-shi 812–0053
e-mail: megmeg@jd6.so-net.ne.jp