MOVEMENT OF ANTECEDENTS AND MINIMALITY

Yuji Takano
Kinjo Gakuin University

1. Introduction

One issue in syntax hotly debated in recent years is how to treat what has traditionally been analyzed in terms of referential dependencies. One instance of this general issue has to do with the treatment of obligatory control. The traditional approach to obligatory control claims that the subject of the control complement clause is a phonetically null pronominal element called PRO and that this PRO is controlled by an element in the higher clause, as shown in (1a). On the other hand, since Hornstein's (1999, 2001) influential work, an alternative approach has been pursued by a number of researchers according to which the controller-PRO relation is replaced by a movement relation, so that the controller moves from the subject of the control complement, as shown in (1b), where the material surrounded by angled brackets shows an unpronounced copy of the moved element.

(1) a. John_i tried [PRO_i to leave].

b. John thinks that <John> he is smart.

This move has also affected the analysis of binding relations. Thus, a number of authors propose that what has traditionally been analyzed in terms of binding be replaced by movement, in such a way that antecedents moves from the position of pronouns or reflexives (Hornstein 2001, Motomura 2001, Kayne 2002, Zwart 2002, Fujii 2007, Lasnik 2007, Miyamoto 2008). According to this proposal, the binding relation between *John* and *he* in (2a) should be captured in terms of movement of *John* from the position of *he*, as shown in (2b).[1]

(2) a. John_i thinks that he_i is smart.

b. John thinks that <John> he is smart.

[1] There are two major proposals about what exactly happens in the subject of the embedded clause. Hornstein (2001) assimilates it to obligatory control, claiming that *John* is merged directly into the subject θ-position of the embedded clause and moves from there, with its copy spelled out as the pronoun *he*. Kayne (2002) proposes an alternative in which *he* and *John* get merged first to form a constituent and this constituent is merged into the embedded subject position, and then *John* moves out of this constituent (in fact, Kayne proposes the same analysis for obligatory control; thus, in this analysis, *John* and PRO in (1) form a constituent, this constituent is merged into the embedded subject position, and *John* moves out of this constituent to the matrix subject position). I will remain neutral on the choice between the two analyses until section 3, where I will in fact argue for Kayne's analysis.

Nanzan Linguistics 9, 193–214

Hornstein (2001) and Kayne (2002) argue for the movement approach to binding shown in (2b) by reducing the effects of binding conditions A, B, and C to properties of movement. One direct consequence of the movement approach, according to Kayne, is that it straightforwardly explains condition C effects of the kind shown in (3a) without appeal to condition C.

(3) a. *He$_i$ thinks that John$_i$ is smart.

b. <John> he thinks that John is smart.

In the movement approach, the sentence in (3a), on the intended coreferential interpretation, involves a derivation illustrated in (3b), where *John* has moved from the matrix clause to the embedded clause. This movement is an instance of lowering, which is generally prohibited in syntax. Thus, the movement approach reduces this kind of condition C effect to the general property of movement. Hornstein and Kayne discuss other effects traditionally attributed to binding conditions A, B, and C, in an attempt to capture them in terms of properties of movement of antecedents.

The purpose of this article is two-fold. First, it defends a movement approach to binding from a novel perspective. Second, it explores a theoretical issue related to the nature of movement of antecedents. The empirical evidence for a movement approach to binding comes from consideration of multiple clefting constructions in Japanese. I will show that to account for their properties, we need to consider antecedents to have moved from the position of the pronouns or reflexives they bind. The result thus lends support to a movement approach to binding. Moreover, I will show that movement of antecedents in those cases shows no minimality effects. This fact raises an interesting theoretical question. I will propose an account of this fact by looking closely at the nature of antecedent movement.

This paper is organized as follows. In section 2, I will show that certain facts about multiple clefting in Japanese provides a new argument in favor of the hypothesis that antecedents move from the position of pronouns/reflexives. In section 3, I will propose an analysis of the absence of minimality effects with A-movement of antecedents by capitalizing on the hypothesis that antecedent movement is movement from a non-θ-position to a θ-position. I will then extend this analysis to A-movement in copy raising constructions in English, which, unlike antecedent movement, does show minimality effects. In section 4, I will conclude the discussion.

2. New Evidence for Antecedent Movement: Multiple Clefting in Japanese

In this section, I provide an argument in favor of a movement approach to binding from a novel perspective, by considering the effects of (phonetically null) pronouns and the

(nonlocal) reflexive *zibun* on multiple clefting in Japanese.[2] Let us begin by looking at properties of cleft constructions in Japanese.

Japanese has a cleft construction that freely allows multiple elements to appear in the focus position. The examples in (4b, c) are cleft sentences formed on the basis of the simple sentence in (4a).

(4) a. Ken-ga Mari-ni hon-o ageta.
Ken-NOM Mari-DAT book-ACC gave

'Ken gave a book to Mari.'

b. Ken-ga hon-o ageta no-wa Mari-ni da.
Ken-NOM book-ACC gave C-TOP Mari-DAT COP

'It is to Mari that Ken gave a book.'

c. Hon-o ageta no-wa Ken-ga Mari-ni da.
book-ACC gave C-TOP Ken-NOM Mari-DAT COP

(Lit.) 'It is Ken to Mari that gave a book.'

In (4b) the dative object appears in the focus position between the topic marker and the copula and in (4c) the subject and the dative object appear there. (4c) is an instance of multiple clefting.

There are two properties of this type of cleft construction relevant for the present discussion. First, this type of cleft construction involves movement of the focus element (Hoji 1987).[3] Second, in multiple clefting, the focus elements must be clausemates (Koizumi 1995,

[2] There are three previous studies based on Japanese that argue for a movement approach to binding. Motomura (2001) first proposes that certain properties of the reflexive *zibun* can be derived in a uniform way under the hypothesis that the antecedent undergoes overt A-movement from the position of *zibun*. Miyamoto (2008) then argues that the effects of bound pronouns on scope interactions between *wh*-phrases and quantifiers in English of the kind discussed by Sloan (1991) can be accounted for under a movement approach to binding (see also Lasnik 2007), and extends the analysis to similar effects caused by Japanese *zibun*. More recently, Abe (2009, 2012) has proposed that the relation between a null subject and its antecedent in Japanese be analyzed in terms of A-movement of the antecedent. Abe (2012) also proposes that in certain cases, A-movement of the antecedent produces a chain where its tail, not its head, is pronounced, yielding "backward binding."

[3] To be precise, Hoji (1987) proposes an analysis in which the focus element is base-generated in the focus position and a null operator corresponding to the focus element moves inside the presuppositional clause. In this analysis, the sentence in (4b) has the structure shown in (i).

(i) $[_{CP}$ Op_i $[_{TP}$ Ken-ga <OP> hon-o ageta] no]-wa $Mari_i$-ni da.

An alternative analysis has been proposed by Hasegawa (1997) and Hiraiwa and Ishihara (2002) according to which the focus element itself moves (leftward) to a focus position, followed by

2000).[4] The contrast between (6) and (7), both of which are formed on the basis of the sentence in (5), shows the second property.

(5) Masao-ga Yumi-ni [Ken-ga Mari-ni hon-o ageta to] itta.
Masao-NOM Yumi-DAT Ken-NOM Mari-DAT book-ACC gave that told

'Masao told Yumi that Ken gave a book to Mari.'

(6) a. [Ken-ga Mari-ni hon-o ageta to] itta no-wa
Ken-NOM Mari-DAT book-ACC gave that told C-TOP
Masao-ga Yumi-ni da.
Masao-NOM Yumi-DAT COP

(Lit.) 'It is Masao Yumi that told that Ken gave a book to Mari.'

b. Masao-ga Yumi-ni [Ken-ga ageta to] itta no-wa
Masao-NOM Yumi-DAT Ken-NOM gave that told C-TOP
Mari-ni hon-o da.
Mari-DAT book-ACC COP

(Lit.) 'It is to Mari a book that Masao told Yumi that Ken gave.'

(7) a. *Masao-ga [Ken-ga Mari-ni ageta to] itta no-wa
Masao-NOM Ken-NOM Mari-DAT gave that told C-TOP
Yumi-ni hon-o da.
Yumi-DAT book-ACC COP

(Lit.) 'It is Yumi a book that Masao told that Ken gave to Mari.'

(leftward) movement (topicalization) of a remnant to a higher topic position. In this analysis, the sentence in (4b) is derived as shown in (ii).

(ii) a. Ken-ga Mari-ni hon-o ageta no da. → movement of focus phrase

b. Mari-ni Ken-ga <Mari-ni> hon-o ageta no da. → topicalization of remnant

c. [$_X$ Ken-ga <Mari-ni> hon-o ageta no]-wa Mari-ni <X> da.

Here I am not committed to either alternative, though I use the term "movement of the focus element" for expository purposes.

[4] There are two major proposals to derive this generalization. One proposal is made by Koizumi (1995, 2000) and Kuwabara (1996), and an alternative is proposed by Takano (2002). For present purposes, it is sufficient to simply assume the generalization.

b. *Yumi-ni [Ken-ga hon-o ageta to] itta no-wa
Yumi-DAT Ken-NOM book-ACC gave that told C-TOP
Masao-ga Mari-ni da.
Masao-NOM Mari-DAT COP

(Lit.) 'It is Masao to Mari that told Yumi that Ken gave a book.'

In (6) the two focus elements come from the same clause, whereas in (7) they come from different clauses. Only the examples in (6) are grammatical. Thus, there is a clausemate restriction imposed on Japanese multiple clefting.

However, there are exceptions to this generalization. One exception has been pointed out by Takano (2002), who observes that (8b) is acceptable, in contrast to (8a).[5]

(8) a. *[Bill-ga a-eru to] omotteiru no-wa John-ga Mary-ni da.
Bill-NOM meet-can that think C-TOP John-NOM Mary-DAT COP

(Lit.) 'It is John Mary that thinks that Bill can meet.'
(John thinks that Bill can meet Mary.)

b. [pro$_i$ a-eru to] omotteiru no-wa John$_i$-ga Mary-ni da.
meet-can that think C-TOP John-NOM Mary-DAT COP

(Lit.) 'It is John Mary that thinks that he can meet.'
(John thinks that he can meet Mary.)

The unacceptability of (8a) is consistent with the clausemate restriction, but the acceptability of (8b) constitutes an apparent counterexample to it. The crucial factor that makes (8b) acceptable seems to be the presence of a phonetically null pronoun (pro) bound by the matrix subject. In fact, the example is acceptable only on the reading on which the pro is bound by the matrix subject *John*; if the pro refers to some other person, the sentence is unacceptable.

Although Takano (2002) discusses only (8b), the effect is quite general (see also Kuno 2007). For example, it is not just a pro that has this effect; the overt reflexive *zibun* also works the same way:[6]

[5] The judgments are relative. The example in (8a) sounds better than those in (7), but I abstract away from this difference, marking (8a) with a star. What is important is the contrast between (8a), which is degraded, and (8b), which is perfectly acceptable. The same reservation holds throughout this article.

[6] In Takano (2002: note 16), I judged (9) to be slightly degraded, as compared with (8b). Although there may be a slight difference between the two in this direction, I believe now that (9) is fairly acceptable and contrasts significantly with (8a). As also noted in Takano 2002: note 16, the sentence becomes unacceptable if we replace *zibun* in (9) with the overt pronoun *kare* 'he':

(9) [Zibun$_i$-ga a-eru to] omotteiru no-wa John$_i$-ga Mary-ni da.
self-nom meet-can that think C-TOP John-NOM Mary-DAT COP

(Lit.) 'It is John Mary that thinks that he can meet.'
(John thinks that he can meet Mary.)

The following examples show the same point.

(10) a. *Ken$_i$-ga [pro$_i$/zibun$_i$-ga iku to] itta no-wa
Ken-nom self-NOM go that told C-TOP
Yumi-ni America-e da.
Yumi-DAT America-to COP

(Lit.) 'It is Yumi to America that Ken told that he would go.'
(Ken told Yumi that he would go to America.)

b. Yumi-ni [pro$_i$/zibun$_i$-ga iku to] itta no-wa
Yumi-DAT self-NOM go that told C-TOP
Ken$_i$-ga America-e da.
Ken-NOM America-to COP

(Lit.) 'It is Ken to America that told Yumi that he would go.'
(Ken told Yumi that he would go to America.)

(11) a. *Yumi$_i$-ni [pro$_i$ iku-beki da to] itta no-wa
Yumi-DAT go-should COP that told C-TOP
Ken-ga Amerika-e da.
Ken-NOM America-to COP

(Lit.) 'It is Ken to America that told Yumi that she should go.'
(Ken told Yumi that she should go to America.)

(i) *Kare$_i$-ga a-eru to omotteiru no-wa John$_i$-ga Mary-ni da.
he-NOM meet-can that think C-TOP John-NOM Mary-DAT COP

(Lit.) 'It is John Mary that thinks that he can meet.'
(John thinks that he can meet Mary.)

However, it seems that (i) is unacceptable for reasons having nothing to do with multiple clefting. It is very hard, to begin with, for *kare* in the presuppositional clause to be interpreted as coreferential with a focus element, as shown in (ii).

(ii) Mary-ga kare-no syasin-o miseta no-wa John-ni da.
Mary-NOM he-GEN picture-ACC showed C-TOP John-DAT COP

'It is to John that Mary showed his picture.'

The cleft sentence in (ii) has a single focus. The sentence is acceptable if *kare* takes a discourse antecedent, but is unacceptable if it takes *John* as its antecedent. The sentence in (i) is probably unacceptable for whatever reason makes coreference between *kare* and *John* impossible in (ii).

b. Ken-ga [pro$_i$ iku-beki da to] itta no-wa
Ken-NOM go-should COP that told C-TOP
Yumi$_i$-ni Amerika-e da.
Yumi-DAT America-to COP

(Lit.) 'It is Yumi to Americal that Ken told that she should go.'
(Ken told Yumi that she should go to America.)

The deviance of (10a) and (11a) can be attributed to the clausemate restriction, given that the two focus elements come from different clauses. The improved status of (10b) and (11b), on the other hand, apparently runs counter to this restriction. The factor distinguishing (10a)/(11a) and (10b)/(11b) is the presence/absence of a binding relation between the focus element from the matrix clause and a pro/reflexive in the embedded clause.[7]

Furthermore, apparent counterexamples to the clausemate restriction are not limited to cases having a pro/reflexive in the subject of the embedded clause. They can also be found in cases where a pro/reflexive is in the object of the embedded clause. The examples in (12) duplicate the patterns in (8) and (9).

(12) a. *[Masao-ga Mari-o suisensu-beki da to] omotteiru no-wa
Masao-NOM Mari-ACC recommend-should COP that think C-TOP
Ken-ga sono kaisya-ni da.
Ken-NOM that company-DAT COP

(Lit.) 'It is Ken to that company that thinks that Masao should recommend Mari.'
(Ken thinks that Masao should recommend Mari to that company.)

b. [zibun$_i$-ga Mari-o suisensu-beki da to] omotteiru no-wa
self-NOM Mari-ACC recommend-should COP that think C-TOP
Ken$_i$-ga sono kaisya-ni da.
Ken-NOM that company-DAT COP

(Lit.) 'It is Ken to that company that thinks that he should recommend Mari.'
(Ken thinks that he should recommend Mari to that company.)

Now observe (13).

[7] We cannot use the reflexive *zibun* in the embedded subject in (11) because *zibun* is subject-oriented and hence cannot have a matrix object as its antecedent.

(13) ?[Masao-ga zibun$_i$-o suisensu-beki da to] omotteiru no-wa
Masao-NOM self-ACC recommend-should COP that think C-TOP
Ken$_i$-ga sono kaisya-ni da.
Ken-NOM that company-DAT COP

(Lit.) 'It is Ken to that company that thinks that Masao should recommend him.'
(Ken thinks that Masao should recommend him to that company.)

This example, like the one in (12b), improves on (12a) though the reflexive is placed in the direct object of the embedded clause.[8]

So far we have seen that the presence of a pronominal (a pro or a reflexive) in an embedded clause that is bound by the matrix element saves the example from violating the clausemate restriction. However, it is not the case that the mere presence of a bound pronominal in an embedded clause ensures this effect.[9] To see this, let us consider (14).

(14) a. *[Masao-ga [Yumi-ga au koto]-o yurusite-kureru to]
Masao-NOM Yumi-NOM meet C-ACC permit-give that
omotteiru no-wa Ken-ga Mari-ni da.
think C-TOP Ken-NOM Mari-DAT COP

(Lit.) 'It is Ken Mari that thinks that Masao will permit Yumi to meet.'
(Ken thinks that Masao will permit Yumi to meet Mari.)

b. ?[Masao-ga [pro$_i$/zibun$_i$-ga au koto]-o yurusite-kureru to]
Masao-NOM self-NOM meet C-ACC permit-give that
omotteiru no-wa Ken$_i$-ga Mari-ni da.
think C-TOP Ken-NOM Mari-DAT COP

(Lit.) 'It is Ken Mari that thinks that Masao will permit him to meet.'
(Ken thinks that Masao will permit him to meet Mari.)

---

[8] There may be a slight difference between (12b) and (13), such that (13) is a little worse than (12b). Here I take the improvement of (13) over (12a) to be an important fact that calls for an explanation. Note also that it is not easy to use a pro in place of *zibun* in (13) due to the intervening fact that coreference between *Ken* and the embedded object pro is hard to get. This is because coreference between the matrix subject and the embedded object pro is not natural in its noncleft counterpart in (i) (see Kuroda (1965), Huang (1984), and Hasegawa (1985) for discussion of this fact in Japanese).

(i) Ken-ga [Masao-ga sono kaisya-ni pro suisensu-beki da to] omotteiru.
Ken-NOM Masao-NOM that company-DAT recommend-should COP that think

'Ken thinks that Masao should recommend him/her/them/etc. to that company.'

It is very hard to interpret the pro in (i) to refer to Ken.

[9] I thank Mamoru Saito for bringing this point to my attention.

Each of these examples contains two embedded clauses. (14a) violates the clausemate restriction and is indeed unacceptable. (14b), on the other hand, has a pronominal bound by the matrix subject in the lower embedded clause and it does improve on (14a). Compare now (14b) with (15).

(15) *[pro$_i$/zibun$_i$-ga [Masao-ga au koto]-o yurusu-beki da to]
self-NOM Masao-NOM meet C-ACC permit-should COP that
omotteiru no-wa Ken$_i$-ga Mari-ni da.
think C-TOP Ken-NOM Mari-DAT COP

(Lit.) 'It is Ken Mari that thinks that he should permit Masao to meet.'
(Ken thinks that he should permit Masao to meet Mari.)

Like (14b), (15) has a pronominal bound by the matrix subject, but here it is in the higher embedded clause and the sentence is much worse than (14b). This shows that the mere presence of a bound pronominal does not save the example from violating the clausemate restriction. In fact, the examples that we have seen so far indicate that the saving effect can be seen only when the bound pronominal is a clausemate with the focus element from an embedded clause. In (8b), (9), (10b), (11b), (12b), (13), and (14b), this situation obtains, but in (15), it does not.

These observations lead us to the following generalization about the clausemate restriction on multiple clefting.[10]

(16) X, Y = focus elements in a multiple cleft
When X is from the matrix clause and Y is from an embedded clause, the sentence is acceptable only if (i) there is a pronominal bound by X and (ii) the pronominal is a clausemate with Y.

The next question is, why should this be the case?

We can derive this generalization straightforwardly under a movement approach to binding, in which the matrix element in question moves from the position of the pronominal it binds. Thus, under a movement approach, the example in (17a) receives the analyses given in (17b, c).

(17) a. Ken$_i$-ga [pro$_i$/zibun$_i$-ga Mari-ni a-eru to] omotteiru.
Ken-NOM self-NOM Mari-DAT meet-CAN that think

'Ken thinks that he can meet Mari.'

b. Ken-ga [<Ken-ga> pro Mari-ni a-eru to] omotteiru

[10] Kuno (2007) puts forward a similar generalization. While Kuno's generalization is restricted to cases where the "pronominal" in (16) is phonetically null (i.e., PRO, pro, or trace), the present generalization also covers cases where it is overt (i.e., *zibun*).

c. Ken-ga [<Ken-ga> zibun-ga Mari-ni a-eru to] omotteiru

In the analyses in (17b, c), the unpronounced copy of *Ken* and *Mari* are clausemates. Thus, there is a stage in the derivation at which *Ken* and *Mari* are clausemates. Then the generalization in (16) falls into place: the apparent exceptions to the clausemate restriction all meet the restriction before movement of the matrix element. Given that movement of antecedents plays an essential role in this account, this result argues for a movement approach to binding.

We cannot achieve the same result under a nonmovement approach to binding, where the relevant matrix element stays in the matrix clause throughout the derivation. In such an analysis, (8b), (9), (10b), (11b), (12b), (13), and (14b) would violate the clausemate restriction and hence their acceptability would be left unaccounted for.

There are two consequences that follow immediately from this proposal. First, the cases falling under (16) show that A-movement out of a CP is possible. In those cases, the antecedent moves out of an embedded clause that is clearly a CP (as evidenced by the presence of an overt complementizer). It is also clear that the antecedent undergoes A-movement, given that it moves to a θ-position in the matrix clause. Bošković (2007) and others cited there claim, on various grounds, that A-movement should be allowed to take place out of a CP. The present proposal lends additional support to their claim.

Another consequence is that movement of antecedents shows no minimality effects. This can be seen clearly in (10b), (13), and (14b), where A-movement of the antecedent crosses an intervening argument (a matrix object in the case of (10b) and an embedded subject in the case of (13) and (14)). It is this property that I will turn to in the next section.

3. Minimality Effects

3.1. The Absence of Minimality Effects with Antecedent Movement

Kayne (2002) observes that movement of antecedents does not obey minimality. This is obvious in cases like (18).

(18) John thinks that Mary likes him.

The movement approach dictates that the coreferential reading arise from movement of *John* from the position of *him* past *Mary*, in apparent violation of minimality.

Japanese multiple clefting leads us to the same conclusion. Let us consider (19), repeated from (10b), (13), and (14b).

(19) a. Yumi-ni [pro$_i$/zibun$_i$-ga iku to] itta no-wa
Yumi-DAT self-NOM go that told C-TOP
Ken$_i$-ga America-e da.
Ken-NOM America-to COP

(Lit.) 'It is Ken to America that told Yumi that he would go.'
(Ken told Yumi that he would go to America.)

b. ?[Masao-ga zibun$_i$-o suisensu-beki da to] omotteiru no-wa
Masao-NOM self-ACC recommend-should COP that think C-TOP
Ken$_i$-ga sono kaisya-ni da.
Ken-NOM that company-DAT COP

(Lit.) 'It is Ken to that company that thinks that Masao should recommend him.'
(Ken thinks that Masao should recommend him to that company.)

c. ?[Masao-ga [pro$_i$/zibun$_i$-ga au koto]-o yurusite-kureru to]
Masao-NOM self-NOM meet C-ACC permit-give that
omotteiru no-wa Ken$_i$-ga Mari-ni da.
think C-TOP Ken-NOM Mari-DAT COP

(Lit.) 'It is Ken Mari that thinks that Masao will permit him to meet.'
(Ken thinks that Masao will permit him to meet Mari.)

In these cases, *Ken* must have originated from the position of the pronominal (otherwise, the sentences would violate the clausemate restriction) and moved to positions where it is assigned a subject θ-role by *itta* 'told' and *omotteiru* 'think.' This means that *Ken* has undergone A-movement. Note that this A-movement of *Ken* crosses A-positions filled by *Yumi* and *Masao*. The movement thus apparently does not obey minimality, contrasting with standard A-movement, which does obey minimality, as in (20).

(20) a. John seems to be likely [<John> to win].

b. *John seems it is likely [<John> to win].

These observations are consistent with Kayne's (2002: 161) suggestion that there are no minimality effects with antecedent movement. But why should this be so?

Note that movement of controllers (under a movement approach to control) does obey minimality. Hornstein (1999, 2001) argues that minimal distance effects of the kind seen in (21b, c) follow from minimality.

(21) a. John told Mary to read the book.

b. John told Mary [<Mary> to read the book]

c. *John told Mary [<John> to read the book]

(21b) is a possible derivation for (21a), but (21c) is not. Hornstein argues that under the movement approach, (21c) is excluded because movement of *John* past *Mary* is blocked by minimality.

Exactly the same effects can be seen in Japanese as well. Thus, for the sentence in (22a), (22b) is a possible derivation, but (22c) is not.

(22) a. Ken-ga Mari-ni sono hon-o yomu yooni itta.
Ken-NOM Mari-DAT that book-ACC read C told

'Ken told Mari to read the book.'

b. Ken-ga Mari-ni [<Mari> sono hon-o yomu yooni] itta

c. *Ken-ga Mari-ni [<Ken> sono hon-o yomu yooni] itta

This shows that in Japanese, as in English, A-movement of controllers obeys minimality.[11]

One possibility that comes to mind is to account for the lack of minimality effects in (19) by appealing to scrambling (see Motomura 2001 for a proposal of this kind in a different context). It is well known that Japanese scrambling shows no minimality effects. If movement of *Ken* in (19) could make use of scrambling before it reached a θ-position in the matrix clause, the apparent minimality violations would be accounted for. However, this is not the right move, given (22). In (22) movement of the controller does obey minimality. If scrambling were available for A-movement to a θ-position, (22c) should be a possible derivation for (22a), contrary to fact. Another problem with this move is that it cannot cover antecedent movement in languages like English without scrambling (see (18)). Thus, we need to seek some other way to account for the lack of minimality effects with antecedent movement.

What makes antecedent movement different from other A-movement? Noting this puzzle, Kayne (2002: note 36) suggests that antecedent movement is not subject to minimality because it is not an instance of attraction. Kayne's suggestion is based on the assumption that A-movement usually involves attraction and that minimality effects follow from the mechanisms of attraction (such as attracting closest possible elements). However, given recent work on movement, the validity of this assumption is not clear. Bošković (2007), for instance, proposes a system in which A-movement does not involve attraction and the sole driving force for A-movement is a Case-related property of the moving DP, not a property of a higher head. Another possibility is put forward by Chomsky (2008), who claims that Internal Merge (i.e., movement) is as free as external Merge (i.e., base generation). On this view, too, it is in principle possible for A-movement to occur without attraction.

[11] See Fujii (2006), Takano (2010), and Takita (to appear) for arguments in favor of a movement approach to control in Japanese.

Given this state of affairs, I pursue Kayne's (2002) intuition (that antecedent movement lacks something common to standard A-movement) in a way that does not rest on attraction. Note first that the observations so far indicate clearly that minimality effects cannot be explained in terms of such notions as A-position and A-movement: with those notions, we cannot distinguish antecedent movement from other A-movement. This point thus strengthens the claim made by Bošković (2007: note 24) that Relativized Minimality should be relativized not with respect to the A/A' distinction but with respect to the features involved (see also Rizzi 2004). I adopt this approach to minimality.

Recall from section 2 that A-movement of antecedents can cross CP boundaries. In the following discussion, I assume with Bošković (2007) and others cited there that A-movement out of a CP proceeds by way of a specifier of C. In other words, "improper movement" of this kind is in principle possible.

Given these assumptions, minimality effects can be seen as arising when an element with some feature crosses another element with the same feature. Under standard assumptions, the feature relevant to minimality for A-movement is a Case feature. Thus, for present purposes, I assume the following generalization about minimality for A-movement (cf. Rizzi 2004).

(23) … X … Z … Y
A-movement of Y to X is blocked if Z intervenes between X and Y, and both Y and Z have Case features.

Following Rizzi (2004), let us assume that the notion of intervention is defined in terms of c-command, as in (27).

(24) Z intervenes between X and Y iff Z c-commands Y and Z does not c-command X.

With this in mind, let us consider what property makes the right distinction between antecedent movement and other A-movement. (25) summarizes what appear to be clear properties of the three types of A-movement under consideration.[12]

(25) a. Standard A-movement
The DP receives a θ-role in its base position and enters into Case licensing in its landing site.

b. Movement of controllers
The DP receives a θ-role in its base position as well as in the course of movement, and enters into Case licensing in its landing site.

[12] Here I use the term "Case licensing" as a cover term for checking, agreement, or valuation.

c. Movement of antecedents
The DP receives a θ-role in the course of movement and enters into Case licensing in its landing site.

We can see that the two properties listed in (25c) cannot distinguish antecedent movement from the other two: the first property is shared by movement of controllers and the second by both standard A-movement and movement of controllers.

This leaves us with a possibility that what makes antecedent movement special has to do with what happens in the base position. The question then boils down to how exactly to treat pronominals under a movement approach to binding. Recall from note 1 that there are two previous proposals about this. These are shown in (26) and (27).

(26) John thinks that he(=<John>) is smart.

(27) John thinks that [<John> he] is smart.

The analysis in (26) is due to Hornstein (2001) and the one in (27) to Kayne (2002). Details aside, Hornstein claims that the antecedent *John* is directly merged into a subject θ-position of the embedded clause and then copied onto a subject θ-position of the matrix clause, with the pronoun *he* inserted, as a last resort, in place of the copy *John* in the embedded subject when direct A-movement from that position is impossible (if direct A-movement is possible, the original position is realized as a reflexive). On the other hand, Kayne proposes that the antecedent first gets merged with the pronoun, forming the constituent [*John he*], that this constituent is merged into a subject θ-position of the embedded clause, and that the antecedent moves out of this constituent to the matrix clause, leaving a copy in the usual way. In the following discussion, I assume that the relevant constituent [*John he*] is a DP.

Let us adopt the widely accepted view that a crucial factor driving A-movement of a DP is that the DP has a Case feature that is yet to be licensed. This favors Kayne's analysis in (27). In this analysis, the DP [*John he*] is licensed for its Case in the embedded clause. Following Kayne, let us further assume that the pronoun *he* is the head of this DP.[13] Then the pronoun is licensed for its Case in the embedded clause, but the antecedent is not. Thus, the antecedent *John* has the property that drives A-movement.

Note also that under Kayne's analysis, the subject θ-role of *smart* is assigned to the DP [*John he*]. Since *he* is the head of this DP, it does not need an independent θ-role. But *John* lacks a θ-role in its base position. In other words, antecedent movement is characterized as movement from a non-θ-position to a θ-position. I propose that this is the crucial factor that

[13] There are a number of possibilities regarding the internal structure of the DP in question. One possibility is that the pronoun *he* is literally the head D of the DP, with possibly a phonetically null NP complement. Alternatively, the pronoun *he* is a noun selected by a phonetically null D. In this case, the DP is an "extended" projection of the pronoun *he* and *he* is the "head" of the DP in the extended sense. Like Kayne (2002), I abstract away from the details of the internal structure of the DP.

distinguishes antecedent movement and other A-movement: the moving DP receives a θ-role in its base position in the case of standard A-movement and controller movement, but it does not in the case of antecedent movement.

To implement this idea, I propose (28).

(28) θ-roles make the Case features of DPs visible to the computation.

The idea here is a reinterpretation of "visibility" to θ-role assignment proposed by Chomsky (1981: chap. 6). Chomsky proposes to derive the Case Filter from the θ-Criterion, claiming that for an argument to receive a θ-role at LF, it must be "visible" at LF and that an argument is visible at LF only if it is assigned Case at S-structure.[14] Thus, on Chomsky's original view, Case has the effect of making DPs visible to θ-role assignment. The proposal in (28) also takes the relation between Case and θ-role seriously, but in a way opposite to Chomsky's view. It works roughly as follows. Given (28), we understand that "Case features" in (23) mean visible Case features. Suppose a DP with a Case feature enters a derivation. Usually this DP is merged into a θ-position. Thus, given (28), the Case feature of this DP is visible to the computation. As a result, minimality effects arise, due to (23), if this DP is to move over another DP with a visible Case feature. This is the situation with standard A-movement and controller movement. A different situation arises with antecedent movement. The antecedent DP with a Case feature is first merged with a pronoun and the resulting DP containing the two is then merged into a θ-position. The antecedent DP lacks a θ-role inside the larger DP headed by the pronoun. Thus, when the antecedent DP undergoes movement, its Case feature is invisible to the computation, due to (28). It is not until it moves to a θ-position that its Case feature becomes visible. In other words, the computation treats the antecedent DP, up to a θ-position, as an element without a Case feature. As a result, the antecedent can move freely up to a θ-position over DPs with visible Case features without causing minimality effects.

Let us look at concrete examples. Under this proposal, nothing special needs to be said about standard A-movement, as in (29).

(29) John seems to be likely [<John> to win]

The Case feature of the DP *John* becomes visible when the DP gets merged into the subject θ-position in the embedded clause. A-movement of *John* from that position to the matrix clause does not cause a minimality violation since it does not cross any element with a Case feature.

Obligatory control cases fall into place, too. Let us consider the grammatical case in (30).

(30) John told Mary [<Mary> to leave]

[14] Chomsky's (1981) original proposal is couched in terms of chains. I abstract away from this aspect of his proposal.

The Case feature of *Mary* is visible in the embedded clause. A-movement of *Mary* to the matrix clause does not cross any element with a Case feature. Thus, the sentence can be derived without causing problems.

Consider the ungrammatical derivation in (31).

(31) *John told Mary [<John> to leave]

Here the Case features of *John* and *Mary* are visible before *John* moves. As a result, movement of *John*, which has a visible Case feature, over *Mary*, which also has a visible Case feature, causes a minimality violation.

Let us consider the superraising case in (32).

(32) *John seems it is likely [<John> to win]

A-movement of *John* crosses the expletive *it*. *John* has a visible Case feature when it crosses the expletive. What about the expletive? Intuitively, the visibility condition in (28) is a condition on elements that require θ-roles. Expletives do not require θ-roles (in fact, they are incompatible with θ-roles). Therefore, they fall outside the condition in (28) and their Case features are inherently visible. Let us make it a concrete proposal, revising (28) to (33).[15]

(33) The Case feature of a DP is visible to the computation iff the DP satisfies the θ-Criterion.

The θ-Criterion relevant to (33) is (34).

(34) An argument must receive a θ-role.

(33) has the same effects as (28) for argument DPs. For expletives, (33) claims that their Case features are visible from the beginning because they satisfy the θ-Criterion vacuously. Given (33), *John* in (32), which has a visible Case feature, moves over an element with a visible Case feature, namely, the expletive *it*, thereby causing a minimality violation.

Let us turn now to movement of antecedents. Let us consider the sentence in (35) on the reading on which *John* is an antecedent of *him*.

(35) John thinks that Mary likes him.

Under a movement approach a la Kayne (2002), this sentence is derived in the following way. First, *John* and *him* are merged, forming the DP [*John him*]. Then the DP [*John him*] is merged into the object of *likes*. After the derivation forms the matrix vP, the antecedent DP *John* moves to Spec,v, where it receives a subject θ-role. Note that in this part of the derivation, *John* moves over *Mary*, which has a visible Case feature. However, due to (33),

[15] This idea was suggested to me by Tomohiro Fujii.

the Case feature of *John* is invisible to the computation when it moves over *Mary*. Its Case feature becomes visible only in Spec,v in the matrix clause. As a result, A-movement of *John* over *Mary* does not cause a minimality violation.

A question arises here. Given that the Case feature of *John* is invisible up to Spec,v in the matrix clause, what drives movement of *John* to that position? There are two possibilities. One is that θ-roles can be a driving force for A-movement, as in Hornstein's (1999, 2001) theory. Since *John* has no θ-role, it moves to a θ-position to receive a θ-role (on this view, we need to assume that the moving element has a driving force; see Bošković 2007 for arguments in favor of this view). Another possibility is to assume with Chomsky (2008) that internal Merge (i.e., movement) is as free as external Merge (i.e., base generation). In this alternative, movement of *John* can take place without a driving force, so that it can move to Spec,v by means of internal Merge.[16] Either way, we can derive the desired results.

In the present analysis, the visibility condition in (33) plays a crucial role in accounting for the lack of minimality effects with antecedent movement. This analysis is made possible under Kayne's (2002) theory, where the antecedent and the pronoun are first merged to from a constituent, which is in turn merged into a θ-position. Thus, the present proposal lends independent support to Kayne's theory.

There is another consequence. The proposed account of the lack of minimality effects in terms of (33) makes crucial reference to whether a given DP has received a θ-role at a given stage of a derivation. This information has to be available through the course of a derivation. Thus, thematic information is not just an interface property but plays a role in the core computation as well.

3.2. The Presence of Minimality Effects with Copy Raising

The analysis that I have just proposed claims that the example in (35) does not violate minimality because the Case feature of *John* is invisible to the computation when it crosses the argument DP *Mary*. The invisibility of the Case feature is ensured by (33). Crucial here is the analysis, due to Kayne (2002), in which the antecedent and the pronominal form a DP that is merged into a θ-position and the antecedent moves out of this DP to a θ-position. In section 3.1, I assumed, following Kayne (2001), that the antecedent does not receive a θ-role inside the DP out of which it moves. This property, coupled with (33), accounts for the lack of minimality effects with antecedent movement.

However, there are cases in which we must assume that an argument receives a θ-role inside the DP out of which it moves. The evidence comes from copy raising constructions

[16] On this view, it is a separate question why elements cannot stay freely in their landing sites (e.g., objects in English cannot appear freely in preverbal positions). The question is related to the issue of what allows free movement such as scrambling. One possibility is that while movement operations are allowed freely by means of internal Merge, properties of landing sites may block moved elements from surfacing there. This will give rise to effects that make movement look like it is not free.

like those in (36).

(36) a. John seems like he is smart.

b. John looks as if he is smart.

Copy raising constructions have the property of having a pronoun in a θ-position and its "antecedent" in a non-θ-position. In the copy raising constructions in (36), *John* is understood to be a subject of *smart* though the subject position of *smart* is occupied by *he* and *John* is in a non-θ-position.

There is evidence that at least in certain cases of copy raising, the surface matrix subject has moved from the embedded subject. Consider the following examples.

(37) a. The shoe looks like it's on the other foot. (Potsdam and Runner 2001)

b. These stories about each other$_i$ sound like they would frighten John and Mary$_i$. (Fujii 2007)

The copy raising example in (37a) has an idiom chunk in the matrix subject position and its "pronominal associate" in the embedded subject position. Potsdam and Runner (2001) report that it is acceptable, though there is some dialectal variation in its acceptability. The availability of the idiom reading in (37a) shows that the matrix subject is interpreted in the embedded subject. A straightforward analysis of this fact will be one where the matrix subject has moved from the embedded subject.[17] Similarly, Fujii (2007) reports that (37b) is acceptable. This indicates that the anaphor *each other* contained in the matrix subject can be bound by the embedded object *John and Mary*. Thus, (37b) patterns with (38a) and not with (38b).

(38) a. Each other$_i$'s mothers seem to please the two boys$_i$.

b.?*John seemed to each other$_i$'s mothers to please the two boys$_i$. (Fujii 2007)

The contrast in (38) indicates that the acceptability of (38a) is due to movement of the matrix subject containing *each other* from the embedded subject, which produces a reconstruction context, making it possible for *each other* to be bound by the embedded object (cf. *It seems that each other's mothers please the two boys.*). (38b) is unacceptable because *each other* stays in the matrix clause throughout the derivation and so can never be bound by the embedded object. Fujii (2007) claims that the fact that (37b) patterns with (38a) shows that in

[17] Potsdam and Runner (2001) propose an analysis in which *John* and *he* are base generated in the matrix clause and in the embedded clause, respectively, and they form a base-generated A-chain. According to their analysis, the presence of the base-generated A-chain accounts for the idiom reading in (37a). In the movement account adopted in the text, we can dispense with the notion of base-generated A-chain.

(37b) the matrix subject has moved from the position of *they* (see also Ura 1998 for an earlier proposal for a movement analysis of copy raising).

In the present perspective, we can account for the movement properties of copy raising by analyzing it as illustrated in (39).

(39) John seems like [<John> he] is smart.

In (39) *John* and *he* form a DP constituent and this DP is merged into a subject θ-position in the embedded clause. Then *John* moves out of the DP to the matrix clause, where its Case is licensed.[18] So far the situation is exactly identical to what happens in the case of standard binding (e.g., *John thinks that [<John> he] is smart*). However, there is an important difference between the two cases. In the case of binding, the antecedent moves to a θ-position, whereas in the case of copy raising, it moves to a non-θ-position. This means that *John* in (39) must receive a θ-role in its base position inside the DP [*John he*]; otherwise, it would receive no thematic interpretation.

Given that the antecedent receives a θ-role in its base position in copy raising, it is predicted, from the perspective of (33), that copy raising will exhibit minimality effects. This is in fact the case. Consider first the following examples from Potsdam and Runner (2001).

(40) a. Bill sounds like Martha hit him over the head with the record.

b. The roach looks to me like Abbie gave it to Myrna.

On the surface, these cases look identical to the cases in (36). The only difference lies in the fact that the pronominal associates appear in subject positions in (36), but in nonsubject positions in (40). Interestingly, however, the apparent copy raising constructions in (40) do not exhibit the reconstruction effect seen in (37b), as the following example shows.

(41) ?*Those stories about each other$_i$ sound like John and Mary$_i$ would fear them.

Fujii (2007) observes that (41) contrasts with (37b) in that the relevant binding is not possible in (41). The ungrammaticality of (41) shows that the matrix subject cannot have moved from the position of its pronominal associate *them*. Fujii claims that this is due to a violation of minimality: A-movement of *those stories about each other* from the position of *them* is blocked by the intervening subject of the embedded clause. This implies in turn that the matrix subject in cases like (40) is necessarily base generated in the matrix subject position and receives a θ-role there, as Potsdam and Runner claim (otherwise; the matrix subject would not receive a θ-role), and also that cases like (36) are ambiguous between a derivation where the matrix subject has moved from the embedded subject and one where the matrix

[18] Here I depart from Fujii (2007). Fujii claims that in the copy raising construction, the raised DP enters into multiple Case licensing, so that its Case is licensed both in the embedded clause and in the matrix clause. I assume, following Kayne's (2001) theory of antecedent movement, that the Case of the raised DP is licensed only in the matrix clause.

subject is base generated in the matrix clause.

Crucial for present purposes is that A-movement of the matrix subject in copy raising does obey minimality, in contrast to A-movement of antecedents. Given that A-movement must originate from a θ-position in the case of copy raising, this difference follows straightforwardly under the present proposal incorporating the visibility condition in (33).

However, the claim that A-movement of the matrix subject in copy raising originates from a θ-position is contrary to Kayne's (2002) proposal, which we adopted in section 3.1, that the antecedent does not receive a θ-role inside the DP out of which it moves. The analysis of the lack of minimality effects proposed in section 3.1 supports Kayne's position, whereas copy raising argues for the contrary. This state of affairs leads us to conclude that Universal Grammar makes both options available. Specifically, I propose (42).

(42) In the DP [X Y], where Y is a pronominal and X is its antecedent,

(i) X does not receive a θ-role in [X Y]; or

(ii) X shares a θ-role with Y.

Recall that the DP [X Y] is merged into a θ-position and Y, being the head of this DP, has a θ-role assigned to the DP. In the case of copy raising, option (ii) must be taken; otherwise, the moved subject would receive no thematic interpretation. As a result of taking this option, copy raising obeys minimality (given the visibility condition in (33)). On the other hand, in the case of antecedent movement, the two options are available in principle. If option (ii) is taken, we predict that antecedent movement will show minimality effects, just like copy raising. However, the other option is also available here. With option (i), coupled with (33), antecedent movement can violate minimality. This is the situation we discussed in section 3.1.[19]

4. Conclusion

In this paper, I have discussed issues related to movement of antecedents in a movement approach to binding. The major claims I have made are summarized in (43).

(43) a. Multiple clefting in Japanese provides new evidence that antecedents move from the position of pronominals.

[19] Copy raising involving idiom chunks shows minimality effects. Compare (37a) with (i) below, which Potsdam and Runner (2001) judge as unacceptable.

(i) *The other foot appears like the shoe is on it.

Given (33), the status of (i) follows from minimality. Idiom chunks, like expletives, do not receive θ-roles and so the Case features of *the shoe* and *the other foot* in (i) are visible to the computation from the beginning. As a result, *the other foot* can never skip *the shoe*.

b. Movement of antecedents does not obey minimality.

c. The antecedent does not receive a θ-role in its base position.

d. θ-roles make the Case features of DPs visible to the computation.

In connection with (43c), I have argued for the hypothesis put forward by Kayne (2002) that movement of antecedents is characterized as movement from a non-θ-position to a θ-position. This is a new kind of A-movement and I have proposed that this property, coupled with the proposal in (43d), is responsible for the interesting effect in (43b).

References

Abe, J. (2009) "Locality of Empty Pronouns," ms., Tohoku Gakuin Univeristy.

Abe, J. (2012) "Backward Anaphora in Japanese: Cases of Condition C Violations," paper presented at *FAJL 6*.

Bošković, Ž. (2007) "On the Locality and Motivation of Move and Merge: An Even More Minimal theory," *Linguistic Inquiry* 38, 589–644.

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, N. (2008) "On phases." In R. Freidin, C. P. Otero, and M. L. Zubisarreta, eds., *Foundational Issues in Linguistics Theory: Essays in Honor of Jean-Roger Vergnaud*, MIT Press, Cambridge, Mass., 133–166.

Fujii, T. (2006) *Some Theoretical Issues in Japanese Control*, Ph.D. dissertation, University of Maryland.

Fujii, T. (2007) "Cyclic Chain Reduction." In N. Corver and J. Nunes, eds., *The Copy Theory of Movement*, John Benjamins, Amsterdam, 291–326.

Hasegawa, N. (1985) "On the So-Called "Zero Pronouns" in Japanese," *The Linguistic Review* 4, 289–341.

Hasegawa, N. (1997) "A Copula-Based Analysis of Japanese Clefts: *Wa*-Cleft and *Ga*-Cleft," *Researching and Verifying an Advanced Theory of Human Language*, *Grant-in-aid for COE Research, Report (1)*, Kanda University of International Studies, 15–38.

Hiraiwa, K. and S. Ishihara (2002) "Missing Links: Cleft, Sluicing and "No Da" Construction in Japanese," *The Proceedings of Humit 2001*, *MIT Working Papers in Linguistics* 43, 35–54.

Hoji, H. (1987) "Japanese Clefts and Chain Binding/Reconstruction Effects," paper presented at *WCCFL* VI.

Hornstein, N. (1999) "Movement and Control," *Linguistic Inquiry* 30, 69–96.

Hornstein, N. (2001) *Move! A Minimalist Theory of Construal*. Blackwell, Oxford.

Huang, C.-T. J. (1984) "On the Distribution and Reference of Empty Pronouns," *Linguistic Inquiry* 15, 531–574.

Kayne, R. (2002) "Pronouns and Their Antecedents." In S. D. Epstein and T. D. Seely, eds., *Derivation and Explanation in the Minimalist Program*, Blackwell, Malden, 133–166.

Koizumi, M. (1995) *Phrase Structure in Minimalist Syntax*, Ph.D. dissertation, MIT.

Koizumi, M. (2000) "String Vacuous Overt Verb Raising," *Journal of East Asian Linguistics* 9, 227–285.

Kuroda, S.-Y. (1965) *Generative Grammatical Studies in the Japanese Language*, Ph.D. dissertation, MIT.

Kuwabara, K. (1996) "Multiple *Wh*-phrases in Elliptical Clauses and Some Aspects of Clefts with Multiple Foci," *Formal Approaches to Japanese Linguistics 2*, *MIT Working Papers in Linguistics* 29, 97–116.

Lasnik, H. (2007) "A Family of Questions," ms., University of Maryland.

Miyamoto, Y. (2008) "A Derivational Approach to Scope Interaction between *Wh*-phrases and Quantifiers," *Canadian Journal of Linguistics* 53, 219–336.

Motomura, M. (2001) "*Zibun* as a Residue of Overt A-movement," *Formal Approaches to Japanese Linguistics 3, MIT Working Papers in Linguistics* 41, 309–325.

Potsdam, E. and J. T. Runner (2001) "Richard Returns: Copy Raising and Its Implications," *CLS* 37, 453–468.

Rizzi, L. (2004) "Locality and Left Periphery." In A. Belletti, ed., *Structures and Beyond: The Cartography of Syntactic Structures*, Oxford University Press, Oxford, 223–251.

Sloan, K. (1991) "Quantifier-Wh Interaction," *More Papers on Wh-movement, MIT Working Papers in Linguistics* 15, 219–237.

Takano, Y. (2002). "Surprising Constituents," *Journal of East Asian Linguistics* 11, 243–301.

Takano, Y. (2010) "Scrambling and Control," *Linguistic Inquiry* 41, 83-110.

Takita, K. (to appear) "VP-scrambling, Linearization Preservation, and the Theories of Control," *NELS* 40.

Ura, H. (1998) "Checking, Economy, and Copy-Raising in Igbo," *Linguistic Analysis* 28, 67–88.