English Linguistics 34: 1-33, 2017.

[Invited Article]

NOTES ON THE LOCALITY OF ANAPHOR BINDING AND A-MOVEMENT

Mamoru Saito

Nanzan University

Abstract

This paper consists of two notes on Minimalist syntax, one on transfer domains and the other on labeling. Assuming Quicoli's (2007) phase-based analysis of the locality of anaphor binding, I first argue that what is transferred upon the completion of CP phase is not the complement TP but the *v*P below when ϕ-feature agreement is absent. This accounts for the lack of NIC effects in languages without agreement and makes Hornstein's (1999) movement analysis of control consistent with the phase theory. In the second note, I consider the {XP, YP} problem discussed in Chomsky (2013). I propose that the dislocation of XP aids labeling of the constituent when and only when the movement is uniform, that is, A-A or A'-A'. I argue that this leads to an explanation for the ban on improper movement and also has a desirable consequence for the proposal in Saito (2016a) to eliminate the θ-criterion.*

Keywords: *phase, labeling, anaphor binding, A-movement, agreement*

1. Introduction

The Minimalist model, as it is pursued in Chomsky (2008, 2013), consists of two major components. One is the operation Merge, which freely forms a constituent $\gamma = \{\alpha, \beta\}$ from two syntactic objects α and β. It accompanies a labeling algorithm to determine the nature (or label) of the formed constituent γ. The other is the theory of phase, which

concerns the units of syntactic derivation and of transfer of information to the C-I and A-P interfaces. In this paper, I present two notes, the first on phase and the second on labeling.

I start the first note in Section 2 with consideration of the locality of anaphor binding. Building on insights of Fiengo (1974), Chomsky (1981) presents Condition (A) of the Binding theory as a syntactic principle that dictates the locality of anaphor binding and A-movement. The latter is constrained in addition by the ECP. However, these principles have no place in the Minimalist model, which aims to eliminate stipulated principles in narrow syntax. The binding conditions are, accordingly, reformulated as the interpretive rules in (1) in Chomsky (1993) and Chomsky and Lasnik (1993).

(1) a. If α is an anaphor, interpret it as coreferential with a c-commanding phrase in D.

b. If α is a pronominal, interpret it as disjoint from every c-commanding phrase in D.

c. If α is an r-expression, interpret it as disjoint from every c-commanding phrase.

A fundamental question that arises from this reformulation concerns the source of the locality in (1a) and (1b). Attempts have been made to explain it in terms of derivational phases and the transfer domains they define, for example, in Quicoli (2008) and more recently in Charnavel and Sportiche (2016).

Against this background, I discuss some examples with lexical anaphors that were captured by Chomsky's (1981) Condition (A) but are not straightforwardly explained by the phase-based analysis. One case concerns the absence of NIC or TSC effects in languages without ϕ-feature agreement. I argue that the phase head C triggers transfer of the complement TP when it carries ϕ-features, but only of *v*P when it does not. Although it is still a tentative hypothesis, I show that it has desirable consequences for the analysis of A-movement. It allows all clausal complements, including ECM complements, to be uniformly CPs, and also fits nicely with the movement theory of control, proposed first by Hornstein (1999).

In the second note in Section 3, I first introduce the labeling algorithm of Chomsky (2013) and its application to Japanese as proposed in Saito (2014). Chomsky hypothesizes that when a head and a phrase merge as in $\gamma = \{H, \alpha P\}$, the head determines the label of the formed constituent. Thus, γ is verbal in nature when $\gamma = \{V, DP\}$. He also proposes that labeling is possible in limited cases when two phrases are merged as in $\gamma = \{\alpha P, \beta P\}$. One case arises when αP is extracted out of γ. In this case, βP is the only element properly contained within γ, and hence can determine its label. The rest of Section 3 is on this particular proposal. I suggest that extraction makes labeling possible in this way only when it forms a uniform chain, A-A or A'-A'. I argue that this leads to a principled account of the ban on improper movement and also aids in the attempt, laid out in Saito (2016a), to eliminate the θ-criterion in favor of labeling. Section 4 concludes the paper.

2. Anaphor Binding and Derivational Phase

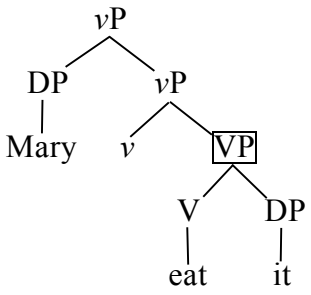
I introduce Quicoli's (2008) phase-based analysis of Condition (A) effects in Section 2.1. In Section 2.2, I discuss some potentially problematic cases and propose that only *v*P is transferred to the C-I interface when a CP phase without ϕ-feature agreement is completed. Section 2.3 is concerned with the implications for the analysis of ECM sentences and the movement theory of control.

2.1. Quicoli's (2008) Proposals

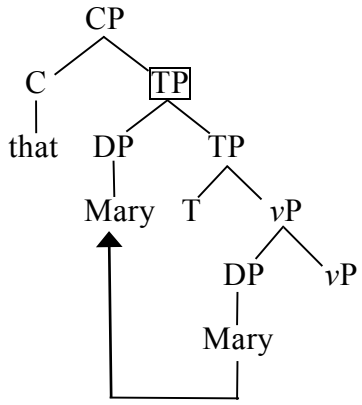
According to Chomsky (2008), for example, syntactic derivation proceeds phase by phase, where phases are defined by phase heads, C and transitive/unergative *v*. Let us consider the embedded clause of (2).

(2) John thinks that Mary ate it

(3) a.



b.



First, the embedded *v*P is built as in (3a) and the complement of the phase head *v*, VP, is transferred to the interfaces. Then, the embedded CP phase is constructed as in (3b). The TP is sent to the interfaces at this point.

The basic idea behind the phase-based analysis of the locality in binding is that an anaphor, for example, must find an antecedent as it is transferred to the C-I interface or within the transfer unit that contains it. Quicoli (2008) takes the former approach and proposes (4).[1]

(4) Information on the reference of an anaphor is sent to the C-I interface along with a transfer domain that includes the anaphor.

Let us consider the concrete examples in (5).

(5) a. John recommended himself

b. John recommended him

The reflexive in (5a) is sent to the C-I interface when the *v*P phase is completed and transfer applies to the shaded complement VP as in (6a).

(6) a. [$_{vP}$ John [*v* [$_{VP}$ recommend himself]]]

b. [$_{vP}$ John [*v* [$_{VP}$ recommend him]]]

As the antecedent *John* is already part of the structure, the information, 'himself = John', can be sent to the C-I interface as well. In the case of the pronoun in (5b), the information, 'him ≠ John', is transferred along with the VP in (6b).

This straightforwardly accounts for the ungrammaticality of (7a).

(7) a. *John thinks that Mary recommended himself

b. [$_{v\text{P}}$ Mary [v [$_{\text{VP}}$ recommend himself]]]

The reflexive is transferred to the C-I interface as part of the embedded VP upon the completion of the embedded *v*P in (7b). As no information on its reference is available at this point, the reflexive fails to receive an interpretation.

Quicoli appeals to the copy theory of movement in order to account for reconstruction effects, basically along the lines of Chomsky (1993). (8a) is grammatical because its *v*P is constructed as in (8b).

(8) a. Which picture of himself did John buy

b. [$_{v\text{P}}$ [which picture of himself] [John [v [$_{\text{VP}}$ buy [which picture of himself]]]]]

The wh-phrase moves to the edge of *v*P, leaving behind a copy in the object position. The complement VP is transferred with the object that contains an instance of *himself*, and information on the reference of the reflexive can be sent to the C-I interface as well. (9) shows that there are in some cases multiple occasions to determine the reference of a reflexive.

(9) Bill wonders which picture of himself John bought

In this example, both *John* and *Bill* can be the antecedent of *himself*. The embedded *v*P has the structure in (10a).

(10) a. [$_{v\text{P}}$ [which picture of himself] [John [v [$_{\text{VP}}$ buy [which picture of himself]]]]]

b. [$_{\text{CP}}$ [which picture of himself] [C [$_{\text{TP}}$ John [T [$_{v\text{P}}$ [which picture of himself] [John [v [$_{\text{VP}}$ ...]]]]]]]]

c. [$_{v\text{P}}$ [which picture of himself] [Bill [v [$_{\text{VP}}$ wonders [$_{\text{CP}}$ [which picture of himself] [C [$_{\text{TP}}$ ...]]]]]]]

The VP can be transferred with the information 'himself = John'. But in this case, it is also

possible to leave the reference of the reflexive pending. The wh-phrase moves the edge of the embedded CP phase as in (10b) and the shaded part of the complement TP is sent to the C-I interface. Then, it moves on to the edge of the matrix *v*P as in (10c). At this point, the information 'himself = Bill' can be transferred along with the shaded part of the matrix VP.

This analysis extends to the Japanese examples of scrambling in (11) from Dejima (1999). (11a) indicates that the subject-oriented reflexive *zibun-zisin* 'self-self' requires a local antecedent.[2]

(11) a. Taroo-ga$_i$ [$_{CP}$ Hanako-ga$_j$ [$_{CP}$ Ziroo-ga$_k$ zibunzisin-o$_{i*, j*, k}$ hihansita to] itta
Taroo-Nom Hanako-Nom Ziroo-Nom self-Acc criticized C said
to] omotteiru (koto)
C think fact
'Taroo$_i$ thinks [that Hanako$_j$ said [that Ziroo$_k$ criticized self$_{i*, j*, k}$]]'

b. Taroo-ga$_i$ [$_{CP}$ Hanako-ga$_j$ [$_{CP}$ zibunzisin-o$_{i*, j, k}$ Ziroo-ga$_k$ hihansita to] itta
Taroo-Nom Hanako-Nom self-Acc Ziroo-Nom criticized C said
to] omotteiru (koto)
C think fact
'Taroo$_i$ thinks [that Hanako$_j$ said [self$_{i*, j, k}$ [that Ziroo$_k$ criticized]]]'

c. Taroo-ga$_i$ [$_{CP}$ zibunzisin-o$_{i, j, k}$ Hanako-ga$_j$ [$_{CP}$ Ziroo-ga$_k$ hihansita to] itta
Taroo-Nom self-Acc Hanako-Nom Ziroo-Nom criticized C said
to] omotteiru (koto)
C think fact
'Taroo$_i$ thinks [self$_{i, j, k}$ [that Hanako$_j$ said [that Ziroo$_k$ criticized]]]'

But when the reflexive is scrambled to the initial position of the most deeply embedded CP as in (11b), the middle subject becomes an eligible antecedent. With further scrambling to the initial position of the middle CP as in (11c), the interpretation of the reflexive is three

ways ambiguous. The matrix, the middle and the most deeply embedded subjects are all possible antecedents for the reflexive.

This state of affairs is predicted by Quicoli's analysis. In (11a), the reference of *zibunzisin* must be determined when the most deeply embedded *v*P phase is completed and its complement VP is transferred, as illustrated in (12).

(12) [$_{vP}$ Ziroo-ga [[$_{VP}$ zibunzisin-o hihans] *v*]]

Then, *Ziroo* must be the antecedent. However, when scrambling applies to the reflexive as in (11b), it can place a copy of the reflexive in the transfer domain triggered by *v* in the middle clause. This is shown in (13).

(13) [$_{vP}$Hanako-ga [*v* [$_{VP}$[$_{CP}$[zibunzisin-o] [[$_{TP}$ Taroo-ga zibunzisin-o hihansita] to]] iw]]]

When the shaded part of the VP is transferred, the information 'zibunzisin = Hanako' can be sent to the C-I interface. Further extension of the binding possibility in (11c) is accounted for in the same way.[3]

2.2. The Absence of NIC Effects in Languages without ϕ-feature Agreement

The phase-based approach to the locality of binding seems quite promising. Yet, as far as I know, it has not reached the point where it successfully accommodates the main data examined, for example, in Chomsky (1981). I discuss a couple of problematic cases in this section and suggest a possible direction toward a solution.

The binding theory of Chomsky (1981) is based on the generalization that the binding domain δ for an anaphor α obtains in the following configurations:

(14) a. [$_{\delta}$ subject [... α ...]] (SSC)

b. [$_{\delta}$ α [$T_{[+AGR]}$ [...]]] (NIC)

The binding domain for an anaphor is the minimal domain that contains a subject

c-commanding the anaphor as illustrated in (14a) or the minimal finite TP with ϕ-feature agreement that contains the anaphor as in (14b). (15a, b), for example, fall under (14a).

(15) a. *John expects [$_{TP}$ the guests to be introduced to himself]

b. *Mary believes [$_{TP}$ the speakers to have been introduced to herself]

The embedded clause is the binding domain for *himself/herself* as it contains a subject that c-commands the reflexive. Thus, the examples are predicted to be ungrammatical. However, it is not obvious how these examples fare with the phase-based approach, given the standard assumption that only C and transitive/unergative *v* are phase heads. Because the embedded clause is in the passive, the matrix *v* is the first phase head in the derivation. Then, *himself* in (15a), for example, should successfully receive interpretation as illustrated in (16).

(16) [$_{vP}$ John [v [$_{VP}$ expect [$_{TP}$ [the guests] to be introduced [the guests] to himself]]]]

The information 'himself = John' can be sent to the C-I interface as the matrix VP is transferred.

Another potential problem arises with the absence of the NIC effect in languages without ϕ-feature agreement. The relevant configuration is (14b), instantiated by (17).

(17) *John thinks [$_{CP}$ that [$_{TP}$ himself will be nominated]]

It was first noted by Yang (1983) that the Korean and Japanese counterparts of examples of this kind are grammatical. Relevant Japanese examples are shown in (18).[4]

(18) a. Taroo-wa [$_{CP}$ [$_{TP}$ zibunzishin-ga suisensareru] to] omotteiru

Taroo-Top self-Nom nominated-will.be C think

'Taroo thinks that he (= Taroo) will be nominated.'

b. Hanako-wa [$_{CP}$ [$_{TP}$ zibunzisin-ga sore-o mita] to] syutyoosita

Hanako-Top self-Nom it-Acc saw C insisted

'Hanako insisted that she (= Hanako) saw it.'

Yang also points out that Chomsky's (1981) binding theory correctly predicts this. As there is no ϕ-feature agreement in Korean and Japanese, T lacks ϕ-features and is [-AGR] in these languages. Hence, (18a, b) are predicted to be grammatical just like (19) although their embedded clauses carry [±past] tense.

(19) John believes [$_{TP}$ himself to be the best candidate]

This account does not carry over to the phase-based analysis if C always triggers the transfer of its TP complement. In (18b), for example, the TP containing *zibunzisin* is transferred upon the completion of the embedded CP phase as in (20).

(20) [$_{CP}$ [[$_{TP}$ zibunzishin-ga [[$_{vP}$ zibunzisin-ga [[$_{VP}$ sore-o mi] *v*]] $T_{[+Past]}$]] to]]

No information on the reference of *zibunzisin* can be sent to the C-I interface, and the reflexive should fail to receive an interpretation.

It has been shown that the examples in (15) and (18) require the phase-based analysis of anaphor binding to be developed further. Let us start with the latter. It seems fairly clear that the contrast between (17) and (18) stems from the presence/absence of ϕ-feature agreement in the embedded TP as proposed by Yang (1983). When there is no ϕ-feature agreement, a reflexive in the embedded subject position can have an antecedent in the matrix clause. This in turn implies that in this case, the embedded subject is outside the transfer domain triggered by C. The difference between TPs with and without ϕ-feature agreement is illustrated below.

(21) a. [$_{CP}$ [C [$_{TP}$ subject [$T_{[+AGR]}$ [$_{vP}$ …]]]]]

b. [$_{CP}$ [C [$_{TP}$ subject [$T_{[-AGR]}$ [$_{vP}$ …]]]]] (order irrelevant)

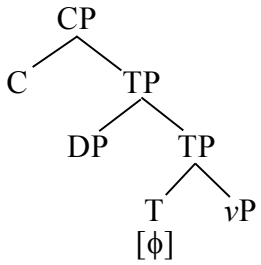
When the CP phase is completed, the complement TP is transferred if T carries ϕ-features, but only *v*P is if T does not agree with the subject. The examples in (18) fall under the latter case. Thus, in (18b), the embedded subject is transferred to the C-I interface when the matrix *v*P is completed as shown in (22).

(22) $[_{vP}$ Hanako-ga $[[_{VP}\ [_{CP}\ [[_{TP}$ zibunzisin-ga $[[_{vP} \ldots]$ $T_{[+Past]}]]$ C]] syutyoos] *v*]]

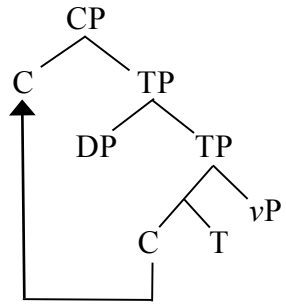
The matrix *v* triggers the transfer of the shaded part of the VP, which contains *zibunzisin*. As the information 'zibunzisin = Hanako' can also be sent to the C-I interface, the example is expected to be grammatical with the intended interpretation.[5]

This analysis raises the obvious question why there should be an asymmetry between CPs with and without ϕ-feature agreement, that is, why the difference in (21) obtains. However, I leave this question for future research and just stipulate a technical implementation here. It is standardly assumed that T and C merge into the structure independently as in (23a). Suppose that this is the case for T with ϕ-features, but that C and T merge first and the former excorporates as in (23b) when T lacks ϕ-features.[6]

(23) a.



b.



Then, the transfer domain can be the maximal phrase that excludes the phase head, where *exclude* is defined as follows:

(24) α *excludes* β $=_{def}$ α does not dominate any instance of β.

The maximal phrase that excludes C is the higher TP in (23a) whereas it is *v*P in (23b).

Again, the transfer domains specified in (23) are yet to be given an explanation.[7] But they do have implications for the analysis of ECM and control sentences. I discuss the

former in the remainder of this subsection and turn to the latter in the next.

The clausal complements in ECM sentences have been assumed to be TPs. This allows, for example, the movement in (25b).

(25) a. John believes [$_{TP}$ Mary to be a genius]

b. Mary is believed [$_{TP}$ Mary to be a genius]

Given the standard assumptions on phases, this movement should be impossible with CP complementation. The TP containing *Mary* should be transferred upon the completion of the embedded CP, as illustrated in (26a).

(26) a. [$_{CP}$ C [$_{TP}$ Mary to be a genius]]

b. [$_{CP}$ Mary [C [$_{TP}$ Mary to be a genius]]]

Then, *Mary* is not accessible for movement into the matrix clause. If it is to escape the embedded clause, it must first move to the edge of the embedded CP as in (26b). But this would result in improper movement of the form A-A'-A.

However, once the transfer domains are defined as in (21), a different picture emerges. The relevant part, (21b), is repeated in (27).

(27) [$_{CP}$ [C [$_{TP}$ subject [$T_{[-AGR]}$ [$_{vP}$ ...]]]]] (order irrelevant)

This states that only *v*P is transferred upon the completion of the CP phase when T lacks ϕ-features. Then, the movement in (25b) should be possible even with CP complementation as illustrated in (28).

(28) Mary ... [$_{CP}$ C [$_{TP}$ Mary [to [$_{vP}$ be a genius]]]]

Thus, the analysis allows clausal complements with T to uniformly be CPs.

This leads to a possible solution for the first set of problematic examples in (15),

repeated below in (29).

(29) a. *John expects [$_{TP}$ the guests to be introduced to himself]

b. *Mary believes [$_{TP}$ the speakers to have been introduced to herself]

If the complements in these examples are CPs and hence phases, the embedded *v*P should be transferred when they are completed. (30) shows how this applies to (29a).

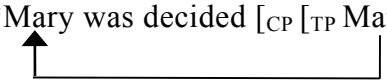
(30) [$_{CP}$ C [$_{TP}$ the guests [to [$_{vP}$ be introduced the guests to himself]]]]

The reflexive *himself* is part of this transfer domain, but information on its reference cannot be sent to the C-I interface. Hence, it fails to receive an interpretation and the ungrammaticality of the example is correctly predicted.

I argued so far that (21), motivated by the absence of NIC effects in languages without ϕ-feature agreement, allows ECM complements to be CPs and also makes it possible to account for the ungrammaticality of (29a, b). It should be noted at this point that raising out of ECM complements as in (25b) has been discussed in conjunction with the impossibility of raising out of control complements as in (31b).

(31) a. Mary decided [$_{CP}$ [$_{TP}$ PRO to go to college]]

b. *Mary was decided [$_{CP}$ [$_{TP}$ Mary to go to college]]



(31b) is straightforwardly ruled out if the embedded TP is transferred upon the completion of the embedded CP. However, the analysis entertained in this section allows this movement, as the transfer domain should be *v*P instead of TP. I come back to this problem in the following section, where I discuss the movement theory of control. As explicitly noted in Hornstein (1999), the theory shares this problem.

2.3. Phases and the Movement Theory of Control

The analysis suggested in the preceding subsection provides indirect support for the

movement theory of control. I first briefly summarize the background of the theory and then discuss how the analysis in the preceding subsection relates to it.

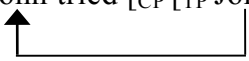
The θ-criterion, stated in (32), was one of the central principles in the theory of Chomsky (1981).

(32) Each θ-role is assigned to exactly one argument and each argument is assigned exactly one θ-role.

It is expected under Full Interpretation that each θ-role must be assigned to an argument and that each argument must receive a θ-role. Each element in the structure must be properly interpreted. What is stipulated in (32) is uniqueness, that is, that there is a one-to-one relation between θ-roles and arguments. However, as Hornstein (1999) notes, it follows from the postulation of D-structure as a pure representation of thematic relations that an argument cannot receive θ-roles in two distinct positions. Let us consider (33).

(33) a. John tried [$_{CP}$ [$_{TP}$ PRO to win the race]]

b. John tried [$_{CP}$ [$_{TP}$ John to win the race]]

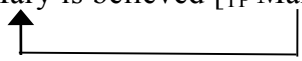


In this example, *John* is construed as the agent of both *try* and *win*. If the example is derived by movement as in (33b), *John* receives two θ-roles in violation of the θ-criterion. But the derivation was excluded independently because of the defining property of D-structure. *John* is in the embedded clause at this level, and hence, D-structure fails to represent the thematic relation of the DP and the matrix verb *try*. It was assumed then that *John* merges in the matrix clause and binds an empty category PRO in the embedded clause as in (33a).

Hornstein points out that it requires a stipulation to rule out (33b) in the minimalist model, which dispenses with the level of D-structure. Nothing in the model prevents a DP from receiving a θ-role at the landing site. As it is conceptually desirable to eliminate the stipulated part of the θ-criterion that an argument can receive at most one θ-role, he pursues

and argues for the movement analysis in (33b). According to his analysis, *John* in (33b) moves to the matrix clause to be valued for Case, just as *Mary* does in (25b), repeated below as (34).

(34) Mary is believed [$_{TP}$ Mary to be a genius]



The only difference is that *John* is assigned an additional θ-role by the matrix *v* in (33b).[8]

The movement analysis of control has a number of attractive features, including the elimination of PRO, whose status was never clear in the theory. But one potential problem was its apparent incompatibility with phase theory. If control complements are CPs, then the embedded subject *John* in (33b) is transferred before it can move into the matrix clause as shown in (35a).

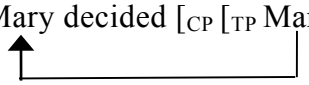
(35) a. [$_{CP}$ C [$_{TP}$ John [to [$_{vP}$ win the race]]]]

b. [$_{CP}$ C [$_{TP}$ John [to [$_{vP}$ win the race]]]]

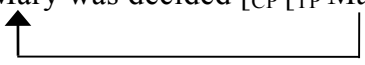
However, if only *v*P is transferred when T lacks ϕ-features, as argued above, the problem no longer arises. *John* is outside the transfer domain as illustrated in (35b). It is then free to move into the matrix clause and receive an additional θ-role. Thus, the proposal on the transfer domains in the preceding subsection fits very well with the movement theory of control.

As the movement theory of control assumes that A-movement is possible across CP, the ungrammaticality of (31b), as opposed to (31a), has been a problem since the theory was proposed. The examples are repeated in (36) with revised structures.

(36) a. Mary decided [$_{CP}$ [$_{TP}$ Mary to go to college]]



b. *Mary was decided [$_{CP}$ [$_{TP}$ Mary to go to college]]



As noted in the preceding subsection, (36b), as it contrasts with (34), poses a problem for the analysis presented in this paper as well.

Following Hornstein (1999), I assume here that the contrast between (34) and (36b) is due to the difference in the tense interpretation of the embedded clause. Stowell (1982) points out that the embedded non-finite clauses in (37a, b) are interpreted differently with respect to tense.

(37) a. John believes [$_{CP}$ [$_{TP}$ Mary to be a genius]]

b. Mary decided [$_{CP}$ [$_{TP}$ Mary to go to college]]

The embedded clause in (37a) lacks tense and the tense of the matrix clause extends to it. Mary being a genius and John believing it to be the case are simultaneous. On the other hand, in (37b), Mary going to the college is future with respect to her decision to do so. Given this, Stowell proposes that *to* in this example expresses unrealized future tense. Martin (2001) develops this analysis and argues that *to* in ECM complements lacks tense whereas *to* in control infinitives carries future tense.

Capitalizing on this difference, I tentatively suggest as part of Full Interpretation that tense defines domains that must be thematically complete. Suppose that an embedded TP carries tense as in (38).

(38) [$_{TP}$ ... tense ... [$_{CP}$ [$_{TP}$ ... tense ...]]]

domain 2 domain 1

The idea is that any DP in A-position must receive an interpretation in the tense domain in which it appears. (37b) satisfies this as *Mary* receives a θ-role in both embedded and matrix clauses. So does (34) because there is only one tense domain in this example as the embedded clause lacks tense. The situation is different in (36b). There are two domains as

in (38) and *Mary* receives a θ-role only in the embedded tense domain. Thus, the condition excludes this example.[9]

There are independent reasons why a condition of this kind is required although it is still a stipulation at this point. One concerns the locality of A-scrambling. As initially noted by Mahajan (1990), clause-internal scrambling can be A-movement. Relevant examples in Japanese, as in (39), are discussed in detail by Tada (1993), Saito (1992) and Miyagawa (2001), among others.

(39) a. [Karera-ga [$_{VP}$ [otagai-no sensei]-o hihansita]] (koto)
they-Nom each other-Gen teacher-Acc criticized fact
'They criticized [each other's teachers].'

b. ?*[[Otagai-no sensei]-ga [$_{VP}$ karera-o hihansita]] (koto)
each other-Gen teacher-Nom they-Acc criticized fact
'[Each other's teachers] criticized them.'

c. [Karera-o [[otagai-no sensei]-ga [$_{VP}$ hihansita]]] (koto)
they-Acc each other-Gen teacher-Nom criticized fact
'*Lit*. Them, [each other's teachers] criticized.'

(39b), as opposed to (39a), is ungrammatical because *karera* 'they' does not c-command the anaphor *otagai* 'each other'. The sentence, however, improves dramatically when *karera* is scrambled to the sentence-initial position, as shown in (39c). This indicates that a scrambled phrase can serve as an A-binder.

Mahajan notes at the same time that the effect in (39c) does not obtain when the antecedent is scrambled across a CP boundary. Thus, no improvement is observed in (40b) over (40a).

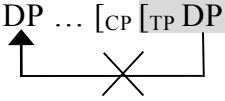
(40) a. *[[Otagai-no sensei]-ga [$_{VP}$ [$_{CP}$ [Tanaka-ga karera-o hihansita] to]
each other-Gen teacher-Nom Tanaka-Nom they-Acc criticized C
itta]] (koto)
said fact
'[Each other's teachers] said that Tanaka criticized them.'

b. *[Karera-o$_i$ [[otagai-no sensei]-ga [$_{VP}$ [$_{CP}$ [Tanaka-ga hihansita] to]
they-Acc each other-Gen teacher-Nom Tanaka-Nom criticized C
itta]] (koto)
said fact
'*Lit*. Them, [each other's teachers] said that Tanaka criticized.'

He concludes that this kind of scrambling is necessarily A'-movement.[10]

Various analyses have been proposed for this asymmetry between scramblings within TP and across CP. But the general consensus is that scrambling to the edge of TP counts as A-movement whereas that to the edge of CP is A'-movement. Given this, the ungrammaticality of (40b) is accounted for if the embedded TP is transferred upon the completion of the embedded CP. This is illustrated in (41).

(41) [… DP … [$_{CP}$ [$_{TP}$ DP [$_{TP}$ subject [$_{v}$P …]]]]]



The scrambled DP cannot move from within the embedded TP to an A-position in the matrix clause if the TP is sent to the interfaces when the embedded CP is constructed. If the DP moves first to the edge of the embedded CP, then further movement to an A-position results in improper movement of the form A-A'-A.

However, I argued in this section that what is transferred upon the completion of a CP phase is not TP but *v*P in the absence of ϕ-feature agreement. If this is correct, *v*P instead of TP should be transferred in (41). Then, A-scrambling out of CP should be possible as in

(42).

(42) [… DP … [$_{CP}$ [$_{TP}$ DP [$_{TP}$ subject [$_{vP}$ …]]]]]

The resulting structure, however, is ruled out independently by the condition on tense domains. In (40b), the embedded clause as well as the matrix clause carry tense. Hence, the scrambled DP must receive a θ-role in both clauses. As scrambling, by definition, is not movement to a θ-position, (42) violates the condition.

3. "Chains" and Labeling

I turn in this section to the mechanism of labeling. In the following subsection, I briefly introduce Chomsky's (2013) labeling algorithm and the proposals in Saito (2014, 2016b) to extend it to Japanese. In Section 3.2, I argue that only XPs in uniform A-chains and uniform A'-chains count as single syntactic objects in the calculation of labels, and show that the hypothesis leads to a principled account for improper movement. Then, in Section 3.3, I demonstrate that it has a desirable consequence for the attempt in Saito (2016a) to eliminate the remaining stipulated part of the θ-criterion in favor of labeling.

3.1. Chomsky's (2013) Labeling Algorithm and its Application to Japanese

The minimal operation, Merge, applies to two objects α and β, and forms a new object γ = {α, β}. Chomsky (2013) hypothesizes that it must accompany an algorithm to determine the nature (or label) of the newly formed object. When a verb and a nominal element merge, the interpretation requires information on whether the formed object is verbal (VP) or nominal (NP), for example. He proposes that this labeling algorithm explains, among other things, why movement (or internal merge) is restricted the way it is and why language has ϕ-feature agreement.

Chomsky (2013) starts the discussion with consideration of the three cases of Merge in (43).

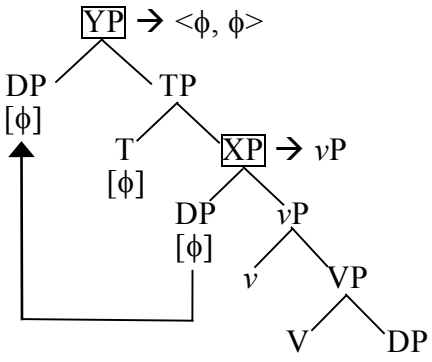
(43) a. $\gamma = \{H, \alpha P\}$

b. $\gamma = \{\alpha P, \beta P\}$

c. $\gamma = \{H_1, H_2\}$

(43a) is straightforward as search into γ immediately yields a unique head, H. In this case, it is assumed that H determines the label of γ. On the other hand, (43b, c) are problematic because the label of γ cannot be determined straightforwardly. Given this, Chomsky makes two concrete proposals to accommodate instances of (43b) that arise in actual derivations. Let us consider the structure in (44).

(44)



Merge applies first to yield {V, DP} and then {*v*, {V, DP}}. These are instances of the unproblematic (43a). But then, the configuration in (43b) arises when the subject DP and *v*P merge. In this case, the DP internally merges with TP later in the derivation after T is introduced into the structure. As a result, the DP occurs in two positions, the Specs of *v* and T. Chomsky proposes that *v*P determines the label of XP at this point because it is the unique element that XP properly contains. The internal merge of the DP with TP again creates an instance of (43b). Here, the DP and (the label of) TP share the same ϕ-features due to ϕ-feature agreement. Chomsky suggests that this feature sharing makes it possible to label YP as <ϕ, ϕ>.

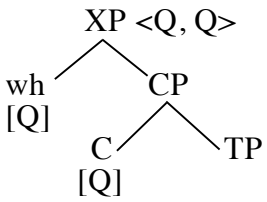
Chomsky points out that this analysis extends to structures created by wh-movement. Let us consider (45):[11]

(45) [$_{YP}$ Which pen [$_{CP}$ do [$_{TP}$ you think [$_{XP}$ which pen [$_{CP}$ that [$_{TP}$ he bought which pen]]]]]]

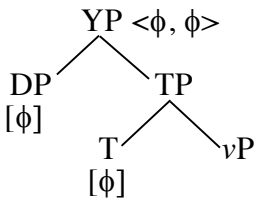
Here, there are two instances of internal merge of a wh-phrase with a CP. In the embedded clause, a wh-phrase merges with a non-question CP. This is allowed because the wh-phrase moves further, and as a result, CP provides the label of XP as the only element properly contained within XP. The one in the matrix is legitimate as the C heads a question with the feature Q, and the formed object, YP, can be labeled as <Q, Q> with feature sharing.

This analysis predicts that internal merge always terminates in a configuration of feature sharing. This is illustrated below.

(46) a. Operator movement



b. NP-movement



In particular, it excludes internal merge at TP and CP without feature sharing unless the merged phrase "moves on." Then, internal merge, as an instance of Merge, is free but is severely restricted by the labeling algorithm. The "last resort" nature of internal merge is thus captured.

Chomsky's (2013) proposals on labeling raise interesting research questions with languages like Japanese. First, Japanese lacks ϕ-feature agreement altogether, at least on the surface. So, it is not obvious how sentences are labeled. Secondly, it is well known that the language allows sentences with multiple nominative subjects. The following example is from Kuno (1973).[12]

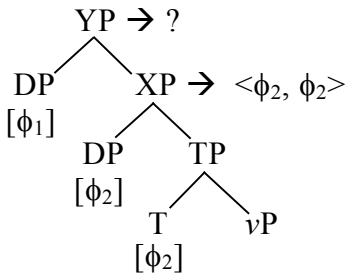
(47) [$_{TP}$ Bunmeikoku-ga [$_{TP}$ dansei-ga [$_{TP}$ heikin-zyumyoo-ga mizikai]]]

civilized.country-Nom male-Nom average-life.span-Nom short.is

'It is in civilized countries that the male population has a short life-span.'

This is ruled out straightforwardly in English. As illustrated in (48), the merger of the higher subject results in failure of labeling because only the lower thematic subject shares ϕ-features with T.

(48)



But then, why is (47) grammatical in Japanese?

Finally, Japanese has scrambling as already discussed. Even when it exhibits A-properties, the scrambled phrase does not participate in ϕ-feature agreement with T. Further, it has been argued that A'-scrambling is not operator movement. (49b) is a typical example that demonstrates this.

(49) a. Minna-ga [$_{CP}$ Hanako-ga dono hon-o eran-da ka] siritagatte iru

all-Nom Hanako-Nom which book-Acc choose-Past Q want.to.know

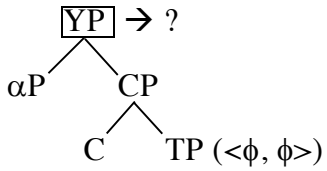
'Everyone wants to know which book Hanako chose.'

b. Dono hon-o minna-ga [$_{CP}$ Hanako-ga dono hon-o eran-da ka] siritagatte iru

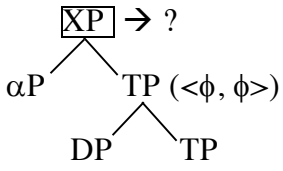
A wh-question is embedded in (49a). In (49b), the wh-phrase, *dono hon-o* 'which book-Acc', is scrambled out of the embedded CP, where it takes scope. Yet, the example is perfectly grammatical and is interpreted exactly as (49a) without scrambling. The movement is not A-movement as it crosses a CP boundary, as discussed in the preceding

section. It cannot be operator movement either because if it were, the scrambled phrase should take scope at the final landing site.[13] Again, scrambling is ruled out, for example, in English by the labeling requirement. If a non-operator object is internally merged with CP or TP, then the formed object fails to be labeled as shown in (50).

(50) a.



b.



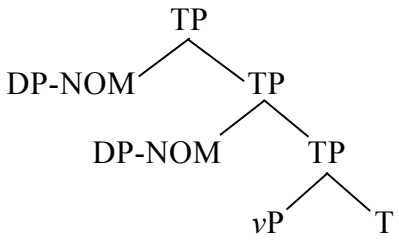
The characteristic properties of Japanese illustrated above suggest that the language employs a labeling mechanism that is not observed in languages with ϕ-feature agreement. The purpose of Saito (2014, 2016b) was to look into this mechanism. The main part of the hypothesis presented there is that suffixal Case markers make phrases invisible for search and as a result, serve as anti-labeling devices. Let us consider (51) for a concrete illustration.

(51) $\gamma = \{\alpha P\text{-Case}, \beta P\}$

The idea is that αP with Case is invisible when the label is calculated for γ, and hence, βP determines the label for γ.[14]

This not only allows Japanese sentences to be labeled without ϕ-feature sharing but also accommodates examples with multiple nominative subjects, as shown in (52).[15]

(52)

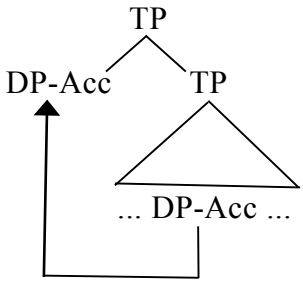


When the lower DP merges with TP, TP determines the label of the formed object because the DP with suffixal Case is invisible. The same mechanism allows the merger of the

higher DP, referred to as a 'major subject' in the literature. Since it accompanies suffixal Case, the middle TP determines the label of the top TP. Thus, sentences with multiple nominative subjects are predicted to be grammatical.

The hypothesis also accounts for why scrambling is possible in Japanese. Scrambling in (53a) creates the configuration in (53b).

(53) a. Sono hon-o Taroo-ga sono hon-o katta
that book-Acc Taroo-Nom bought
'Taroo bought that book.'

b. 

The object formed by this internal merge is successfully labeled. As DP-Acc is invisible for labeling, the lower TP provides the label for the newly formed TP. This analysis is extended to PP scrambling in Saito (2014). PPs in the subject position and within nominal projections must appear with suffixal Case as shown in (54).

(54) a. Koko-kara-ga huzi-san-ni noboriyasui
here-from-Nom Mt. Fuji-Dat climb.easy.is
'It is from here that one can easily climb Mt. Fuji.'

b. Taroo-no Yooroppa-e-no ryokoo
Taroo-Gen Europe-to-Gen trip
'Taroo's trip to Europe'

c. Hanako-no Tookyoo-kara-no syuppatu

Hanako-Gen Tokyo-from-Gen departure

'Hanako's departure from Tokyo'

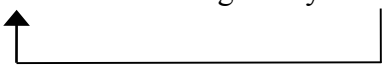
As PPs require Case in these contexts, I suggested that those within VP are accompanied by null suffixal Case as well. This accounts for PP scrambling as in (55).

(55) a. Hanako-ga tosyokan-kara hon-o karidasita

Hanako-Nom library-from book-Acc checked.out

'Hanako checked out a book from the library.'

b. Tosyokan-kara Hanako-ga tosyokan-kara hon-o karidasita

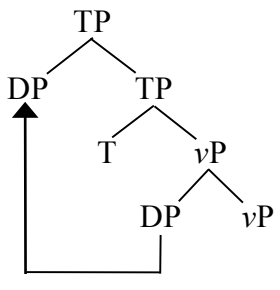


3.2. Improper Movement and Labeling

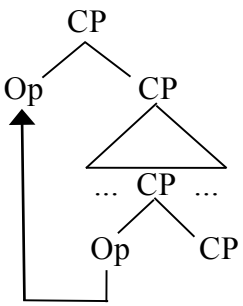
The main proposal of this section concerns precisely when movement, that is, internal merge, makes labeling possible. As seen above, Chomsky (2013) considers two cases, a DP moving out of {DP, *v*P} and a wh-phrase moving out of {wh, CP}. A uniform A-chain is formed in the former case, and a uniform A'-chain in the latter. I suggest in the remainder of this paper that internal merge aids labeling in this way only when it creates a uniform chain. I argue in this section that this hypothesis leads to a principled explanation for the ban on improper movement.

The two cases Chomsky (2013) considers are illustrated in (56).

(56) a.



b.



In (56a), the internal merge of DP makes it possible for the lower *v*P to provide the label for the higher *v*P. The reasoning behind this is that the syntactic object, DP, occurs in two positions, forming a "chain" in the classical terminology, and hence, the higher *v*P does not contain the DP.

It is interesting in this connection that Chomsky (1991) lists (57a-d) as syntactic objects allowed by Full Interpretation.

(57) a. uniform A-chains: A-A-A

b. uniform A'-chains: A'-A'-A'

c. uniforms head chains: H-H-H

d. operator-variable chains: A'-A

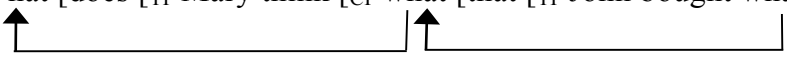
Let us take a closer look at (57a, b). If an A-chain, as in (57a), has more than one member, the element forming the chain must be in ϕ-feature sharing configuration at the head position in languages like English for the structure to be properly labeled. And if the element is an argument and not an expletive, it receives a θ-role in at least one of the positions. In reality, as an argument is initially merged into a θ-position, it receives a θ-role at the tail of the chain. But it can receive additional θ-roles in higher positions as was seen in the discussion of control structures. *Mary* in (58) receives three θ-roles altogether.

(58) Mary wants to try to open the door

Then, in a non-trivial A-chain, a single syntactic object appears in multiple positions and can receive a θ-role in any of the positions.

A uniform A'-chain has a similar property. Let us first consider (59).

(59) [$_{CP}$ What [does [$_{TP}$ Mary think [$_{CP}$ what [that [$_{TP}$ John bought what]]]]]]



The higher two instances of *what* form a uniform A'-chain. The wh-phrase is interpreted as

an operator ‘for which x: x a thing’ at the head position. But this is required for labeling as discussed above. Unless the wh-phrase internally merges with a question CP in the last step, the newly formed constituent fails to be labeled. The situation is different in the Japanese (49b), repeated below in (60) with more precise structure.

(60) [$_{CP}$ Dono hon-o [$_{CP}$ minna-ga [$_{CP}$ dono hon-o [$_{CP}$ [$_{TP}$ Hanako-ga dono hon-o

which book-Acc all-Nom Hanako-Nom

eranda] ka]] siritagatte iru]]

chose Q want.to.know

‘*Lit*. Which book, everyone wants to know Hanako chose.’

(= Everyone wants to know which book Hanako chose)

In this example, the wh-phrase *dono hon* ‘which book’ internally merges with a question CP first and then moves on to the initial position of the matrix, non-question CP. Here, the two landing sites form a uniform A’-chain, and the wh-phrase receives interpretation as an operator at its tail. This is allowed because the matrix CP is successfully labeled, thanks to the suffixal accusative Case on the wh-phrase. Then, aside from the labeling requirements, an operator can be interpreted at any position in a uniform A’-chain.[16]

It is worth mentioning that if the syntactic object in a uniform A’-chain is not an operator, it need not receive an interpretation at all. This would parallel the case of an expletive in an A-chain. Thus, a non-operator can be A’-scrambled as in (61).

(61) [$_{CP}$ Sono hon-o [$_{CP}$ minna-ga [$_{CP}$ sono hon-o [$_{CP}$ [$_{TP}$ Hanako-ga sono hon-o

that book-Acc all-Nom Hanako-Nom

eranda] to]] omotte iru]]

chose C think

'*Lit*. That book, everyone thinks that Hanako chose.'

(= Everyone thinks that Hanako chose that book.)

The matrix CP is successfully labeled, again because of the suffixal accusative Case on the scrambled object. As the scrambled object is not interpreted as an operator, scrambling has no effect on the interpretation in this case.

It was argued so far that a single syntactic object appears in a uniform A-chain and can be interpreted as an argument in any position within the chain. Similarly, a non-trivial uniform A'-chain is formed when a single object occurs in multiple A'-positions. The object can be interpreted as an operator in any of those positions. The situation is a little different in operator-variable chains. Let us consider (59) again. The initial internal merge of *what* forms an operator-variable chain. As internal merge is copying, it can be said that a single syntactic object occurs in two positions, one A and the other A'. However, unlike the case of uniform chains, it is not that the syntactic object can be interpreted in either position. A wh-phrase is interpreted as an operator in A'-positions and as a variable in A-positions, as illustrated in (62).

(62) *what* (A') ... *what* (A') ... *what* (A') ... *what* (A) ... *what* (A) ... *what* (A)

for which x: x a thing x

Then at least the interpretive component views *what* in A'-positions and *what* in A-positions as distinct objects. A possibility arises here that syntax does the same, and if this is the case, internal merge from an A-position to an A'-position should not aid labeling

in the way that those in (56) do.[17]

Although Chomsky (1991) does not consider internal merge from an A'-position to an A-position as an operation forming a legitimate syntactic object, the operation itself should be possible if internal merge is a subcase of free Merge, as assumed in Chomsky (2008, 2013). But just like the operator-variable case, it arguably does not create a configuration in which a single syntactic object occurs in two positions. Let us look at the relevant structure in (63).

(63) $\underbrace{\alpha\,(A) \ldots \alpha\,(A) \ldots \alpha\,(A)}_{\text{argument interpreted}} \ldots \underbrace{\alpha\,(A') \ldots \alpha\,(A') \ldots \alpha\,(A')}_{\text{operator interpreted}}$

It is not that α can receive interpretation in any of the six positions. If it is an argument, it must receive a θ-role in at least one of the A-positions and not be interpreted at all in the A'-positions. If it is an operator, it should be interpreted as such in one of the A'-positions and receives no interpretation, like expletives, in the A-positions. It is possible then that α in the A-positions and α in the A'-positions are construed as distinct syntactic objects. This leads to the hypothesis that only uniform A-A movement and A'-A' movement aid labeling as illustrated in (56). I argue in the remainder of this subsection that this hypothesis leads to a principled explanation for the ban on improper movement.

I simply assumed in the discussion in Section 2 that improper movement of the form, A-A'-A, is illicit. But as far as I know, an issue remains as to how it is to be excluded.[18] Improper movement is often discussed in relation to examples of "super raising," examined in Lasnik and Saito (1984) and Chomsky (1986). Relevant examples are shown in (64).

(64) a. *John seems [$_{CP}$ that [$_{TP}$ it is likely [$_{CP}$ [$_{TP}$ John to win the election]]]]

b. *Mary seems [$_{CP}$ that [$_{TP}$ it was told Mary [$_{CP}$ that [$_{TP}$ John would be here]]]]

The indicated derivations are excluded by the phase theory. When the CP complement of *seem* is completed, the shaded TP is transferred. Hence, the raising fails. But this presupposes that the movement cannot proceed through the edge of the CP. If the raising in (64b), for example, can take place as in (65), the example cannot be ruled out in this way.

(65) *Mary seems [$_{CP}$ Mary [that [$_{TP}$ it was told Mary [$_{CP}$ that [$_{TP}$ John would be here]]]]]

This particular example is excluded independently by the condition on tense domains suggested in Section 2.3. *Mary* appears in A-positions in two independent tense domains as shown in (66).

(66) Mary seems [$_{CP}$ Mary [that [$_{TP}$ it was told Mary [$_{CP}$ that [$_{TP}$ John would be here]]]]]

domain 3 domain 2 domain 1

Then, the condition requires *Mary* to be θ-marked in domains 2 and 3. (65) is ruled out as the DP fails to receive a θ-role in domain 3. More relevant then are examples of control that violate the locality requirements as in (67b).

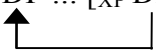
(67) a. Mary thinks [$_{CP}$ that [$_{TP}$ John wants [$_{CP}$ [$_{TP}$ John to attend the meeting]]]]

b. *Mary thinks [$_{CP}$ Mary [that [$_{TP}$ John wants [$_{CP}$ [$_{TP}$ Mary to attend the meeting]]]]]

In (67b), *Mary* receives θ-roles in the two tense domains it occurs in, the most deeply embedded TP and the matrix TP. Then, improper movement in this case still needs to be excluded.[19]

The hypothesis that only uniform chains aid labeling readily serves the purpose here. What is problematic in (67b) is the last step of the movement, illustrated in (68).

(68) ... DP ... [$_{XP}$ DP [$_{CP}$ that ...



As the DP originates in an A'-position and is internally merged in an A-position, it is visible when search applies to XP. As XP contains two phrases, DP and CP, and there is no feature sharing between the two, XP fails to be labeled. Thus, improper movement necessarily leads to failure in labeling.

3.3. On the Elimination of the θ-Criterion in Favor of Labeling

A piece of evidence was presented that the extraction of XP from γ = {XP, YP} does not make it possible to label γ when XP originates in an A'-position and moves to an A-position. In this subsection, I argue that the same conclusion holds when the movement is from an A-position to an A'-position.[20] The argument is a little involved, and is based on the proposal in Saito (2016a) to eliminate the θ-criterion in favor of labeling. I first briefly go over this proposal.

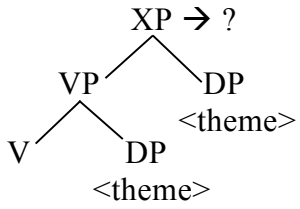
I introduced Hornstein's (1999) movement theory of control in Section 2.3. If it is correct, the stipulation in the θ-criterion that an argument can receive at most one θ-role should be dispensed with because it does not hold in control sentences. In Saito (2016a), I argued that the remaining stipulation that a θ-role can be assigned to at most one argument should also be dispensed with. The argument is based on the fact that there is much overlap in the effects of this stipulation and the labeling requirement. Let us consider the straightforward cases of θ-criterion violations in (69).

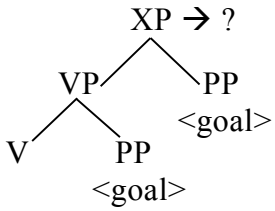
(69) a. *Mary hit the head John

b. *Mary went to Germany (three times) to Europe

There are two theme arguments in (69a) and two goal arguments in (69b). But independently of the θ-criterion, the predicates in these examples fail to be labeled as

illustrated in (70).

(70) a. 

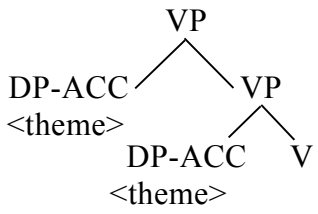
b. 

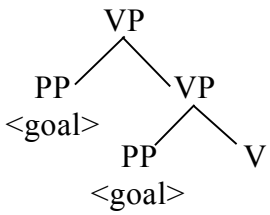
Then, the θ-criterion is not necessary to account for the ungrammaticality of these examples.

This leads to the hypothesis that the stipulated part of the θ-criterion should be dispensed with altogether. And Japanese provides an interesting testing ground for the hypothesis. Recall the proposal in Saito (2014) that suffixal Case in Japanese serves as an anti-labeling device. That is, suffixal Case in (71) makes XP invisible for labeling and as a result, allows YP to determine the label of γ.

(71) $\gamma = \{XP\text{-}Case, YP\}$

If this is correct, the Japanese counterparts of (70a, b) should have no problem with labeling, as illustrated in (72).

(72) a. 

b. 

If the θ-criterion holds independently of labeling, then these structures should be illicit in Japanese. On the other hand, the hypothesis to eliminate the θ-criterion in favor of labeling predicts that these structures are allowed in the language.

Kuroda (1988) already discusses examples that bear out the prediction. One of his examples is shown in (73).

(73) a. ??Masao-ga Hanako-o hoho-o butta

Masao-Nom Hanako-Acc cheek-Acc hit

'Masao hit Hanako on the cheek.'

b. [$_{CP}$ Masao-ga Hanako-o butta no]-wa hoho(-o) da

Masao-Nom Hanako-Acc hit C-Top cheek-Acc is

'It is on the cheek that Masao hit Hanako.'

(73a) is marginal because there is a somewhat mysterious surface constraint in Japanese that prohibits multiple occurrences of accusative Case within a simple sentence. But one can avoid its effect by dislocating one of the accusative phrases as in (73b). Kuroda concludes then that the core grammar of Japanese allows examples like (73a).

He further shows that examples of this kind are allowed only when the two accusative phrases are interpreted as the theme objects. He contrasts (73) with (74).

(74) a. *Masao-ga Hanako-o yubi-o ni-hon otta

Masao-Nom Hanako-Acc finger-Acc two-Cl broke

'Masao broke two of Hanako's fingers.'

b. *[$_{CP}$ Masao-ga Hanako-o otta no]-wa yubi-o ni-hon da

Masao-Nom Hanako-Acc broke C-Top finger-Acc two-Cl is

'*Lit*. It is two fingers that Masao broke Hanako.'

This example is totally ungrammatical because one can break fingers but not a person, as confirmed in (75).

(75) a. Masao-ga Hanako-no yubi-o ni-hon otta

Masao-Nom Hanako-Gen finger-Acc two-Cl broke

'Masao broke two of Hanako's fingers.'

b. *Masao-ga Hanako-o otta

Masao-Nom Hanako-Acc broke

'*Lit*. Masao broke Hanako.'

(73a), then, indeed instantiates the configuration in (72a).

Argument doubling of this kind, which Kuroda (1988) observes rather casually, is found extensively in Japanese. The goal argument is doubled in (76).

(76) a. ??Hanako-ga Yooroppa-e Doitu-e san-kai(-dake) itta

Hanako-Nom Europe-to Germany-to three-Cl-only went

'*Lit.* Hanako went (only) three times to Germany to Europe.'

b. [$_{CP}$ Hanako-ga Yooroppa-e itta no]-wa Doitu-e san-kai(-dake) da

Hanako-Nom Europe-to went C-Top Germany-to three-Cl-only is

'*Lit*. It is (only) three times to Germany that Hanako went to Europe.'

c. Yooroppa-e-wa Hanako-ga Doitu-e san-kai itta

Europe-to-Top Hanako-Nom Germany-to three-Cl went

'*Lit*. To Europe, Hanako went to Germany three times.'

The marginality of (76a) indicates that the surface constraint against multiple accusatives applies to postpositions as well. Its effect is avoided by cleft in (76b) and by PP topicalization in (76c).[21]

An example with doubling of the source argument is shown in (77).

(77) a. ??Amerika-kara Nyuuyooku-kara(-dake) tegami-ga todoita

the U.S.-from New York-from-only letter-Nom arrived

'*Lit*. A letter arrived (only) from New York from the U.S.'

b. [$_{CP}$ Amerika-kara tegami-ga todoita no]-wa Nyuuyooku-kara(-dake) da

the U.S.-from letter-Nom arrived C-Top New York-from-only is

'*Lit*. It is (only) from New York that a letter arrived from the U.S.'

c. Amerika-kara-wa Nyuuyooku-kara tegami-ga todoita
the U.S.-from-Top New York-from letter-Nom arrived
'*Lit*. From the U.S., a letter arrived from New York.'

(76) and (77) show that Japanese allows the structure in (72b).

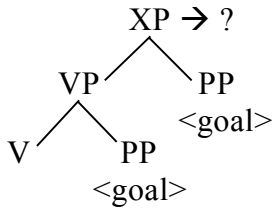
The Japanese examples discussed above indicate that argument doubling is possible as long as the structure is successfully labeled. This shows in turn that the θ-criterion should be eliminated and that the examples in (69), repeated below in (78), are ungrammatical because of failure in labeling.

(78) a. *Mary hit the head John

b. *Mary went to Germany (three times) to Europe

However, Masao Ochi (personal communication) points out that one further pattern should be examined to make the analysis complete. The structure in (70b), repeated in (79), fails to be labeled.

(79)



But the relevant examples remain ungrammatical even when the higher PP is topicalized, as shown in (80).

(80) a. *To Europe, Mary went to Germany three times

b. *From the U.S., a letter arrived from New York

If the movement of XP out of γ = {XP, YP} enables YP to provide the label for γ, then these examples are expected to be grammatical, at least for those who generously accept topicalization.

Here, the hypothesis presented in the preceding subsection makes it possible to extend

the labeling analysis to these examples. The hypothesis was that only uniform movement from an A-position to an A-position or from an A'-position to an A'-position accommodates labeling in this way. The topicalization in (80) is from an A-position to an A'-position, and creates an operator-variable chain.[22] Thus, the predicates in (80a, b) should fail to be labeled.

4. Conclusion

In this paper, I suggested possible directions to pursue in the investigation of phase and labeling. Although it is still controversial whether the locality of anaphor binding is to be explained by phase theory, it seems to be a promising hypothesis. I tried to contribute to this endeavor in the first note. There, I took the absence of NIC effects in languages without ϕ-feature agreement seriously, and suggested that *v*P, and not TP, is transferred to the CI interface when the CP phase lacks ϕ-feature agreement. I showed that this makes it possible to assume that ECM complements are CPs, like other types of clausal complements, and also to capture the SSC effects in those complements. I argued further that the suggestion fits well with Hornstein's (1999) movement theory of control.

The second note was on Chomsky's (2013) hypothesis that the extraction of XP makes it possible for YP to provide the label for γ = {XP, YP}. I suggested that this is possible when and only when the movement of XP creates a uniform chain of the form, A-A or A'-A'. I showed that this leads to an explanation for the ban on improper movement in terms of labeling. I argued in addition that the suggestion provides the missing piece in the effort to eliminate the θ-criterion in favor of labeling. The discussion here was also exploratory. But I hope it too showed that a new possible direction for research arises when the theory is examined against data from both English and Japanese.

References

Akmajian, Adrian and Frank Heny (1975) *An Introduction to the Principles of Transformational Syntax*, MIT Press, Cambridge, MA.

Bošković, Željko (1997) *The Syntax of Nonfinite Complementation: An Economy Approach*, MIT Press, Cambridge, MA.

Bošković, Željko (2007) "On the Locality and Motivation of Move and Agree: An Even More Minimalist Theory," *Linguistic Inquiry* 38, 589-644.

Charnavel, Isabelle and Dominique Sportiche (2016) "Anaphor Binding: What French Inanimate Anaphors Show," *Linguistic Inquiry* 47, 35-87.

Chomsky, Noam (1976) "Constraints on Rules of Grammar," *Linguistic Analysis* 2, 303-351.

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

Chomsky, Noam (1986) *Barriers*, MIT Press, Cambridge, MA.

Chomsky, Noam (1991) "Some Notes on Economy of Derivation and Representation," *Principles and Parameters in Comparative Grammar*, ed. by Robert Freidin, 417-454, MIT Press, Cambridge, MA.

Chomsky, Noam (1993) "A Minimalist Program for Linguistic Theory," *The View from Building 20: Essays in Honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 1-52, MIT Press, Cambridge, MA.

Chomsky, Noam (2008) "On Phases," *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta, 291-321, MIT Press, Cambridge, MA.

Chomsky, Noam (2013) "Problems of Projection," *Lingua* 130, 33-49.

Chomsky, Noam (2015) "Problems of Projection: Extensions," *Structures, Strategies and*

Beyond: Studies in Honour of Adriana Belletti, ed. by Elisa Di Domenico, Comelia Hamann and Simona Matteini, 3-16, John Benjamins, Amsterdam.

Chomsky, Noam and Howard Lasnik (1993) "The Theory of Principles and Parameters," *Syntax: An International Handbook of Contemporary Research*, ed. by Joachim Jacobs, et al., 506-569, Mouton de Gruyter, Berlin.

Dejima, Mayumi (1999) *A Study on Remnant-Scrambled Sentences in Japanese and their LF Representations*, M.A. thesis, Nanzan University.

Fiengo, Robert (1974) *Semantic Conditions on Surface Structure*, Doctoral dissertation, MIT.

Fukui, Naoki (1993) "A Note on Improper Movement," *The Linguistic Review* 10, 111-126.

Heycock, Caroline (2008) "Japanese *-wa*, *-ga* and Information Structure," *The Oxford Handbook of Japanese Linguistics*, ed. by Shigeru Miyagawa and Mamoru Saito, 54-83, Oxford University Press, New York.

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

Hornstein, Norbert and Maria Polinsky, eds. (2010) *Movement Theory of Control*, John Benjamins, Amsterdam.

Kato, Shizuka (2016) "Condition (A) and Complex Predicates," *Nanzan Linguistics* 11, 15-34.

Kuno, Susumu (1973) *The Structure of the Japanese Language*, MIT Press, Cambridge, MA.

Kuroda, S.-Y. (1988) "Whether We Agree or Not," *Linguisticae Investigationes* 12, 1-47.

Lasnik, Howard and Mamoru Saito (1984) "On the Nature of Proper Government," *Linguistic Inquiry* 15, 235-289.

Mahajan, Anoop (1990) *The A/A-bar Distinction and Movement Theory*, Doctoral dissertation, MIT.

Martin, Roger (2001) "Null Case and the Distribution of PRO," *Linguistic Inquiry* 32, 141-166.

May, Robert (1981) "Movement and Binding," *Linguistic Inquiry* 12, 215-243.

Miyagawa, Shigeru (2001) "The EPP, Scrambling, and *Wh*-in-Situ," *Ken Hake: A Life in Language*, ed. by Michael Kenstowicz, 293-338, MIT Press, Cambridge, MA.

Nakamura, Masaru (1986) *Sokubaku-Kankei* (Binding Relations), Hituzi Syobo, Tokyo.

Quicoli, A. Carlos (2008) "Anaphora by Phase," *Syntax* 11, 299-329.

Rizzi, Luigi and Ur Shlonsky (2007) "Strategies of Subject Extraction," *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, ed. by Uli Sauerland and Hans-Martin Gärtner, 115-160, Mouton de Gruyter, Berlin.

Rooth, Mats (1992) "A Theory of Focus Interpretation," *Natural Language Semantics* 1, 75-116.

Saito, Mamoru (1989) "Scrambling as Semantically Vacuous A'-movement," *Alternative Conceptions of Phrase Structure*, ed. by Mark R. Baltin and Anthony S. Kroch, 182-200, University of Chicago Press, Chicago.

Saito, Mamoru (1992) "Long Distance Scrambling in Japanese," *Journal of East Asian Linguistics* 1, 69-118.

Saito, Mamoru (2001) "Movement and θ-roles: A Case Study with Resultatives," *Proceedings of the Second Tokyo Conference on Psycholinguistics*, ed. by Yukio Otsu, 35-60, Hituzi Syobo, Tokyo.

Saito, Mamoru (2014) "Case and Labeling in a Language without ϕ-feature Agreement," *On Peripheries: Exploring Clause Initial and Clause Final Positions*, ed. by Anna Cardinaletti, Guglielmo Cinque and Yoshio Endo, 269-297, Hituzi Syobo, Tokyo.

Saito, Mamoru (2016a) "Labeling and Argument Doubling in Japanese," ms., Nanzan University, to appear in *Tsing Hua Journal of Chinese Studies* 47.2.

Saito, Mamoru (2016b) "(A) Case for Labeling: Labeling in Languages without ϕ-feature Agreement," *The Linguistic Review* 33, 129-175.

Stowell, Tim (1982) "The Tense of Infinitives," *Linguistic Inquiry* 13, 561-570.

Tada, Hiroaki (1993) *A/A-bar Partition in Derivation*, Doctoral dissertation, MIT.

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

Yang, Dong-Whee (1983) "The Extended Binding Theory of Anaphors," *Language Research* 19.2, 169-192.

FOOTNOTES

* This paper developed out of a research seminar at the University of Connecticut, held in March, 2016. I would like to thank the participants, including Mona Anderson, Željko Bošković and Diane Lillo-Martin, for helpful discussion. I also benefited from discussions with Kenshi Funakoshi, Hisa Kitahara and Masao Ochi. Finally, thanks are due to two anonymous reviewers for comments and to Naohiko Kurokawa for editorial help. The research leading to the completion of this paper was supported in part by the JSPS Grant-in-Aid for Scientific Research (C) #16K02647 and the Nanzan University Pache Research Subsidy I-A-2 (2016).

[1] Charnavel and Sportiche (2016), on the other hand, take the latter approach. Here, I base my discussion on Quicoli (2008) as it is more straightforward, but as far as I can tell, the choice is not crucial.

[2] See, for example, Nakamura (1986) for detailed discussion on the locality imposed on *zibunzisin* and *otagai* 'each other'. As often noted in the literature, the judgments are not always uniform and there are examples in which these anaphors appear to allow "long-distance" binding. However, Kato (2016) presents evidence that these anaphors indeed require local antecedents as Nakamura proposed. Charnavel and Sportiche (2016) note that while reflexives with animate antecedents can be logophors, those with inanimate antecedents exhibit the locality required of anaphors. *Zibunzisin* and *otagai* normally take animate antecedents. But Kato examines examples where their antecedents are inanimate and demonstrates that the expected locality is observed clearly in those cases.

[3] I assumed in (13) that the scrambled object *zibunzisin-o* internally merges with CP. It is argued in the following subsection that the same effect obtains even if the "landing site" is TP instead of CP.

[4] Kato (2016), mentioned in Footnote 2, confirms this, presenting examples with inanimate

antecedents. Her analysis is that the embedded subjects in (18) move to a position within the embedded C projection, thus escaping the transfer domain triggered by the embedded C. I pursue an alternative in this paper.

[5] The analysis may extend to examples such as (ia, b), also discussed in Chomsky (1981), although the precise structure of these examples is not as clear as that of (18a, b).

(i) a. John prefers for himself to be nominated

b. They want (very much) for each other to succeed

[6] As Hisa Kitahara points out, this is reminiscent of the classical analysis, for example, in Akmajian and Heny (1975), where *(for)-to* is generated in C and *to* replaces T.

[7] An anonymous reviewer raises the question whether a contrast similar to (21) is expected in the *v**P phase. The contrast should be that VP is transferred with object agreement but only the complement of V is without. This is an important question, and the answer depends on how (21) is to be explained.

[8] Further arguments for this analysis can be found in Takano (2010) and the articles in Hornstein and Polinsky (2010), among others. Independently of the movement theory of control, it has been proposed in pursuit of the minimalist model that movement into θ-positions takes place. See, for example, Bošković (1997), Saito (2001), and the references cited therein.

[9] The condition is reminiscent of the tensed S condition (TSC) of Chomsky (1976), but is assumed here as a generalization to be derived from Full Interpretation.

[10] Mahajan (1990) notes as a counter-example to this generalization that A-scrambling out of control complements is possible. Takano (2010) examines the relevant cases in detail and shows that once the movement theory of control is adopted, the apparently problematic examples can be explained away under the assumption that A-scrambling is strictly clause-bound.

[11] The movement proceeds through the edges of the *v*Ps as well. I ignore this for ease of exposition in (45) and also in subsequent examples where it is not important.

[12] When the matrix predicate is individual-level, the sentence-initial nominative phrase is interpreted as the focus, as shown in the translation of (47). This holds also for the subsequent examples although I do not always indicate it in the translations. See Kuno (1973) and Heycock (2008) for detailed discussion on this interpretive property of matrix sentences with individual-level predicates.

[13] See Saito (1989) for detailed discussion on the non-operator nature of A'-scrambling.

[14] The hypothesis, I assume, can be stated in more precise terms on the assumption that Case in Japanese is a weak head in the sense of Chomsky (2015).

[15] In Saito (2014, 2016b), I adopt Bošković's (2007) proposal that Case valuation takes place independently of ϕ-feature valuation. His Case valuation mechanism automatically allows multiple occurrences of nominative Case (and other Cases) as exemplified in Japanese.

[16] It may be expected that an operator can be interpreted in more than one position in an A'-chain, just as an argument can receive θ-roles in multiple A-positions. I assume that this is excluded because of the ban on vacuous quantification.

[17] More precisely, the *what*s in (62) are all occurrences of a single syntactic object, but only its features as an argument are visible in A-positions and only its features as an operator matter in A'-positions.

[18] Various proposals have been made to explain why improper movement is disallowed. May (1981), for example, proposes that A'-movement leaves a variable behind and as a result, improper movement is excluded by Condition (C) of the binding theory. Fukui (1993), on the other hand, defines movement as applying uniformly to A-positions or to A'-positions in order to exclude improper movement. But these accounts seem

incompatible with later developments in the theory, the former with the copy theory of movement and the latter with free Merge.

[19] It may be thought that the derivation is ruled out because the lowest copy of *Mary* is transferred without its Case valued. The Case of the DP is valued only in the matrix clause. However, if one adapts the movement theory of control, as I do here, then it must be assumed that a DP can be transferred to the C-I interface without Case valuation. This happens in standard examples like (i).

(i) Mary tried [$_{CP}$ [$_{TP}$ Mary [to [$_{vP}$ Mary [v [$_{VP}$ fax the car]]]]]]

By hypothesis, the embedded *v*P containing a copy of *Mary* is transferred upon the completion of the embedded CP phase. The assumption makes sense if Case valuation is required for the phonetic realization of a DP, as originally proposed by Jean-Roger Vergnaud. The copy of *Mary* in the transfer domain of (i) is not phonetically realized.

[20] It should be noted that this is inconsistent with Chomsky's (2015) account for the *that*-trace effect and necessitates an alternative approach to the phenomenon. See Rizzi and Shlonsky (2007) and the references cited therein for detailed discussion more generally on the restrictions on subject extraction.

[21] I argue in Saito (2016a) that argument doubling is subject to a semantic condition, which is that the lower argument should be construed as a focus and the higher one should serve to specify the set of alternatives for the focus in the sense of Rooth (1992). This is speculated to follow from Full Interpretation, as it requires every phrase to have a unique semantic role in the overall interpretation. This is why the lower argument has to be placed in focus as in (76b) or the higher argument should be topicalized as in (76c) when one of the doubled arguments is displaced.

[22] As the moved element is a PP, the term 'A-position', which is usually employed for DPs, is used in a broad sense here. The topicalization forms an operator-variable chain probably

with the reconstruction of P.