

# Identity as a Chain-Formation Condition

Akiko KOBAYASHI\*

## 1. Introduction

This paper investigates how various kinds of *wh*-movements are allowed/ disallowed in the Minimalist Program. In Kobayashi (1998a) I claim that the main previous analyses, the Minimal Link Condition (MLC), Kitahra's (1997) Chain Formation Condition (CFC), and Kuno and Takami's (1993) functional analysis are unable to deal with a variety of multiple *wh*-question data.<sup>1)</sup> In Section 2, I propose that chain-formation conditions should be re-defined and that the legitimacy of *wh*-movement should be guaranteed by one of those chain-formation conditions. In Section 3, I show how well this analysis explains the relevant data. Section 4 concludes the discussion.

## 2. New Chain-Formation Rules

In this section I provide chain-formation conditions. Before considering the conditions, however, it is necessary to make clear which minimalist principles and assumptions this analysis relies on.

### 2.1 Principles and Assumptions

In spite of a number of researches within the minimalist framework, the assumptions proposed there are still various and sometimes incompatible. Following are the principles and assumptions that are germane to this paper:

- ( i ) Chomsky's (1995) MLC
- ( ii ) Rizzi's (1990) Relativized Minimality
- (iii) Kuroda's (1968) decomposition of *wh*-phrases
- (iv) Chomsky's (1993) 'copy + deletion' theory

Chomsky's (1995) MLC determines which element is to be attracted. Note that it differs from Rizzi's (1990) Relativized Minimality, or, Chomsky's (1993) MLC, which prevents an element from moving over its potential landing site. It is derived from the economy consideration (Shortest Move) and therefore is violable for convergence. Although most minimalist works have usually chosen between these two MLC's for their purposes, here I



position that remains at LF should be FF(WH), whereas the part of *whose picture* in the base-generated position that remains at LF should be FF(Pro) (and features concerning its meaning, of course). Then the LF representation of (1a) will be like the following:

- (2) [<sub>CP</sub> *whose picture* did [<sub>TP</sub> John see *whose picture*]]  
 LF:     FF(WH)                             FF(Pro), meaning

In the next subsections, I will advance new chain-formation conditions and show how they account for the relevant data on the basis of these minimalist principles and assumptions.

## 2.2 Chain-Formation Conditions

In this section, I suggest chain-formation conditions. The most important point of the analysis is that a chain is formed only when a moved element is ‘identical’ with its base element from which the movement starts.<sup>2)</sup> If this requirement is met, a chain is automatically formed at that point of derivation. Let us call it a ‘derivational chain’ (D-chain) and define it as follows:

- (3) A D-chain is formed only if  
 a. a moved element is a full copy of its base element, and  
 b. the movement is legitimate.

Note that (3) entails that no D-chain is formed when covert movement applies to a base-generated element. For instance, when an object DP covertly shifts its FF to a Case position, a D-chain is not formed since its FF is not a full copy, but just a ‘partial’ copy of the DP. If they do not form a chain, how should the covert movement avoid the nonconvergence at LF?

Here another chain-formation is needed as a relief measure. I propose ‘LF-chain’ formation defined as follows:

- (4) At the point of interpretation, an LF-chain can be formed only if a moved element and its base element are nondistinct.

Rigidly, this is not a chain-*formation* since LF representation undergoes no operations: the two elements are *interpreted* as nondistinct, if they share their FF.

Now let us turn to the case mentioned above. At LF representation, Delete is applied to object DP and then there remain FF(Pro) and meaning features at the position.<sup>3)</sup>

- (5)    .... FF(DP)    .... DP  
       |            |  
 LF: FF(Pro)    FF(Pro), meaning

At this point the two elements share FF(Pro) and therefore are interpreted as nondistinct.

Since this interpretation has the same effect as the D-chain-formation in (3), the LF representation satisfies FI and converges.

Thus we obtain LF-chain-formation as a remedy for LF-nonconvergence. In this connection, let us make clear when this remedy takes place. In the previous section I have adopted Chomsky's (1993) 'copy + deletion' theory, according to which Delete should apply to each position to interpret that position. For instance, *what* in an operator position deletes all of its features except FF(WH). Then it is interpreted as an operator. Note that, however, if this *wh*-phrase has no variable to bind, FF(WH) exhibits vacuous quantification. To save the representation from crashing, the remedy, i.e. LF-chain-formation, should take place. Specifically, FF(WH) in the operator position seeks any other element that contains the same FF(WH). If there is one, then the two elements are interpreted as nondistinct, and the operator avoids vacuous quantification. It must be noted that LF-chain-formation takes place *only after* Delete applies to a lexical element: every full lexical item has a potential to be interpreted if it does not constitute a nontrivial chain, and an interpretation problem to be remedied arises only after Delete applies. As for the case above, for example, we do not know that *what* in the operator position is a problem until we try to interpret it as an operator through the application of Delete.

Now let us consider when LF-chain-formation applies to FF(LI). FF(LI) is different from a full lexical item in that FF(LI) can be interpreted only when it forms a nontrivial chain, since it stands as a partial copy of some LI. In other words, we know that FF(LI) which stands alone bears a problem for interpretation and therefore requires a remedy, even before Delete is applied. Then let us assume that FF(LI) can form an LF-chain before Delete. Take the following situation for an illustrative example of this procedure: object

*what* covertly shifts its FF copy to the Case position. This shift does not create a D-chain since *what* and  $FF(\textit{what})$  are not identical. At LF representation, Delete applies to the base-generated *what* since it is a full LI, whereas  $FF(\textit{what})$  withholds the application of Delete since it cannot receive an interpretation unless it forms a nontrivial chain, as assumed above. (6) shows the LF representation of the relevant part:<sup>4)</sup>

$$(6) \quad \dots \quad \text{FF}(\textit{what}) \quad \dots \quad \textit{what}$$

$$\text{LF: } \text{FF(WH)} + \underline{\text{FF(Pro)}} \quad -\text{FF(WH)}, \underline{\text{FF(Pro)}} \quad (\text{meaning})$$

Here  $FF(\textit{what})$  forms an LF-chain to ensure a proper interpretation. It is possible since the two elements share  $FF(\text{Pro})$ . Hence  $FF(\textit{what})$  can form a legitimate chain and avoid nonconvergence at LF.

To sum up, we have the following generalization as to the timing of the application of LF-chain-formation:

- (7) a. A full lexical item forms an LF-chain after the application of Delete.  
 b.  $FF(\text{LI})$  (with a trivial chain) forms an LF-chain before the application of Delete.

In the next section let us see how these chain-formation rules, i.e. D-chain-formation in (3), LF-chain-formation in (4), and the timing of LF-chain-formation in (7), work for the purpose of explaining various types of multiple *wh*-interrogative data.

### 3. An Explanation

Let us start by considering simple sentences such as (8a-c).

- (8) a. \**What* did *who* buy *t*?  
 b. ?\**What* did you tell *who* to read *t*?  
 c. ?\**Who* did you give *what* to *t*?

These deviant examples are correctly excluded by Chomsky's (1995) MLC.

Now let us turn to the explanation of the special property of *how/ why*: why are such *wh*-adjuncts not allowed to remain in situ, as shown by the deviance of (9)?

(9) \**Who t* came to the party *how/ why*?

Overt *wh*-movement of *who* poses no problem: it forms a D-chain since it moves an identical copy of the base element. The problem has to do with covert movement of *how/ why* to the operator position. When this covert raising takes place, a D-chain is not formed since the moved element (FF(*how/ why*)) is not identical with the base-generated element (*how/ why*). LF-chain-formation is therefore required but turns out to be impossible. Consider (10), the LF representation of the relevant part of (9):

(10) FF (*how/ why*) + C ..... *how/ why*  
 LF: FF (WH) (no FF remained)

Note that we have postulated that the only FF that adjunct *wh*-phrases bear is FF(WH), and in (7), that Delete applies to LI but not to FF(LI) before LF-chain-formation. Consequently, the base-generated element deletes its only FF, i.e. FF(WH), whereas FF(*how/ why*) maintains its FF(WH) as it is. Since the two elements share no FF, an LF-chain is not formed. Therefore, (10) does not satisfy FI and (9) crashes.

Unlike the case of *how/ why*, covert raising of argument *wh*-phrases can form a chain with no problem. Consider the convergent derivation (11):

(11) *Who t* bought *what*?

Overt *wh*-movement of *who* forms a D-chain. On the other hand, covert raising of *what* creates no D-chain since FF(*what*) and *what* are not identical. (12) shows the movement paths of FF(*what*) in the covert syntax:

(12) [<sub>CP</sub> *who* FF (*what*) + C [<sub>TP</sub> ... FF (*what*) + Vb [<sub>VP</sub> *t*, *what*]]]

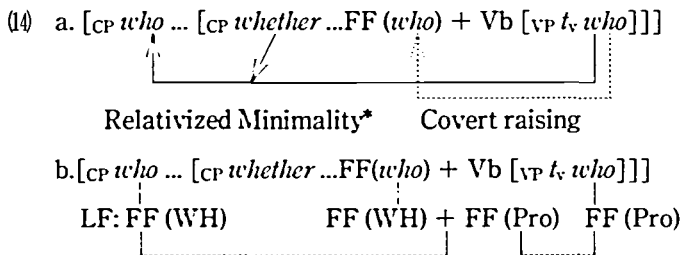
LF: FF (WH) + FF (Pro) FF (WH) + FF (Pro) FF (Pro)

It should be noted that covert raising of FF(*what*) takes place twice, and that the two positions, i.e. the operator position and the Case position contain the identical elements (FF(*what*)). Thus they form a D-chain. Since those FFs have already formed a nontrivial chain, Delete applies to each FF position: the operator deletes its FF(Pro) part, and the Case position deletes its FF(WH) part. At this point the latter FF(*what*) consists of the same FF, i.e. FF(Pro), as the base-generated element *what*, to which, too, Delete has applied and deleted its FF(WH) part. Therefore the two elements are united by LF-chain-formation, and (12) satisfies FI and converges. It is a correct prediction.

As a second step, let us consider how we can account for those sentences which Kitahara (1997) deals with under his CFC analysis. In fact, argument-adjunct/ subject-object contrast in long-distance extraction can also be accounted for under the present chain-formation conditions. Let us consider (13):

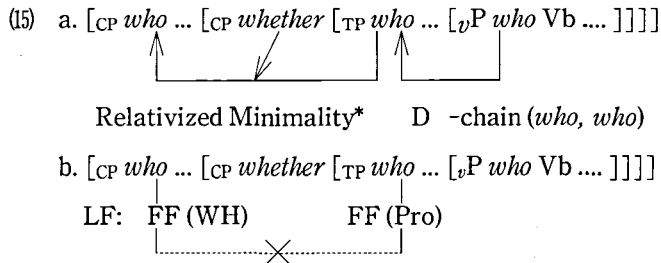
- (13) a. \**Who* do you wonder *whether t* can help us?  
 b. ??*Who* do you wonder *whether we* can help *t*?  
 c. \**How* do you wonder *whether we* can help Bill *t*?

From the assumption that *whether* is not attracted by the matrix C (cf. Kobayashi (1998b)), all of the long-distance movements in (13) observe the MLC. However, they do not form a D-chain since they violate the Relativized Minimality by moving over the potential landing sites. Then it is expected that the derivation converges if the failed D-chain-formation is made up for by LF-chain-formation, and that the derivation crashes if not. Firstly, let us consider example (13b), where the object *wh*-phrase has undergone a long-distance movement. Its movement paths and LF representation are shown in (14a) and (14b), respectively:



Being an object, *who* undergoes two movements: one to the operator position and the other to the Case position. Neither movement creates a D-chain, since the former violates the Relativized Minimality, and the latter moves a non-identical copy. At LF representation (14b), according to (7), *who*'s in the operator and the base-generated positions delete their irrelevant features before LF-chain-formation, whereas FF(*who*) in the Case position stands as it is. At this point, *who* in the operator position and FF(*who*) in the Case position can form an LF-chain since they share FF(WH). However, note that they are not fully identical because FF(*who*) is composed of two FF clusters but only one of them is used for identification. The LF-chain formed under the partial identity is less perfect, and degrades grammaticality. After this first LF-chain-formation, Delete applies to FF(*who*) and deletes its FF (WH) part that is irrelevant to interpretation. Then the last two elements in (14b) share the same FF, i.e. FF(Pro), and form an LF-chain. This time the LF-chain has no problem since the identity is fully guaranteed. Therefore, (14b) consists only of legitimate chains and converges. A slight degradation of grammaticality of (13b) is attributed to the fact that the identity between the elements in the operator and the Case positions is not fully guaranteed.

On the other hand, the LF representations of (13a, c) never satisfy FI. Let us consider the derivation and the LF (13a). Its movement paths and relevant LF representation are shown in (15a) and (15b), respectively:

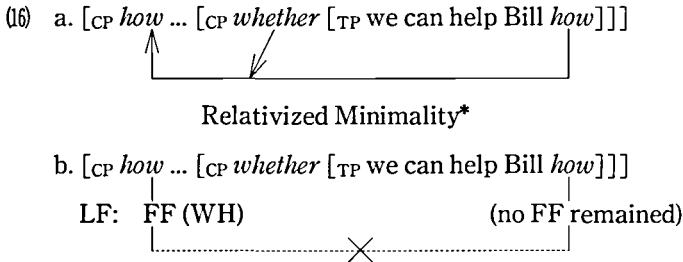


Since overt A-movement of *who* to the subject position moves an identical copy of base-generated *who* in [Spec, *vP*], it forms a D-chain. The second overt movement, i.e. long-distance *wh*-movement, does not form a D-chain since it violates the Relativized Minimality. At LF representation, both *who*'s in the operator and the subject positions delete their irrelevant features before LF-chain-formation, according to (7). Then, as (15b)



shows, the two elements do not share any features, and LF-chain-formation is impossible. Therefore (15b) does not satisfy FI and crashes. The deviant status of (13a) thus obtains.

A similar account holds for (13c). Its derivational paths and LF representation are shown in (16a) and (16b), respectively:



Overt *wh*-movement of *how* does not form a D-chain since it violates the Relativized Minimality. At LF representation, LF-chain-formation is also impossible, since, as (16b) shows, after Delete applies, the two elements do not have any features in common: in the operator position there remains only FF(WH), and in the base-generated position *how* deletes its sole FF, i.e. FF(WH), and there remains no FF there. Therefore (16b) crashes and the deviant status of (13c) obtains.

In this way, the present chain-formation analysis together with the FI principle can capture the subject-object asymmetry (13a, b), and the adjunct-argument asymmetry (13b, c).

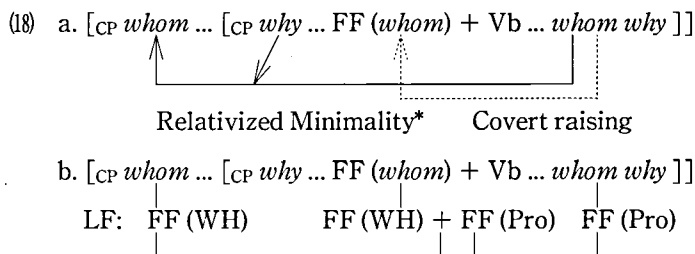
This chain-formation analysis can also provide an account for the second type of complex multiple *wh*-interrogatives, where both of two *wh*-phrases are forced to move overtly. Consider examples (17):

- (17) a. ??*Whom* do you wonder *why* John will invite *t* ?  
 b. \**Why* do you wonder *whom* John will invite *t* ?

One possible explanation is that this contrast in grammaticality is derived from the MLC: since *why* is closer to the embedded C than *what*, *why* should move to the embedded operator position, as (17a) shows. However, the fact is that *wh*-phrases with different scope do not compete for attraction.<sup>51</sup> Then *whom* and *why* in the examples above should not compete

for attraction: the embedded C can ‘see’ only those *wh*-phrases with embedded scope feature, and the matrix C can ‘see’ only those with matrix scope feature. Then in (17b), the attraction of *whom* by the embedded C does not cause an MLC violation since *why* bears a matrix scope feature and is irrelevant to the attraction from the embedded C. If so, what makes example (17b) deviant?

First, let us consider the convergent (17a). Its derivational paths are shown in (18a) and its LF representation in (18b):



Overt movement of *why* does not pose a problem: it moves an identical copy of the base-generated element and forms a D-chain. As for *whom*, it undergoes two movements. One is to the operator position in the overt syntax, which violates the Relativized Minimality and creates no D-chain. The other is to the Case position in the covert syntax, which creates no D-chain, either, since it does not move an identical copy. At LF representation, being overt elements, both *whom*'s in the operator and the base-generated positions delete their irrelevant features before LF-chain-formation. On the other hand, FF(*whom*) in the Case position stands as it is. At this point, i.e. in (18b), the two *wh*-phrases in the operator and the Case positions form an LF-chain since they are nondistinct. After this chain-formation, Delete applies to FF(*whom*) and deletes its FF(WH) part. Then FF(*whom*) share the same FF content with *whom* in the base-generated position. Then an LF-chain is formed between them, too. Consequently, (18b) satisfies FI and converges. Thus we obtain the convergent (17a): its somewhat degraded status is due to the fact that the first LF-chain has been formed under the ‘partial’ identity.

Now let us consider deviant sentence (17b), where *whom* undergoes a local movement, while *why* undergoes a long movement. Its movement paths and LF representation are shown in (19a) and (19b), respectively:

- (19) a. [<sub>CP</sub> *why* ... [<sub>CP</sub> *whom* [<sub>TP</sub> ..... *whom why* ]]]
- 
- Relativized Minimality\*
- b. [<sub>CP</sub> *why* ... [<sub>CP</sub> *whom* [<sub>TP</sub> ..... *whom why* ]]]
- 
- LF: FF (WH) (no FF remained)

Overt movement of *whom* poses no problem: it legitimately moves an identical copy, and then forms a D-chain. As for *why*, it violates the Relativized Minimality, and hence does not form a D-chain. At LF representation, being overt elements, both *why*'s in the operator and the base-generated positions delete the irrelevant features before LF-chain-formation. Consequently LF-chain-formation is impossible since the two elements share no FF. Thus the LF (19b) crashes, and hence sentence (17b) results in deviance.

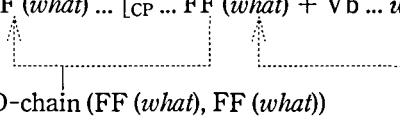
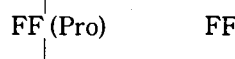
From what has been said above, we can see that the present analysis successfully and uniformly accounts for various types of multiple *wh*-interrogative data. With these chain-formation conditions, we regain the explanatory power of the ECP.

Before concluding the discussion, it should be noted that the present analysis has an extra empirical merit that the previous analyses did not have. Consider the contrast between the following examples, taken from Lasnik and Saito (1992:126):

- (21) a. *Who* said that John bought *what*?  
 b. ?*Who* said that *who* bought the book?

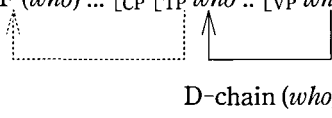
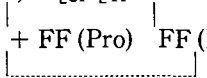
In the GB framework, neither Superiority nor the ECP provides a proper account for this subject-object asymmetry. In fact, Lasnik and Saito leave this contrast unexplained.

Let us consider how the present analysis deals with the contrast in (20). We disregard the movement of matrix *who* since it poses no problem: it moves an identical copy of the base-generated element, thereby forming a D-chain. The other *wh*-phrase in each example, i.e. object *what* in (20a) and subject *who* in (20b), shows a different process as to chain-formation. The movement paths of *what* in (20a) are shown in (21a) and its LF representation in (21b):

- (21) a.  $[_{CP} \textit{who} \text{FF}(\textit{what}) \dots [_{CP} \dots \text{FF}(\textit{what}) + \text{Vb} \dots \textit{what} ]]$   
  
 D-chain (FF (*what*), FF (*what*))
- b.  $[_{CP} \textit{who} \text{FF}(\textit{what}) \dots [_{CP} \dots \text{FF}(\textit{what}) + \text{Vb} \dots \textit{what} ]]$   
 LF:  $\text{FF}(\text{Pro}) \quad \text{FF}(\text{Pro})$   


Being an object, *what* undergoes two covert movements: one is to the Case position, and the other to the operator position. The latter movement forms a D-chain since the two elements in the operator and the Case positions are identical, whereas the former one does not, since FF(*what*) in the Case position and *what* in the base-generated position are not identical. At LF representation, Delete applies to all the *wh*-elements: it applies to *what* in the base-generated position since, according to (7), LI undergoes Delete before LF-chain-formation. Delete also applies to FF(*what*) since it has already formed a nontrivial chain. At this point, as shown in (21b), the last two elements share the same FF content, and LF-chain is formed with no degradation in grammaticality. Hence the LF (21b) satisfies FI and converges. The full grammatical status of (20a) thus obtains.

On the other hand, subject *who* in (20b) has difficulty in forming a chain. The movement paths of *who* are shown in (22a) and the LF representation in (22b):

- (22) a.  $[_{CP} \textit{who} \text{FF}(\textit{who}) \dots [_{CP} [_{TP} \textit{who} \dots [_{VP} \textit{who} \text{Vb} \dots ]]]]$   
  
 D-chain (*who*, *who*)
- b.  $[_{CP} \textit{who} \text{FF}(\textit{who}) \dots [_{CP} [_{TP} \textit{who} \dots [_{VP} \textit{who} \text{Vb} \dots ]]]]$   
 LF:  $\text{FF}(\text{WH}) + \text{FF}(\text{Pro}) \quad \text{FF}(\text{Pro})$   


The embedded subject *who* takes two movements: one is to the subject position in the overt syntax, and the other to the matrix operator position in the covert syntax. The former creates a D-chain since it moves an identical copy of the base-generated element, while the latter does not since it moves a non-identical copy. At LF representation, *who* in the subject position deletes the irrelevant features before LF-chain-formation while FF (*who*) in

the operator position stands as it is, according to the generalization of (7). Then the LF-chain-formation between them is possible because they have FF(Pro) in common. However, the chain is not perfect since the identity of the two elements is not fully guaranteed: although FF(*who*) in the operator position consists of two FF clusters, only one of them is used for the identification. Therefore, a convergent, yet slightly degraded sentence (20b) obtains.

#### 4. Conclusion

In this paper I have shown that the chain-formation conditions fitted in the minimalist framework correctly account for a wide variety of *wh*-question data in a uniform way. Legitimate chains are formed by one of the two procedures:

- ( i ) D-chains are formed under *identity*.
- ( ii ) LF-chains are formed (or, interpreted) under *nondistinctness*.

The difference (identity vs. nondistinctness) should be derived from the assumption that LF-chain-formation is a relief measure to save the otherwise uninterpretable LF representation.

It should be noted that the present analysis has a conceptual merit, too: it gives a possible answer to the redundancy problem. As noted by Brody (1995), if all possible movements automatically created well-formed chains, then the legitimacy of derivation and that of representation would be highly interrelated. However, under the present analysis, a chain is formed by independent conditions: some legitimate operations of Move satisfy the necessary and sufficient conditions for chain-formation, and some do not. Thus, redundancy does not occur since movement is regulated by derivational constraints (eg. MLC, Economy) and chain-formation is regulated by representational constraints (i.e. the identity between an element and its copy that is made when moving).

#### NOTES

\* I would like to express my deep appreciation to Kunihiro Iwakura, Mitsunobu Yoshida, and Hiromu Sakai for their invaluable comments and suggestions. Needless to say, responsibility for the texts rests entirely upon me.

- 1) See Kobayashi (1998a) for a detailed review of their analyses.
- 2) Since the Minimalist Program holds 'Inclusiveness' on its procedural basis, nothing but

identity between the moved elements guarantees the chain-formation. With this regard I consider this requirement conceptually valid.

- 3) Delete might not be applied to DP if it does not contain other formal features than FF(Pro). In any way, what is important here is that Object Shift does not form a D-chain for lack of identity, and that some rescue measure is required at LF representation.
- 4) Chomsky (1995:265) claims that 'Move F automatically carries along FF(LI), the set of formal features of LI.' This paper follows his claim and assumes that FF(*what*) takes FF(WH) as well as FF(Pro) to its Case-checking position, although FF(WH) part is obviously unnecessary for Case-checking. Since this pied-piping is automatic, economy consideration is irrelevant.
- 5) Consider the following example:

(i) *Who* wonders *what who* bought *t*? (Lasnik and Saito (1992:118))

This is grammatical if the embedded subject *who* takes a matrix scope. If *wh*-phrases of different scope competed for attraction, the embedded C would attract *who* instead of *what* and (i) would never be generated.

## REFERENCES

- Bošković, Ž. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.
- Brody, M. 1995. *Lexico-logical grammar: A radically minimalist theory*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, eds. by K. Hale and S. Keyser. Cambridge, Mass.: MIT Press: 1-52.
- Chomsky, N. 1995. *The Minimalist program*. Cambridge, Mass.: MIT Press.
- Kitahara, H. 1997. *Elementary operations and optimal derivations*. Cambridge, Mass.: MIT Press.
- Kobayashi, A. 1998a. On multiple Wh-interrogatives in English. *Studies in European and American Culture* 5: 69-82.
- \_\_\_\_\_. 1998b. On multiple WH-interrogatives in English: Reformulation of the ECP. A paper read at the 117th conference of the Linguistic Society of Japan.
- Kuno, S. and K. Takami. 1993. *Grammar and discourse principles: Functional syntax and GB*

*theory*. Chicago: The University of Chicago Press.

Kuroda, S.-Y. 1968. English relativization and certain related problems. *Language* 44: 244-268.

Lasnik, H. and M. Saito. 1992. *Move  $\alpha$* . Cambridge, Mass.: MIT Press.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.