Word Order and Its Variations in Korean: A TAG's Approach

Sung-Kook Han

It is generally required that a grammar formalism explain the word-order variation since such the variation is universal phenomenon of all natural languages and that the complex patterns cannot be realized by reodering the terminals. This phenomenon is especially important to nonconfigurational languages that are relatively free in word order. We show how TAG can handle the word-order and its variation in Korean. We derive a new property called Scramble-α innated in scrambling. The domain of scrambling can be realized within the elementary trees of TAG and can be localized under TAG formalism. We propose a new adjoining constraint suitable for the description of word order of Korean. We show that TAG generates the syntactic structures as well as ordering precedence at the same time. We also show that long-distance scrambling is another type of cross-serial dependencies and analyzed with the same principle that local scrambling uses in TAG.

1. Introduction

The prominent characteristics of natural languages can be visualized with word order. The functions of word order, however, receive less attention than other linguistic phenomena in spite of close relationships to grammatical structures. Word order has been simply regarded as subsidiary element that depends on syntactic, semantic and pragmatic structures, especially, in so-called free word-order languages like Korean. Though word order does not have a great influence on syntactic structures, the word order should be treated properly as other grammatical components since it has also grammatical significance as linguistic components.

It is well known that natural languages allow for word order variation

though they show some differences with respect to the amount of order variation permitted.[1] Word-order variation in general has been regarded as a concomitant that carrys some kinds of information about the status of the content of an utterance in the discourse. Recently word-order and its variation have received a great deal of attention in Generalized Phrase Structure Grammar (GPSG) from the point of view of syntactic components, which is formalized by immediate dominance/linear precedence (ID/LP) format (Gazdar (1985)). This claims that word order in natural language can be expressed by the generalization in grammar formalism.

Many linguistic claims concerned with word order and its variation, however, are concentrated on the explanation of the apparent constraints on syntactic structures without considering the underlying principles for word order[2]. Many approaches on scrambling such as movement analysis and S-adjunction also have not provided the satisfactory explanation to enlight the linguistic principles of scrambling though these may elucidate some aspects of scrambling. The aim of this paper is to extract the base principles of word order that make scrambling possible and to show how Tree Adjoining Grammars (TAG) can capture these structures. TAGs proposed by Joshi (1975, 1987b) are midly context-sensitive in the generative capacity and realize the extended domain of locality by factoring recursion and dependencies. It is expected that such properties of TAGs are appropriate to describe the diverse variations of word order.

In Section 2, we will review the well-known syntactic behaviors and some arguments against scrambling in Korean so that we understand the linguistic properties innated in scrambling. We will discuss why syntactic structures of Korean can be freely scrambled in general. In Section 3, after

[1] We have used the term, word-order variation, without formal definition. The word-order variation, roughly, can be defined as the variation of surface order among constituients which does not accompany or require the morphological variation of lexical items and the order-dislocated constituients do not occupy the grammatically licensed positions. Thus the term, word-order variation, used in this paper is distinguished from topicalization or wh-movement which accompany morphological or syntactic transformation. Sometimes "scrambling" is used more frequently in some linguistic literature. We will also use "scrambling" with "word order variation" without any distinction.

[2] In Hale (1983), the ordering is one of the important factors to characterize linguistic structures of languages whether they are configurational or not.

briefly presenting the fundamental formalisms of TAGs, we will propose a new constraint suitable for the description of free order languages. We will apply the elementary trees with this constraint to describe the syntactic structures of Korean. In Section 4, we will evaluate TAG's description of word order by means of analyzing the controversial phenomenon called long-distance dependencies. We will show that local and long-distance scrambling can be handled with the same principle. We will summarize TAG's approaches to word order variations in the conclusion.

2. Word Order of Korean

In some languages like Latin, Warlpiri and Finnish known as W*-languages, word order is regarded as being "free" (Hale (1983), Karttunen (1985)). Korean is also free in word order in some senses. We will review some important properties of word order so that we can see what the word order implies, especially in Korean.

2.1. Some Facts of Word Order and Scrambling

Although we can acknowledge the notion of preferred or unmarked word order for Korean, the word order in principle is considerably free compared with English and other configurational languages. Let us consider the following examples which are all variants by scrambling, but they have the identical syntactic structures[3].

(1) a. John-i ku sonyen-eykey chayk-ul cwu -ess -da.
-NOM that boy -DAT book -ACC give-PAST-DEC
"John gave a book to that boy."

[3] The abbreviations used are here as follows:

Particle					
NOM	nominative	DAT	dative	ACC	accusative
PADV	adv. particle	TOP	topic		
Ending					
PAST	past	PRE	present	EDEC	declarative
EDET	determinative	CONJ	conjunctional	COP	copular

b. John-i chayk-ul ku sonyen-eykey cwu -ess -da.
c. ku sonyen-eykey John-i chayk-ul cwu -ess -da.
d. ku sonyen-eykey chayk-ul John-i cwu -ess -da.
e. chayk-ul John-i ku sonyen-eykey cwu -ess -da.
f. chayk-ul ku sonyen-eykey John-i cwu -ess -da.

All variants in (1) have the same syntactic structure and represent the same meaning[4]. One of the notable properties of free order languages is characterized with the highly development of particles or the affixation to represent the grammatical functions. All arguments can keep their grammatical functions (GFs) according to these postpositional case marker in spite of scrambling their positions in a sentence. The free order languages like Korean, that is, already have some facilities to make them independent of syntactic configurationality. There is no necessities for arguments to move some grammatically licensed positions so that they can obtain GFs. Since the roles of arguments can be uniquely identified by the attached particles no matter where they are located, the movement seems to be inadequate if we consider how GFs are given to nominal agruments. It is plausible to assume that scrambling is not Move-α. Since the scrambling shown in the above simplex sentences do not accompany any grammatical variations and empty categories occurred in Move-α. So maximal arguments such as NP and PP can be freely scrambled as long as predicate is located at the end of arguments within maximal projection. We can realize that scrambling is possible since the grammatical roles of arguments given by predicate are definitely distinguishable even though they are scrambled.

(2) Any arguments can be scrambled, not moved from the predefined positions.

As we can see in (1), it seems that scrambling does not require the grammatical mechanism to interpret the variants of constituients order or is not controlled by any grammatical principles. Since all possible permutations are equivalently acceptable, scrambling in Korean seems to be independent of other syntactic phenomena. In other words, any arguments can be scrambled - not

[4] It may represent the minor differences in the represented meaning under a certain circumstance at the time of discourse. However, their syntactic and semantic structures are in principle identical.

moved from licensed positions unless it violates some linguistic constraints that may be concerned with scrambling. It is obvious, however, that there may be some linguistics principles that constrain the freedom of scrambling. We already know many linguistic factors such as pronominal coreference and topicalization can constrain the scrambled arguments even in a simplex sentence. The scrambling phenonenon, nevertheless, is not an attendant phenomenon dependent on these factors. The problem of pronominal coreference in fact is for pronoun itself, not the general properties of scrambling[5]. Though this sort of constraint can constraine the freedom of scrambling, the nature of scrambling in (2) may be universal. Let us call the assertion in (2) "Scramble-α" in contrast to Move-α in English.

Now let us consider another problem provoked by scrambling.

(3) a. [John-i ssu] -n chayk-ul Mary-ka sa -ss -da.
-NOM write-EDET book-ACC -NOM buy-PAST-DEC
"Mary bought a book that John wrote."
b. Mary -ka [John-i ssu]-n chayk-ul sa -ss -da.
-NOM -NOM write-EDET book -ACC buy-PAST-DEC
c. na-nun [ecey John-i moim -ey ka-ta] -ko mit-nun-ta.
-NOM yesterday -NOM meeting-PADV go-DEC-COMP believe-DEC
"I believe that John went to the meeting yesterday."
*d. ecey na-nun [John-i moin -ey ka-ta]-ko mit-nun-ta.
yesterday I metting go -COMP believe

The above examples explicitly show that there is a certain domain of locality in the freedom of scrambling. Many assertions have been presented whether scrambling is clause-bound or not[6]. Whether or not scrambling is

[5] Considering the strong overt characteristics of pronoun in Korean and Japanese, the binding domain of pronoun may be determinded by linear precedence rather than the binding theory (Mohanan (1981)).

[6] Scrambling is assumed to be clause-bound in Muraki (1979), Tonoike (1980), McCawley (1976). On the other hand, in Havada (1977) and Kuno (1978) it is argued to the contrary. Long-distance scrambling seems to provide some evidences regarding that scrambling is not clause-bound. We will deal with this problem in Section 4 within our framework of scrambling.

clause-bound may be obvious when we consider the fact that arguments should receive their grammatical roles from predicate before scrambling. This means that arguments cannot drift away from the domain specified by predicate. If arguments break away from their bounded domain, they cannot maintain their θ-roles anymore and then make a sentence unacceptable. Arguments should be related with predicate to obtain their θ-roles and exist within the domain of the dominant predicate. No matter which particles are given to represent θ-role, arguments can be equally existed at any positions within the domain of predicate without reciprocal interference[7].

Although the above claim does not constitute positive evidence for boundness of scrambling, it seems plausible that one argument is related with one predicate. In this case, the relationship between argument and predicate is not determined with configurational manners, but represented explictly with particles in Korean. In other words, the subcategorization frame to offer grammatical functions to arguments does not equip any mechanisms to specify the ordering between arguments, that is, flat. Scrambling thus can be the primitive phenomenon, like Move-α in English, orginated from the base structure. We can express the convincing arguments on the locality of scrambling and subcategorization frame of predicate as follows: [8]

(4) There is only VP node that has the flat subcategorization frame to offer the thematic role to argument.

One may argue that the examples shown above are too simple to illustrate the whole aspects of scrambling and that the assertions in (2) and (4) may be absurd or sterotyped. They might wonder how we can solve the configurationalities of syntactic structures as shown in Weak Crossover and pronominal coreference with the flat subcategorization frame. Now let

[7] This means the syntactic structures of Korean is flat in some senses. Note that there is no reserved position in which subjects are structually distinct from non-subject arguments, and thus that argument positions are indeterminate-scrambled at anytime. In this paper, the "flatness" means that there are no ordering constraints in subcategorization frame unlike English.

[8] In general, a VP node implies a node that dominates the verb (with or without object NPs) while not dominating subject NP. In our discussion, however, VP node regards as a maximal projection of V including subject NP.

us examine detaily how scrambling is realized in the surface structure while we concentrate what scrambling is like in order to show Scramble-α.

A number of arguments on free word order phenomenon have proposed for a movement analysis (Saito (1987)). It is suggested specifically that scrambling is like QR (Quantifier Raising) in that it involves adjuction to S. According to this analysis, multiple scrambling as shown in (5) that the direct object and the indirect object both precede the subject is analyzed as (6).

(5) a. Mary-ga John-ni sono hon -o watasi-ta [Japanese]
-NOM -DAT that book-ACC hand -PAST
"Mary handed that book to John"
b. sono hon-o John-ni Mary-ga watasita
c. John-ni sono hon-o Mary-ga watasita

(6) [sono hon-o_i [John-ni_j [Mary-ga t_i t_j watasita]]]

It is argued that if scrambling involves adjunction to S, the unacceptablity provoked by pronominal coreference can be solved by binding theory : A pronoun cannot c-command its antecedent.

(7) a. * kare-ga [[Mary-ga John-ni okut-ta] tegami-o] mada
yonde inai [Japanese]
he -NOM -NOM -DAT send-PAST letter-ACC yet
read not
"He has not read the letter Mary sent to John"
b. [[Mary-ga John-ni okut-ta] tegami-o] kare-ga made yonde inai
"The letter Mary sent to John, he has not read."

In (7), the pronoun kare *he* does not c-command its antecedent *John*. Since (7-b) is not in violation of binding condition, it is definitely acceptable. By the way, the movement analysis of scrambling is not sufficient enough to generalize the scrambling phenomenon[9]. Assuming that scrambling is adjunction to S by QR, the movement will creat a new adjoined position which is an A-bar

[9] One of the intension of Move-α analysis is to show the existence of VP, that is, to show the configurationality of language. See Imai (1987) for detail.

position. In Korean, however, the moved phrases always keep their θ-role in terms of case particle attached to nominal. Wherever phrases will be moved by scrambling, their moved position are always A-position. Furthermore, scrambling does not require external principle different from that of A-bar movement which is forced by some external principle like the necessities of interpretation in the case of a quantified NP. In other words, there are no *triggering conditions* for scarmbling in syntactic structures of Korean. And the movement analysis will bring about the debate about the unmarked word order. In Korean, the GFs shown in English such as [NP, S] or [NP, VP] does not have different weight. The assumption about an unmarked or underlying order reveals that it does not match with other grammatical consequences. There is no necessity that subject should reserve its position in a sentence since grammatical role of arguments can be represented by the attached particle. Therefore all nominal arguments dominated by predicate are equally treated and can be equally scrambled. Then, there remains only one way to explain the scrambling. Let us remind that VP has a subcategorization frame for NP. Since all arguments are associated with predicate to obtain GFs and these are independently accomplished, arguments join VP in the manner of VP-adjunction.

(8) Scrambling is realized in the form of VP-adjunction.

It is well-known fact that the ellipsis of arguments is quite usual in free order languages even if it is a subject. This implies that all arguments are associated with VP like adjuncts[10]. From the point of view of adjoining, it is implied that there is no distintion between arguments and adjuncts, which join as VP-adjunction.

2.2. Subcategorization Frame of Korean

The scrambling can be realized in principle by VP-adjunction with the arguments that have the thematic roles given by predicate. The fundamental role of particles is to provide such the grammatical relations for arguments by the association with them. The particles play a crucial role of indicating

[10] Arguments receive the thematic roles from VP, but adjuncts do not. NPs with kernel particles are regarded as arguments in this paper.

and differentiating GFs of the arguments in Korean.

(9) a. Mary-ka chayk-ul ilk-nun-da.
-NOM book -ACC read-PRE-DEC
"Mary reads a book."
b. chayk-ul Mary-ka ilk- nun-da.
-ACC -NOM

The arguments, Mary and chayk *book* can preserve their GFs, subject and (direct) object, with the presence of particles, -ka and -ul, respectifully, even though they are scrambled.

We may assign a certain thematic role (θ-role) to each arguments or constituents by means of the typologies of particles. However, it turns out that the thematic role which is represented with particle is highly redundant. In spite of attachment of same particle to arguments, the GFs of nominal arguments are quite different. This implies that there is not always one -to-one correspondance between the thematic roles and particles. Although it is certainly true that the particle as case marker may be used to indicate GFs, the equivocal typological features between marked particles and the thematic roles reveal that the thematic roles are not depended only on particles entirely. Since the thematic role can be realized in various particles, another issue on which particles can play a role of case markers are raised. For the reason that the extensive researches on case particles have been concentrated the encoding semantic roles on surface sentence, their functional analysis may not be consistent with subcategorial features of particle. The concept of kernel particles for defining subcategorial structure of predicate is proposed in Han (1988). The kernel particles can be used to mark the θ-role in the base component. The issues on which particles are used in lexicon are important to scrambling language in order to determine the structure of the base component, for the only way to stipulate subcategories and assign the thematic role is the presence of the particles, not configurational structures. We classify these particles as following groups :

(10) Group 0 : -un, -nun, (-ka, -i) Group 1 : -ka, -i, (-un, -nun)
Group 2 : -ul, -lul, -l Group 3 : -eyse
Group 4 : -eykey Group 5 : -ey
Group 6 : -lo Group 7 : -wa, -kwa

Thus the subcategorization frame, for instance, will be represented in the form of (11).

(11) a. ilk-ta : {NP1, NP2, NP3, NP4, NP5, NP6}
read
b. ka-ta : {NP1, NP3, NP4,. NP5, NP6}
go
c. kyelhonha-ta : {NP1, NP3, NP5, NP7}
marry

The subcategorization frame of course is the unordered set on account of the underlying linguistic properties of scrambling. This scheme of subcategorization of predicate is intutively plausible for the case of Korean. We can realize that the orginal properties for scrambling are already inhered in the base component like this manner, in Korean. We also point out some linguistics aspects concerned with this subcategorization frame, that is, this scheme shows the incremental property in terms of VP-adjunction[11].

(12) a. ka-n -ta
go-PRE-EDEC
b. tosekwan-ey ka-n -ta
library -P5 go-PRE-EDEC
c. bus-lo tosekwan-ey ka-n -ta
-P6 library -P5 go-PRE-EDEC
d. John-i cip -eyse bus-lo tosekwan-ey ka-n -ta
-P1 home-P3 -P6 library -P5 go-PRE-EDEC
"John goes to the library from house by bus."

The incremental property in generation is important to understand Scramble-α in Korean. This shows that there may be many X-bar levels in VP.

[11] If we remind scrambling is VP-adjunction, we can derive the phrase structure rule of Korean as follows:
a. S → VP
b. VP → NP VP
The recursive rule (b) shows the monotonical incrementality in the generation of a sentence. This ruls also implies that syntactic structures is strictly binary as Saito (1987) argues.

3. TAG's Description of Word Order

In this section we will present the overview of TAG formalism and we will consider what functions TAG have to hold for the description of scrambling. We also explain how TAG capture Scramble-α.

3.1. Overview of TAG Formalism

TAG introduced by Joshi (1975, 1987) is proved that TAG is more powerful than CFG, but midly so. TAG turns out to have enormous fruitfulness for the description of natural languages by the realization of the recursion and the extended domain of locality. The linguistic relevance of TAG shows that TAG can describe various syntactic structures more naturally than other formalisms on account of its extended domain.

A tree adjoining grammar (TAG) $G=(I, A)$ is specified by a finite set of elementary trees. The trees in I and A are called the initial trees and the auxiliary trees, respectively. A tree can be an initial tree if it has a form in Fig. 3-1(a) : that is, the root and internal nodes of tree are nonterminals and the frontiers are all terminals. A tree is an auxiliary tree if it is of the form in Fig. 3-1(b) : that is, the root node of a tree is nonterminal labelled X and the frontier nodes are terminals except one which is nonterminal labelled X, called a foot node. The auxiliary trees that support the unique recursions are used to construct more complex sentences from the skeletal sentential structures.

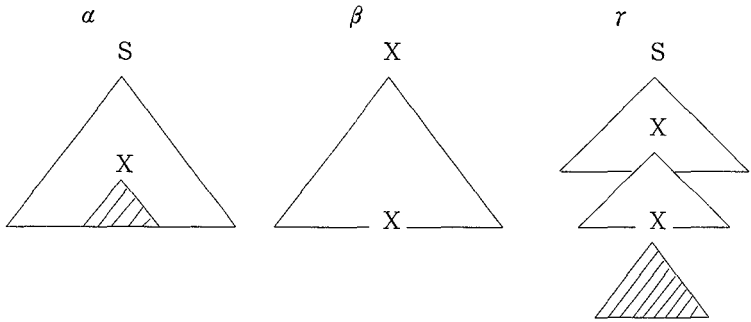

Fig. 3-1. TAG and Adjoining Operation

The major operation, adjoining or adjunction, in TAG is defined as follows (Joshi (1987)).

Let α be a tree (α is any tree) containing a node n labeled by X and let β be an auxiliary tree whose root is also labeled by X. Then the adjunction β to α at node n will be the tree γ built as follow:

1. The subtree of α dominated by n, called as t, is excised, leaving a copy of n behind.
2. The auxiliary tree β is attached at n and its root node is identified with n.
3. The subtree t is attached to the foot node of β and the root node n of t identified with the foot node of β.

The resulting tree γ called a derived tree, is shown in Fig. 3-1. In TAG, all mathemathematical power comes from the adjoining operations. On adjoining operation, there are some important constraints that make TAG more powerful. In TAG with local constraints, each nonterminal nodes can have one of the following constraints to specify adjoining operation.

- selective adjoining (SA): Only one of the specified auxiliary trees can be adjoinable at the addressed node n.
- Obligatory adjoining (OA): At least one tree out of all adjoinable tree must be adjoined at n.
- null adjoining (NA): No auxiliary trees are adjoinable at this node.

These three adjoining constraints are widely used for the correct derivation of syntactic structures and the prevention from overgenerating ungrammatical structures. From the formal aspect of view these constraints provide relationships between trees. Since TAG is a tree generating system, the formal objects of system are trees and adjoining constraints defined bewteen trees act as relations of them. Linguistically, on the other hand, adjoining constraint represents the association between linguistic objects that have the extended domain of locality. Thus adjoining constraint is significant in the derivation of syntactic structures.

When we consider the formal functions of adjoining constraints, we can find some similarity betweem them[12]. The handling of trees under adjoining constraints in fact is quite restricted by virture of linearity in successive

[12] SA will be regraded as a special case of OA with null tree and NA as OA with only null tree, respectively.

adjunctions. Though the above constraints might be enough to define tree set for the derivation of syntactic structures of English, they are insufficient for the free order languages like Korean. As we already know, the syntactic structures of these languages are derived by agglutinating the various components simultaneously rather than expanding the base structure hierachically. In this case, adjoining constraints in TAG may not give the proper way to agglutinate the linguistic objects since adjoining constraints show sequential property in the real application. We define another constraints called conservative adjoining (CA) for the purpose of realizing subcategorization frame of free order languages. The newly proposed constraint CA is informaly defined as follows : A subset of specified trees can be adjoinable at the addressed node. Note that the introduction of CA neither increase the formal power of TAG nor require any extension of the functions in TAG[13]. Since the subset does not impose on the order between elements, the constraint CA is suitable for the Scramble-α and VP node that has flat subcategorization frame. Above all, notice that this constraint is consisted with the subcategorization frame given in (11).

The fundamental TAG formalism is extended to capture linguistic properties of natural language in the efficient manner. Especially, a version of adjoining operation, called multicomponent adjoining, is introduced under which a set of trees is adjoined to a given elementary tree. It is shown that multicomponent adjoining is effective in the analysis of extraposition (Kroch (1987)). We will also use multicomponent adjoining for the analysis of scrambling, but the concepts of linking used here is different from others.

3.2. Scrambling in TAG

The subcategorization frame to fulfill Scramble-α is expressed with the following two elementary tree in TAG.

The above tree corresponds to subcategorization frame of verb ka-ta *go*. This tree contains all claims for free word order and scrambling we argue in Section 2. Note that we use multicomponent adjoining with link between

[13] The introduction of CA is purely for linguistic and pratical matters. We can realize CA in terms of tree group selectively constrained with subset of specified adjoinable trees. For example, a tree with $CA(\alpha 1, \alpha 2)$ is equivalent to a group of trees with $SA(\alpha 1)$, $SA(\alpha 2)$ an $SA(\alpha 1, \alpha 2)$ roughly.

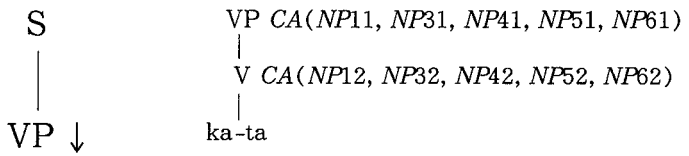

Fig. 3-2. Elementary Trees for Scrambling

NP and V node. For NP, it should be related with predicate to obtain the thematic role and its appearence in a sentence is accomplished by means of VP-adjunction. Multicomponent adjoining on VP and V represents these VP-adjunction and projection, respectively. In these elementary trees, we have to pay more attention to the role of VP node. Since the subcategorization frame is determined in V node, VP node only plays a role to represent the adjunction site of VP-adjunction specified in V node. VP node of free order languages does not have definite weight as in English except that it is adjunction site. So we simply represent VP node with CA constraint as VP# since other source for multicomponent adjoining is fixed to occur at V node. The auxiliary trees for VP-adjunction and projection will be given as in Fig. 3-3.

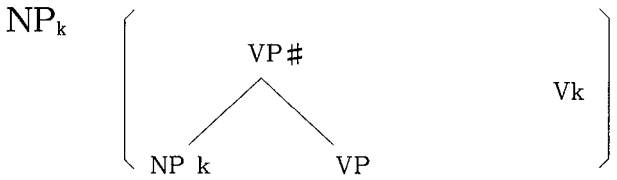

Fig. 3-3. Auxiliary Tree for Multicomponent Adjoining

All linguistic requisites for scrambling are existed is in this tree such as the binding domain for NP and word order. Since TAG has the extended domain of locality, all constraints can be localized in this manner. The coindexing between NP and V characterize the governed domain of NP rather than filler-gap dependencies in the usual context[14]. It also plays a remarkable role of linear precedence, NP > V, to specify word order as we

[14] In free order languages, filler-gap dependencies are unusual except some endocentric constructions such as relativization. The presence and the trace of empty categories are complex on account of ellipsis and drop of arguments which are very common in these languages.

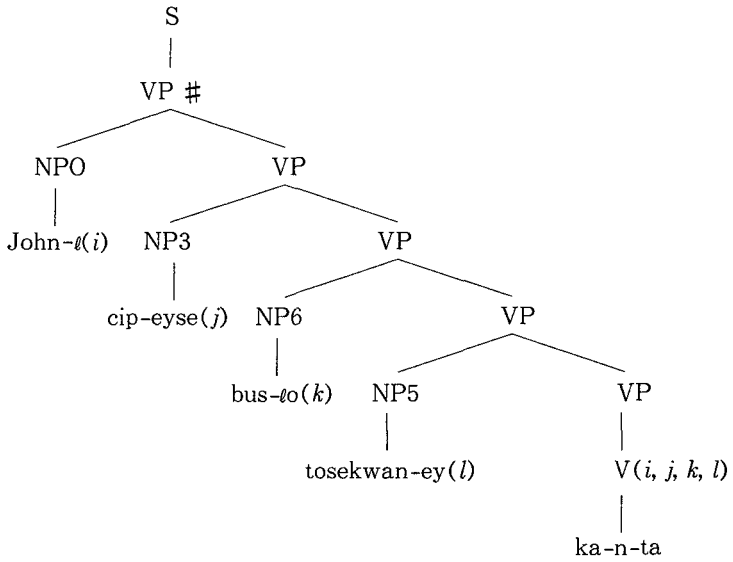

Fig. 3-4. Derived Tree

will see below. Linear precedence in scrambling is localized by VP-adjunction in the elementary trees of TAG.

The derivation tree in Fig. 3-4 represents the syntactic structure of sentence (12). We find that this tree furnish some important linguistic implication on syntactic structures of Korean. The VP node with CA constraints reveals VP-bar property in X-bar theory. When we consider the coordination of Koran, we can easily verify that the various types of VP nodes are in Korean. Another characteristics of this tree is strictly binary as Saito argues in Japanese (Saito (1987)). It may come from the fact that Korean is agglutinative in the generation. The coindexing represents the binding domain as we mention above. We should pay attention to another function of coindexing, which plays a role of linear precedence in word order. On account of the extended domain in TAG, we can localize linear precedence of word order and this can be implemented with subcategorization frame at the same time as in Fig. 3-3. With the extended domain of TAG, these two relations are captured in an elementary tree at the same time. Scramble-α with (2) and (8) thus is nicely explained using elementary trees. In Fig. 3-4, we can infer that 24 (4!) variants can be generated by means of linear precedence. In TAG's description of Korean, we can represent all the vari-

ants derived from trees as well as their syntactic structures at the same time.

4. Long-distance Scrambling

Since TAG is midly context-senstive, it is already shown that TAG can represent cross-serial dependencies in a unique way (Joshi (1987)). We will show the idiosyncrasy in scrambling, called long-distance scrambling, is also another type of cross-serial dependencies.

The free order languages generally have abundant inflectional verbal endings to conjoin simplex sentences. The long-distance scrambling observed in Korean is related with these verbal ending, especially nonfinal verbal endings regraded as complementizer[15]. Before considering the long-distance scrambling, let us review the limitation of scrambling.

(13) a. John-i Bill-ekey [Mary-ka mek-n] kwaca-ul cu -ies -ta.
-P1 -P4 -P1 eat-EDET cake -P2 give-PAST
-EDEC
*b. John-i t [Mary-ka Bill-ekey mek-n] kwaca-ul cu-ess-ta.
-P1 -P1 -P4 eat-EDET cake -P2 give
"John gave Bill a cake that Mary was eating."

The arguments in (13-a) cannot be freely scrambled and (13-b) is ungrammatical since it violates the subcategorization frame of mek *eat*. The subcategorization frame of predicate is in general locally defined and cannot be overlapped. If they are overlapped, we cannot specify the dominant domain for arguments explicitly and then it causes a sentence unacceptable. Even if overlapped without any conflict in thematic roles, a sentence will be ambiguous. As long as arguments can be locally governed within the domain of predicate, they can be scrambled. Although any arguments can be scrambled by Scramble-α, the government relations must be always

[15] Since the long-distance scrambling receives less attention within grammar formalisms, its linguistic properties are not known well except that long-distance scrambling is more restricted than the usual scrambling. This type of scrambling might be related with some sort of verbal endings.

kept. This is the reason why long-distance scrambling is more restricted. In the following, the long-distance scrambling means that the argument can be preposed beyond the matrix subject, as shown in (14).

(14) a. sakwa-lul$_i$ John-i [Mary-ka t_i mek-ess-ta] -ko Bill-ekey malhay-ss-ta.
apple-P2 -P1 -P1 eat-PAST-EDEC-COMP -P4 say-PAST-EDEC
"John said to Bill that Mary ate an apple."

b. ku maul -ey$_i$ John-i [Mary-ka t_i sa-n-ta] -ko saykakha-n-ta.
that village-P5 -P1 -P1 live -COMP think -EDEC
"John thinks that Mary lives in that village".

The most general approaches to explain this phenonenon is Move-α and topicalization. In Whitman (1982), it is argued that scrambling is exactly like topicalization if we assume that the object NP in (14-a) is preposed to sentence-initial position by scrambling. This assumption also leads S-adjunction analysis of scrambling. This analysis tries to extend the dominant domain of predicate in any ways. The preposed position sakwa-lul *apple* cannot be the extension of domain of its dominant predicate mek- *eat* since *each domain of predicates is kept independeptly without overlapping even in* complex sentences. The preposed position in free order languages cannot be defined by the matrix predicate. Now we will show how the preposed position can be directly defined by its won predicate in TAG formalism.

It is usual that we assume the complement stucture like Fig. 4-1 since the embedded clause becomes adverbial to predicate in Korean. Note that this elementary tree has a adjunction site at root node, similiar with Fig. 3-3.

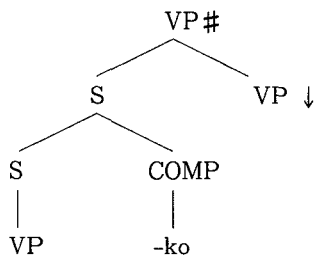

Fig. 4-1. Complement Structure

Now let us examine how the enlarged domain shown in long-distance scrambling can be handled with the local domain. The embedded clause of (14-a), [sakwa-lul Mary-ka mek-ess-ta] will be represented as in Fig. 4-2.

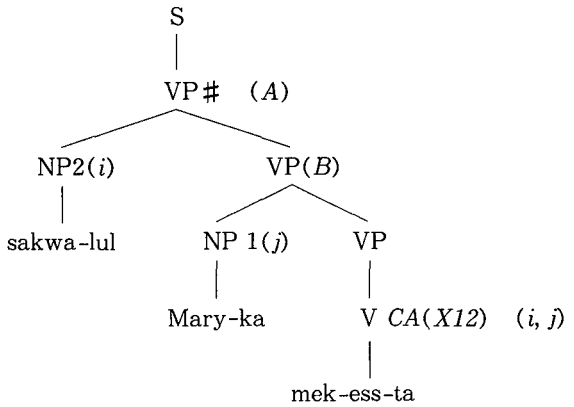

Fig. 4-2. Derived Tree of Embedded Clause

As the elementary tree for complementation shown in Fig. 4-1 is adjoined at node A, we will obtain the following derived tree by some successive multicomponent adjoining for matrix clause.

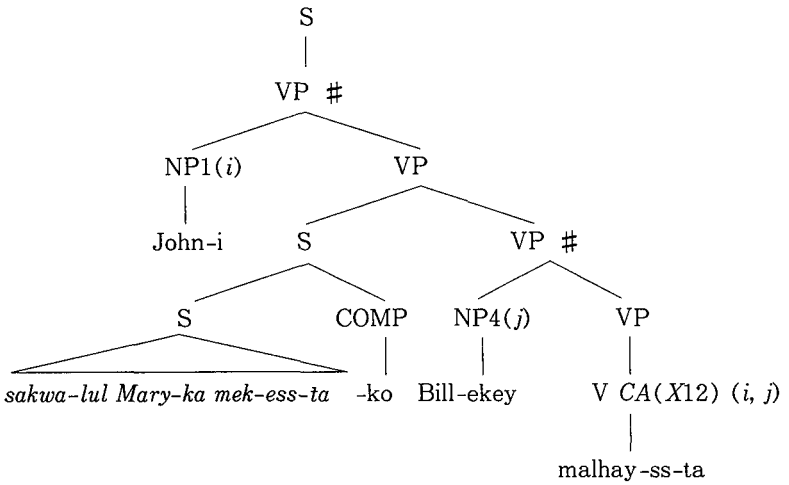

Fig. 4-3. Derived Tree of Embedded Sentence

This tree corresponds to simple scrambling that regards the emdedded clause as one adverbial NP. If we impose the linear precedence between S and COMP in Fig. 4-1, we can generate all the possible variant sentences from this derived tree[16]. In TAG the derived tree can represent syntactic structure as well as ordering constraints at the same time. By the way, the derived tree has two adjunction sites. Although it reveals nondeterministic in adjoining arguments, VP-adjunction is fully controlled by subcategorization frame in V node[17]. The multiple adjunction sites can be understood as the primitive property of Scrambling-α. On the other hand, if we adjoin the tree of Fig. 4-1 at node B of Fig. 4-2, then the long-distance scrambling tree is derived.

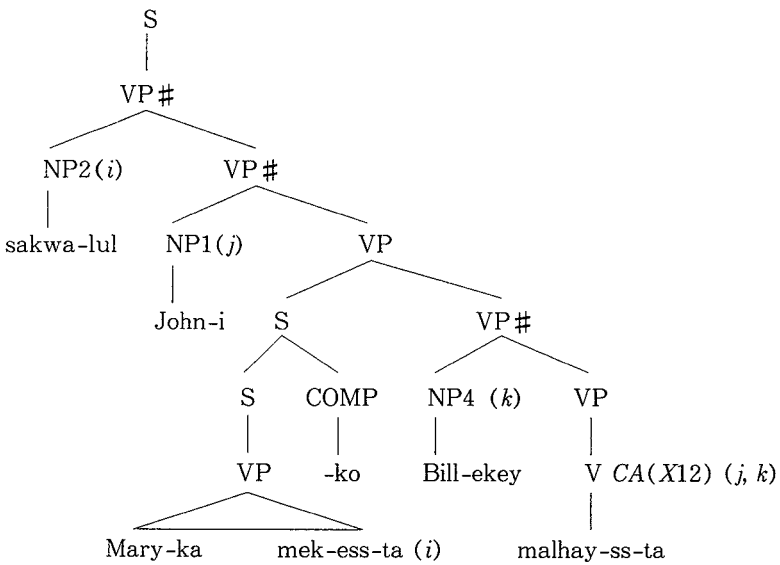

Fig. 4-4. Derived Tree of Long-distance Scrambling

[16] The right daughter node of VP regards as one complete constituient that cannot be interuded. In Korean, only the right constituients of VP can particpate in total ordering.

[17] The multiple adjunction sites reminescent the coindexing in unification-based formalisms.

The preposed NP is also governed within the domain of its predicate. Note that NP2 sakwa-lul is still bounded by its dominant predicate mek-. As we can see, the long-distance scrambling is an adjunction site problem in TAG. In other words, scrambling can be analyzed with the same principle whether it is local or long-distance scrambling. With our formalism, we can explain the ambiguity shown in long-distance scrambling. In the long-distance scrambling, the ambiguity is propotional to the number of the preposed arguments[18].

(15) hakkyo-eyse sakwa-lul John-i
school-P3 apple-P2
[Mary-ka mek-ess-ta]-ko Bill-ekey malhay-ess-ta.
eat -COMP say
"John said to Bill that Mary ate an apple in the school."

In summary, the long-distance scrambling is a matter of adjoining node in TAG. When we consider the government relations between arguments and its dominant predicate, it is obvious that long-distance scrambling is one sort of cross-serial dependencies.

5. Conculsion

The word-order is the skeletal structure to constitute the linguistic structures and the word-order variation is one of important operational characteristics. There are many evidences that we cannot regard word-order variation as a dependent on pragmatic or discoursal process. The variation should be captured on syntactic basis since natural languages have partially ordered syntactic structures.

We present the TAG's description to handle word-order and its variation of Korean on syntactic basis. We review the general properties of scrambling. Many evidences concerned on scrambling suggest that scrambling cannot be analyzed in terms of movement. We show that any arguments can be scrambled as long as they can keep their thematic roles in a sen-

[18] This seems to have close relationships with multiple adjunction site.

tence. Scrambling-α seems to be plausible to understand scrambling phenomenon. Scrambling requires the structures that can represent the binding domain and subcategorization frame at the same time. Since TAG has the extended domain of locality, the word order variation can be realized on the local trees. To represent subcategorization frame, we propose a new adjoining constraint CA, which is effective to realize VP-adjunction of scrambling. Since the binding relation is also linear precedence for word order in our analysis, TAG can represent both syntactic structure and precedence relation in the same tree. We discuss that long-distance scrambling is the same phenomenon that provokes local scrambling. We show that this is one sort of cross-serial dependencies and a matter of adjunction site in TAG.

Though we do not persuit the semantic constraints shown in scrambling, we expect these can be properly handled in the elementary trees of TAG. In general, all languages have the partial relations between syntactic categories. In this case, the extended domain of locality will provide more concrete base for word-order and its variation.

Acknowledgement: I appreciate Dr. Aravind K. Joshi to support me to continue the research at his institute and discuss theoritical issuses on TAG formalism. I also appreciate Language Research Institute, Seoul National University, that organized SICONLP-90. This paper is the revised version of my paper presented at SICONLP-90.

References

Baltin, M. R. (1987) 'Degree Complements,' in Huck, G. J. and Ojeda, A. E., (1987) *Syntax and Semantics*, Vol. 20, Academic Press Inc.

Chomsky, N. (1982) 'Some Concepts and Consequences of the Thoery of Government and Binding,' *Linguistic Inquiry Monograph* 6, Cambridge, Mass.: The MIT Press.

Gazdar, G., Klein, E. and Pullum G. K. (ed.) (1983) *Order, Concord and Constituency*, Foris Publications.

Gazdar, G., Klein, E., Pulluum, G. and Sag, I. (1985) *Generalized Phrase Structure Grammar*, Harvard Univ. Press, Cambridge.

Hale, K. (1983) 'Warlpiri and the Grammar of Non-configurational Lan-

guages,' *Natural Language and Linguistic Theory*, 1. 1.

Han, S. K. (1988) 'On the Kernel Particles of Korean,' *Proc. of lst Conference of Cognitive Science*, Seoul.

Hoji, H. (1987) 'Weak Crossover and Japanese Phrase,' In T. Imai et al. eds., *Issues in Japanese Linguistics*, Dordrecht: Foris.

Imai, T. and Saito, M. (1987) *Issues in Japanese Liguistics*, Foris Publications.

Joshi A. K. et al. (1975) *Tree Adjunct Grammars*, J. Compt. Syst. Sci.

Joshi, A. K. (1987) 'An Introduction to Tree Adjoining Grammars,' in *Mathematics of Language* edited by A. Manaster-Ramer, John Benjamins Publ. Co., PA.

Karttunen, L. and Kay, M. (1985) 'Parsing in a Free Word Order Language,' in *Natural Language Parsing* edited by Dowty, D. R., Karttunen, L. and Zwicky A. M., Cambridge University Press.

Kroch A. S. and Joshi A. K. (1985) 'The Linguistic Relevance of Tree Adjoining Grammar,' *CIS Technical Report MS-CIS-16*, Univ. of Pennsylvania.

Kroch A. S. and Joshi A. K. (1987) 'Analyzing Extraposition in a Tree Adjoining Grammar,' in *Syntax and Semantics*, Vol. 20.

Kuno, S. (1973) *The Structure of Japanese Language*, Cambridge, Mass., The MIT Press.

McCawley, J. D. (1982) 'Parenthetical and Discontinuous Constituent Structure,' *Linguistic Inquiry* 13, pp 91-106.

Mohanan, K. P. (1982) 'Infinite, Subjects, Goverment and Abstract Case,' *Linguistic Inquiry*, 13.

Ojeda, A. E. (1988) 'A Linear Precedence Account of Cross-Serial Dependencies,' *Linguistics and Philosophy*, Vol. 11, No. 4.

Saito, M. (1987) 'Three Notes on Syntactic Movement in Japanese,' in Imai. T. and Saito, M. (1987).

Department of Computer Engineering
Won Kwang University
344-2, Shinyong-dong, Iri
Chen-buk 570-749
Korea