FEATURE INHERITANCE AND VACUOUS MOVEMENT

AKIHIKO SAKAMOTO

University of Tsukuba

It has been argued in the literature (George (1980), Chomsky (1986), Agbayani (2000, 2006), among others) that a movement operation without an effect on PF output can be suspended. This is known as the vacuous movement hypothesis. In this paper, it is shown, based on recent minimalist frameworks (Chomsky (2000, 2001, 2004, 2008)), that the vacuous movement hypothesis is reducible to an explanatory system in which the mechanism of feature inheritance operates not only on the Agree feature but also on the edge feature, which draws on clausal typing (Cheng (1997a)). The proposed system is also shown to have favorable consequences and implications.*

Keywords: inheritance of the edge feature, the clausal typing hypothesis, the VMH effect, the VP-internal subject hypothesis, vacuous movement

1. Introduction

A series of works in the literature (George (1980), Chomsky (1986), Agbayani (2000, 2006), among others) has elaborated the vacuous movement hypothesis (henceforth, VMH), which states that a movement operation without an effect on PF output can be suspended (see also Abe and Hornstein (2010) and Mikami (2011)). The VMH as presented in the seminal work of Chomsky (1986) is formulated as follows:¹

* An earlier version of this paper was presented at the 29th Conference of the English Linguistic Society of Japan held at Niigata University in November 2011 and later published as Sakamoto (2012). I express my sincere gratitude to Yukio Hirose, Nobuhiro Kaga, Masaharu Shimada, Naoaki Wada, Kunio Nishiyama, and two anonymous EL reviewers for their invaluable comments and suggestions. I am also grateful to Tatsuhiro Okubo for his helpful discussions. Needless to say, any remaining errors and shortcomings are my own.

¹ This paper does not consider Agbayani’s (2000, 2006) version of the VMH. His analysis is based on the theory of feature movement, which I do not adopt in this paper. See Shimada (2008) for convincing counterarguments to Agbayani’s (2006) analysis.

© 2012 by the English Linguistic Society of Japan
(1) Vacuous movement is not obligatory at S-structure. This formulation indicates that vacuous movement can be delayed until LF. To put it differently, if the phonological effects of movement operations in overt syntax and in covert syntax are the same, movement preferably occurs in covert syntax. This manifests as a prominent difference between subject and object wh-constructions:

(2) a. Who saw John?
b. Who did John see?

In (2a), wh-movement to Spec-C is suspended at S-structure because covert wh-movement to Spec-C yields an equivalent effect on PF output. The situation differs in (2b), in which the wh-phrase moves from object position to Spec-C at S-structure, producing an effect on PF output. The VMH in (1) thus states that wh-subjects can remain in their original position at S-structure.

The plausibility of the VMH is bolstered by the observation that island effects weaken in wh-subject constructions:

(3) a. He is the man to whom I wonder [who knew [which book to give t_i]].
b. He is the man to whom I wonder [who John told [which book to give t_i]].

According to Chomsky (1986: 50), (3a) is more acceptable than (3b).² The former includes relativization from within the island in which the wh-phrase functions as a subject. In contrast, the latter involves relativization outside of the island in which the wh-phrase functions as an object. If wh-subjects can stay in situ at S-structure, as guaranteed by (1), then embedded Spec-C in (3a) would serve as an escape hatch for successive-cyclic movement. On the other hand, embedded Spec-C in (3b) is occupied by the object wh-phrase who. This offending element interferes with the movement of to whom. This is why the contrast in acceptability is observed between (3a) and (3b).³

That vacuous movement of wh-subjects to Spec-C is suspended at S-

² Because Chomsky (1986) mentions his acceptability judgments for (3) only in the text, I do not assign any judgment marks to these sentences.

³ Here, following Chomsky’s (1986) argument, I introduce the contrast in acceptability between (3a) and (3b) as deriving from one of the principles constituting core grammar, more specifically, the VMH. However, in this paper, the contrast in (3) is considered a manifestation of some sort of subject-object asymmetry and not directly related to core grammar (see section 2.2.2 for detailed discussion).
structure would mean that it applies at LF. The existence of this sort of LF-movement is independently evidenced by the ungrammaticality of (4).

(4) *How do you wonder [who fixed the car]? Chomsky (1986: 49) attributes the ungrammaticality of this sentence to an ECP (Empty Category Principle) violation. The LF-movement of who to embedded Spec-C, which satisfies semantic selection for wonder (cf. Grimshaw (1979)), eliminates the intermediate trace of how, and the proper government of t becomes impossible. The result is an ECP violation.

The VMH is reminiscent of the principle of Procrastinate, which regards covert operations as less costly than overt operations. If an operation need not be overt to meet some condition, the operation should apply at LF (cf. Chomsky (1995a: 69, 198)). This view is not conceptually preferable because it demands comparison between a derivation with an overt operation and a derivation with a covert operation (cf. Collins (1997)). In addition, within recent minimalist frameworks, movement operations reduce to Merge, which applies only at narrow syntax (Chomsky (2000, 2001, 2004, 2008), among others). This means that there occurs no LF-movement in derivation. These minimalist tenets exclude the VMH, which premises the existence of LF-movement. The VMH must be reconsidered in the context of structure-building computation at narrow syntax.

Given these considerations, a specific question automatically arises as to what derives the VMH effect, in which subjects refuse overt vacuous wh-movement but behave as if they reside in Spec-C at LF. There are three logical possibilities for vacuous movement at narrow syntax: (i) vacuous movement is not permitted at all; (ii) vacuous movement is optional, as stated in the VMH; or (iii) vacuous movement is regulated by some principle in a minimalist framework. This paper shows that (iii), from which the VMH effect follows, is the case.

This paper is organized as follows. Based on recent minimalist frameworks (Chomsky (2000, 2001, 2004, 2008)), section 2 proposes a computational system in which feature inheritance can be extended and applied to the edge feature, which draws on clausal typing (Cheng (1997a)), and shows that the VMH effect follows from this system. Section 3 discusses some consequences of the proposed system. Section 4 provides some concluding remarks.

2. Proposals: Explaining VMH Phenomena in a Minimalist Framework

Let me begin by introducing the analysis of subject wh-constructions in
Chomsky’s (2008) system that is to be adapted in this paper. Chomsky (2008) proposes that the phasal heads, C and v*, are responsible for all syntactic operations such as Merge and Agree. In other words, every operation applies in parallel at a phase level. This means that C and v* have both the Agree feature (AF), which is derivationally inherited from C to T and from v* to V, and the edge feature (EF), which enables a lexical item to be merged. The inherited AF yields A-movement, and simultaneously, the EF creates A′-movement. As an example of this derivation, let us consider (5).

(5) \[ CP \text{ Subj } C_{[E_F]} \ [TP \text{ Subj } T_{[A_F]} \ [v^*_P \text{ Obj } \ [v^*_P \text{ <Subj> } v^*_{[E_F]} \ [VP \text{ Obj } V_{[A_F]} \text{ <Obj>}] ]]]

(5) represents the schematic derivation of subject wh-constructions based on Chomsky’s (2008) system. At the phasal v*P, the object DP is raised to Spec-V by the inherited AF and to outer Spec-v* by the EF. Similarly, at the phasal CP, the subject DP is raised to Spec-T by the inherited AF and to Spec-C by the EF.

In the following discussions, by adopting some fundamental principles, such as feature inheritance, Agree, Merge, Transfer, and the Phase Impenetrability Condition (PIC), that are postulated within recent minimalist frameworks (Chomsky (2000, 2001, 2004, 2008)), I propose that feature inheritance also operates on the EF.4 Under this proposal, whether vacuous movement occurs or not is determined in a principled manner. More specifically, applicability of EF inheritance unambiguously determines the occurrence or nonoccurrence of vacuous movement. The VMH effect follows from this system. As a first step toward our main proposal, section 2.1 is devoted to clarifying the inherent functions of the EF that are related to the process of clausal typing (Cheng (1997a)).

---

4 The PIC is formulated as follows:

(i) In phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

This formulation dictates that the complement of a phase, undergoing a Transfer operation to the interfaces, be invisible to any syntactic operations involving Agree and Merge. The PIC has often been adopted in the literature, but the point at which Transfer occurs remains controversial. This paper adopts weak interpretation of the PIC (Chomsky (2001)), under which it is not until the higher phase is created that the lower phase transfers its complement into the interfaces. The weak version of the PIC is confirmed by quirky subject constructions in Icelandic (cf. Taraldsen (1995), Sigurðsson (1996), Nomura (2005), etc.), in which T is allowed to establish an agreement relation with nominative subjects in the object position by crossing the transitive verbal domain. See Richards (2011) for a detailed outline of issues surrounding the PIC.
2.1. Exploring the Inherent Functions of the EF

As mentioned above, Chomsky (2008) argues that C and v*, which are phasal heads, have the EF and that T and V derivationally receive the AF from C and v*, respectively. It seems uncontroversial for the AF to consist of tense-features and φ-features, which are valued by Agree (see Chomsky (2000, 2001)). To point to the importance of EF inheritance proposed in section 2.2, we need to gain a deeper understanding of the intrinsic functions of the EF. Chomsky (2008: 139) describes the EF as a feature that permits a lexical item to be merged, which explains unbounded iteration of Merge. This description falls short, however. Although there seems little doubt that lexical items as input to Merge in the Lexicon all bear the EF (cf. Chomsky (2008: 139)), there is a certain difference between the EF of phasal heads and the EF of nonphasal heads.5 To highlight this difference, I put forward the hypothesis that only the EF of the phasal heads serves as a target for feature valuation at narrow syntax.6 Under this hypothesis, we need to demonstrate the way in which such valuation is implemented.

Let us first discuss the EF of C. C is at the edge of a clause, a terminal position for Merge. In this position, the clause type is determined. It is then reasonable to suppose that the EF of C is relevant to determining the clause type. To crystallize this idea, let us consider Cheng’s (1997a) clausal typing hypothesis, which is proposed on numerous typological grounds:

(6) Every clause needs to be typed. In the case of typing a wh-question, either a wh-particle in C0 is used or else fronting of a wh-word to the Spec of C0 is used, thereby typing a clause through C0 by Spec-head agreement.

According to Cheng, clausal typing is a property that is met at S-structure, unlike semantic selection, which is satisfied at LF (see Cheng (1997a: sec-

5 See Narita (2011: chapter 3) for a variety of consequences deriving from the hypothesis that the EF is a property of lexical items.
6 It is worth noting here that I do not intend the introduction of two types of the EF and that the EF serves consistently as a feature that prompts structure extension. When structure extension proceeds to a certain extent, that is, to the level of phasal formation, it should be taken for granted that the computational system has a phase identify its semantic property, to the extent that the phasal domain as a derivational unit is transferred into, and interpreted at, the LF interface, keeping to Full Interpretation (see also the discussions below in the text). Given that structure extension is executed by Merge and that Merge is motivated by the EF, it is then natural to consider that such a phasal semantic property is identified by valuing the EF in conjunction with Merge. That only the EF of the phasal heads behaves as a target for feature valuation at narrow syntax is merely a natural consequence of the phase-based model.
FEATURE INHERITANCE AND VACUOUS MOVEMENT

...tion 2.2)). Associating our present idea with clausal typing, we then find that the EF of C is valued by means of the merger of a given element to the edge. What is guided by this line is that the EF of C is transparent or underspecified at the initial state with respect to its value. If the element with an interpretable wh-feature values the EF of C through its merger to the edge, the relevant clause is typed as [+wh] at narrow syntax. Otherwise, the clause is typed as [−wh] at narrow syntax. The typed clause will be interpreted as interrogative or declarative at the LF interface (see section 3.2, in which clausal typing works on semantic selection for verbs).

Naturally, a question arises as to what types of features constitute a natural class of possible feature valuation. This paper also argues that the EF of C can be assigned some value contributing to topic interpretation at LF, which I designate as [±topic] for simplicity. Phasal CP works as a topic sentence if a topical element values the EF of C via Merge. In any case, what sorts of features value the EF of C is an empirical matter.

Let us turn to the EF of v*. v*, unlike C, is not at the edge of a clause, so the EF of v* does not have to do with clausal typing. What then does the EF of v* assume? We can speculate that the EF of v* is, so to speak, associated with ‘predicate typing.’ It has thus far been proposed that the EF of C determines its value via wh-hood or topichood with nominal elements. In other words, the EF of C makes reference to the “discourse-related properties” (Chomsky (2008: 140)) of nominal elements. Antithetically, it makes no sense at all that the EF of v* refers to the very same...
properties because the verbal domain reflects argument structure. It is thus natural to reason that the EF of $v^*$ discerns θ-roles (or θ-features) of nominal elements. According to this reasoning, the EF of $v^*$ is valued via the information of argument structure at narrow syntax, and this EF value materializes as predicates such as transitives, unaccusatives, and unergatives and contributes to interpretation at LF. Technical details of ‘predicate typing’ are not obvious immediately. Further exploration of them is entrusted to future research.

To wrap up this subsection, it has been proposed, adopting the insights in Chomsky (2008) and Cheng (1997a), that the EF of the phasal heads, unlike the EF of nonphasal heads, serves as a target for feature valuation at narrow syntax and that such valuation determines phasal interpretation at LF. Based on this characterization of the EF, section 2.2 demonstrates that EF inheritance occurs from phasal heads to nonphasal heads at narrow syntax, showing that the VP-internal subject hypothesis successfully fits into, and the VMH effect immediately follows from, the system of EF inheritance.

2.2. EF Inheritance from Phasal Heads to Nonphasal Heads

In the preceding subsection, we have inspected inherent functions of the EF of phasal heads rather than nonphasal heads and shown that it is valued by the merger of a given element to the relevant edge. Along with the Strong Minimalist Thesis (SMT) (cf. Chomsky (2000, 2001, 2004, 2008)), which holds that the faculty of language is a perfect computational system that contains nothing more than what makes efficient the linkage between the computational system and the performance systems (i.e. the conceptual-intentional system and the sensorimotor system), I propose that this EF property requires both the AF and the EF to undergo inheritance from phasal heads to nonphasal heads:10, 11

9 In the present argument, EF valuation in the verbal domain warrants (non)transitivity of verbs. That is, the initial state does not distinguish between transitive $v^*$ and nontransitive $v$. These are purely notational variants, and this distinction means nothing more than it.

10 Chomsky (2008: 157) suggests the possibility that the EF of C is inherited to T and tries to relate it to the EPP feature of T. Pursuing this possibility, Kitada (2011) explores some empirical consequences.

11 It should be noted here that the present view in this paper is apparently incompatible with the one in Chomsky (2008), according to which it is essential in deriving the A/A’ distinction to apply only AF inheritance to the syntactic computation, in conformance with
(7) \[
\begin{array}{c}
{\text{CP C [TP Subj T}_{[\text{AF][EF}]}} [v^* <\text{Subj}> v^* [\text{VP Obj V}_{[\text{AF][EF}] <\text{Obj}>}]])]
\end{array}
\]

(7) is the schematic derivation of subject \textit{wh}-constructions that involve simultaneous inheritance of the AF and the EF from C to T and from \(v^*\) to V. In this case, unlike in (5), the subject DP and the object DP move only to Spec-T and only to Spec-V, respectively. Utilizing inheritance of the EF is tantamount to avoiding a superfluous derivational step, namely, a redundant merger to the edge of C and \(v^*\).\textsuperscript{12} To put it another way, EF inheritance enables feature valuation to be implemented in one derivational position. For that reason, inheritance of the EF is computationally optimal and should apply to all cases by default. In the following two subsections, I corroborate the plausibility of our present claim.

2.2.1. Incorporating the VP-internal Subject Hypothesis into the Computational System

The current view that inheritance of the EF is computationally optimal poses a further inevitable question of how far inheritance of the EF can extend. If inheritance of the EF has computational efficiency, it should iterate without restriction. Taking into consideration the VP-internal subject hypothesis (VPISH) (Fukui and Speas (1986), Kitagawa (1986), Kuroda (1988), among others), however, repetitive inheritance of this sort turns out to be unfavorable, since subject DPs are generally assumed to move from Spec-\(v^*\) to Spec-T for the purpose of feature valuation.

Suppose now that \(v^*\) receives the EF via T from C. Then, there is no EF on T that raises subject DPs from their original position to Spec-T. This situation conflicts with the derivation that is to be envisaged under the VPISH, in which no repetitive inheritance occurs and subject DPs move out of their original position to value the EF inherited by T. Consideration

the SMT. However, that AF inheritance and EF inheritance apply simultaneously does not imply straightforwardly that the A/A’ distinction disappears. In effect, Kitada (2011) demonstrates that both A-properties and A’-properties emerge as a result of simultaneous application of AF inheritance and EF inheritance. Under this view, A/A’-properties derive not from syntactic positions but from syntactic features. The idea that the A/A’-distinction is made by syntactic features is conceptually desirable in accordance with the bare phrase structure theory (cf. Chomsky (1995a: 249, 1995b)), in which there is no room to postulate syntactic positions. Our present view thus complies with the SMT, although we leave aside closer scrutiny of A/A’-properties with subject \textit{wh}-constructions for now. I am indebted to one of the reviewers for bringing my attention to the significance of discussing the issue of the A/A’ distinction.

\textsuperscript{12} See Shimada (2008: 535) for a similar but somewhat different consideration of economy.
of the intrinsic natures of the EF and $v^*$ provides us with a convincing argument that EF inheritance may not be inherited without limit. The EF demands that something be merged into the edge, and $v^*$ does not require any additional element at the edge because $v^*$ inherently undertakes an external argument. It is concluded from the interplay between their internal characteristics that $v^*$ cannot receive the EF from the higher phasal head.\footnote{Richards (2007) confirms on theoretical grounds that $v^*$ cannot receive the AF from the higher phasal head.}

The discussion thus far illuminates a novel potential challenge for the VMH. Recall that vacuous movement can be suspended at structure-building computation under the VMH. This is equivalent to stating that subject DPs need not undergo movement from Spec-$v^*$ to Spec-T at narrow syntax, which yields no effect on PF output. Movement of this sort, however, is obligatory due to feature valuation. Suspending Spec-$v^*$-to-Spec-T movement of subject DPs at narrow syntax is never permitted under the VPISH. Thus, the VMH is at odds with the VPISH. Our system is more adequate in that it is free of this kind of problem and is able to incorporate the VPISH.\footnote{One reviewer points out, based on Takahashi’s (1994) arguments that subject DPs can stay in situ, that the argument here does not fall into place. However, an aspect of the matter under consideration lies in the point that, when subject DPs move from Spec-$v^*$ to Spec-T under the VPISH, such movement is string-vacuous. This type of movement follows unproblematically from our system, but not from the VMH. The account of the phenomenon noted by the reviewer might be ascribed to the copy theory of movement (Chomsky (1995a)), which allows the lower copy of a moved element to be interpreted and pronounced at the interfaces (cf. Nunes (2004), Mikami (2010), etc.). This paper yields pursuit of this possibility to another occasion.}

In summary, it is irrelevant to our system whether movement operations have an effect on PF output. Our system allows of movement with no effect on PF output. Differently from principles such as the VMH, occurrence of movement crucially depends on applicability of EF inheritance. The EF of C can be inherited by T but not by $v^*$. Subject DPs are required to be raised vacuously from Spec-$v^*$ to Spec-T based on need to value the EF inherited by T. In the next subsection, I show that the computational system postulating EF inheritance from C to T is also capable of capturing the VMH effect, in which subjects refuse overt vacuous wh-movement but behave as if they reside in Spec-C at LF (see section 1).
2.2.2. Deriving the VMH Effect

Let us start by considering how the proposed system derives subject *wh*-constructions:\textsuperscript{15}

(8)  
  a. Who saw John? \( (= (2a) \)  
  b. \[CP C [TP Who T_{[AF][EF]} [v^{*P} <who> v^* [VP John saw_{[AF][EF]} <John>]]] \]

In (8b), the derivation for (8a), the AF and the EF are both inherited. At the phasal \( v^*P \), the inherited AF establishes a \( \varphi \)-feature agreement relation with John in the object position, and the inherited EF is valued by the merger of John to Spec-V. At the subsequent stage, the phasal CP, the inherited AF enters into a \( \varphi \)-feature agreement relation with the subject *wh*-phrase, which originates in Spec-\( v^* \), and the inherited EF is valued as \([+wh]\) via the merger of who containing an interpretable *wh*-feature to Spec-T. Because subject *wh*-phrases inherently have both subjectionhood and *wh*-hood, this derivation converges at the interfaces without any discrepancy.

It is important to note here that the derivation given here properly derives the VMH effect. In (8b), the EF inheritance from C to T prevents the subject *wh*-phrase from undergoing overt vacuous movement to Spec-C. Furthermore, the movement of this element to Spec-T values the inherited EF as \([+wh]\), and the valued EF will be interpreted as interrogative at LF. This produces the same effect that covert vacuous *wh*-movement to Spec-C does. Consequently, EF inheritance from C to T successfully derives the VMH effect.

Let us now turn to object *wh*-constructions:

(9)  
  a. Who did John see? \( (= (2b) \)  
  b. *[CP C [TP John T_{[AF][EF]} [v^{*P} <John> v^* [VP who see_{[AF][EF]} <who>]]]]
  c. \[CP Who C_{[EF]} [TP John T_{[AF]} [v^{*P} <John> v^* [VP <who> see_{[AF][EF]} <who>]]]]\]

Given that inheritance of the EF is computationally optimal, as stated above, the object *wh*-construction in (9a) should be derived in a similar manner. However, if EF inheritance applies to the derivation of object *wh*-constructions, then it creates an illicit structure, as illustrated in (9b), in which John, rather than who, values the EF inherited at the phasal CP. This structure is ill-formed as a result of the violation of the requirement of Full

\textsuperscript{15} For the sake of convenience, this paper continues to put aside the derivational position of (auxiliary) verbs.
Interpretation (FI). FI states that every constituent of the interfaces contributes to interpretation (cf. Chomsky (1995a), Narita (2011), etc.), whereby interpretable formal features are required to value uninterpretable formal features. In the case of (9b), *who*, though bearing an interpretable *wh*-feature, does not value any uninterpretable formal feature. In structure (9b), *John* is base-generated in Spec-*v* and is configurationally higher than *who*, which is raised from object position to Spec-*V* by the EF inherited at the phasal *v*P. As a result, *John* is the only object that can value the EF inherited at the phasal CP. This makes (9b) illicit, yielding an FI violation.\(^{16}\)

The illicitness of (9b) leads to the consideration that (9a) is generated through the derivations in (9c), where the EF of C is not inherited by T. In this case, the merger of *John* to Spec-T and the merger of *who* to Spec-C apply in a parallel fashion, obeying cyclicity under the weak version of the PIC, according to which completion of the higher phase instructs the lower phase to transfer its complement into the interfaces (see also note 4). The merger of *who* values the EF of C as \([-wh]\), and the valued EF is interpreted as interrogative at LF. Structure (9c) thus obtains a licit status, observing the requirement of FI.

It should be noted here that the derivations of subject *wh*-constructions and object *wh*-constructions proposed here predict that (3a) as well as (3b) will exhibit *wh*-island effects:

\[(10) \text{ a. } \text{He is the man to whom I wonder } [\text{CP C [TP who T[AF][EF]} [\text{*vP <who> knew which book to give <to whom>]}]]\]
\[(10) \text{ b. } \text{He is the man to whom I wonder } [\text{CP who C[EF]} [\text{TP John T[AF]} [\text{v*P <John> told <who> which book to give <to whom>}}]]\]

As shown in (10a), the EF inheritance from C to T occurs in the subject *wh*-construction as the complement clause of *wonder*. The EF of C makes C meaningful in the sense that its value determines the status of phasal CP (see section 2.1). Thus, EF inheritance from C to T amounts to stating that C itself is inherited to T. In other words, EF inheritance amalgamates C with T. This amalgamation renders the complement clause of *wonder*...
in (10a) a *wh*-island, in parallel with the situation in (10b), which has the island headed by the indirect object *wh*-phrase in the complement clause of *wonder*. The structures in (10) both involve relativization from their complement clauses that function as *wh*-islands, hence the deviance of (3).

This view on the difference between (3a) and (3b) is different from Chomsky’s. Recall that Chomsky (1986) attributes the subtle contrast in acceptability between (3a) and (3b) to the VMH as one of the principles constituting core grammar (see section 1). Our system, however, equally rules out both sentences for the reason stated above and counts the subtle contrast in (3) just mentioned as a manifestation of some sort of subject-object asymmetry. In fact, some previous studies (Cheng (1997a), Richards (2001), Agbayani (2006), etc.) consider there to be no grammatical difference between extraction from the island where the *wh*-phrase functions as a subject (cf. (3a)) and extraction from the island where the *wh*-phrase functions as an object (cf. (3b)). Our present conclusion conforms to this perspective.17

Summarizing this section, we have proposed a computational system that involves inheritance of the EF, which draws on clausal typing (Cheng (1997a)), and considered the derivations of subject *wh*-constructions and object *wh*-constructions.18 We have also established that this system incor-

17 The following grammatical contrast provided by Chomsky (1986: 48–49) should also be treated in our system:

(i) a. What, do you wonder [who saw *t*]?  
   b. *How, do you wonder [who fixed the car *t*]? (=4)  
Sentence (ia), which is parallel to (3a) in that it contains extraction from within the island headed by a subject *wh*-phrase, displays a weaker *wh*-island violation than (ib). Within a preminimalist framework, the contrast in (i) is ascribed to the ECP. In (ib), the trace of *how* is not properly governed for the reason already stated in section 1. The trace in (ia), on the other hand, is properly governed because it is an argument of the embedded verb.

However, we can no longer utilize the ECP as a tool for explaining the contrast in (i) because it does not get along with minimalist tenets. It is expected from the present system that sentences (ia, b) are equivalently excluded as a violation of minimality, as with the situation in (3), and that the contrast in acceptability between them should be attributed to a difference in referentiality between the relevant *wh*-phrases, no matter what general principle the ECP reduces to in a minimalist framework. Because this paper does not have enough space to discuss the reducibility of the ECP, I leave a full account of the contrast in (i) for future research.

18 In the case of object *wh*-constructions, EF inheritance from C to T by default results in an FI violation (cf. (9b)), so the syntactic computation is forced to select the derivation without this inheritance (cf. (9c)). As one reviewer notes, this state of affairs seems contain a serious look-ahead property insofar as the convergence at the interfaces triggers or blocks the application of a syntactic operation (i.e. EF inheritance) that is to be optimal in
incorporates the VPISH with no problem and explained why subject *wh*-constructions display the VMH effect. In section 3, we discuss some consequences of the system proposed here.

3. Consequences

Under the system proposed in section 2, EF inheritance from C to T raises subject *wh*-phrases to Spec-T at narrow syntax and causes them to remain in the same position at LF. In contrast, the VMH does not prohibit subject *wh*-phrases from moving to Spec-C at LF. None of the cases that we will consider in this section is easy to capture under any system that allows for LF-movement to Spec-C, which is in conflict with minimalist tenets. These cases are derived naturally from the system proposed here.

3.1. Superiority Effects

The system proposed here is based on the premise that EF inheritance is computationally optimal and applies to all cases by default. Under this system, when EF inheritance produces a well-formed derivation, other derivations are not allowed, even if such possible derivations are well-formed. The existence of superiority effects confirms this (cf. Chomsky (1973, 1981, 1995a, 2008), Lasnik and Saito (1992), etc.):

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

line with the SMT. However, we can provide object *wh*-constructions with the derivation that involves EF inheritance from C to T but dispenses with this problem of globality:

(i) [CP Who C [TP John T [AF][EF] [νp <John> ν* [νp <who> se<AF>>]]]]

As stated above in the text, EF inheritance amalgamates C with T. C and T can jointly license ‘multiple Specs’ after this amalgamation, given that they each have the EF intrinsically. In (i), this C-T relation licenses the simultaneous mergers of John and who. These mergers meet the requirements of both the AF and the EF inherited from C to T. In this case, who, which has an interpretable *wh*-feature, properly values the EF as [+wh], thus circumventing an FI violation. To the extent that this derivation is on the right track, the look-ahead property noted by the reviewer does not occur.

Note incidentally that (8b), which is the derivation for subject *wh*-constructions, cannot license ‘multiple Specs.’ In (8b), the requirements of the features inherited from C to T are fulfilled solely by the merger of who since the material is a subject DP that concurrently serves as subjecthood and *wh*-hood. In any case, we can elude the potential look-ahead property that object *wh*-constructions have along the lines of the present statement. However, the rest of this paper does not mention a derivation of the sort given in (i) because consideration of such a derivation is beyond the scope of this paper (we can also assign embedded object topicalization in (17a) a derivation similar to (i)).
The grammatical asymmetry between these sentences illustrates that in multiple *wh*-constructions, object DPs, crossing subject DPs, cannot occupy the sentence-initial position. The observed effect follows from our system:

(12) a. \[CP \text{C [TP Who \text{T[AF][EF]} [v^*P <who> v^* [VP what saw[AF][EF] <what>]]]}\]

b. \[CP \text{What C[EF] [TP who \text{T[AF]} [v^*P <who> v^* [VP <what> see[AF][EF] <what>]]]}\]

The derivation with EF inheritance, given in (12a), generates the sentence in (11a). At the phasal v^*P, the object DP establishes a \(\phi\)-feature agreement relation with the inherited AF and values the inherited EF via Merge. At the phasal CP, the subject DP enters into a \(\phi\)-feature agreement relation with the inherited AF and values the inherited EF as \(+wh\) via Merge. This derivation is impeccable with no FI violation.

The convergent derivation in (12a) leaves no room for any other derivations. Consider the derivation in (12b) without EF inheritance from C to T. This derivation might observe the requirement of FI, with the merger of *who* to Spec-T and the merger of *what* to Spec-C applying in parallel. Nevertheless, the sentence generated from this derivation, as shown in (11b), is not acceptable.\(^{19}\) The grammatical asymmetry in (11) cannot be easily explained without our system that regards EF inheritance as computationally optimal.\(^{20}\)

### 3.2. Semantic Selection

Furukawa and Fukuda (2009) point out a paradigm of semantic selection for verbs that challenges any analysis based on LF-movement:

(13) I wonder [who saw what]. (Chomsky (1986: 52))

(14) a. *Who wondered [John saw what]?

b. Who wondered [what John saw]?

(Furukawa and Fukuda (2009: 271))

The verb *wonder* requires an interrogative complement at the level of se-

\(^{19}\) Compare this result with the case of single object *wh*-constructions (see (9)), in which our system admits the derivation without EF inheritance from C to T since the derivation with EF inheritance from C to T crashes yielding an FI violation. Nevertheless, both cases are easily compatible with our system.

\(^{20}\) Consideration of multiple *wh*-constructions in German, where no superiority effect is observed, leads Chomsky (2008) to claim that English is allowed to have a derivation for generating the sentence in (11b). However, I believe that English cannot have such a derivation for the reason discussed here. The origin of the lack of superiority effects in German is unclear, so I leave this issue open for future work.
mantic representation (see Grimshaw (1979)). If *wonder* in (13) could satisfy semantic selection via LF-movement of *who* to embedded Spec-C, the grammatical contrast in (14) would remain a mystery. Specifically, it is unclear why sentence (14a) fails to obtain a grammatical status, with *what* moving to embedded Spec-C at LF.

Our system gives a principled explanation for the paradigm:

\begin{align*}
(15) & \quad \text{I wonder} [\text{CP C} [\text{TP who} T_{[\text{AF}]\text{[EF]}]} \left[\text{v}^{*\text{p}} \text{<who>} \text{v}^{*} \left[\text{VP what saw}_{[\text{AF}]\text{[EF]}} \text{<what>}\right]\right]]
\end{align*}

\begin{align*}
(16) & \quad \text{a. *Who wondered} [\text{CP C} [\text{TP John} T_{[\text{AF}]\text{[EF]}]} \left[\text{v}^{*\text{p}} \text{<John>} \text{v}^{*} \left[\text{VP what saw}_{[\text{AF}]\text{[EF]}} \text{<what>}\right]\right]]
\end{align*}

\begin{align*}
& \quad \text{b. Who wondered} [\text{CP what} C_{[\text{EF}]} [\text{TP John} T_{[\text{AF}]} \left[\text{v}^{*\text{p}} \text{<John>} \text{v}^{*} \left[\text{VP what saw}_{[\text{AF}]\text{[EF]}} \text{<what>}\right]\right]]]
\end{align*}

The embedded CP of (15) corresponds to subject *wh*-constructions. The derivation at the phasal CP is crucial here. At the phasal CP, the inherited AF establishes a φ-feature agreement relation with the subject *wh*-phrase, which originates in Spec-*v*-, and the inherited EF is valued as [+wh] by the merger of *who* involving an interpretable *wh*-feature to Spec-T. The valued EF will be interpreted as interrogative at LF. Consequently, by taking its proper complement, *wonder* in (15) fulfills semantic selection.\(^{21}\)

This account of semantic selection readily captures the difference in grammaticality between (16a) and (16b). In (16a), which accords with the illicit derivation of object *wh*-constructions (see (9b)), *John*, rather than *what*, occupies embedded Spec-T. Because the element *John* has no interpretable *wh*-feature, the EF inherited at the phasal CP is valued as [−wh] and interpreted as declarative at LF. As a result, *wonder* in (16a) fails to take its proper complement, thus yielding an FI violation.

By contrast, *wonder* in (16b), which is parallel to the licit derivation of object *wh*-constructions (see (9c)), can take its proper complement. In (16b), *what*, rather than *John*, values the EF of C as [+wh] via Merge. The valued EF is interpreted as interrogative at LF, thereby satisfying semantic selection for *wonder* and the requirement of FI.

Under the VMH, *wonder* in (13) needs to fulfill semantic selection via

\[^{21}\] Semantic selection is generally assumed to obtain on the basis of sister relation. Given this assumption, one reviewer wonders how our system establishes this relation in (15). As discussed in section 2.2.2, EF inheritance amalgamates C with T, and thus, there is no semantically and phonologically null C after EF inheritance from C to T. Consequently, the matrix verb in (15) properly enters into sister relation with the embedded interrogative clause, thereby satisfying semantic selection.
LF-movement of who to embedded Spec-C because the overt counterpart at narrow syntax is suspended. If this is the case, the grammatical contrast in (14) would not fall into place for the reason mentioned above. In this way, the discussion here has further ensured the validity of our computational system.

3.3. Anaphor Binding in Embedded Topicalization

Next, we consider the grammatical asymmetry of anaphor binding in embedded topicalization between (17a) and (17b), which is given by Lasnik and Saito (1992) and later discussed by Agbayani (2000, 2006) in the context of vacuous movement.

(17) a. John\textsubscript{i} thinks that himself\textsubscript{i}, Mary likes \textit{t}\textsubscript{i}.
    b. *John\textsubscript{i} thinks that himself\textsubscript{i}, \textit{t}\textsubscript{i} likes Mary.

In (17a), the anaphor himself, which is topicalized from object position in the embedded clause, can take the matrix subject John as its antecedent. On the other hand, in (17b), which involves potential embedded subject topicalization corresponding to embedded object topicalization, there is no anaphor binding relation between John and himself.\footnote{Sentence (i) from Lasnik and Saito (1992: 198), though somewhat marginal, illustrates that embedded anaphoric subjects can undergo topicalization in principle. (i) ??John, thinks that himself, Mary said \textit{t} won the race. In this sentence, the embedded anaphoric subject, occupying the topicalized position, is permitted to be bound by its matrix antecedent. Thus, the ungrammaticality of (17b) does not mean that its embedded anaphoric subject cannot be subject to topicalization.}

Here, we adopt the hypothesis that Condition A is an “anywhere” condition (Belletti and Rizzi (1988)), under which any level of representations (i.e. D-structure and S-structure) can meet binding requirements for anaphoric elements. If this hypothesis is reinterpreted within the phase-based system postulated in this paper, it then follows that anaphoric elements can be licensed phase by phase at narrow syntax.

Let us now consider how the proposed system explains the grammatical asymmetry of anaphor binding in (17):\footnote{For the purpose of exposition, I assign the embedded clauses of (17) complex CP structures proposed by Rizzi (1997). Such structures, if any, should not be represented at narrow syntax because they involve pragmatic information (see also Chomsky (2008)).}
As depicted in (18), the derivations for (17), the shaded domain represents where Transfer applies, namely, the complement of phases. According to the weak version of the PIC, which this paper has posited, completion of the higher phase instructs the lower phase to transfer its complement into the interfaces (see note 4). At the embedded CP of (18a), in which there is no EF inheritance from Top to T in parallel with the licit derivation of object \textit{wh}-constructions (cf. (9c)), \textit{Mary} moves to embedded Spec-T, and simultaneously, \textit{himself} moves to embedded Spec-Top.\footnote{Our system is forced to have the illicit derivation of embedded object topicalization in (i) by default.}


The illicitness of this derivation, which involves EF inheritance from Top to T, is ascribed not simply to a Condition A violation but also essentially to an FI violation, in parallel with the situation surrounding object \textit{wh}-constructions in (9) (see note 18 for the possibility that this sort of derivation can be licit forming ‘multiple Specs’ in its embedded clause). In plain words, this derivation is illicit because it is not associated with topic interpretation. In (i), \textit{himself} is taken to have some kind of interpretable formal feature relevant to topic interpretation at LF, but this feature fails to value the EF inherited at the embedded CP. This failure produces an FI violation, thereby rendering (i) illicit. The computational system thus initiates structural assembling without EF inheritance from Top to T, as depicted in (18a).

One might, however, disagree with the claim that a topicalized element has some kind of interpretable formal feature contributing to topic interpretation at LF for valuing the EF inherited at the embedded CP. This claim is equivalent to stating that topicalization is responsible for determining the clause type. I believe that this statement is not necessarily incorrect because in-situ languages such as Japanese have topic markers. It is natural to consider that in-situ languages such as Japanese value the EF at phasal CP with topic markers and that movement languages such as English value the EF at phasal CP via internal Merge of topical phrases. This consideration leads to the conclusion that topicalization is relevant to identifying the clause type, and a topicalized phrase has some kind of interpretable formal feature contributing to topic interpretation at LF for valuing the EF at phasal CP.
cause the anaphoric object *himself*, occupying the edge of TopP, evades this transferred domain, the element is allowed to be bound by its antecedent in matrix Spec-v*. Thus, in (17a), the coreferential reading obtains between *John* and *himself*.

In contrast, sentence (17b) can establish no such coreferential reading. This is because *himself* in (18b) cannot evade the transferred domain in question. The embedded CP of (18b) is parallel to the derivation of subject *wh*-constructions (cf. (8b)), so EF inheritance from Top to T does occur. This inheritance makes *himself* stay in the lower TP complement. Completion of the higher v*P instructs the lower TopP to transfer its TP complement containing *himself* into the interfaces, hence no anaphor binding relation is established between *John* and *himself*.

Under the system that posits the VMH, it is not easy to capture the asymmetric behavior of anaphor binding because the VMH does not prevent the anaphoric subject in (18b) from moving to Spec-Top at LF. If *himself* in (18b) is raised to Spec-Top at LF, a Condition A violation might be circumvented. This situation does not actually occur, however. Again, it has been confirmed that our computational system is more plausible.

3.4. Across-the-Board Movement Phenomena

Our ensuing consideration turns to across-the-board (ATB) movement phenomena of *wh*-phrases. It has been argued in the literature that ATB movement is feasible under structural parallelism between conjuncts (cf. Ross (1967), George (1980), Goodall (1987), Bošković and Franks (2000), etc.). To see how ATB movement phenomena are derived, I present the following example from George (1980: section 5.3):

(19) They removed the prisoner, who(m) the judge has sentenced and *(who(m)) the warden will execute.

In (19), each conjunct has a *wh*-phrase that undergoes movement to the parallel position, Spec-C. In this case, the *wh*-phrase in the second conjunct can be elided. Such situations are known as ATB movement phenomena.

ATB movement phenomena provide the proposed system with further plausibility. Again, the relevant data are cited from George (1980: section 5.3):

(20) a. They removed the prisoner, who has lost his appeal and *(who(m)) the warden will execute.

b. They removed the prisoner, who(m) the judge has sentenced and *(who) will now appeal.

Sentence (20a) includes the *wh*-subject in the first conjunct and the *wh*
object in the second conjunct. In (20b), the situation is the other way around. As our system predicts, these sentences do not license ATB movement:

(21) a. They removed the prisoner, \([\text{CP } \text{who} \text{ T}_{\text{AF}} [\text{EF} [\text{TP} \text{who(m)} \text{ v* } \text{VP} \text{who(m)} \text{ will execute}_{\text{AF}} [\text{EF} \text{who(m)}]]]]\) and \([\text{CP } \text{who(m)} \text{T}_{\text{AF}} [\text{EF} [\text{TP} \text{the warden} \text{ v* } \text{VP} \text{who(m)} \text{ will execute}_{\text{AF}} [\text{EF} \text{who(m)}]]]]\)

b. They removed the prisoner, \([\text{CP } \text{who(m)} \text{C}_{\text{EF}} [\text{TP} \text{the judge} \text{ v* } \text{VP} \text{who(m)} \text{ has sentenced}_{\text{AF}} [\text{EF} \text{who(m)}]]]]\) and \([\text{CP } \text{C TP } \text{who(m)} \text{T}_{\text{AF}} [\text{EF} [\text{TP} \text{who(m)} \text{ will now appeal}_{\text{AF}} [\text{EF} \text{who(m)}]]]]\)

As depicted in (21a), the wh-phrase in the first conjunct is raised to Spec-T by the EF inherited at the phasal CP; the wh-phrase in the second conjunct is raised to Spec-C by the EF of C. As just mentioned, the situation in (21b) is reversed. In either case, each conjunct has a wh-phrase that occupies a different position. Since ATB movement occurs under structural parallelism between conjuncts, these derivations fail to undergo ATB movement.

Recall again that the VMH permits subject wh-phrases to move to Spec-C at LF. If subject wh-phrases may be raised to Spec-C at LF, each conjunct in (21a) and (21b) would have a parallel structure in which the wh-phrase in each conjunct occupies Spec-C. This situation incorrectly predicts that the derivations in (21a) and (21b) are allowed to undergo ATB movement. However, our system does not have any difficulty in predicting applicability of ATB movement.

3.5. Parasitic Gaps

Finally, I show that the proposed system is also applicable in explaining parasitic gap phenomena. Before proceeding, let me clarify what parasitic gap phenomena are and how previous works have explained them:

(22) a. Which articles did you file \(t\) without reading \(pg\)?

(Engdahl (1983: 5))

b. \([\text{CP Which articles did you file } t [\text{CP Op} \text{ without PRO reading } pg]]\)

(23) a. *John was killed \(t\) by a tree falling on \(pg\).

(Engdahl (1983: 13))

b. *[\text{TP John was killed } t [\text{CP Op} \text{ by a tree falling on } pg]]

Example (22a), a typical parasitic gap sentence, has often been assumed to be derived as in (22b). The gap signified by \(t\) is referred to as a real
gap, which is created by the movement of *which articles*, whereas the gap represented as *pg* is known as a parasitic gap (see Engdahl (1983)). A parasitic gap is generally analyzed as a trace of null operator movement and interpreted as referentially identical to the operator of a real gap by the algorithm of chain composition (see Chomsky (1986)).

It has been proposed in the literature that chain composition properly occurs under the parallelism in chain types between a real gap chain and a parasitic gap chain (cf. Kim and Lyle (1996), Karimi (1999), Sakamoto (2011), etc.). In (22b), the real gap chain is established by the movement of *which articles* to Spec-C in the root clause; the parasitic gap chain is formed by the null operator movement to Spec-C in the adjunct clause. The parallelism in chain types appropriately drives chain composition, thereby rendering the structure in (22b) interpretable. This licenses the parasitic gap in (22a).

In contrast, example (23a) does not license the parasitic gap. As shown in (23b), the structure for (23a), there is no parallelism in chain types between the real gap chain and the parasitic gap chain. The real gap chain is created by the movement of *John* to Spec-T in the root clause, but the parasitic gap chain is produced by the movement of *Op* to Spec-C in the adjunct clause. The asymmetry in chain types does not motivate chain composition, hence the uninterpretability of (23b). This does not tolerate the occurrence of the parasitic gap in (23a).

Our system explains the (in)consistency of parasitic gaps with subject *wh*-constructions, as observed in (24), in tandem with the analysis just introduced.

---

25 Chain composition is formulated as follows:

(i) If \( C = (\alpha_1, \ldots, \alpha_n) \) is the chain of the real gap, and \( C' = (\beta_1, \ldots, \beta_n) \) is the chain of the parasitic gap, then the “composed chain” \( (C, C') = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) \) is the chain associated with the parasitic gap construction and yields its interpretation.

To see how (i) serves as an interpretational algorithm for parasitic gaps, let us take (22b) as an instance. In (22b), the real gap chain is \( C = (\text{*which articles*}, t) \), and the parasitic gap chain is \( C' = (\text{Op}, \text{pg}) \). As a result of chain composition, we obtain \( (C, C') = (\text{*which articles*}, t, \text{Op}, \text{pg}) \). The formed chain has *which articles* as its head and *pg* as its tail, whereby the parasitic gap is interpreted as referentially identical to the operator of the real gap.

Here, I do not pursue reducibility of chain composition to some other general principle because this issue is beyond the scope of this paper. For attempts to develop Chomsky’s (1986) original insight in light of minimalism, see Nunes (2004) and Kasai (2010), among others.
As illustrated in (24), when real gaps correspond to subject wh-phrases, in some cases, parasitic gaps are refused, and, in other cases, parasitic gaps are licensed. Our system assigns these cases the following structures:

\[
\begin{align*}
(25) \quad a. \quad & \text{a man } [_{CP} C [_{TP} who T_{[AF][EF]} [_{vP} <\text{who}> \text{ looks old } [_{CP} Op \text{ whenever } + C_{[EF]} I \text{ meet } pg]]]] \\
b. \quad & \text{a man } [_{CP} who C_{[EF]} [_{TP} [_{CP} Op \text{ whenever } + C_{[EF]} I \text{ meet } pg] [_{TP} who T_{[AF]} [_{vP} <\text{who}> \text{ looks old }]]]]
\end{align*}
\]

In (25a), the root clause, parallel to subject wh-constructions, involves the movement of who to Spec-T. The adjunct clause, comparable to object wh-constructions, contains the movement of Op to Spec-C under the assumption that whenever appears on C. The asymmetry in chain types prohibits chain composition from being carried out, thus rendering (25a) uninterpretable. This precludes the availability of the parasitic gap in (24a).

Expression (24b), in contrast, licenses the parasitic gap. In this case, the root clause does not permit the EF of C to be inherited by T because of the nature of the EF, which demands Merge to the edge. At the phasal CP of (25b), the adjunct clause is at the edge of T via external Merge.\footnote{Adjunct clauses (e.g. because (restrictive), when, after, before, since (temporal), while (successive), and so that (purposive) clauses) can occur outside of verbal domains (see Nakajima (1982)). Observe the following: (i) John came here before I arrived, but Mary did so after I arrived. As observed in (i), adjunct clauses can be excluded from verbal domains under do so substitution. I thus assume that the whenever clause in (25b) is externally merged into the TP domain.} The occurrence of this material at the edge makes EF inheritance from C to T impossible (see also section 2.2.1). Such impossibility does not hold true for AF inheritance by definition. As a result, who, an operator of the real gap, is raised both to Spec-C and to Spec-T simultaneously. Note that Op, an operator of the parasitic gap, is raised to Spec-C by the EF of C, as with (25a). The parallelism in chain types renders chain composition available and the structure in (25b) interpretable.\footnote{The A-chain independently created in (25b), which does not consist of an operator-variable relation, is regarded as invisible to chain composition, given the ill-formedness of (23b).} This licenses the parasitic gap in (24b).

It is important to note here that previous works (e.g. Chomsky (1982,
1986), Engdahl (1983), etc.) have often attributed the grammatical contrast of the sort observed in (24) to the anti-c-command condition:

(26) A parasitic gap may not be c-commanded by the real gap. According to (26), the expression in (24a) is deemed ungrammatical because $t$ c-commands $pg$ within the adjunct clause; the expression in (24b), which does not have such a configurational relation, is considered grammatical.

However, the condition in (26) encounters some counterevidence. More concretely, there are cases in which the anti-c-command condition incorrectly excludes the occurrence of parasitic gaps (see Engdahl (1984), Kiss (1985), Brody (1995), Sakamoto (2011), among others). The relevant sentences are presented in (27a) and (28).

(27) a. Which man did the police warn $t$ [that they would arrest $pg$]?  
   b. *The police warned him$_i$ [that they would arrest John$_i$].  
   c. The police warned everybody$_i$ [that they would arrest him$_i$].  
      (Kiss (1985: 42, 45))

(28) ?Who did Bill believe $t$ visited you [without you having invited $pg$]?  
      (Brody (1995: 83))

In (27a), the real gap occupying the matrix object position c-commands the parasitic gap within the complement clause. This c-command relation is clear from (27b, c). In (27b), coreferential interpretation is not established between him and John. In (27c), the pronoun him can be interpreted as a variable bound by the operator everybody. These interpretational patterns indicate that the matrix object position c-commands into the complement clause. Similarly, the real gap in (28) c-commands the parasitic gap, with the adjunct clause adjoined to the embedded verbal domain. Despite the fact that the parasitic gaps are c-commanded by the real gaps, sentences (27a) and (28) still remain grammatical. It thus follows from the observation in (27) and (28) that the anti-c-command condition is not qualified as a licensing condition of parasitic gaps.

We can unproblematically capture these offending cases without making recourse to the anti-c-command condition:

(29) a. $[\text{CP} \text{Which man } C_{[EF]} [\text{TP the police warn } <\text{which man}> [\text{CP Op that } + C_{[EF]} [\text{TP they would arrest } pg]]]]$

   b. $[\text{CP} \text{Who } C_{[EF]} [\text{TP Bill believe } <\text{who}> \text{ visited you } [\text{CP Op without } + C_{[EF]} [\text{TP you having invited } pg]]]]$

In (29a), the derivation for (27a), the real gap chain is parallel to the parasitic gap chain because the operator of each chain moves to Spec-C. This parallelism in chain types makes chain composition viable. The same holds
for (29b), which is the derivation for (28).\textsuperscript{28} Sentences (27a) and (28) thus license the parasitic gaps. Not only do the cases given in (27a) and (28) exclude the anti-c-command condition as an independent licensing condition of parasitic gaps, but they also lend additional support to our system.

4. Concluding Remarks

In this paper, I have reconsidered the VMH, a principle stating that a movement operation without an effect on PF output can be suspended (George (1980), Chomsky (1986), Agbayani (2000, 2006), among others), within recent minimalist frameworks (Chomsky (2000, 2001, 2004, 2008)). It has been argued here that some phenomena that have been explained by the VMH follow as a natural consequence from the system proposed in this paper. This system holds that the EF of the phasal heads (i.e. C and $v^*$), unlike that of nonphasal heads, is assigned a certain value via Merge to the relevant edge by drawing upon clausal typing (Cheng (1997a)) and that applicability of EF inheritance unambiguously determines the occurrence or nonoccurrence of vacuous movement. I have also established that the proposed system receives empirical support in various ways.

We can obtain some additional results by reducing the VMH to the proposed system, namely, by ceasing to view the VMH as an independent principle. First, it becomes no longer problematic to postulate movement of phonologically null elements such as null operators (cf. Chomsky (1986), Browning (1987), and references cited therein), which has no effect on PF output. Our system allows of movement with no effect on PF output (see section 2.2.1).\textsuperscript{29}

Then, a significant issue emerges as to how our system excludes “vacuous application of scrambling” (Hoji (1985), Takano (1996)) in Japanese. To understand the significance of this issue, let us observe a scopal effect in scrambling:

\textsuperscript{28} Here, I look upon without as a complementizer, rather than a preposition, that introduces a gerundive clause behind. This view is not novel. In fact, Huang (1982: 83–86) shows that prepositions can function as complementizers in English and Chinese.

\textsuperscript{29} As stated in section 3.5, parasitic gaps are often assumed to be derived as a result of null operator movement (cf. Chomsky (1986)). Movement of null operators does not induce an effect on PF output because they have no phonological feature despite the fact that they are raised from their original position to Spec-C (see e.g. (22b)). The existence of such movement cannot constitute a problem for the system proposed in this paper, in which the VMH is virtually inoperative.
(30) a. Dareka-ga daremo-o semeta.
someone-Nom everyone-Acc criticized
“Someone criticized everyone.” (unambiguous)
b. Daremo-o dareka-ga <daremo-o> semeta.
everyone-Acc someone-Nom criticized
“Everyone, someone criticized.” (ambiguous)

(Hoji (1985: 342))

(30a) exemplifies the basic word order in a Japanese transitive sentence. This sentence is scopally unambiguous, with the quantified subject only taking wide scope over the quantified object. In (30b), the quantified object is scrambled over the quantified subject. In this case, scopal ambiguity is observed between the quantified subject and the quantified object.

The observed effect of scrambling allows us to expect that the two quantified phrases in (30a) exhibit scopal ambiguity if they undergo multiple scrambling in a string-vacuous fashion, as shown in (31).

(31) [Dareka-ga [daremo-o [<dareka-ga> [<daremo-o> semeta]]]]

(31) illustrates that multiple scrambling of the two quantified phrases produces the same word order as in (30a). Unlike (30a), (31) should have ambiguous scope interpretation because scrambling is applied to the quantified object, as observed in (30b). However, such scopal ambiguity is not obtained in (31). Based on facts of this sort, Hoji (1985: 352) reaches the following conclusion:

(32) A syntactic adjunction operation cannot apply if it does not change the order of the overt lexical string.

Relating (32) to the VMH as proposed by George (1980), Hoji states that vacuous application of scrambling is not allowed in Japanese (see Hoji (1985: section 3.5) for details).

Hoji’s (1985) conclusion seems to be in conflict with the proposed system, since this system does not prevent vacuous application of movement (cf. section 2.2.1) and scrambling is generally grouped as a kind of movement (cf. Hoji (1985), Saito (1985)). Although I have no immediate resolution to this conflict, it should be noted that scrambling has been widely analyzed as optional movement (see e.g. Fukui (1993)), which is not driven by need to value some feature. If scrambling is indeed an instance of optional movement, then it differs in nature from what we have discussed in section 2.2.1, that is, vacuous movement of subject DPs from Spec-\(v^*\) to Spec-T, which is driven by need to value the EF inherited by T. It seems to be highly promising to pursue the possibility that the conflict between our analysis and Hoji’s (1985) analysis may be resolved by reducing it to
the difference in movement types, but I leave this issue open for future investigation.

REFERENCES


Mikami, Suguru (2011) *Parametric Variations in A-Movement between Subject-Prominent Languages and Focus-Prominent Languages*, Doctoral dissertation, University of Tsukuba.


[received April 1 2012, revised and accepted July 21 2012]
Doctoral Program in Literature and Linguistics
University of Tsukuba
1−1−1 Tennodai, Tsukuba
Ibaraki 305−8571
e-mail: mo_ttsu@yahoo.co.jp