Abstract

I make two proposals in this article: (a) an economy condition on the operation Copy, which states that Copy should apply to as small an element as possible, and (b) the “two types of head movement” hypothesis, which states that Universal Grammar allows head movement via substitution as well as head movement via adjunction. I argue that with these proposals, we can not only explain two generalizations about what I call headless XPs, but also attribute crosslinguistic variation in the applicability of these generalizations to parameters that are responsible for the availability of multiple specifiers.

Keywords: Copy, head movement, ellipsis, substitution/adjunction, multiple-Specs

1 Introduction

This article is concerned with what I call headless XPs. By this term, I mean a phrase whose head has moved out of it. For example, the XP in (1) is a headless XP.

(1) \[ Y_P [Y X-Y] [X_P t_X Z_P] \]

Notice that it does not matter whether the lower copy of the head, which is expressed as \( t_X \) in (1), is actually pronounced or not. Even if the lower copy is pronounced, the XP in (1) is headless because the head X has moved out of it.
In particular, I am concerned with the following two generalizations about headless XPs: Takano’s Generalization (TG, Takano 2000) and Lasnik’s Generalization (LG, Lasnik 1999). The former states that headless XP-movement is prohibited, and the latter states that headless XP-ellipsis is prohibited.\(^1\) TG and LG are illustrated by the unacceptability of English examples like (2a) and (2b), respectively.

\[(2) \quad \text{a.} \quad \ast \text{It’s } [\text{VP a book } t_1 \text{ to Mary}]_2 \text{ that John gave}_1 t_2. \]
\[\text{b.} \quad \ast \text{Meg gave Ken advice, and Ken will give Sam } [\text{VP give } t_1 \text{ to Sam}]_2 \text{ advice}. \]

The cleft sentence in (2a) involves headless VP-movement, and in (2b), a sort of pseudo-gapping sentence involves headless VP-ellipsis (sections 2.1–2.2).

The aim of this article is to explain TG and LG and crosslinguistic variation in the applicability of these generalizations. In order to account for TG, first I propose an economy condition on Copy requiring Copy to apply to as small an element as possible (section 3.1). In section 3.2, I argue that the same account can be applied to LG if we adopt a modified version of Johnson’s (2001) movement approach to VP-ellipsis.

Furthermore, I point out that there are languages where TG and/or LG do not hold (section 4), as illustrated by a Japanese example in (3) and a Hebrew example in (4).\(^2\)

\[(3) \quad [\text{CP John-ga } t_1 \text{ ageta}_2 \text{ no]-wa } [\text{VP Mary-ni hon-o ni-satu } t_2]_1 \text{ da.} \]

\[
\text{John-NOM gave C-TOP Mary-DAT book-ACC two-CL be} \\
\text{(Lit.) ‘It is Mary two books that John gave.’} \\
\text{‘John gave Mary two books.’}
\]
a. Q: (Ha-’im) Miryam hisi’a et Dvora a-makolet?

Q  Miryam drive.PAST3FSG ACC Dvora to.the-grocery.store

‘(Did) Miryam drive Dvora to the grocery store?’

b. A: Ken, hi hisi’a₁ ₁₁ {et Dvora a-makolet}.

yes she drive.PAST3Fsg

‘Yes, she drove Dvora to the grocery store.’

According to Koizumi (1995), Japanese cleft sentences like (3) are instances of headless VP-movement. Goldberg (2005) argues that Hebrew sentences like (4b), in which elements within the VP are elided, are derived by headless VP-ellipsis. I argue that this crosslinguistic variation in the applicability of TG and LG can be explained and attributed to parameters that are responsible for the availability of multiple specifiers if the “two types of head movement” hypothesis holds (section 5). The hypothesis states that Universal Grammar (UG) allows head movement via substitution as well as head movement via adjunction. Although this hypothesis runs contrary to the widely held assumption that head movement is always performed via adjunction, this is the null hypothesis, as I will argue in section 5.1.

2 Constraints on Headless XPs

2.1 Headless XP-Movement: Takano’s Generalization

Takano (2000) argues that the following generalization holds:

(5)  Takano’s Generalization

XP-movement is prohibited if the head of XP has moved out of XP.
He attributes examples like (6) to the generalization in (5).

(6) a. *[Ihr
her ein Buch \( t_1 \)]_2 gab_1 Hans \( t_2 \).
   her a book gave Hans
   ‘Hans gave her a book.’

b. *[Het boek aan Marie \( t_1 \)]_2 gaf_1 Jan (waarschijnlijk) \( t_2 \).
   the book to Marie gave Jan probably
   ‘Jan (probably) gave the book to Marie.’

c. *It’s [a book \( t_1 \) to Mary]_2 that John gave\( t_1 \) \( t_2 \). (Takano 2000 : 145)

The German example in (6a) and the Dutch example in (6b) are derived by headless VP-movement, in which the finite verb moves out of the VP. The English example in (6c) also involves movement of a headless VP whose head has undergone head movement out of the VP, given the Larsonian shell analysis of ditransitive clauses (Larson 1988, Chomsky 1995c). All of these examples involve headless XP-movement. Therefore, the unacceptability of these examples is consistent with TG as stated in (5).

Moreover, Wurmbrand (2004) argues that the following paradigm in German can be attributed to TG (the Headless Fronting Constraint in her terms):

(7) a. weil Hans seinen Bruder angerufen hat
   since Hans his brother up.called has
   ‘since Hans phoned his brother’

b. Gestern rief der Hans \([vP/VP seinen \text{ Bruder an } t_V]\)
   yesterday called the.NOM Hans his.ACC brother up
‘Hans phoned his brother yesterday.’

c. \[ _{vP} \text{Seinen Bruder angerufen} \text{ hat nur der Hans} _{t_{vP/VP}} \]

[ his.ACC brother up.called] has only the Hans

‘Only Hans phoned his brother.’

d. *\[ _{vP} \text{Seinen Bruder an} _{t_V} \text{ rief der Hans gestern} _{t_{vP/VP}} \]

[ his.ACC brother up] called the Hans yesterday

‘It was yesterday that Hans phoned his brother.’ (Wurmbrand 2004 : 7)

All of these examples involve particle-verb constructions. The acceptability of (7b) indicates that a particle can be stranded when a verb moves to C via verb-second movement. The acceptability of (7c) illustrates that fronting of a V/vP that includes the object, the verb, and the particle is allowed. Since V-movement stranding a particle is generally allowed and fronting of V/vP that include a particle is allowed, as shown in (7b) and (7c), respectively, the unacceptability of (7d) must be attributed to the fact that the headless V/vP is fronted. Thus, the unacceptability of (7d) suggests that TG holds in German.\(^5\)

2.2 Headless XP-Ellipsis: Lasnik’s Generalization

Lasnik (1999) considers the following generalization on headless XPs, which is similar to TG:\(^6\)

\(8\) \textit{Lasnik’s Generalization}

XP-ellipsis is prohibited if the head of XP has moved out of XP.
LG amounts to saying that headless XP-ellipsis is prohibited. Some facts concerning English pseudogapping are attributed to the generalization. Lasnik argues that pseudogapping sentences like (9a) involve VP-ellipsis. According to this analysis, the sentence in (9a) has the underlying representation in (9b).\footnote{Lasnik 1999 : 147}

(9)

\begin{enumerate}
\item Mary hasn’t dated Bill, but she has Harry.
\item Mary hasn’t dated Bill, but she has Harry \{VP \_dat\_t\}. \text{(Lasnik 1999 : 147)}
\end{enumerate}

In (9a), the object \textit{Harry} (A-) moves out of the VP while the V \textit{dated} stays in situ. Thus, VP-ellipsis yields the pseudogapping sentence. Lasnik argues that the reason why V can stay in situ in the pseudogapping sentence is that V’s strong feature that drives V-movement is eliminated by VP-ellipsis, avoiding a PF crash. This explains why V \textit{can} stay in situ in the pseudogapping sentence.\footnote{However, it does not \textit{prohibit} Vs from moving. Thus, this analysis cannot account for the sentence in (10a), where the direct object of a double object construction is elided. This sentence can be derived by VP-ellipsis if the V \textit{give} and Sam move out of the VP and \textit{advice} stays does not, as shown in (10b). Notice that if the V stays in situ, a legitimate pseudogapping sentence is obtained, as shown in (10c).}

(10)

\begin{enumerate}
\item *Meg gave Ken advice, and Ken will give Sam advice.
\item *Meg gave Ken advice, and Ken will give Sam \{VP \_give \_t\_Sam advice\}.
\item ?Meg gave Ken advice, and Ken will Sam \{VP \_give \_t\_Sam advice\}.
\end{enumerate}

If LG holds in English, the unacceptability of (10a) is unproblematic for Lasnik’s analysis of pseudogapping, since the headless VP is elided in (10a).
Furthermore, the fact that null objects are not allowed in English can be partially attributed to LG. As shown in (11), English does not allow null objects.

(11) *John solved the problem, and Mary solved, too.

One of the possible representations for (11) is (12).

(12) *John solved the problem, and Mary solved pro, too

The second conjunct in (12) contains pro. This representation is ruled out by the standard assumption that English does not have pro. However, this assumption is not sufficient to rule out all apparent null objects in English since the null object sentence in (11) can have another possible representation without pro. Given that V moves to v and an object can stay in situ in English, headless VP-ellipsis yields a null object sentence, as shown in (13).

(13) *John solved the problem, and Mary \[
{\text{[}}_{\text{VP}}{\text{ solved }}{\text{[VP }}{\text{solve the problem}}{\text{]}}{\text{], too.}}
\]

In the second conjunct in (13), the verb moves to v and the object stays in the VP, resulting in a headless VP. If the headless VP is elided, the null object sentence obtains. Thus, in order to capture the fact that null objects are not allowed in English, we need to rule out VP-ellipsis in English. Notice that vP-ellipsis is possible in English, as shown in (14).

(14) John solved the problem, and Mary did \[
{\text{[}}_{\text{VP}}{\text{ solve the problem}}{\text{], too.}}
\]

It is unclear why vP can be elided but VP cannot. This difference between vP and VP can be attributed to LG since VP is headless while vP is “headed.” Under LG, “headed” vPs can be elided but headless VPs cannot.

As pointed out by Lasnik (1999), LG is also consistent with the fact that sluicing is prohibited when I moves out of IP, as shown in (15b).
(15)  

a. Mary saw someone. 

b. Who (*did)? 

c. Who did \{Mary, \textit{see}\}?  

(Lasnik 1999 : 158)

Given that sluicing involves IP-ellipsis, this fact falls under LG. This is because, as illustrated in (15c), the headless IP must be elided in order to derive the sluicing example with the remnant auxiliary.

3 Explanation

In this section, I propose an economy condition on Copy, according to which Copy must apply to as small an element as possible. In section 3.1, I will show that this condition can explain TG. In section 3.2, I will argue that LG can also be explained by the same condition if we adopt a revised version of Johnson’s (2001) movement approach to VP-ellipsis.

3.1 Explanation for Takano’s Generalization: Economy Condition on Copy

TG follows from an economy condition on Copy, which states, in effect, that Copy should apply to as small an element as possible. Before formulating a precise definition of this condition, let us see how this idea can account for TG informally.

3.1.1 Informal Explanation

Consider the following structure, where a probe $\alpha$ and a potential goal $X$ have a feature [F], which is the driving force of Agree, and XP also has [F] since XP is a projection of X.
In (16), X undergoes head movement to Y, yielding a headless XP. Crucially, X adjoins to Y, Y projecting a segment not a category. Thus, given the distinction between categories and segments and the definition of c-command based on that distinction (X c-commands Y iff every category that dominates X also dominates Y and X does not dominate Y), X c-commands XP and conversely. That is, a mutual c-command relation holds between X and XP.

As a result, both X and XP can be a candidate for Agree in terms of minimality. Given the Copy Theory of Move and the widely held view that Agree is a prerequisite for Move, this means that both X and XP are candidates for Move. However, X is smaller than XP in some sense. Thus, the economy condition on Copy prevents the bigger element, XP, from moving. As such, headless XP movement is prohibited by the economy condition on Copy: headless XP movement is impossible since a smaller element X is a candidate for Move.\(^\text{10}\)

A crucial assumption in this explanation is that the computational system does not discriminate head movement from phrasal movement in the relevant respect since they compete against each other in which movement is performed. In the case of headless XPs, the possibility of head movement of X prohibits XP from moving. I take this assumption to be desirable, in the context of a research program that aims to assimilate head movement as far as possible to phrasal movement. An interesting issue under this research program is the
question of how we can derive differences between head movement and phrasal movement from independent factors. One important difference is concerned with locality: head movement seems to be restricted to the next higher head (the Head Movement Constraint) while phrasal movement can target a position farther away. In the next section, I will show that this difference is derived from an economy condition on Copy and structural intervention.

3.1.2 Formal Explanation

Now, in order to implement this general idea into the formal theory of grammar, we need to formulate several definitions. The following is one of the plausible implementations.

First, I define Agree Domain (AD) based on Agree, as in (17). AD is a set whose members can be candidates for other operations such as Copy (and maybe Value).

(17) \textit{Agree Domain (AD)}

At a given stage of a derivation where $\alpha$ is a probe;

a. $\text{AD}(\alpha) = \{ x : \text{Agree}(\alpha, x) \}$

b. $\text{Agree}(\alpha, \beta)$ iff

(i) $\alpha$ c-commands $\beta$ and

(ii) $\alpha$ matches $\beta$ and

(iii) there is no $\gamma$ that structurally intervenes between $\alpha$ and $\beta$.

The statement in (17a) amounts to saying that AD of $\alpha$ is a set of potential goals of Agree. As shown in (17biii), a locality condition is incorporated into the definition of Agree. Structural intervention is defined as follows:

(18) \textit{Structural Intervention}
\(\gamma\) structurally intervenes between \(\alpha\) and \(\beta\) iff

a. every category that dominates \(\gamma\) also dominates \(\beta\) and

b. some category that dominates \(\beta\) does not dominate \(\gamma\) and

c. \(\alpha\) c-commands \(\gamma\) and

d. \(\gamma\) matches \(\alpha\) and \(\beta\) and

e. \(\gamma\) is not a complement of \(\alpha\).

The statement in (18a) amounts to stating that either a c-commanding element or a dominating element can be an intervener ((18a) is adopted from Tanaka (2004 : 925)). Thus, this definition has the combined effect of Relativized Minimality or the Minimal Link Condition (MLC) (Rizzi 1990, Chomsky 1995c) and the A-over-A Principle (Chomsky 1964).

Take the following schematic structures for examples:

(19) a. \([XP\ \alpha\ [YP\ \gamma\ [ZP\ \beta\ \ldots]]]\]

b. \([XP\ \alpha\ [YP\ \ldots[\gamma\ \beta\ \ldots]]]\]

The structure in (19a) illustrates MLC effects. In (19a), \(\gamma\) could structurally intervene between \(\alpha\) and \(\beta\) since every category that dominates \(\gamma\) also dominates \(\beta\) (XP and YP). (19b) illustrates A-over-A effects. \(\gamma\) could structurally intervene between \(\alpha\) and \(\beta\) in this configuration too: every category that dominates \(\gamma\) also dominates \(\beta\).

I need to incorporate the A-over-A Principle into this definition because otherwise an economy condition on Copy would prohibit all phrasal movement if we have only the MLC as a locality condition. We will see why this is so after I introduce a formal definition of an economy condition on Copy. See Fukui 1997 and Fukui 1999 for the detailed discussion.
about the A-over-A principle under the minimalist perspective.

The statement in (18b) is required in order to prevent an element $\gamma$ from intervening between a probe and $\gamma$ itself. For example, in (19a), it is trivially true that every category that dominates $\gamma$ (i.e. XP and YP) dominates $\gamma$. Thus, without (18b), $\gamma$ structurally intervenes between $\alpha$ and $\gamma$. This means that nothing can be a goal of Agree because everything intervenes between a probe and itself. The statement in (18b) overcomes this self-intervining problem since there is no category that simultaneously dominates and does not dominate something, which is a clear contradiction.

The statement in (18e) realizes some kind of notion of Anti-Locality (cf. Abels 2003, Grohmann 2003), saying that an element that is too close to the probe is not an intervener. We will see why this condition is required after introducing a formal definition of an economy condition on Copy.$^{11}$

Given these assumptions, take schematic structures like (20a) and (20b).

(20)  
   a. $[\alpha(F) [\gamma(F) [\beta(F) \ldots]]]$  
   b. $[\alpha(F) [\gamma(F) \beta(F)]]$

In (20a), $\text{AD}(\alpha)$ is $\{\gamma\}$. $\beta$ is not in $\text{AD}(\alpha)$ since $\gamma$ structurally intervenes between $\alpha$ and $\beta$. Thus, other operations like Copy can apply to $\gamma$ but not $\beta$. On the other hand, in (20b), $\text{AD}(\alpha)$ is $\{\beta, \gamma\}$ since $\beta$ does not intervene between $\alpha$ and $\gamma$ nor $\gamma$ does not intervene between $\alpha$ and $\beta$: although every category that dominates $\gamma/\beta$ also dominates $\beta/\gamma$, there is no category that dominates $\gamma/\beta$ and does not dominate $\beta/\gamma$. As a result, in (20b), unlike in (20a), Copy can apply to either $\beta$ or $\gamma$. The goal that undergoes Copy is determined by
other factors including economy conditions.

I propose one such economy condition, the economy condition on Copy (ECC), which is formulated as follows:\textsuperscript{12}

(21) \textit{Economy Condition on Copy (ECC)}

At a given stage of a derivation where \( \alpha \) is a probe;

Copy(\( \beta \)) only if

a. \( \beta \in \text{AD}(\alpha) \) and

b. \( \gamma \notin \text{AD}(\alpha) \) such that \( \gamma \preceq \beta \).

This condition says that the copying of an element is prohibited if there is another smaller element in the relevant AD. For example, in (20b), if \( \gamma \) is smaller than \( \beta \), copying of \( \beta \) is impossible. In order to make this condition precise, we need to define smallness. The following is one possible definition of smallness, based on the notion of \textit{term} of Chomsky 1995:

(22) \textit{Smallness}

a. \( \alpha \prec \beta \) iff \( \alpha \in \text{sub-term of } \beta \) (ST(\( \beta \)))

b. \( \text{ST}(\beta) = \{x: x \text{ is a term of } \beta \text{ and } x \neq \beta\} \)

(23) a. \( K \) is a term of \( K \).

b. If \( L \) is a term of \( K \), then the members of the members of \( L \) are terms of \( K \).

(Chomsky 1995c : 247)

According to these definitions, \( \alpha \) is smaller than \( \beta \) if and only if \( \alpha \) is a term of \( \beta \) unless \( \alpha = \beta \).
With these definitions in mind, let us return to TG. Consider again the schematic structure in (16), repeated below as (24).

(24)

\[
\begin{array}{c}
\alpha[F] \\
Y \\
X[F] Y[t_X] ZP
\end{array}
\]

In (24), \( AD(\alpha) = \{X, XP\} \) since \( X \) does not structurally intervene between \( \alpha \) and \( XP \) nor does \( XP \) structurally intervene between \( \alpha \) and \( X \): although every category that dominates \( X/XP \) also dominates \( XP/X \), there is no category that dominates \( X/XP \) and does not dominate \( XP/X \). Thus, Copy can apply to either \( X \) or \( XP \). However, \( X \) is smaller than \( XP \) according to the definition of smallness in (22) because \( X \) is a term of \( XP \) and \( X \neq XP \). Therefore, \( Copy(XP) \) is prohibited by the ECC. This is why headless phrases cannot be moved. Notice in passing that although Copy(\( X \)) and movement of \( X \) in (24) (i.e. excorporation of \( X \)) are allowed in principle, the resulting representation is filtered out by some morphological constraints such as Baker’s that prohibits word internal traces (Baker 1988). That is, the movement of \( X \) itself is a legitimate operation in syntax but the resulting head \( Y \) violates Baker’s morphological condition since \( Y \) contains the trace of \( X \).

Before closing this section, let us see why we need to incorporate the A-over-A principle into the definition of structural intervention (see (18a)) and assume that the complement of a probe is not an intervener (the condition in (18e)). First, consider the following configuration, which illustrates movement of subjects to SpecTP:
In (25), T is a probe because it has an uninterpretable $\phi$-feature, $[u\phi]$. Suppose that we have only the MLC as a locality condition. Then, DP, $D'$, and D are contained in AD(T) since D has the matching feature $[\phi]$ and $D'$ and DP are projections of D. However, given that intermediate projections are syntactically invisible, following Chomsky (1995a) (see also Speas 1990, Travis 1984), only DP and D are contained in AD(T). Notice that D is in AD(T). This is because T c-commands D and there is no intervener between them with respect to the MLC (there is no c-command relation between DP and D). Thus Copy(DP) is prohibited in terms of the ECC since D is smaller than DP (D is in SD(DP)). Therefore, the ECC in (21) predicts that there is no phrasal movement, which is obviously an undesirable consequence.

On the other hand, with the definition of structural intervention in (18), in which the A-over-A principle as well as the MLC are incorporated, D is not contained in AD(T). The DP structurally intervenes between T and D according to this definition: (i) every category that dominates DP also dominates D, (ii) some category that dominates D does not dominate DP, (iii) T c-commands DP, (iv) DP matches T and D with respect of the $\phi$-feature, (v) and DP is not a complement of T. Therefore, Copy can apply to DP even if DP is bigger than D.

Next, let us consider the non-complement restriction on the interveners. We need this restriction since head movement would never occur otherwise. To see this, let us consider the following configuration, which illustrates V-to-\(v\) movement:
In (26), V cannot move to v without (18e): VP intervenes between v and V. On the other hand, with (18e), VP is not considered an intervener since VP is a complement of the probe v. Then AD(v) is {VP, V}. V is smaller than VP. Therefore, V rather than VP moves to v.

Notice that the Head Movement Constraint (HMC: Travis (1984)) and the Anti-Locality Principle can be deduced from our system thanks to this non-complement restriction. The HMC states that head movement of X to Y cannot skip an intervening head Z. For example, in the following schematic structure, X cannot move to Y:

\[
\text{(27)} \quad [\ YP\ Y(F)\ ZP\ Z[\ XP(F)\ X(F)\ .\ .\ .] \ ]
\]

Under the present system, this fact can be explained as an A-over-A effect. AD(Y) is \{XP\}. Crucially, X is not contained in AD(Y) because XP, which is not a complement of Y, structurally intervenes between Y and X. The definition of structural intervention in (18) predicts that head movement occurs only if the projection of the goal is a complement of the probe, thus deriving the HMC.

The most standard version of the Anti-Locality Principle states that a complement of a head X cannot move to SpecXP. Thus, in (26), VP cannot move to SpecvP. This can be explained by the non-complement restriction in the definition of structural intervention and the ECC. That is, movement of VP (actually Copy of VP) is prohibited by the ECC because AD(v) contains a smaller element V (VP does not intervene between v and V because VP is a complement of v). Thus, Anti-Locality effects can be reduced to the ECC.

To sum up, in this section, I proposed a formal system that implements the idea of an
economy condition on Copy, and illustrated that TG can be reduced to the ECC under this system. Furthermore, I showed that HMC effects and Anti- Locality effects can also be deduced under the proposed system.

3.2 Explanation for Lasnik’s Generalization: Movement Approach to VP-Ellipsis

TG and LG look similar. Thus, a natural question that arises is whether they are rooted in the same principle (i.e. whether they can be uniformly explained). In this subsection, I answer this question in the affirmative, at least in the domain of VP-ellipsis. In particular, I argue, in the spirit of Johnson’s (2001) movement approach to VP-ellipsis and an idea in Aelbrecht and Haegeman 2011, that VP-ellipsis involves VP-movement to either SpecTopP in the CP periphery or SpecTopP in the vP periphery.\textsuperscript{15} If the idea that ellipsis reduces to movement somehow is correct, LG can be explained by the ECC in the same manner as TG.

3.2.1 VP-Ellipsis Is Derived Through VP-Topicalization: Johnson 2001

Johnson (2001) argues, in the spirit of Lobeck 1995, that sentences involving VP-ellipsis like the second conjunct in (28a) are derived by topicalizing VP and eliding the moved VP, as illustrated in (28b). Thus, according to this analysis, in order for a VP to elide, it must first topicalize.

\begin{align*}
(28) & \quad \text{a. John ate a cake, and Bill did, too.} \\
& \quad \text{b. } \{\text{\textit{v}P \textit{eat a cake}}\}, \text{ Bill did } t_1, \text{ too.}
\end{align*}

This analysis is conceptually appealing since under this analysis, we do not need to assume an additional operation specific to ellipsis: the same operation that is used for copy
deletion can be utilized for ellipsis. This analysis has also an empirical motivation because it can straightforwardly capture similarities between VP-ellipsis and VP-topicalization in the licensing conditions. VP-ellipsis is licensed by auxiliary verbs like *do*, *have*, *be*, and infinitival *to*, as illustrated in the following sentences, which are drawn from Johnson 2001:

(29) a. José Ybarra-Jaegger likes rutabagas, and Holly does ∆ too.
    b. José Ybarra-Jaegger ate rutabagas, and Holly has ∆ too.
    c. José Ybarra-Jaegger is eating rutabagas, and Holly is ∆ too.
    d. Mag Wildwood wants to read Fred’s story, and I also want to ∆.

In the absence of such auxiliary verbs, VP-ellipsis is prohibited, as the following examples indicate (drawn from Johnson 2001):

(30) a. I can’t believe Holly Golightly won’t eat rutabagas. I can’t believe Fred *(won’t), either.
    b. Sally Tomato started running down the street, but only after José started *(to).

Johnson shows that the same licensing condition is applied to VP-topicalization, as illustrated by the following examples (drawn from Johnson 2001):

(31) Madame Spanella claimed that . . .
    a. *eat rutabagas*, Holly wouldn’t *t.*
    b. *eaten rutabagas*, Holly hasn’t *t.*
    c. *eating rutabagas*, Holly should be *t.*
    d. *eating rutabagas*, Holly’s wants to *t.*

(32) Madame Spanella claimed that . . .
a. *would eat rutabagas, Holly t.

b. *hasn’t eaten rutabagas, Holly t.

c. ?*eating rutabagas, Holly started t.

This distributional similarity between VP-ellipsis and VP-topicalization can be straightforwardly captured by Johnson’s movement approach to VP-ellipsis since the former is derived through the latter under this approach. Whatever rules out (32) also rules out (30).

Despite its advantages, however, we cannot adopt Johnson’s analysis in its original form because there are a number of counterexamples to it, as Aelbrecht and Haegeman (2011) point out, which are reviewed in the next section.16

3.2.2 TopP in a VP periphery: Aelbrecht and Haegeman 2011

Aelbrecht and Haegeman (2011) argue against Johnson’s analysis, showing that the distribution of VP-topicalization is much more restrictive than that of VP-ellipsis (that is, there are a number of cases where VP-ellipsis is allowed even if VP-topicalization is not). Some counterexamples are listed below:

(33) Wh Complements

a. *I knew that one student presented this article in my class but I can’t recall now [which of the students [present this article] did]

b. I knew that some students presented this article in my class but I couldn’t recall [which of the students didn’t] (Aelbrecht and Haegeman 2011 : 8)

(34) Adverbial Clauses

a. *Mary wanted to move to London, and after [move to London] she did, her
life changed entirely.

b. Mary wanted to move to London, and after she did, her life changed entirely.

(Aelbrecht and Haegeman 2011: 10-11)

(35)  **Factive Complements**

a. *John intends to make a table, and we’re afraid that [make one] he will.

b. John intends to make a table, and we’re afraid that he will.  (Aelbrecht and Haegeman 2011: 12)

As illustrated in the above examples, VP-ellipsis is allowed in the environments where VP-topicalization is not, such as Wh complements, adverbial clauses, and factive complements (see Aelbrecht and Haegeman 2011 for other counterexamples to Johnson’s analysis). This discrepancy between VP-ellipsis and VP-topicalization is unexpected under Johnson’s analysis, in which VP-ellipsis is derived through VP-topicalization: again whatever rules out VP-topicalization should also rule out VP-ellipsis in his analysis.

Aelbrecht and Haegeman (2011) suggest that we can overcome the problem with Johnson’s approach while maintaining its core idea if we take a low vP periphery into consideration. Assuming that there is a TopP in the vP periphery, along the line of Jayaseelan (2001) and Belletti (2001, 2004, 2009), they argue that VP-ellipsis involves VP-movement to SpecTopP in the vP periphery rather than in the CP periphery. Then the counterexamples to Johnson’s original analysis lose their force under this alternative analysis if we assume that what causes the ungrammaticality in the counterexamples above is not VP-movement itself but VP-movement to the CP periphery.
Given this discussion, I propose the following hypothesis:

(36) a. VP-topicalization is derived through verbal phrase-movement to SpecTopP in the CP periphery.

b. VP-ellipsis is derived through verbal phrase-movement to SpecTopP in either the CP periphery or vP periphery.

This hypothesis accounts for the fact that the distribution of VP-topicalization is more limited than that of VP-ellipsis since VP-ellipsis can be derived in two ways while VP-topicalization can only be derived in one. For example, the second conjunct clause in (37a) can be derived in two ways, as illustrated in (37b) and (37c).

(37) a. John solved the problem, and Mary did too.

b. $[\text{TopP } \{\text{solve the problem}\} \text{ TP Mary did } t ]$

c. $[\text{TP Mary did } [\text{TopP } \{\text{solve the problem}\} \text{ t } ] ]$

In (37b), the elided verbal phrase is moved to SpecTopP in the CP periphery while in (37c), it is moved to SpecTopP in the vP periphery. On the other hand, the embedded clause of the second conjunct in (38a) can only be derived in one way like (38b).

(38) a. John solved the problem, and I think that [solve the problem] Mary will too.

b. $[\text{TopP } \{\text{solve the problem}\} \text{ TP Mary will } t ]$

The deference between VP-ellipsis and VP-topicalization in the possible landing site can be explained by the ECC (Anti-Locality effects) and Lasnik’s (1999) assumptions about V-movement, which are (i) that V has a strong feature that must be eliminated until the derivation reaches PF and (ii) that the strong feature of V can be eliminated either by
moving to \( v \) or being elided by deletion. First, VP-topicalization cannot be derived through VP-movement to SpecTopP in the \( vP \) periphery. This is because in the derivation illustrated in (39), \( V \) does not move to \( v \).

(39) \*[[TP Mary will [\( \text{TopP} \) [\( \text{VP} \) solve(F) the problem] [\( \text{Top} \) \( \text{Top} \) [\( vP \) \( v \) \( t_{vP} \) ]]]]]

This representation is an illicit PF representation since the strong feature of \( V \) (F) is not eliminated. On the other hand, in the derivation in which VP-topicalization is derived through \( vP \)-movement to SpecTopP in the \( vP \) periphery, the strong feature of \( V \) is eliminated since \( V \) can move to \( v \). This is illustrated in the following:

(40) \*[[TP Mary will [\( \text{TopP} \) [\( vP \) solve(F)-v the problem] [\( \text{Top} \) \( \text{Top} \) [\( t_{vP} \) ]]]]]

However, notice that \( vP \)-movement to SpecTopP in the \( vP \) periphery is a typical movement that violates the Anti-Locality Principle. As we saw in section 3.1.2, this kind of super-local movement is prohibited by the ECC. Therefore, there is no licit derivation in which VP-topicalization is derived through verbal phrase-movement to SpecTopP in the \( vP \) periphery.

On the other hand, the derivation in which \( vP \)-movement to SpecTopP in the CP periphery is unproblematic with respect to both \( V \)'s strong feature and the Anti-Locality Principle, as illustrated in the following:

(41) [[\( \text{TopP} \) [\( vP \) solve(F)-v the problem] [\( \text{Top} \) \( \text{Top} \) [\( \text{TP} \) Mary will \( t_{vP} \) ]]]]

The derivation in which VP moves to SpecTopP in the CP periphery is ruled out for the same reason as the case of VP-movement to SpecTopP in the \( vP \) periphery (i.e., \( V \)'s strong feature remains uneliminated). Therefore, the only possible derivation for VP-topicalization is the one in which \( vP \) moves to SpecTopP in the CP periphery. This is
why VP-topicalization can be derived only through verbal phrase-movement to SpecTopP in the CP periphery.

In contrast, VP-ellipsis can be derived through verbal phrase-movement to SpecTopP in either the CP periphery or vP periphery. This is because V’s strong feature can be eliminated along with the deletion of VP even if V does not move to v. VP-ellipsis cannot be derived through vP-movement to SpecTopP in the vP periphery as in the case of VP-topicalization. This is due to the ECC (Anti-Locality effects), as illustrated in the following:

\[(42) \quad *[_{TP} \text{Mary will } [_{TopP} [_{vP} \text{solve}(<F>) \text{the problem}] [_{Top'} \text{Top } t_{vP} ] ] ]\]

In contrast with VP-topicalization, VP-ellipsis can be derived through VP-movement to SpecTopP in the vP periphery even if V does not move to v to eliminate its strong feature F. This is because the strong feature can be eliminated along with the deletion of VP, as can be seen in the following:

\[(43) \quad [_{TP} \text{Mary will } [_{TopP} [_{vP} \text{solve}(<F>) \text{the problem}] [_{Top'} \text{Top } t_{vP} ] ] ]\]

Therefore, VP-ellipsis, unlike VP-topicalization, can be derived through verbal phrase-movement to SpecTopP in the vP periphery. As for movement to SpecTopP in the CP periphery, both VP and vP can be moved. Both derivations are unproblematic in terms of the ECC and V’s strong feature, as illustrated in the following:

\[(44) \quad a. \quad [_{TopP} [_{vP} \text{solve}(<F>) \text{the problem}] [_{Top'} \text{Top } [_{TP} \text{Mary will } [_{vP} v t_{vP} ] ] ] ]\]
\[b. \quad [_{TopP} [_{vP} \text{solve}(<F>) \text{the problem}] [_{Top'} \text{Top } [_{TP} \text{Mary will } t_{vP} ] ] ]\]

Thus, VP-ellipsis can be derived through verbal phrase-movement to SpecTopP in either the CP or vP periphery.
In this way, the hypothesis in (36) can be derived under the present system, in which the ECC and Lasnik’s assumptions about V-movement are adopted. Furthermore, in Funakoshi 2011b, I provide empirical evidence in favor of this hypothesis over Johnson’s original movement analysis of VP-ellipsis and other non-movement approach to VP-ellipsis.

Given this analysis of VP-ellipsis, we can explain LG in the case of V/vP-ellipsis by the ECC and structural intervention (MLC effects). In order to do so, we need to show that both headless V/vP-movement to the CP periphery and headless V/vP-movement to the vP periphery are ruled out.

First, let us see how V/vP-movement to SpecTopP in the CP periphery is ruled out when the moved phrase is headless. There are five possible configurations to be considered, depending on (i) what is elided (VP or vP) and (ii) where V moves (v, T, or Top). Suppose that V moves to v and stops there. Then the following configuration results after the higher Top is merged with TP:

(45)

In this configuration, a headless VP obtains. This headless VP cannot be moved to Spec-TopP since AD(Top) contains both V and VP, neither of which intervene for the other.
According to the ECC, Copy cannot apply to VP because of the presence of a smaller element V in AD(Top).

If the [V-v] complex further moves to T, a headless VP and vP obtain, as seen in (46).

(46)

Headless vP movement to SpecTopP is ruled out by the ECC since AD(Top) is \{v, vP\}, prohibiting Copy from applying to vP by the ECC. Furthermore, the headless VP cannot be moved to SpecTopP because VP is not contained in AD(Top) due to the structural intervention by V (every category that dominates V also dominates VP (TopP and TP) and some category that dominates VP does not dominate V (vP)).

Finally, if the [V-v-T] complex further moves to Top, the following configuration obtains:
Here, the headless vP cannot be moved since AD(Top) does not contain vP due to the structural intervention by v: Top c-commands v, every category that dominates v also dominates vP (TopP), and some category that dominates vP does not dominate v (TP). The headless VP cannot be moved either for the same reason: V structurally intervenes between Top and VP.

In this way, both headless VP-movement and headless vP-movement to SpecTopP in the CP periphery are ruled out either by the ECC or by the structural intervention (MLC effects) regardless of the final landing site of V.

Let us next consider whether headless V/vP-movement to SpecTopP in the vP periphery is ruled out. In this case, headless V/vP-ellipsis can be derived in three ways, depending on (i) what is elided (VP or vP) and (ii) where V moves (v or Top). When V moves to v and stops there, a headless VP results, as illustrated in the following structure:
In this configuration, the headless VP cannot be moved since AD(Top) is \{V, VP\}: V does not intervene between Top and VP nor VP does not intervene between Top and V since there is no category that dominates V/VP and does not dominate VP/V. The ECC requires that Copy apply to V rather than VP. Therefore, the headless VP cannot be moved in this case.

Next, let us consider the case in which the [V-v] complex moves to Top. In this case, there are two headless XPs, vP and VP, as shown in the following structure:

The headless vP cannot be moved because AD(Top) is \{v, vP\}, and the ECC prohibits Copy from applying to vP. The headless VP cannot be moved due to the structural intervention by V: AD(Top) does not contain VP since V structurally intervenes between Top and VP (Top c-commands V, every category that dominates V also dominates VP (TopP) and some
category that dominates VP does not dominate V (vP)).

Therefore, headless v/VP-movement to SpecTopP in the vP periphery is prohibited either by the ECC or by the structural intervention (MLC effects) regardless of the final landing site of V as in the case of headless v/VP-movement to SpecTopP in the CP periphery. Under the hypothesis in (36), then, we can explain why headless V/vP-ellipsis is prohibited because v/VP-ellipsis is derived through v/VP-movement to SpecTopP either in the vP periphery or in the CP periphery.

If this line of analysis can be extended to the ellipsis of other categories such as IP and NP, and if movement is required to license ellipsis in general, as Johnson (2001) suggests, we get the full explanation of LG. However, the explanation of the prohibition of headless V/vP-ellipsis might be sufficient because almost all the data that were accommodated under LG in section 2.2 involve headless VP-ellipsis except the matrix sluicing data in (15). Notice that the matrix sluicing data is consistent with not only LG but also Merchant’s (2001) Sluicing-COMP generalization that “in sluicing, no non-operator material may appear in COMP” (Merchant 2001: 62). Thus, it is conceivable that the prohibition of headless IP-ellipsis will be explained by an independent principle that explains the Sluicing-COMP generalization.

3.3 Interim Summary

To summarize the discussion so far, I proposed an economy condition on Copy (the ECC) and showed that it can explain TG (that is, why headless XPs cannot be moved). Furthermore, I argued that if we adopt a movement analysis of VP-ellipsis, which originated from
Johnson’s (2001) idea and modified by Aelbrecht and Haegeman (2011), LG (a subpart of it at least) can be reduced to the ECC.

In what follows, I will turn to cross-linguistic variation in the applicability of TG and LG and argue that if UG allows two types of head movement, the cross-linguistic variation can be reduced to parameters that are responsible for the availability of multiple-Specs.

4 Exceptions to Takano and Lasnik’s Generalizations

If the account given in the last section is correct, it is expected that TG and LG should hold in all languages. In this section, however, it is shown that there are languages where TG and/or LG do not hold. That is, languages vary in the availability of headless XP-movement/ellipsis.

4.1 Exceptions to Takano’s Generalization


(50) \[ [_{CP} \text{John-ga } t_1 \text{ ageta}_2 \text{ no]-wa } [_{VP} \text{Mary-ni hon-o ni-satu } t_2 ]_1 \text{ da}. \]

John-NOM gave COMP-TOP Mary-DAT book-ACC two-CL be

(Lit.) ’It is Mary two books that John gave.’

’John gave Mary two books.’
Koizumi (1995) argues that sentences like (50) involve headless verbal phrase-movement, which I assume is vP. Thus, as illustrated in (51), the verb age ‘give’ moves to v, and v containing the verb moves to T. It yields the headless vP containing the indirect object Mary-ni ‘Mary-DAT’ and the direct object hon-o ni-satu ‘book-ACC two-CL’. The multiple cleft sentence in (50) obtains if the resulting headless vP moves to the focus position.

(51) \[ \text{TP John-NOM} \ [ \text{vP}_1 \text{John} \ [ \text{VP Mary-DAT two books-ACC} \ t_v \ ] t_v ] \ [ \text{T give-v-T} ] \]

If this analysis is correct, it indicates that headless vP-movement is possible in Japanese. That is, TG does not hold in Japanese.

Takano (2000) also points out, attributed to an LI reviewer, that according to Huang (1997), certain action sentences in Chinese involve headless verbal phrase-movement. If this is correct, Chinese does not obey TG either.21

Furthermore, Landau (2006) argues that verbal phrase-fronting sentences in Hebrew like (52a) involve headless verbal phrase-movement, which I assume is vP-movement.

(52) a. liknot et ha-praxim, hi kanta.

to-buy ACC the-flowers she bought

‘As for buying the flowers, she bought.’

b. \[ \text{vP}_1 \text{liknot}_1 \text{et} \ \text{ha-praxim}_2 \ [ \text{TP hi kanta}_1 \ t_2 ] \]

to-buy ACC the-flowers she bought

‘As for buying the flowers, she bought.’ (Landau 2006)

In (52), liknot ‘to-buy’ moves to v and v moves to T, as illustrated in (53), yielding a headless vP. If the resulting headless vP moves to SpecCP and the copies of liknot ‘to-buy’
are pronounced both in the head and the tail of the chain, the sentence in (52) is derived. Both copies of the verb are pronounced for phonological reasons (see Landau 2006 for details).

\[(53) \text{[CP } \text{vP t}_\text{she} \text{ [t}_\text{v} \text{ [VP t}_\text{buy} \text{ the-flowers } ] ] [C'} \text{C [TP she [T'} \text{T buy-v-T} \text{ t}_\text{vP } ] ] ]\]

Although the head of the fronted vP can be “seen”, this is an instance of headless vP-movement in the sense that the head moves out of the vP. Thus, TG does not hold in Hebrew.

To sum up, Japanese and Hebrew (and possibly Chinese) have constructions that involve headless vP-movement. This means that TG does not hold in these languages. This appears to be problematic to the present account of TG since headless vP-movement should be prohibited due to the ECC.

4.2 Exceptions to Lasnik’s Generalization

Recall the discussion about null objects in English in section 2.2, in which I argued that in order to rule out null object sentences, headless VP-ellipsis has to be prohibited. Thus, the ECC (combined with the assumption that pro is not available in English) can explain the fact that null objects are not allowed in English.

However, there are languages that allow null objects, unlike English. For example, the following languages allow null objects (and/or null VP internal elements): Japanese, Chinese, Korean (Huang 1988, 1991, Otani and Whitman 1991), Hebrew (Doron and Heycock 1999, Goldberg 2005), Swahili (Goldberg 2005), Persian (Sailor 2010), Serbo-Croatian (Lasnik 1997), and Irish (McCloskey 1991, Goldberg 2005). Some examples are given below:
    John-TOP self-GEN letter-ACC discarded Mary-also discarded
    (Lit.) ‘John threw out letters of himself. Mary also threw out’ [Japanese (Otani
    and Whitman 1991 : 346-7)]

(55) a. Q: (Ha-’im) Miryam hisi’a et Dvora a-makolet?
    Q Miryam drive.PAST3Fsg ACC Dvora to.the-grocery.store
    ‘(Did) Miryam drive Dvora to the grocery store?’
    b. A: Ken, hi hisi’a [e].
       yes she drive.PAST3Fsg
    ‘Yes, she drove Dvora to the grocery store.’ [Hebrew (Goldberg 2005 : 53)]

(56) Juma a-li-beb-a m-toto, na Kumau a-li-beb-a [e] pia.
    Juma 1Su-Past-CARRY-FV 1-child and Kamau 1Su-Past-CARRY-FV too
    ‘Juma carried a child, and Kamau carried a child too.’ [Swahili (Goldberg 2005 :
    57)]

(57) Naysan ketaab-ro ba deghat khoond, Nasim ham khoond [e].
    Naysan book-OBJ with caution read Nasim also read
    ‘Naysan read the book carefully, and Nasim also did (read the book carefully).’
    [Persian (Sailor 2010 : 7)]

(58) Ivan piše rad pažldivo, a njegov asistent čita [e].
    Ivan writes paper carefully and his assistant reads
‘Ivan writes paper carefully and his assistant is reading it carefully.’ [Serbo-Croatian (Lasnik 1997 : 180)]

(59) a. Q: Ar chuir tú isteach air?  INTERR.COMP put.PAST you in.on.it
‘Did you apply for it?’

b. A: Chuir [e].
   put[PAST]
‘Yes, I apply for it.’ [Irish (McCloskey 1991 : 272)]

These languages are radical pro-drop languages, in which objects as well as subjects can be dropped without any agreement on predicates. The authors listed above argue that these languages utilize what I call headless v/VP-ellipsis in order to derive some instances of null object (and/or VP internal elements) sentences. For instance, Goldberg (2005) analyzes the Hebrew sentence in (55b) as having the following underlying structure:

(60) [TP she [T drive-v-T] [vP t_she [t_v [VP ACC-Dvora t_v to.the-grocery.store ] ] ] ]

In (60), V moves to v and v moves to T. As a result of this head movement, a headless vP and a headless VP are derived. If either of these headless phrases is elided, the elliptical sentence in (55b), in which both the direct and indirect object are null, is derived.

If the headless v/VP-ellipsis analysis of null object constructions in these languages is correct, it means that LG does not hold in these languages. However, null object constructions could be analyzed in multiple ways. They could at least be analyzed as constructions involving noun phrase ellipsis, pro, or headless v/VP-ellipsis. Goldberg (2005) provides
evidence that some instance of null object sentences in Hebrew, Swahili, and Irish should be derived by headless v/VP-ellipsis. For example, she argues that the Hebrew sentence that we just considered cannot be analyzed as either pro or noun phrase ellipsis. Consider the following example, which is drawn from Goldberg 2005: 45:

(61) Karmela natna et ha-sefer le-Xagit, ve-Yosef zarak

Karmela give.PAST3Fsg ACC the-book to-Chagit and-Yosef throw.PAST3Msg

et ha-kadur.

ACC the-ball

‘Karmela gave the book to Chagit, and Yosef threw the ball.’

The second conjunct can only mean that Yosef threw the ball but not that Yosef threw the ball to Chagit. This suggests that locative arguments can neither elide nor be pro by themselves in Hebrew. On the other hand, as shown in (55b), a locative argument can be null when a direct object is also null. Thus, the sentence in (55b) cannot be analyzed as either a sentence in which the direct object and the locative phrase are independently elided or a sentence in which two pros appear. Only a headless v/VP-ellipsis analysis can capture the acceptability of this sentence. See Goldberg 2005 for similar arguments for the availability of headless v/VP-ellipsis in Swahili and Irish. Furthermore, see Sailor 2010 and Lasnik 1997 for Persian and Serbo-Croatian, respectively. If their arguments are correct, headless v/VP-ellipsis is allowed in Hebrew, Swahili, Irish, Persian, and Serbo-Croatian.

The analysis of null object constructions in Japanese is controversial because a number of researchers such as Oku (1998), Saito (2004, 2007), Goldberg (2005), Takahashi
(2008a,b), among many others, argue that null object (or null VP internal element) constructions in these languages should be analyzed as noun phrase (or argument)-ellipsis. As for Japanese, however, Funakoshi (2011a) provides empirical evidence that some instances of the null object (or null VP internal element) constructions in Japanese involve headless v/VP-ellipsis. In what follows, I review one of those arguments that Japanese allows headless v/VP-ellipsis.

In Japanese, a focus particle -dake ‘only’ can attach to a PP. It has been observed that dake can appear either after a postposition or between the complement of a postposition and the postposition, as shown in the following:

(62)  
   a. Mary-to-dake ‘Mary-with-only’, Mary-dake-to ‘Mary-only-with’  
   b. Mary-ni-dake ‘Mary-to-only’, Mary-dake-ni ‘Mary-only-to’  
   c. basu-de-dake ‘bus-with-only’, basu-dake-de ‘bus-only-with’

Shoji (1986) observes that the order between -dake and a postposition affects the scope pattern of -dake. When -dake is before a postposition (hereafter the internal dake), -dake can take narrow scope under -(rar)e ‘can’, as shown in (63a). In contrast, when -dake is after a postposition (hereafter the external dake), the sentence only allows the wide scope of -dake, as shown in (63b) (drawn from Funakoshi 2011a: 66).²²

(63)  

   John-TOP Mary-only-with play-can-PRES

   ‘John can play only with Mary’ (only > can, can > only)

John-TOP Mary-with-only play-can-PRES

‘John can play only with Mary’ (only > can, *can > only)

The sentence in (63a) can mean either that the only person who John can play with is Mary (he cannot play with others: the only > can reading) or that John can play with Mary alone (without playing with others: the can > only reading). Only the former reading is available in (63b).

Shoji (1986) accounts for this difference between the internal dake and the external dake by assuming that PP with the external dake is forced to move to a focus position outside of vP, leading to the wide scope reading of -dake while PP with the internal dake can stay in-situ, resulting in the narrow scope reading of -dake.

With this in mind, consider the following examples (drawn from Funakoshi 2011a: 67):

(64) John-wa Mary-dake-to asob-e-ru. Bill-mo [e] asob-e-ru

John-TOP Mary-only-with play-can-PRES Bill-also play-can-PRES

(Lit.)‘John can play only with Mary. Bill can play, too.’ (*only > can, can > only)

The sentence in (64) involves a null PP whose antecedent is PP with the internal dake. Crucially, this sentence only allows the narrow scope reading of dake. More strikingly, sentences involving null PPs that take as their antecedents PPs with the external dake are totally unacceptable, as shown in the following (drawn from Funakoshi 2011a: 67):
These null PP sentences cannot be considered to involve pro because pro is DP but not PP. Then, these sentences should be analyzed either as PP-ellipsis sentences or headless v/VP-ellipsis sentences. Suppose, following Shoji 1986, that PPs with the external dake must move out of vP while PPs with the internal dake can stay within vP. Then, the headless v/VP-ellipsis analysis can account for the fact that the null PPs with the internal dake must take scope under can and the null PPs with the external dake are unacceptable.

Suppose that movement of dake-phrases is overt movement to SpecFocP in the vP periphery (cf. Miyoshi and Hoshi 2007). Then, the internal dake sentence in (64) could have the following two representations since PPs with the internal dake can either stay within vP or move out of vP:

(66) a. Bill-also [\{v/VP Mary-only-with \{\n/\}_{\text{play}}\} play-can.

b. Bill-also [\{FocP Mary-only-with \{v/VP \}_{\text{play}}\} ] play-can.

As shown in (66a) and (66b), headless v/VP-ellipsis yields the null PP sentences only when the PP stays within vP. Therefore, under the headless v/VP-ellipsis analysis, the lack of the wide scope reading of null PP with the internal dake can be explained. On the other hand, this fact remains a mystery under the PP-ellipsis analysis.

Under the headless v/VP-ellipsis analysis the null PP sentences with the external dake in (65) only have a structure like (67) since PPs with the external dake must move out of
Thus, the headless v/VP-ellipsis cannot derive the null PP sentence. On the other hand, if PP-ellipsis is possible, this fact remains a mystery. In this way, these null PP sentences suggest that headless v/VP-ellipsis is available in Japanese. See Huang 1988, 1991, Otani and Whitman 1991, and Takahashi 2008a for arguments for headless v/VP-ellipsis analysis of null objects in Korean and Chinese.

In addition to the above languages, Chickasaw (Munro 1999), Thai, Indonesian, Vietnamese (Simpson 2005, Huang 2000), and Hungarian (Huang 2000) are also radical pro drop languages (in particular, they allow null object sentences). If at least some null object sentences in these languages must be analyzed as headless v/VP-ellipsis, like the languages mentioned above, LG does not hold in these languages either.

To summarize this section, we have seen that TG does not hold in Japanese and Hebrew (and possibly Chinese) and that LG does not hold in Japanese, Chinese, Korean, Hebrew, Swahili, Persian, Serbo-Croatian, and Irish (and possibly Chickasaw, Thai, Indonesian, Vietnamese, and Hungarian too). Interestingly, Japanese and Hebrew (and Chinese) do not obey either TG or LG. This fact suggests that these generalizations are rooted in the same principle, supporting the unified account of these generalizations that I propose. The question of whether the other languages in which LG does not hold also do not obey TG remains for future research.

The question that I have to answer now is why the generalizations hold in some lan-
guages such as English but not in others like Hebrew. In the next section, I answer this question by arguing that TG and LG only hold in languages that do not allow head movement via substitution (they only allow head movement via adjunction).

5 Explanation for the Cross-linguistic Variation

5.1 "Two Types of Head Movement" Hypothesis

It has been generally assumed (with the exception of Rizzi and Roberts 1989, Roberts 1991, 1994) that head movement, unlike phrasal movement, has only one mode, adjunction. However, I argue that there are two types of head movement, proposing the following hypothesis:

(68) "Two Types of Head Movement" Hypothesis

UG allows head movement via substitution (HMS) as well as head movement via adjunction (HMA).

More specifically, my claim is that although HMA is basically always available, HMS is also an available option under certain circumstances. Thus, either HMA or HMS can occur when HMS can occur.

The two types of head movement are illustrated in the following schematic structures:

(69) Head Movement of X to Y via Adjunction

```
       YP
      /  \
     Y    XP
    /  \
   X    tX
    \   /  \
     Y  ZP
```
In (69), X moves to Y by adjunction, which is standard head movement. On the other hand, in (70), X moves to Y but via substitution, projecting Y into Y'. HMS described in (70) corresponds to Pesetsky’s (2007) Under-Merge and Uriagereka’s (2010) Sub-Merge (Pesetsky 2007, Uriagereka 2010). Takano (2007) also proposes the same type of operation for phrasal movement (complement-forming movement in his terms).

I argue that (68) is the null hypothesis. UG has two independently motivated structure-building operations, as Chomsky (1995b, 2000, 2004) argues, adjunction/substitution in GB-terms, Pair-merge/Set-Merge in Chomsky’s (2000, 2004) terms. Thus, unless we make stipulations to the contrary, we should postulate the two types of the operations not only for phrasal movement but also for head movement.

However, one might wonder if the configuration that HMS creates would violate any principles. Before seeing how the hypothesis in (68) can capture the cross-linguistic difference in question, I would like to address this question. First, HMS of X to Y in (70) results in the complement of Y being a head (X) rather than maximal projection. This is problematic under X-bar theory since it requires that every complement and specifier be a maximal projection (cf. Chomsky 1986). However, under the Bare Phrase Structure theory, this ceases to be problematic since under this theory, notions like minimal projections (heads) and maximal projections (phrases) are considered entirely relational notions:
Given a phrase marker, a category that does not project any further is a maximal projection XP and one that does not project at all is a minimal projection $X^0$; any other is an $X' \langle \ldots \rangle$. (Chomsky 1995a: 61)

In (70), the higher copy of $X$ is simultaneously minimal and maximal since it does not project further and does not project at all even if the lower copy of $X$ is a minimal projection. Therefore, (70) is unproblematic under the Bare Phrase Structure theory. Although I have used X-bar theoretic notions like adjunction, substitution, categories, and segments, the present system is compatible with the Bare Phrase Structure theory, as we will see in section 5.3.

Second, in (70), HMS of $X$ leads to the situation where the original complement of $Y$, XP, becomes a specifier of $Y$, given a relational definition of complement and specifier:

(72) The head-complement relation is the ‘most local’ relation of XP to a terminal head $Y$, all others within $YP$ being head-specifier (apart from adjunction $\langle \ldots \rangle$). (Chomsky 1995a : 63)

Thus, the potential problem with the structure in (70) is that XP becomes a specifier of $Y$ even though $Y$ subcategorizes for XP. However, as Takano (2007) argues, this is unproblematic if we adopt the derivational approach to syntactic relations (cf. Epstein et al. 1998). Under the derivational approach, $Y$’s subcategorization requirement is derivationally satisfied when $Y$ is merged with XP before HMS of $X$ to $Y$. Therefore, HMS of $X$ to $Y$ does not affect the relation already derivationally satisfied between $Y$ and XP. In fact, as I will show in section 5.2, the fact that HMS turns an original complement into a derived specifier
plays a crucial role for a reduction of the cross-linguistic variation in the applicability of
TG and LG to a multiple-Spec parameter.

Finally, HMS creates a configuration in which the moved head does not c-command
its original position, violating both the Proper Binding Condition (PBC) and the Extension
Condition. The former requires that traces be bound (Fiengo 1977, May 1977, Saito 1989)
and the later states that Merge must always expand the tree. As for the PBC, Takano
(2007) points out that its empirical validity has been questioned, based on examples like
the following:28

(73) [How t₁ likely to win]₂ is John₁t₂?

As for the Extension Condition, as Bošković and Lasnik (1999), Lasnik (2006), Richards
(1999), and Richards (2004) point out, all the derivations that the Extension Condition
rules out can also be ruled out by Featural Cyclicity, which requires that a strong feature
be checked as soon as possible after being introduced into the derivation (Richards 1999:
127). Given this redundancy and empirical evidence for Featural Cyclicity over the Ex-
tension Condition, these authors argue that we should adopt the former, dispensing with
the latter. Notice that HMS of X to Y in (70) does not violate Featural Cyclicity if we,
following Chomsky 1995b, interpret “as soon as possible” in its definition as before the
maximal projection of the head bearing the strong feature gets embedded. Therefore, HMS
is unproblematic under the model in which Featural Cyclicity instead of the Extension
Condition is adopted.
Given these discussions, I conclude that (i) the "two types of head movement" hypothesis is conceptually well-motivated since it is the null hypothesis and (ii) it is also theoretically tenable under a certain model of syntax in which the Bare Phrase Structure theory, the derivational approach to syntactic relations, and Featural Cyclicity are adopted. Furthermore, in what follows, I argue that this hypothesis is also empirically motivated by illustrating that we can explain why some languages like English do not allow headless XP-movement/ellipsis while other languages like Hebrew do.

Notice that the explanation for TG/LG given in section 3 was solely based on the structure in which HMA occurs (see the structure in (16), repeated below as (74)).

(74)

In this configuration, XP does not structurally intervene between $\alpha$ and X nor does X structurally intervene between $\alpha$ and XP since there is no category that dominates XP/X but does not dominate X/XP. Thus, $\text{AD}(\alpha)$ contains both XP and X. Then headless XP movement is ruled out by the ECC since the smaller element X blocks Copy(XP).

Given the "two types of head movement" hypothesis in (68), the question that arises is whether the same explanation for TG/LG can also apply to the structure in which HMS occurs. In fact, this is not the case. To see this, consider the following schematic structure, in which X head-moves to Y but in this case, the head movement is performed via substitution:
Notice that in this case, X is not in AD(α) since XP intervenes between α and X: (i) α c-commands XP, (ii) XP matches with α and X with respect to [F], (iii) every category that dominates XP also dominates X, (iv) and some category (Y’) that dominates X does not dominate XP. Therefore, Copy(XP) is allowed even if X is smaller than XP. Given this consideration, the following condition follows:

(76) XP-movement/ellipsis can be headless only if X head-moves to Y, the next higher head, via substitution.

To see how (76) works, let us see more concrete examples. For example, if headless vP-movement/ellipsis is allowed in a language, v should be able to head-move to T via substitution in the language, as illustrated in (77).

(77)
In (77), V head-moves to v via adjunction and then the [V-v] complex head-moves to T via substitution. In this configuration, vP intervenes between v and a probe α. Thus, v cannot be a member of AD(α), allowing headless vP-movement/ellipsis even if v is smaller than vP.29

Likewise, if a language allows headless VP-movement/ellipsis, HMS of V to v is allowed in the language, as illustrated in (78). In (78), VP structurally intervenes between α and V. Thus, V is not in AD(α) and headless VP-movement/ellipsis is allowed.

(78)

Before proceeding to the next section, a remark is in order with regard to the Phase Impenetrability Condition (PIC) (Chomsky 2000). I don’t assume the ‘standard’ version of the PIC that states that a complement domain of a phase head is inaccessible to operations outside the phase. This version of the PIC prohibits the headless VP from moving in (78) since VP is the complement of v, a phase head. Therefore, I adopt a slightly weaker version of the PIC that is proposed by Chomsky (2001). According to this version, a complement domain of a phase head is inaccessible to operations at the next higher phase. Given this, VP-movement in (78) is unproblematic since the next higher phase, CP, has not been com-
pleted. See Asarina 2011 for independent evidence in favor of this weaker version of the PIC over the stronger version.

5.2 Multiple-Specs

Given the condition in (76), the question that immediately arises concerns when HMS is allowed. I argue that the (im)possibility of HMS of X to Y is reduced to the possible number of specifiers that Y can take. To see this, recall that HMS turns the original complement into a derived specifier, as I mentioned in section 5.1. Given this, HMS of X to Y creates the multiple-Spec configuration if Y has another (original) Spec. This is illustrated in (79).

(79) \[
\begin{array}{c}
\text{YP} \\
\text{WP} \\
\text{Y'} \\
\text{XP} \\
\text{X} \quad \text{Y} \\
\text{ZP}
\end{array}
\]

In (79), X head-moves to Y via substitution, turning the former complement XP into a specifier of Y. Note that in (79), Y has another Spec, WP. Thus, Y has multiple-Specs in this structure, WP and XP. If Y does not allow multiple-Specs, the configuration in (79) should be ruled out. Given these considerations, the following condition follows: 30

(80) Head movement of X to Y, the next higher head, can be performed via substitution only if either

a. Y allows multiple specifiers or

b. Y does not have an element in its Spec.
For example, let us consider when HMS of \( v \) to T is possible. When a language allows multiple TP Specs, \( v \) can head-move to T via substitution in the language even if T has a subject in its Spec, as illustrated in the following:

\[(81) \]

(81)

```
TP
  \-------- Subj
  \-------- T'
    \-- vP
      \-- T
        \-- T_{subj}
          \-- v'
            \-- VP
              \-- t_v
                \-- Obj
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
```

On the other hand, if a subject does not move to SpecTP in a language, HMS of \( v \) to T is possible in the language even if T does not allow multiple-Specs. This is illustrated in (82).

\[(82) \]

(82)

```
TP
  \-------- T'
    \-- vP
      \-- T
        \-- Subj
          \-- v'
            \-- VP
              \-- t_v
                \-- Obj
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
    \-- t_v
```

Combining this condition in (80) with the condition on headless XP-movement/ellipsis in (76), the following condition obtains by transitivity:

\[(83) \]

(83) XP-movement/ellipsis can be headless only if either

a. Y, the next higher head, allows multiple-Specs or

b. Y does not have an element in its Spec.
Given this condition, we expect (i) that in a language in which headless v/VP-movement/ellipsis is allowed, either multiple TP/vP Specs are allowed or subjects do not move to SpecTP and (ii) that in a language in which neither multiple TP/vP Specs are not allowed and SpecTP cannot be empty, headless v/VP-movement/ellipsis are not allowed.

The latter prediction is borne out in languages like English. English allows neither multiple TP/vP Specs nor empty SpecTP and does not allow headless v/VP-movement/ellipsis, as we saw in section 2.  

The former prediction is confirmed in the languages where TG and LG do not hold. Following Ura 1994, I assume that T allows multiple-Specs in languages that have multiple subject constructions. Japanese, Korean, Chinese, Persian, Indonesian, Vietnamese (Ura 1994), Hebrew (Doron and Heycock 1999), Thai (Kumashiro and Langacker 2003), and Chickasaw (Munro 1999) have multiple subject constructions. Some examples are given below:

(84) Zoo-ga hana-ga nagai.
    elephant-NOM nose-NOM long
    'Elephants’ noses are long.'  
    [Japanese (Ura 1994 : 34)]

(85) Ruti yeS la savlanut.
    Ruti there-is to-her patience
    'Ruti has patience.'  
    [Hebrew (Doron and Heycock 1999 : 71)]

(86) Muhmud ketâb-aš gom ŝod.
    Mahmud-NOM book-3sg-NOM got lost
'It is Mahmud that his book got lost.' [Persian (Ura 1994 : 33)]

(87) Chán tháaw too
I foot big
‘I have a big feet.’ [Thai (Kumashiro and Langacker 2003 : 2)]

(88) Hattak-at ofi’-a im-alhkania.
man-NOM dog-NOM DAT-forget
‘The man forgets the dog.’ [Chickasaw (Munro 1999: 254)]

Furthermore, according to Richards (1997), also in Serbo-Croatian and Hungarian T allows multiple-Specs (IP absorption languages in his terms). Lastly, based on his analysis of possessor raising, Ura (1994) argues that in Swahili and Korean V (\(v\) under the phrase structure that we adopt in this paper) allows multiple-Specs.

If this is correct, these languages, which allow headless \(v/VP\)-movement/ellipsis, as I argued in the last section, allow multiple \(vP\) Specs and/or multiple TP Specs. This is what my analysis predicts.

Irish, which allows headless \(vP\)-ellipsis, can be considered a subject in-situ language in accord with what (83) predicts. Irish has a VSO basic order in finite clauses. McCloskey (1991) argues that VSO order in Irish is derived by verb movement to T and subject in-situ.\(^{32}\)

Furthermore, it is expected that in Irish, subjects are obligatorily elided along with objects (i.e. an object cannot be elided by itself). To see this, let us consider the structure in (82) above. In (82), \(v\) head-moves to T via substitution and Subj stays in Spec\(vP\). Thus,
headless \(\nu P\)-ellipsis is allowed. Notice that \(\nu P\)-ellipsis in (82) is impossible due to the ECC since \(\nu P\) does not intervene between the probe (Top above TP) and V. Thus, if \(pro\) is not available for null objects in Irish, a headless \(\nu P\) must be elided in order to derive a null object sentence. The \(\nu P\) contains an in-situ subject. Therefore, it is predicted that subjects are obligatorily elided along with objects in Irish. This is indeed the case (cf. McCloskey 1991 and Goldberg 2005). In the Irish example in (59), the subject as well as the object is elided.

The present analysis makes a prediction about island effects. When a head X moves to Y via substitution, the complement XP becomes a specifier of Y. This means that elided headless XPs are always in the specifier position since XP-ellipsis can be headless only if X head-moves to Y via substitution (see (76)). Given that specifiers are islands (cf. Huang 1982, Nunes and Uriagereka 2000), it is predicted that elements in elided headless XPs cannot be extracted. This prediction is indeed borne out in Japanese.

In Japanese, a complement CP can be elided, as shown in the following example:

(89) Mary-wa [John-ga ringo-o tabeta to] omotteiru, kedo Sue-wa
     Mary-TOP John-NOM apple-ACC ate COMP think but Sue-TOP
     \(\underline{John-ga \ ringo-o \ tabeta to}\) omottei-nai.
     John-NOM apple-ACC ate COMP think-NEG
     ‘Mary thinks John ate an apple, but Sue does not think John ate an apple.’

In the second conjunct of the sentence in (89), the complement clause is elided. According to headless \(\nu P\)-ellipsis analysis of null objects, this is derived by applying headless ellipsis.
to the matrix vP that contains the complement clause. As shown in (90), the matrix verb moves out of the matrix vP.

(90)  Sue-TOP [vP [CP John-NOM apple-ACC ate COMP] tthink] think-NEG

If the resulting headless vP is elided, sentence (89) is derived.

Moreover, Japanese has long distance scrambling, as shown in (91).

(91)  Ringo-o Mary-wa [John-ga t1 tabeta to] omotteiru.
      apple-ACC Mary-NOM John-NOM ate COMP think

‘Mary thinks John ate apple.’

In (91), the object in the embedded clause, ringo-o ‘apple-ACC’, is scrambled out of the embedded clause. However, as Tanaka (2008) observes, scrambling out of an elided complement CP is prohibited, as shown in (92).

(92)  *Ringo-o1 Mary-wa [vP [John-ga t1 tabeta to] t2] omotteiru, kedo
      apple-ACC Mary-TOP John-NOM ate COMP think but
      orange-ACC Sue-TOP John-NOM ate COMP think-NEG

‘Mary thinks John ate an apple, but Sue does not think John ate an orange.’

In the second conjunct in (92), the object in the embedded clause, orenzi-o ‘orange-ACC’ is scrambled out of the headless vP, which is elided, yielding an unacceptable sentence. This is what the present analysis of headless XP-ellipsis predicts, lending it strong support. Notice that the island effect does not obtain without ellipsis in (92). This is expected under the present analysis because HMA is always possible when HMS is available. HMS is
forced only when headless XP-movement/ellipsis has taken place as in (92).  

To sum up, from the “two types of head movement” hypothesis and my analysis of headless XP-movement/ellipsis, it follows that the availability of headless XP-movement/ellipsis is related to the availability of multiple-Specs: the former entails the latter. Japanese, Korean, Chinese, Persian, Hebrew, Swahili, Serbo-Croatian, and Irish allow some instances of headless XP-movement and/or ellipsis. Indonesian, Vietnamese, Thai, Hungarian, Chickasaw are radical pro drop languages, which might be analyzed as involving headless v/VP-ellipsis. All of these languages except Irish allow some instances of multiple-Specs (multiple vP Specs and/or multiple TP Specs). Irish allows SpecTP to be empty. On the other hand, languages like English allow neither multiple TP/vP Specs nor empty SpecTP. Thus, they do not allow headless v/VP-movement/ellipsis either. In this way, we can attribute the cross-linguistic variation in the availability of headless XP-movement/ellipsis to the parameters that are responsible for the availability of multiple Specs.

5.3 Bare Phrase Structure Theory and Unbounded Merge

Before closing this section, a remark is in order with regard to X-bar theoretic notions like adjunction, substitution, segments and categories, which I have used through this paper. It is not very clear how to define the distinction between segments and categories and that between substitution and adjunction under a current theory of Bare Phrase Structure (BPS). However, this does not mean that my analysis is incompatible with BPS. In this subsection, I suggest one of the possible ways to reconcile the present analysis with BPS.

The crucial assumption adopted in this paper is that a node that is created as a result
of adjunction of X to Y does not dominate X while a node that is created as result of substitution of X to Y does. The first one is May’s theorem that “adjuncts are not dominated by the categories to which they are adjoined” (May 1989: 92). This is a defining property of adjunction. If there are at least two modes for the structure-building operation, that is, substitution/Set-Merge and adjunction/Pair-Merge, as Chomsky (1995a, 2000, 2004) argues, then any phrase structure theory must be able to capture this property in one way or another.

One promising way of capturing this property of adjunction under BPS theory is to adopt a proposal developed by Chametzky (2000), Hornstein and Nunes (2008), Hornstein (2009), and Uriagereka (2002). A core idea of their proposals is that substitution yields a labeled constituent while adjunction yields a non-labeled constituent. I call the former operation L(abeled)-Merge and the latter operation Concatenate. Following Hornstein 2009 and Hornstein and Nunes 2008, I assume that Concatenate (adjunction) is a simple operation that combines two syntactic objects while L-Merge, which is the same operation as Chomsky’s Set-Merge, is a composite operation consisting of Concatenate and Label. That is, an operation whereby one of the two inputs to Concatenate names the resulting constituent. This is illustrated in the following (the underlined symbol is a label): 35

(93)    a. Concatenate(X, Y) = \{X, Y\} 

        b. L-Merge(X, Y) = Label(Concatenate(X, Y)) = \{X, \{X, Y\}\}

Under this version of BPS theory, the structures created by the two types of head movement can be derived in the following way, using Concatenate and L-Merge (the subscript
numbers are just for expository purposes):

(94)  *Head Movement of X to Y via Concatenate (Adjunction)*

a.  L-Merge(X, ZP): \{X, \{X, ZP\}\} (= XP)

b.  L-Merge(Y, XP): \{Y, \{X, \{X, ZP\}\}\}

c.  Copy(X)

d.  Concatenate(Y, X): \{Y, \{X, \{X, ZP\}\}\} (= YP)

e.  L-Merge(\alpha, YP): \{\alpha, \{\alpha, \{Y, \{X, \{X, ZP\}\}\}\}\}

(95)  *Head Movement of X to Y via L-Merge (Substitution)*

a.  L-Merge(X, ZP): \{X, \{X, ZP\}\} (= XP)

b.  L-Merge(Y, XP): \{Y, \{X, \{X, ZP\}\}\}

c.  Copy(X)

d.  L-Merge(Y, X): \{Y, \{X, \{X, ZP\}\}\} (= YP)

e.  L-Merge(\alpha, YP): \{\alpha, \{\alpha, \{Y, \{Y, X\}\}, \{X, \{X, ZP\}\}\}\}

Under this framework, the difference between HMA and HMS is that in the former, head movement is realized by Concatenate, combining two heads without applying Label to the result while in the later, this is done by L-Merge, combining two heads followed by Label.

Given this, let us consider how we can define the notion of dominance based on set theoretic notations of phrase structure, which BPS theory assumes. In a tree diagram, dominators are always nodes (or a certain kind of nodes, namely categories) and dominatees are everything under the dominator node. What corresponds to nodes in set theoretic notations like (94) and (95) is a label. Given this consideration, it seems natural to define the
dominance relation as in the following:

\[ \text{(irreflexive) Dominance} \]
\[ \alpha \text{ dominates } \beta \ \text{iff } \alpha \text{ is a label and} \]
\[ \text{a. } \beta \text{ is a sister of } \alpha \text{ or} \]
\[ \text{b. } \beta \text{ is a descendant of a sister of } \alpha. \]

The crucial point is that only labels can be a dominator in this definition. Sisters and descendants are defined as follows:

\[ \text{Sister} \]
\[ \beta \text{ is a sister of } \alpha \ \text{iff } \beta \text{ is another member of a set } A \text{ such that } \alpha \text{ is a member of } A. \]

\[ \text{Descendant} \]
\[ \text{a. A member of } \gamma \text{ is a descendant of } \gamma. \]
\[ \text{b. If } \delta \text{ is a descendant of } \gamma, \text{ then a member of } \delta \text{ is also a descendant of } \gamma. \]

Given these definitions, it follows that Copy of a headless XP is ruled out by the ECC when HMA of X to Y occurs while it is not when HMS of X to Y occurs. In (94), neither X₂, which underwent HMA, nor \{X, \{X₁, ZP\}\}, which is a headless XP, structurally intervene between the probe \(\alpha\) and the other. What dominates X₂ are \(\alpha\) and \(\underline{Y}\) (X₂ is a descendant of both the sister of \(\alpha\) and the sister of \(\underline{Y}\)). What dominates \{X, \{X₁, ZP\}\} are also \(\alpha\) and \(\underline{Y}\). Therefore, there is nothing that dominates one of them and does not dominate the other.

This means that AD(α) contains both X₂ and \{X, \{X₁, ZP\}\}. Then the ECC prevents Copy from applying to \{X, \{X₁, ZP\}\} since it is bigger than X₂.

In the representation in (95), derived by HMS of X to Y, the dominators for X₂ are \(\alpha\).
\(Y_2\), and \(Y_1\) while those for \(\{X, \{X_1, ZP\}\}\) are only \(\alpha\) and \(Y_2\): \(\{X, \{X_1, ZP\}\}\) is neither a sister nor a descendant of a sister of \(Y_1\). Therefore, \(\{X, \{X_1, ZP\}\}\) structurally intervenes between the probe \(\alpha\) and \(X_2\). As a result, \(AD(\alpha)\) does not contain \(X_2\), and Copy is able to apply to \(\{X, \{X_1, ZP\}\}\).

In this way, it is possible to make the present analysis compatible with BPS theory without using X-bar theoretic notions nor notations. Although the specific implementation might turn out to be wrong, the important point that I would like to make here is that the present analysis is not inherently incompatible with BPS theory.

Finally, I would like to point out that under this model of BPS, we can reduce a multiple-Spec parameter to a parameter that has to do with the restriction on the number of applications of L-Merge, without resorting to the notion of specifiers. Suppose that UG has the following parameter:

\[(99)\]
\[\begin{array}{ll}
\text{a. } \text{Languages like English: L-Merge can apply to a single head at most two times.} \\
\text{b. } \text{Languages like Japanese: L-Merge can apply to a single head without limit.}
\end{array}\]

Given this parameter, both English and Japanese allow the following structure, in which \(X\) takes a complement, a Spec, and an adjunct:

\[(100)\]
\[\{WP, \{X, \{YP, \{X, \{X, ZP\}\}\}\}\}\] (Spec = YP, Compl = ZP, Adjunct = WP)

\[\begin{array}{ll}
\text{a. } \text{L-Merge}(X, ZP): \{X, \{X, ZP\}\} (=A) \\
\text{b. } \text{L-Merge}(A, YP): \{X, \{YP, \{X, \{X, ZP\}\}\}\} (=B) \\
\text{c. } \text{Concatenate}(B, WP): \{WP, \{X, \{YP, \{X, \{X, ZP\}\}\}\}\}\}
\end{array}\]
This is because L-Merge applies to a single head X only two times: the third operation is Concatenate rather than L-Merge. On the other hand, only languages like Japanese allow the following structure, where X takes a complement and two Specs:

\[(101) \quad \{X, \{WP, \{X, \{YP, \{X, \{X, ZP\}\}\}\}\}\}\] (Specs = WP and YP, Compl = ZP)

- a. L-Merge(X, ZP): \(\{X, \{X, ZP\}\}\) (=A)
- b. L-Merge(A, YP): \(\{X, \{YP, \{X, \{X, ZP\}\}\}\}\) (=B)
- c. L-Merge(B, WP): \(\{X, \{WP, \{X, \{YP, \{X, \{X, ZP\}\}\}\}\}\}\)

This is because L-Merge applies to X more than two times. The difference between L-Merge and Concatenate in this respect (i.e. the application of the former is restricted while that of the latter is free) seems natural since Concatenate is a general cognitive operation not specific to language while L-Merge involves Label, which is a language-specific operation.

The application of L-Merge seems more costly than that of Concatenate in a sense.

Under this conception of the multiple-Spec parameter, when a head Y has an element in its Spec, HMA of X to Y is allowed in both languages like English and Japanese while HMS of X to Y is allowed only in languages like Japanese, as illustrated below:

\[(102) \quad \{Y, \{WP, \{Y, \{X, \{X, ZP\}\}\}\}\}\]

- a. Copy(X) & Concatenate(Y, X) \hspace{1cm} (HMA of X to Y)
  \(\{Y, \{WP, \{Y, \{X, \{X, ZP\}\}\}\}\}\)
- b. Copy(X) & L-Merge(Y, X) \hspace{1cm} (HMS of X to Y)
  \(\{Y, \{WP, \{Y, \{Y, \{X, ZP\}\}\}\}\}\)

Before head movement takes place, as shown in (102a), L-Merge applies to Y two times,
leaving Y with a Spec WP and a complement \{X, \{X, ZP\}\}. Therefore, HMA of X to Y can take place in both types of languages while HMS of X to Y can take place only in the Japanese type languages: HMS of X to Y requires the third L-Merge apply to Y.

This view of the multiple-Specs parameter fits well in Fukui’s (2011) conception of multiple-Specs or multiple subjects (see footnote 28). Fukui (2011) argues that a number of differences between English and Japanese, including the availability of multiple subjects, can be reduced to a more fundamental difference between them that “‘unbounded Merge’ is in full force in the syntax of Japanese” (Fukui 2011: 90) while Merge is bounded somehow in English. If this is correct, we can conclude that the cross-linguistic difference in the applicability of TG and LG is eventually reduced to the availability of unbounded (L-)Merge.

6 Concluding Remarks

In this paper, I have proposed two things: the economy condition on Copy and the ”two types of head movement” hypothesis. It was argued that (i) the former gave an explanation to TG and LG and (ii) with the former and the latter, the cross-linguistic variation in the applicability of these generalizations can be attributed to parameters that are responsible for the availability of multiple Specs. Furthermore, I have suggested that the availability of multiple Specs can be reduced to the availability of unbounded (L-)Merge if we adopt a certain model of BPS theory in which Concatenate and L-Merge are available as basic structure-building operations.
References


Aelbrecht, Lobke, and Liliane Haegeman. 2011. VP ellipsis is not licensed by VP topicalization. Ms. GIST University Ghent.


Doron, Edit, and Caroline Heycock. 1999. Filling and licensing multiple specifiers. In


ing and verifying an advanced theory of human language, ed. Kazuo Inoue, 61–84.

Chiba: Kanda University of International Studies.


Ussery, 603–614. GLSA Publications.


Uriagereka, Juan. 2010. To escape an island: submerge and resurface. Handout of a talk at Formal Approaches to Slavic Linguistics 19.


Department of Linguistics, University of Maryland

1401 Marie Mount Hall, College Park, Maryland 20742-7505

funakosh@umd.edu

Notes

An earlier version of the material in this article was presented to audiences at Kwansei Gakuin University, Osaka University, and the University of Maryland. I would like to express my thanks for insightful comments and suggestions by Tonia Bleam, Chris LaTerza, Masao Ochi, Masahiko Takahashi, Hiroyuki Tanaka, and Hiroyuki Ura. I am especially grateful to Alex Drummond, Sayaka Goto, Norbert Hornstein, Bradley Larson, Howard Lasnik, Jeffrey Lidz, Terje Lohndal, Yuji Takano, and two *LI* reviewers for reading earlier versions, and making valuable suggestions for improvement. This work was partially supported by a Fulbright grant.

1Lasnik (1999), however, immediately rejects the generalization on the basis of the fact
that there are languages in which V is moved out of VP, while VP-ellipsis is allowed. I will account for exceptions to LG in section 4.

Abbreviations: ACC = Accusative, CL = Classifier, COMP = Complementizer, DAT = Dative, F = Feminine, FOC = Focus, FV = Final Vowel, GEN = Genitive, M = Masculine, NEG = Negation, NOM = Nominative, Obj = Object, PRES = Present, PRG = Progressive, PAST = Past, SG = Singular, Subj = Subject, TOP = Topic Marker, 3 = 3rd person

As a reviewer correctly points out, although the English example in (6c) is consistent with TG, it does not support TG. This is because this sentence might be ruled out independent of TG: clefting of VP is not allowed in English, as shown in the following sentence:

(i) *It’s give a book to Mary that I might.

Given this interfering factor, the following example might be more appropriate to illustrate TG:

(ii) *The book to Mary, John gave.

One of the possible representations of this sentence, in which the book and to Mary independently undergo topicalization, is illicit because multiple topicalization is prohibited in English:

(iii) *[DP the book] [PP to Mary], John gave tDP tPP

However, in order to account for the unacceptability of the sentence in (ii), we need to rule out another possible representation for the sentence, which is given in (iii a).

(iii) a. *[VP The book t1 to Mary]2, John gave1 t2.

b. [VP Give the book to Mary], John did t.
As shown in (iiia), sentences like (ii) can be derived by headless VP-movement. Given that vP-topicalization is allowed in English, as shown in (iiib), it is unclear why (iiib) is acceptable while (iii) is not. TG can accommodate this fact since VP-topicalization in (iii) involves headless XP-movement.

4 Takano (2000) gives an explanation for his generalization, which is based on Chomsky’s (1995b: 304) condition that only the head of a chain enters into the operation Attract/Move. Although this can account for TG, the status of Chomsky’s condition is unclear under the Copy Theory of Movement. If a trace is indeed a copy of the moved element, it is not clear why a trace (a copy) is invisible for the operation Attract/Move. But more importantly, this approach remains silent about exceptions to TG, which I will discuss in section 4.

5 Kasai (2001: 101), who attributes the observation to Akira Watanabe, points out that there is a counterexample in English to TG.

(i) a. They all said that John was shrewd and shrewd he was.

b. *They all said that John was shrewd and shrewd he looks.

The unacceptability of (ib) indicates that an AP cannot be fronted by themselves. Thus, it must be the case that in (ia), the moved category is a VP rather than an AP. The fronted category in (ia) can be analyzed as a VP if the verb was moves out of the VP and the resulting headless VP is fronted. Notice that the same kind of derivation is impossible in (ib) since V, unlike be, does not move to T. Thus, the acceptability of (ia) could be a counterexample to TG.
However, a reviewer suggested me that the acceptability of (ib) improves if the copula verb in the first conjunct clause is replaced by verbs like *look* or *seem*. Thus, sentences like the following are much better than (ib):

(ii) *They all said that John looked/seemed shrewd and shrewd he looks/seems.*

This suggests that AP can be fronted. Furthermore, the reviewer pointed out that the following sentence makes the same point:

(iii) *They all said that John might be late and late he probably will be.*

Unless non-finite *be* undergoes short head movement, it seems to be the AP that is fronted in (iii). If this is correct, we cannot conclude that it is a headless VP rather than an AP that is fronted in (ia) since the AP is a possible candidate for fronting. I propose another possible explanation of (ia) in Funakoshi (2010).

Roberts (1998) and Potsdam (1997) propose a similar generalization, which amounts to saying that headless VPs cannot antecede VP-ellipsis. Lasnik (1997) rejects their generalization because of the exceptions to it.


This proposal has a precursor in Chomsky’s (1995) principle concerning pied-piping:

(103) *F carries along just enough material for convergence.*  
(Chomsky 1995: 262)

This principle makes the amount of material that is pied-piped (Copied in my terms) mini-
mized just as our economy condition does.

10 This informal explanation and a formal one that I will propose in section 3.1.2 crucially rely on X-bar theoretic notions like categories, segments, and adjunction. This, however, does not mean that I adopt X-bar Theory as a phrase structure theory. As is obvious from seeing that I assumed the Copy Theory in the explanation, I adopt a Bare Phrase Structure Theory in this paper. In section 5.3, I will show that we can reconcile these X-bar theoretic notions with a Bare Phrase Structure Theory if we adopt an idea developed by Chametzky (2000), Hornstein (2009), Hornstein and Nunes (2008), and Uriagereka (2002) to the effect that adjunction is a kind of unlabeled Merge. However, until then, I keep using the X-bar theoretic notions and notations just for expository purposes.

11 A reviewer criticizes the notions of AD and structural intervention in terms of minimalist perspective, pointing out that I bring in additional operations. However, notice that in the definitions of AD and structural intervention, I do not utilize any other operations in addition to standardly assumed operations or relations such as Agree, c-command, domination, and match. Moreover, AD is just a formal implementation of a tacitly assumed view that Agree is a prerequisite for Move. Thus, Agree-as-a-prerequisite-for-Move presupposes some formal apparatus like this.

12 The economy condition in (21) is compatible with Ura’s view on economy that (Generalized) Pied-Piping, like other syntactic operations, is subject to the general economy condition (Ura 2001). This runs counter Tanaka 2004, which argues that (Generalized) Pied-Piping (Copy in my terms) is cost-free, meaning that (Generalized) Pied-Piping is not
constrained by any economy conditions.

As Yuji Takano (pers. comm.) points out to me, it has to be guaranteed that in (24), XP cannot move to the edge of YP in order to maintain the present explanation of TG. This is because if XP were be able to move to the edge of YP, XP would structurally intervene between $\alpha$ and X. As a result, AD($\alpha$) only contains XP. Then, Copy should be able to apply to XP because there is no smaller element than XP in AD($\alpha$). This undesirable, super local movement of XP to the edge of YP can be ruled out as an Anti-Locality effect. In the end of this section, I will show that this kind of Anti-Locality effect can be reduced to the ECC.

As Howard Lasnik (pers. comm.) points out, I cannot say that intermediate projections are completely invisible in syntax because intermediate projections must be visible in terms of the computation of the dominance relation, which will be crucial for my explanation for cross-linguistic variation in applicability of TG and LG, as we will see in section 5. Thus, I had better state that intermediate projections cannot be candidates for syntactic operations such as Agree.

Donati (2003) also explores the hypothesis that ellipsis can be reduced to movement. See also Fitzpatrick (2006), Hornstein (2009), Kayne (2005), Rizzi (2005), and Szczeginiak (2005) for relevant discussion.

I thank a reviewer for drawing my attention to Aelbrecht and Haegeman 2011

Actually, they also suggest two other alternatives in addition to the $\nu$P periphery analysis. See Aelbrecht and Haegeman 2011 for details.

This analysis is also compatible with the following examples that a reviewer provides
as potential counterexamples to Johnson’s analysis:

(i)   a. Did someone call 911? Yes, John did [[call 911] and [(then) waited patiently]].
       b. John’s having been arrested surprised me but [Mary’s having been [arrested]]
           was completely expected.

In (ia), the elided VP is contained in the VP coordination structure. If VP-ellipsis is de-
derived through VP-movement to the CP periphery, as Johnson argues, it is unclear why this
sentence is not ruled out by the Coordinate Structure Constraint. On the other hand, under
the vP periphery analysis, this sentence can be derived without extracting VP out of the
coordinate structure. Suppose that the coordinated categories are TopPs in a vP periphery
in (ia). Then, as illustrated in the following structure, VP-movement to the lower SpecTopP
does not violate the Coordinate Structure Constraint:

(ii)  [[&P [TopP [VP call 911] [ Top [vP vtVP ]] and [TopP Top [vP waited patiently]]]]

Likewise, the sentence in (ib) would violate the Subject Island Constraint under Johnson’s
analysis but not under the vP periphery analysis.

19 The Phase Impenetrability Condition (Chomsky 2000) could redundantly rule out VP-
movement in (45) since VP is the complement of v, a phase head. As I will discuss in
section 5.1, however, I do not adopt this version of the Phase Impenetrability Condition.

20 Top c-commands v since every category that dominates Top also dominates v (TopP)
and Top does not dominate v. Notice that Top does not dominate v because only one
segment of Top dominates v.

21 Takano (2000) does not give an explanation for these exceptions although he suggests
that the constructions in question should be analyzed in a different way. Indeed, Takano (2002) provides an alternative analysis of multiple cleft sentences in Japanese, in which the moved phrase in sentences like (50) is not a headless vP.

22 Shoji (1986) judges sentences like (63a) as unambiguous, allowing only the narrow scope of *-dake*. However, Futagi (2004) finds the sentence ambiguous, allowing both the narrow scope and the wide scope of *-dake*. My judgment agrees with Futagi’s.

23 The undeletability of the external *dake* is observed by Futagi (2004). However, she argues that the internal *dake* cannot be null either. Thus, according to her judgment, sentences like (64) are also unacceptable. However, I have a robust intuition that the sentences are acceptable under the narrow scope reading of *-dake*. All Japanese speakers that I have consulted agreed with my judgment.

24 I need to assume that *dake*-phrase movement is overt because this paper presupposes the PF-deletion analysis of ellipsis rather than the LF-copy analysis (cf. Merchant 2001).

25 In section 5.3, I will demonstrate how to implement this idea in a Bare Phrase Structure Theory without using X-bar theoretic notions like substitution and adjunction. There, substitution and adjunction are reinterpreted into labeled Merge (L-Merge) and unlabeled Merge (Concatenate).

26 The higher copy does not inherit the minimal/maximal status from the lower copy since minimal projections and maximal projections, as I said, are relational notions rather than inherent properties that categories have.

27 As a reviewer points out, this argument does not hold if head movement is derived
through what Bobaljik and Brown (1997) call Interarboreal Movement. This is because under this approach, X has to move to Y before Y is merged with XP, whereby Y’s subcatégorization requirement is satisfied, in order to satisfy the Extension Condition (Chomsky 1993). However, as I discuss in the following discussion, I adopt Featural Cyclicity (Chomsky 1995c, Richards 1999) as a way of deriving cyclicity, discarding the Extension Condition. Therefore, I do not have to assume that head movement is derived through Interarboreal Movement.

Lasnik and Saito (1992), however, argue that the sentence in (73) does not necessarily constitute a counterexample to the PBC because likely can ambiguously take a raising complement and a control complement. They provide the following example as evidence for their claim, attributing the observation to Anthony Kroch:

(i) *[how likely t₁ to be a riot]₂ is there₁ t₂

In this sentence, the matrix subject is expletive there. Thus, likely in this sentence, unlike the one in (73), is unambiguously a raising verb. See Müller 1996 and Takano 1994 for relevant discussion.

Although I am agnostic on linearization mechanisms in this paper, it might be the case that I cannot adopt the Linear Correspondence Axiom (LCA: Kayne 1994) or other mechanism which predicts that Specs always precede Heads. According to the present analysis, SVO languages like Hebrew allow structures like (77), in which v head-moves to T via substitution. The choice between HMA and HMS is completely optional when HMS is available. It is then wrongly predicted that objects in Hebrew can optionally precede or
follow verbs if Specs always precede Heads. That is, vP containing an object becomes a specifier of T after HMS of v to T applies. I thank Alex Drummond, Yuji Takano, and a reviewer for bringing this issue to my attention.

A reviewer correctly points out that this does not follow if we adopt a certain view on a multiple-Spec parameter according to which a ban on multiple-Specs for a head H in a language L is reduced to the fact that H in L is not allowed to have the features that would introduce more than one Spec. For example, under Ura’s (1994, 2000) conception, the head that allows multiple-Specs, unlike the head that does not, has the ability to check a certain feature of elements in its Specs multiple times. Under this conception of multiple-Specs, the condition in (80) does not follow. This is because the derived Spec does not have to be licensed as a Spec by the relevant feature since it is introduced as a complement prior to HMS. Given this, the reviewer argues that in order to maintain my argument, I need to adopt a representational view on a ban on multiple-Specs. However, as I will discuss in detail in section 5.3, in addition to fitting in the representational view, the present analysis fits in Fukui’s (2011) view on multiple-Specs (or multiple subjects), according to which in languages like Japanese, unlike English, the number of the application of Merge (substitution or Set-Merge) is unbounded.

I assume that in successive cyclic movement, A’-movement to an intermediate landing site is performed via adjunction rather than substitution. For example, when an object wh-phrase moves to SpecCP, the intermediate movement to the edge of vP is via adjunction. Therefore, even if a subject occupies SpecvP, the successive cyclic movement
of the object wh-phrase does not yield a multiple-Spec configuration. I thank Yuji Takano and a reviewer for bringing this issue to my attention.

32 As a reviewer points out, in a more recent paper (McCloskey 1996), McCloskey argues that subjects do not stay in-situ in Irish but move to somewhere between TP and vP. Even if this is correct, however, my prediction is not obviated because SpecTP is empty also under the new analysis.

33 As Tanaka (2008) points out, extraction out of elided material is not prohibited in general since in English, extraction out of an elided vP is possible, as illustrated in (i).

(i) I know which book John read, but I don’t know which one1 Bill did [vP read t1].

A structurally more similar sentence to the Japanese one is also acceptable, as shown in (ii).

(ii) Bob’s book, I argued John read, but Bill’s book1, I didn’t [vP argue [John read t1]].

This is because the elided vPs in (i) and in (ii) are not headless but “headed”.

34 However, there are several other languages that we have not discussed that have been argued to have headless verbal phrase-ellipsis such as Chingoni, Kikuyu (Ngonyani and Githinji 2006), Finnish (Holmberg 2001), European and Brazilian Portuguese (Cyrino and Matos 2002, Martins 1994), and Tagalog (Richards 2003). It is not clear to me whether they can be analyzed either as languages allowing multiple-Specs or as languages in which subjects do not move to SpecTP. If they cannot, I need to provide an explanation for why they have headless verbal phrase-ellipsis. One of the possible approaches to this problem is to reject the headless verbal phrase-ellipsis analysis of the relevant constructions in these
languages, arguing that the alleged headless verbal phrase-ellipsis should be analyzed in some other ways. I leave this issue for future research.

35 I assume that some independent principle or algorithm, such as Chomsky’s (cf. Chomsky 2008), determines what becomes a label.