Head Movement and Labeling*

Jeong-Shik Lee
(Wonkwang University)

Lee, Jeong-Shik. (2014). Head Movement and Labeling. *The Linguistic Association of Korea Journal, 22*(3), 1-17. In this paper, in line with Carstens, Hornstein, and Seely (2013), I advocate the traditional concept of head movement that can be maintained in terms of category and relevant features of the moving item and the target, contra Chomsky (2013) and Cecchetto and Donati (2010). This is made possible by incorporating the notion "label" into the definition of "extension" and "command" relation. The major results of the discussion are: A moved head can bind its trace in terms of label-command, head movement may not violate the Extension Condition under the revised "extension" defined in terms of label, locality can be measured in terms of label-crossing. Thus, there remains little anomaly with head movement.

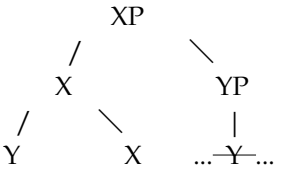
Key Words: head movement, labeling, label-command, label-crossing, Extension Condition

1. Introduction

The phenomenon described as head movement has drawn lots of attention in syntax mainly from the theoretical perspectives. This movement has been represented as a head-adjunction structure under the traditional X'-theoretic framework given in (1).

* This paper was presented at the conference held at Chosun University on May 24, 2014. I thank the audience for questions and comments. I also thank three anonymous reviewers for their comments and suggestions. Any remaining error, however, is mine.

(1)



The head Y of YP, complement of the head X, has moved to the head X via adjunction.

Here, the moved Y, strictly speaking, does not c-command its trace since the first branching node that dominates the moved Y is just X. Further movement of Y out of X, which constitutes a case of excorporation, is not allowed in general: head movement is not successive-cyclic, unlike successive-cyclic XP movement (both A and A'-movement).[1] So it has remained unclear how this head movement can be allowed in syntax.

In the Minimalist Program (Chomsky 1993 and afterwards), structure building is subject to the Extension Condition (hereafter, EC) to the effect that Merge applies only to the root, the top node, in a way to extend the structure. Consequently, EC derives the c-command condition on movement. The head movement shown in (1), however, apparently does not extend the structure, thus violating the EC. In addition, head movement is claimed to have no LF effects.

One prominent alternative to syntactic head movement was to propose that this movement is a PF-operation (Chomsky 2001, Boeckx and Stjepanovic 2001). On this view, then, head movement does not have to apply successive-cyclically, extend the structure, involve c-command relation, and bring any LF effects, as desired.

It has been observed, however, that not all head movement occurs at PF.[2] Roberts (2010: 10), for instance, argues that there are cases where head movement leaves semantic effects, as illustrated in (2).

---

1) According to Roberts (1991), excorporation still remains as an option. Roll-up cases of head movement are due to morphological reasons: excorporation is not allowed if the adjoined head Y forms a morphological unit with the head X; otherwise, it may be allowed.

2) Of course, this does not necessarily mean that head movement cannot take place at PF (see Scoorlemmer and Temmerman 2012 for arguments in favor of PF approach). But see also Matuchansky (2006: Appendix) for objection to the hypothesis that head movement occurs at PF.

(2) a. *Which one of them does anybody not like?
b. Which one of them doesn't anybody like?

T-to-C head movement carrying the negation with it in (2b) can license a subject NPI (Negative Polarity Item), in contrast with (2a), in which the negation remains in its place. Most of all, given that head movement constitutes a case of merge (i.e., Internal Merge), this movement should not be taken to apply at PF: it is an operation in narrow syntax.

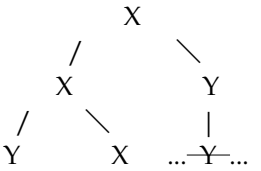
Now if head movement is a syntactic operation, the main task that emerges is to solve the problem of EC violation. In this paper, I will offer a solution in terms of the recent theory of labeling (cf. Chomsky 1995, 2013, Hornstein 2009, Cecchetto and Donati 2010, among others) and thereby attempt to maintain a traditional view of head-movement differently from Matuchansky (2006), Cecchetto and Donati (2010), and Chomsky's (2013).

2. Accommodating Head Movement in Syntax

If head movement is to be accommodated in syntax, the apparent violation of the EC must be evaded, and also the resulting c-command problem must be settled in one way or another. In the traditional X'-theory, head movement, being an X^{o}-movement, is an adjunction process to a head, an X^{o}-category; A-/A'-movement, being an XP-movement, apply to a maximal projection, an XP-category. This is an important generalization in that two types of movement have to be distinguished on a descriptive level under the Structure Preservation Hypothesis of Emonds (1976).

Under the Bare Phrase Structure (Chomsky 1995), relational information regarding category levels, for example, X' and XP, does not exist any longer. Thus (1) is now represented as given in (3).

(3)



Above, the top node, previously XP, is just labeled as X, which is determined by a labeling algorithm such that the head X, with Y as its complement, projects its label, and the topmost X is interpreted as XP. Under this approach, however, the problem of violation of the EC in the case of head movement still persists.

Cecchetto and Donati (2010) propose the following algorithm for labeling.

(4) The label of a syntactic object {a, b} is the feature(s) which act(s) as a Probe of the merging operation creating {a, b}.

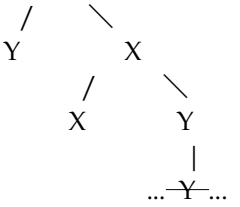
According to them, a lexical item (LI) with its own EF (Edge Feature) is a Probe by definition, which always activates the algorithm in (4), and its categorial feature can provide the label. When two LIs get merged as in (5) below, the transitive verb *saw* selects the direct object *John* as its complement, while *John* does not select *saw*. So the categorial feature of the head *saw* projects to provide the label.

(5) {saw, John}

To put it simply, *saw* has more features (i.e., an EF and a selection feature) than *John* with an EF only, and thus *saw* wins over *John* in providing the label of the merged complex atom in (5).

With this background, in line with Matuchansky (2006), Donati (2006), among others, Cecchetto and Donati (2010) contend that a head can merge to the root of the structure in syntax, as in (6).

(6)



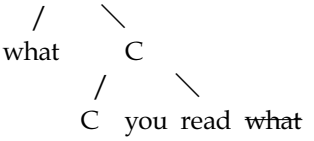
It is seen that the resulting structure obeys the EC; after the head movement of Y, head conflation between Y and X is obtained by an additional affixation process, called m(orphological)-merger, thereby yielding the same structure as in (3). Thus it is seen that X eventually determines the label.[3)]

Interestingly, Cecchetto and Donati (2010) further suggest that their algorithm may also allow the moving LI to provide the label (see also Donati 2006). The following data is related to this.

(7) a. I wonder what you read.
b. I read what you read.

The embedded clause will have a following derivational stage:

(8)



Under their system, head movement and phrasal movement are not distinguished on the basis of the Probe-Goal relation they establish. Thus, in (8) a single LI *what* is internally merged to the edge of a clause via head movement.

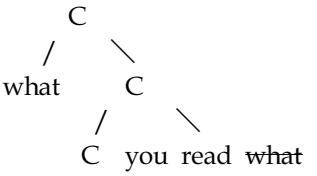
---

3) Although Matuchansky (2006) provides independent reasons for m-merger (e.g., m-merger with phrasal movement, m-merger without movement, head movement without m-merger), I wonder what can prevent m-merger (without movement) between T and V in examples like (ia) from taking place to produce the ill-formed (ic), assuming that verbal inflection is not obtained by phonological cliticization.

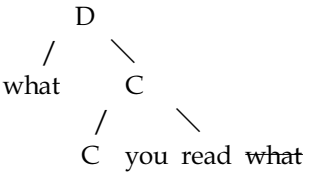
(i) a. John [T-ed [V see what]] => m-merger produces *see-ed* 'saw'
b. John [saw what]
c. *What (do) John saw? (Cf. What did John see?)

Since it is a Probe with its EF as an LI, its movement results in conflict between the two Probes. Cecchetto and Donati contend that this conflict manifests as an ambiguous labeling; that is, in (8) either *what* or C determines the label, as shown in (9).

(9) a.



b.



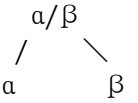
(9a) is the embedded clause in (7a), and (9b) in (7b).[4)]

In the next section, however, I raise some conceptual questions against Cecchetto and Donati's system and also observe some empirical difference in the Korean sentences corresponding to the English ones in (7a,b). I will show that the problem of EC violation still hangs around the head movement.

3. Against Cecchetto and Donati (2010)

Cecchetto and Donati's (2010) proposal, despite its superficial success, was built on some suspicious assumptions. To simply restate it, the labeling conflict arising in the configuration in (10) is resolved by allowing either α or β to provide the label:

(10)



4) Chomsky (2013: 46) also discusses this analysis for data like (7a,b).

Thus labeling appears to be ambiguous. However, I should like to point out that it remains unclear what makes β silent when α wins over β, and conversely what makes α silent when β wins over α. In other words, it is unclear what are meant by "win over" and "silent." Recall that according to Cecchetto and Donati, an LI has feature(s) that lead(s) to the label. Thus, the above kind of ambiguous labeling actually suggests that when α projects the label, the feature(s) of β should be invisible, and vice versa. But feature(s) of α or β cannot be something that can be inert or optional. They should be just there; otherwise, α or β could not be an LI.

It could also be reminded that the familiar structural ambiguity of a linear string like (11a) is captured by different bracketing as given below:

(11) a. X Y Z
b. [[X Y] Z]
c. [X [Y Z]]

The structure of (11b) is different from that of (11c). This possibly suggests that (10) also represents two different structures: in English, (i) when α is internally merged to β, the label of the resulting complex atom is provided by β, like in (9a) but unlike in (9b), (ii) when α first-merges with β, the label of the resulting complex atom is provided by α, being a head.[5] Recall also that in the recent minimalist program (Chomsky 2000, 2001), the notion of optionality is rephrased as follows: a phase head like v* either has a relevant feature like [+EPP] or it does not have this feature, viz., [-EPP], and thus, v*[+EPP] and v*[-EPP] belong to a different numeration. This understanding can also suggest that the ambiguous labeling in (10) eventually fails since it is unlikely that an LI can either have feature(s) or it does not.

From the empirical perspective, the Korean counterparts to (7a,b) from English clearly distinguish indirect *wh*-questions from free-relatives with respect to syntactic structure, as illustrated in (12):

5) Thus it remains to be discussed how (7b) can be represented in terms of labeling. I will return to this matter below shortly.

(12) a. Na-nun [ney-ka mwues-ul ilk-ess.nun-ci]
I-Top you-Nom what-Acc read-Past-Comp
kwungumha-ta.
wonder-Dec
'I wonder what you read.'
b. Na-nun [ney-ka e_i ilk-un-kes$_i$-ul] ilk-ess-ta.
I-Top you-Nom read-Past-KES-Acc readPast-Dec
'I read what you read.'

The indirect question is formed by the in-situ *wh*-expression *mwues* 'what' with the question Comp *-ci*. The free relative can be formed by the formal noun *kes* associated with the gap in the relative clause. Relevant to the current discussion is that in the indirect *wh*-question in (12a), the complement is realized as a CP headed by the Comp *-ci*, and in the free relative in (12b), the complement is realized as a DP determined by the noun *kes*. This fact then weakens the argument advanced by Donati (2006) and Cecchetto and Donati (2010). Adger (2013: 11-12) also points out the same problem by discussing data from Scottish Gaelic in which indirect questions are different from free relatives in the structure.

Thus what is to be questioned is the possibility represented in (9b). Given the conceptual and empirical difficulties raised above, I would rather adhere to the traditional conception that an internally merged LI does not provide the label. This amounts to saying that free relatives in English have a different structure like (13a) and surfaces as (13b) via a certain process combining *it* and *[+Wh][Rel]* or *it-that* into a morphological unit *what* (see Adger 2013, Chomsky 2013 for related discussion):

(13) a. [D it [C[+Wh][Rel$_i$] [you read e_i]]
b. (I read) what you read.

Here it is suggested that the free relative head *what* in English is morphologically complex (e.g., *it-that*), and that the reduced determiner *it* projects a label (Chomsky 2013: 46). What this analysis could mean is that the label of the free relative in English is not provided by the moved *what*, as seen

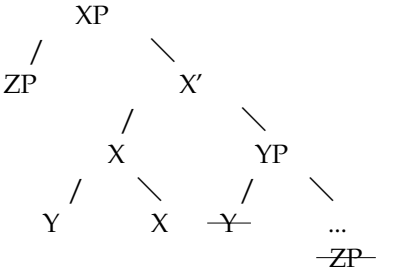
in (9b), contra Cecchetto and Donati (2010). As mentioned above, ambiguous labeling also remains suspicious. We are then led to go back to the initial problem of the head movement regarding the violation of the EC. Next section will discuss this matter.

4. Rethinking the Extension Condition

In this section, I attempt to basically maintain the EC while allowing the head movement without a violation of this condition. This will require relaxation of the notion "extension" in structure building in terms of the recent notion “label.”

Let us first consider the following traditional X′-structure:

(14)

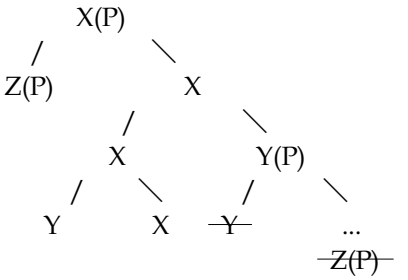


Under the above X′-structure, a maximal category ZP has moved to the Spec of the maximal category XP (or it could have adjoined to XP), and a head Y has adjoined to a higher head X. It may also be assumed that a head has its own Spec, say, a minimal Spec (and the ZP position may be called a maximal Spec). Then Y could move to the Spec of the head X via head movement.

As mentioned in section 2, thus, X^o-head-movement applies to a head element X^o and XP-movement, either A- or A′-movement, applies to a maximal projection XP, which has been widely recognized as Structure Preservation Condition since Emonds (1976). Under the above structure, as noted in section 1, head movement brings about problems concerning the EC, counter-cyclicity, etc. So it turns out that satisfying the Structure Preservation Condition undesirably leads to a violation of the EC.

In what follows, I will attempt to resolve this conflict by relaxing the notion "extension" a little bit in terms of labeling. Under the Bare Phrase Structure, now, the structure in (14) can be represented as in (15) (in which (P) is just added to indicate that the preceding category is interpreted as a maximal category):

(15)



Here the head X provides the label X and the head Y the label Y. Now notice that the phrasal ZP movement and the head Y-movement both target the same label X. This observation could deal with the problem at hand. The core idea is that the EC can be satisfied as long as an operation applies to the top label. In (15), X is the top label, and here the ZP-movement as well as the Y-movement applies to the top label X, thus possibly leading to the satisfaction of the EC. From a different angle, it is noted that the phrasal ZP-movement expanded the top category (i.e., X' => XP) and the head Y-movement expanded the top head category (i.e., X) by creating an adjunction structure. Note also that both cases of movement are completely consistent with the Structure Preservation Condition.

Setting aside the case of External Merge since it always extends the structure, for the case of Internal Merge I now propose a revised notion of "extension" of the structure that can be understood as follows:

(16) Any operation that targets the same category as the top label extends the structure.

Thus, both cases of movement in (15) extend the structure in that they target the top label. As a consequence, cases of Richards' tucking-in can be accommodated within this approach. As for the c-command problem, mentioned in section 1,

the notion of label may also be used in defining the command relation for the binding of the trace left by head-movement (l-command stands for label-command). Below is the proposed definition of the notion "l-command."

(17) α l-commands its trace if the label that dominates α dominates its trace.

In (15) the label X dominating Y also dominates its trace, hence Y l-commands its trace for binding.[6] So l-command is basically close to m-command in a sense. But since the notion "m-command" holds in the traditional X'-framework, the present "l-command" defined under the Bare Phase Structure is to be preferred at least from the structural point of view.

Under the current approach, excorporation still remains as an option (see Roberts 1991 for cases of excorporation). Further head movement of Y in (15) can apply successive-cyclically to a higher head as long as morphology allows.

As a result, the current approach brings forth a unified labeling algorithm in a simple manner: A head always projects a label for External Merge and Internal Merge. Cecchetto and Donati's (2010) labeling algorithm requires additional stipulation for ambiguous cases like free relative clauses in English. In section 3, I also presented an empirical evidence that shows that their analysis cannot extend to Korean free relatives (and Scottish Gaelic free relatives not introduced here).

Finally, it will be worth comparing Matuchansky/Cecchetto and Donati's position with the current one with regard to head movement. In the former, head movement targets the root, applying to the spine of the tree; additional m-merger is assumed; EC and c-command are thereby strictly maintained, but m-merger is an exception to c-command and the EC. Alternatively, in the current latter position, head movement does not target the root, applying to a head; additional m-merger is unnecessary; EC is relaxed in terms of label, and c-command is replaced by l-command.

---

6) Of course, the domain of l-command encompasses that of c-command. Further consequences of adopting the notion "l-command" remain to be explored.

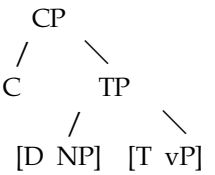
5. Chomsky's (2013) "Problems of Projection"

Chomsky (2013) in his article "Problems of Projection" (henceforth, POP) reiterates structure-dependency of grammatical processes, as illustrated with the following fact with Yes/No question formation in English:

(18) a. Can eagles that fly swim?
b. Is it the case that [eagles that fly] can swim?
c. #Is it the case that eagles that swim can fly?

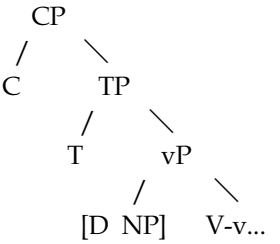
In (18a) the raised *can* must relate to the matrix verb *swim* and cannot link to the linearly closer *fly*, and thus, (18a) has the paraphrase in (18b), not (18c) (see also Carstens, Hornstein, and Seely 2013 for more related discussion). This fact leads Chomsky to claim that grammatical operations like T-to-C movement is structure-dependent, based on hierarchical relations instead of linear proximity. Chomsky further assumes that T-to-C movement depends on locality independent of category. Thus, as seen in the structural input given below, this raises a question of why this inversion moves T rather than (a subpart of) the expression in Spec, TP, viz. D(P), both being equi-distant from C.

(19)



POP's solution to this question relies on a timing approach: Under the VP-internal Subject Hypothesis only T can raise to C because the C-T relationship is established when EA (external argument, i.e., Subject) is still in situ, as seen below.

(20)



Carstens, Hornstein, and Seely (2013) (henceforth, CHS), however, point out a number of problems with this kind of timing approach. According to Chomsky (2013), minimal search should find D and v as they are equi-distant from T, and thus, either D-to-T or v-to-T should be allowed.[7] CHS, in favor of the traditional approach to head-movement, argue for an analysis that allows only v-to-T movement sensitive to categorial and other features of the target and the moving item, not just to locality. Their data include the following contrast in *Wh*-questions:

(21) a. Which boys (*did) eat the pizza?
b. Which pizza *(did) the boys eat?
c. Which boys *(did) you say t ate the pizza?

This fact shows that local subject *Wh*-questions disallow T-to-C movement but all other direct *Wh*-questions require it, indicating that something other than locality is involved in motivating and constraining head-movement.

Under POP's approach, in (20) EA-raising to Spec, TP takes place before V-v-to-T movement. Recall that in POP only T can raise to C in (20) because C-T relationship is established when EA is still in situ. Thus it needs to be assumed that T-to-C and EA-raising both precede V-v-to-T. CHS point out that this is not only counter-cyclic but also incompatible with the fact that V-v-to-T feeds T-to-C movement in languages like Spanish:[8]

---

7) Both labeling and identification of candidates for Internal Merge are based on minimal search in POP.

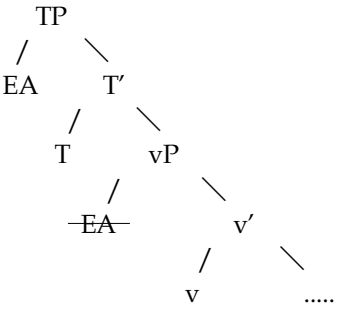
8) CHS (in section 2.4) discuss more examples of head-movement from Irish and Xhosa that present problems with POP.

(22) a. Que querian esos dos?
what wanted those two
'What did those two want?'
b. [CP what want+T+C [TP [DP those two]<T>[vP
<EA> <v> ...]]]

In short, CHS (p. 13) claim that head movement is driven and constrained by factors other than pure locality; selection for category and other features work in concert with hierarchical locality considerations to dictate what moves.

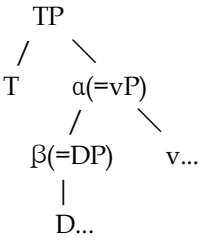
Under the current analysis, a head always provides a label, being a selector of some sort. In a configuration like (23) the head v provides the label vP. When the head T merges with its complement vP, it also provides the label TP; when EA is internally merged to Spec, TP from out of Spec, vP, the moved EA does not provide a label.

(23)



Thus, I assume that unlike in Chomsky (2013), v-to-T is a shorter movement than D-to-T in terms of label-crossing, as seen below.

(24)



It is seen that the label α is directly associated with T while the label β is only indirectly associated with it. Thus, v-to-T crosses only one label α, but D-to-T crosses two labels, α as well as β (see also Hornstein 2009 for a similar path account in terms of node crossing). This implies that locality can be measured in terms of label-intervention and importantly leads to the assumption, with Chomsky (2008) and Hornstein (2009), among others, that labeling is required for Merge (as well as interpretation).

6. Summary

In this paper, in line with Carstens, Hornstein, and Daniel Seely (2013), I advocated the traditional concept of head movement that can be maintained in terms of category and relevant features of the moving item and the target, contra Chomsky (2013) and Cecchetto and Donati (2010). This was made possible by incorporating the notion "label" into the current system. The results of the discussion are summarized below.

(25) a. A head always provides a label,
b. A moving LI does not provide a label,
c. A head reserves its Spec for successive cyclic head movement,
d. Head movement targets a head, and phrasal movement targets a specifier,
e. Head movement is syntactic,
f. Head movement obeys Structure Preservation Hypothesis,
g. A moved head can bind its trace by way of label-command,
h. Locality can be measured in terms of label-intervention,
i. An operation that targets the same category as the top label may extend the structure. Thus, head movement may not violate the Extension Condition.
j. Tucking-in is allowed.

Crucially, with the traditional head movement kept intact, more restrictive c-command is replaced by more permissive l-command (see (25g)), and the

Extension Condition is somewhat relaxed in terms of "label" (see (25i)). Under the current label approach, thus, there remains little anomaly with head movement.

References

Adger, D. (2013). *A syntax of substance*. Cambridge, Mass.: MIT Press.

Boeckx, C., & Stjepanovic, S. (2001). Head-ing toward PF. *Linguistic Inquiry, 32*(2), 345-355.

Carstens, V., Hornstein, N., & Seely, T. D. (2013). *Head movement in Problems of Projection*. Ms. University of Missouri, University of Maryland, and Eastern Michigan University.

Cecchetto, C., & Donati, C. (2010). On labeling: Principle C and head movement. *Syntax, 13*(2), 1-34.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 1-52). Cambridge, Mass.: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.

Chomsky, N. (2000). Minimal inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howrd Lasnik* (pp. 89-156). Cambridge, Mass.: MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1-52). Cambridge, Mass.: MIT Press.

Chomsky, N. (2008). On phases. In R. Freiden, C. P. Otero, & M. Zubizarreta (Eds.), *Foundational issues in linguistic theory. Essays in honor of Jean-Roger Vergnaud* (pp. 291-321). MIT, Cambridge, MA.

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

Donati, C. (2006). On *Wh*-head movement. In L. Cheng & N. Corver (Eds.), *Wh-movement: Moving on* (pp. 21-46). Cambridge, Mass.: MIT Press.

Emonds, J. (1976). *A transformational approach to English syntax: Root, structure-preserving and local transformations*. New York: Academic Press.

Hornstein, N. (2009). *A theory of syntax: Minimal operation and universal grammar*. Cambridge: Cambridge University Press.

Matuchansky, O. (2006). Head movement in linguistic theory. *Linguistic Inquiry, 37*(1), 69-107.

Roberts, I. (1991). Excorporation and minimality. *Linguistic Inquiry, 22*(1), 209-218.

Roberts, I. (2010). *Agreement and head-movement: Clitics, incorporation, and defective goals*. Cambridge, Mass.: MIT Press.

Scoorlemmer, E., & Temmerman, T. (2012). Head movement as a PF-phenomenon: Evidence from identity under ellipsis. In J. Choi, E. A. Hogue, J. Punske, D. Tat, J. Schertz, & A. Trueman (Eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics* (pp. 232-240). Somerville, MA: Casacadilla Proceedings Project.

Jeong-Shik Lee
English Department, Wonkwang University
460 Iksan-daero, Iksan
Jeonbuk 570-749, S. Korea
Phone: 063-850-6873
Email: jslee@wku.ac.kr

Received on July 10, 2014
Revised version received on August 28, 2014
Accepted on September 5, 2014