



HAL
open science

Verb Cluster, Non-Projectivity and Syntax-Topology Interface in korean

Jihye Chun

► **To cite this version:**

Jihye Chun. Verb Cluster, Non-Projectivity and Syntax-Topology Interface in korean. DepLing2013, Aug 2013, Prague, Czech Republic. pp.51-59. halshs-00867290

HAL Id: halshs-00867290

<https://shs.hal.science/halshs-00867290>

Submitted on 28 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Verb Cluster, Non-Projectivity, and Syntax-Topology Interface in Korean

Jihye Chun

MoDyCo – UMR7114

University of Paris Ouest Nanterre La Défense
France

chunjihye@gmail.com

Abstract

This article proposes a simple modeling of Korean word order within the framework of the topological dependency grammar – the first topological modeling for this language – a system of formal rules accounting for the correspondence between the dependency tree of a sentence and an ordered constituent structure. We show that a fairly small number of linearization rules can account for the word order facts of Korean, considered to be a language with a relatively free order. These rules will be described, especially the non-projectivity phenomenon based of the notion of “verb cluster”, a cohesive topological constituent, which appears in a syntax-topology interface.

1 Introduction

First of all, let us consider the following examples. The difference between these declarative sentences is the placement of the two verbs, *ta-ko* ‘take’ and *ka-ss-da* ‘go’, marked in bold:

- (1) a. 영이가 엄마 차를 **타고**
Yeongi-ka eomma cha-leul **ta-ko**
Yeongi-NOM mother car-ACC take-VM
시골에 **갔다**
sigol-e **ka-ss-da**
country-LOC go-P-DEC¹
‘Yeongi went to the countryside taking her mother’s car’
- b. 영이가 시골에 엄마 차를
Yeongi-ka sigol-e eomma cha-leul
Yeongi-NOM country-LOC mother car-ACC

¹ ACC: accusative, ADV: adverb, C: copula, DAT: dative, DEC: declarative, HON: honorification, LOC: locative, NEG: negation, NM: nominalization, NOM: nominative, P: past, PRES: present, REL: relative, TOP: topic, VM: verbal morpheme with which verb dependents are combined (note that Korean is an agglutinative language).

타고 **갔다**
ta-ko **ka-ss-da**
take-VM go-P-DEC

‘Yeongi went to the countryside taking her mother’s car’

These two sentences are acceptable and natural. On the other hand, when the nominal dependent of *ta-ko* ‘take’ is extracted, we remark that there are restrictions on the placement of verbs. Let us observe the following examples in which the nominal dependent of *ta-ko* ‘take’ is extracted:

- (2) a. 영이가 시골에 **타고** **간**
Yeongi-ka sigol-e **ta-ko** **ka-n**
Yeongi-NOM country-LOC take-VM go-REL
엄마 차는 검은색이다
eomm cha-neun keomeunsaek-i-da
mother car-TOP black-C-DEC
‘the mother’s car that Yeongi took for going to the countryside is black’
- b. ?* 영이가 **타고** 시골에 **간**
Yeongi-ka **ta-ko** sigol-e **ka-n**
Yeongi-NOM take-VM country-LOC go-REL
엄마 차는 검은색이다
eomm cha-neun keomeunsaek-i-da
mother car-TOP black-C-DEC

As illustrated above, when the dependent verb is separated from its governor (example 2b), it is not possible that the nominal dependent of this dependent verb is extracted. On the other hand, when the dependent verb is placed next to its governor (example 2a), it is possible that the nominal dependent is extracted. This fact leads us to think about the correlation between extraction and constraints on the placement of verbs.

In this paper, we are interested in this restriction of the linear position of verbs in the case of extraction, and we’d like to propose a simple modeling for this linguistic fact. To do this, we think that it is necessary to present a

general description of Korean word order. Then, we will discuss the selection of suitable framework for word order variation of this language (section 2). We will propose to be placed in a topological approach based on Gerdes & Kahane (2001). In section 3, we will define a topological structure for Korean, based on its word order property. We will also develop a topological dependency grammar. In section 4, we will show that our grammar is fully capable of establishing the correct linear order, with non-projectivity phenomena illustrated in the examples above. We will show the utility of the notion of the “verb cluster”, a cohesive topological constituent which controls non-projectivity phenomena.

2 Word Order Variation of Korean

Korean is known as a language in which word order is relatively free (Chung 1998, Choi 1999, Kim & Lee 2001 etc). First of all, let us consider the following examples in which the verb *ju-* ‘give’ has three nominal dependents: *Yeongi-ka* ‘Yeongi’, *Cheolsu-eke* ‘to Cheolsu’, and *chaek-eul* ‘a book’. The order variation of these constituents permits six possible orders:

- (3) a. 영이가 철수에게 책을 주었다
 Yeongi-ka Cheolsu-eke chaek-eul ju-eoss-da
 Yeongi-NOM Cheolsu-DAT book-ACC give-P-DEC
 ‘Yeongi gave a book to Cheolsu’
- b. 영이가 책을 철수에게 주었다
 Yeongi-ka chaek-eul Cheolsu-eke ju-eoss-da
 Yeongi-NOM book-ACC Cheolsu-DAT give-P-DEC
 ‘Yeongi gave a book to Cheolsu’
- c. 철수에게 영이가 책을 주었다
 Cheolsu-eke Yeongi-ka chaek-eul ju-eoss-da
 Cheolsu-DAT Yeongi-NOM book-ACC give-P-DEC
 ‘Yeongi gave a book to Cheolsu’
- d. 철수에게 책을 영이가 주었다
 Cheolsu-eke chaek-eul Yeongi-ka ju-eoss-da
 Cheolsu-DAT book-ACC Yeongi-NOM give-P-DEC
 ‘Yeongi gave a book to Cheolsu’
- e. 책을 영이가 철수에게 주었다
 chaek-eul Yeongi-ka Cheolsu-eke ju-eoss-da
 book-ACC Yeongi-NOM Cheolsu-DAT give-P-DEC
 ‘Yeongi gave a book to Cheolsu’
- f. 책을 철수에게 영이가 주었다
 chaek-eul Cheolsu-eke Yeongi-ka ju-eoss-da
 book-ACC Cheolsu-DAT Yeongi-NOM give-P-DEC
 ‘Yeongi gave a book to Cheolsu’

As illustrated in these examples, the verb occurs at the end of these sentences (i.e. Korean is a verb final language), while the nominal elements of the main verb are freely placed.

This linear behavior of nominal elements in the examples above brings about two questions: First, is the idea of *standard* word order SOV pertinent in the case of Korean? Second, how could we represent word order variation in a simple and elegant way?

Greenberg (1963) proposed classifying word order types of languages from a typological point of view in terms of their *basic order* such as SVO, SOV, VSO, VOS, OSV or OVS. This proposition implies a fixed or at least clearly dominant order, which does not apply to Korean word order variation as illustrated above.² Moreover, following Ross (1967), Korean word order variation has been discussed in terms of “scrambling” (Han 1998, Chung 1998, Choi 1999 etc.), which demands the concept of *movement*. We believe that this operation could make such a representation very complicated for Korean where the communicative structure plays an important role, unlike English or French where word order largely depends on the syntactic function.

According to Choi (1999) and Kim & Lee (2001), there are several factors intervening in Korean word order variation: grammatical morphemes, communicative structure, syntactic functions etc. From this point of view, in our study, we do not suppose a *standard* word order contrary to X-bar syntax in which syntactic function and constituency are represented in a same structure. We believe that it is more convincing to separate different levels of information, for representing word order variation in a simple way. We thus propose to use the framework of Dependency Grammar where syntactic function and constituency are separately represented (Tesnière 1959, Mel’čuk 1988). More precisely, we base ourselves on Gerdes & Kahane (2001), broadly inspired by the classical topological model first introduced in the description of German. They integrated this model into the framework of dependency grammar, elaborating a syntax-topology interface. Note that topology is an intermediate level between a dependency-based syntactic structure and a prosodic structure.³ That is to say,

² Note that Korean also has an unmarked order which is communicatively neutral.

³ We are based on the Meaning-Text model (Mel’čuk 1988) which posits multiple strata of representations

word groupings in topology are strongly related to prosodic units.

In this paper, keeping the issue raised in the introduction in mind, we describe the correspondence between unordered syntactic structures and ordered constituent structures on the basis of the Korean topological dependency grammar we propose. We will remark that linearization rules are simpler than we thought for a language considered as a relatively free order language. This will be described precisely in the section 4 with the case of non-projectivity phenomena. This description will be a solution of the question raised in the introduction.

3 Syntax-Topology Interface in Korean

In this section, we propose a topological model for Korean, based on its word order property. We also define the Korean topological dependency grammar, which accounts for all possible word order variations. This grammar will allow for describing the correspondence between a given dependency tree and an ordered topological structure.

3.1 Topological Structure for Korean

The basic idea of the topological model is to “consider that a sentence is a template-like sequence of different fields each being able to host different types of constituents” (Gerdes & Kahane 2007). These different types of constituents correspond to “domains”. The internal structure of domains is a sequence of “fields”. Choi (1999) and Gerdes (2002) insist on the fact that constraints on word order in Korean resemble those of German. However, we do not follow the architecture of the topological structure of German, in which the superior domain directly contains five fields (Vorfeld, left bracket, Mittelfeld, right bracket, Nachfeld). Korean is often considered as a discourse-oriented language (Kim 2003), or a topic-prominent language (Li & Thompson 1976). In other words, the communicative structure plays an important role in the organization of sentences in Korean. Furthermore, all the elements are not obligatorily present. That is to say, it is not necessary to produce the elements that speakers understand in a given context (i.e. frequency of zero anaphora, cf. Kim 2003). This could make interpretation of the structure

related by explicit interfaces. We are interested especially in an interface where linearization takes place.

of sentences complicated, and we could have at least two interpretations: 1) elements in a sentence are under subcategorization of the main verb, and 2) they are simply repetition of elements of the antecedent sentences in a context.

Moreover, the topological behaviors of the *neun*-phrase are very interesting in that it can appear in different linear order depending on the communicative structure (cf. Chun 2013). Consider the following examples in which the two *neun*-phrases appear:

(4) a. 그 이야기는 저는 들었어요
 keu iyaki-neun jeo-neun deul-eoss-eo-yo
 this story-TOP me-TOP hear-P-DEC-HON
 ‘for that story, it’s me who heard that’

b. 저는 그 이야기는 들었어요
 jeo-neun keu iyaki-neun deul-eoss-eo-yo
 me-TOP this story-TOP hear-P-DEC-HON
 ‘for me, it’s that story that I heard’

As illustrated in the gloss of the examples above, their interpretation differs from the linear position of the *neun*-phrase: in (4a), *keu iyaki-neun* ‘that story’ is interpreted as a topic, while in (4b), this is interpreted as a focus contrastive. This means that there is a particular linear position of elements of a sentence. Furthermore, the first constituent containing the *neun* marker in each sentence tends to be separated from the following element with a high prominence in prosody (Seong & Song 1997, Hwang 2002).

These points we mentioned above lead us to introduce two syntactic modules: the macro-syntax and the micro-syntax following Blanche-Benveniste (1990). The latter is for elements in “proper” syntactic level, while the former contains detached elements which don’t fall under subcategorization. We believe that the introduction of two syntactic modules (macro- and micro-syntax) allows us to better understand the overall organization of sentences on different levels: syntactic level, communicative level and prosodic level etc.

Let us now present our architecture of the Korean topological structure (cf. Figure 1). The two modules, the macro- and micro-syntax are integrated into our model as a macro-domain and a micro-domain respectively. Note that for a sequence of fields in the macro-domain, we introduce the term “kernel”⁴ interpreted as

⁴ We borrow this term from Gerdes *et al.* (2005).

“noyau” in French following Blanche-Benveniste (1990). The macro-domain is composed of three fields: the pre-kernel field, the kernel field, and the post-kernel field. The kernel field “receives the other verbal dependents, especially all the elements that saturate the verbal valency” (Gerdes *et al.* 2005). The pre-kernel field and post-kernel field contain detached elements such as topicalized or dislocated elements. In the example (5), there is one more element apart from the elements under subcategorization of the main verb, namely, *na-neun* ‘me’ interpreted as a topic. This means that this element is placed in another topological position than in the kernel field:

(5) **나는** 내가 많이 발전했지
 na-neun nae-ka manhi baljeonha-eoss-ji
 me-TOP me-NOM much progress-P-DEC
 ‘for me, I have much progressed’

We thus propose the pre-kernel field before the kernel field.

Note that Korean nominal dependents can appear after the main verb. In the example (6), the subject is placed after the main verb *ha-neunde* marked in bold:

(6) 다시 생각을 해야 **하는데** 너는
 dasi saengkak-eul ha-eoya ha-neunde neo-neun
 again thought-ACC do-VM do-although you-TOP
 ‘you have to think again’

Therefore, we need one more field behind the kernel field, i.e. the post-kernel field.

The micro-domain has two fields: the principal field and the head field. This domain hosts the elements governed by the principal verb. This principal verb is placed in the head field, and its dependents occupy the principal field.

In our architecture of the topological structure, there is a particular verbal grouping of words, namely, a verb cluster, in that this is not a simple verbal constituent, but a constituent which intervenes in the case of extraction. This is our main problem in this paper, and we will discuss that in detail in the section 4. The verb cluster is composed of four fields: the dependent verb (dep-V) field, the adverb (ADV) field, the negation (NEG) field and the verb (V) field.⁵ The verb cluster forms a very rigid verbal constituent with great cohesion, which

⁵ Note that the verb cluster is not a domain such as macro- and micro-syntax.

tends to form one prosodic unit.⁶ The order of these four fields is fixed:

(7) 요즘 영이 공부 잘 안 해
 yoseum Yeongi kongbu jal an ha-eo
 these days Yeongi-TOP study well NEG do-DEC
 ‘these days, Yeongi doesn’t study well’

Note that certain constructions do not permit insertion of adverb or negation⁷ between two verbs, contrary to the example (7):

(8) a. 영이가 철수를 도와 주었다
 Yeongi-ka Cheolsu-leul dou-a ju-eoss-da
 Yeongi-NOM Cheolsu-ACC help-VM do a favor-P-DEC
 ‘Yeongi helped Cheolsu with favor’

b. * 영이가 철수를
 Yeongi-ka Cheolsu-leul
 Yeongi-NOM Cheolsu-ACC
 도와 안 주었다
 dou-a an ju-eoss-da
 help-VM NEG do a favor-P-DEC
 ‘Yeongi didn’t help Cheolsu with favor’

We now propose the Korean topological structure with three embedded levels:

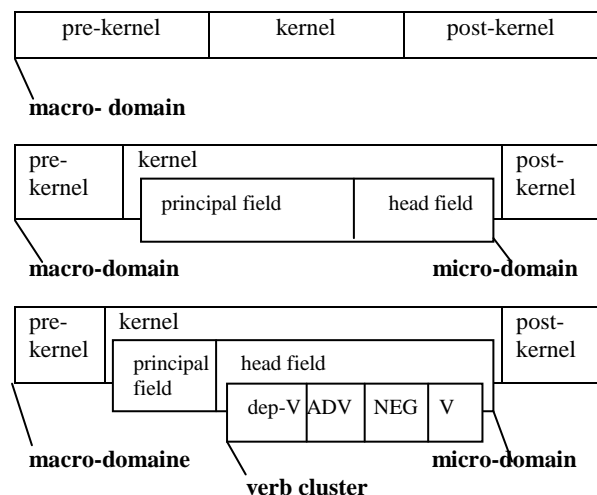


Figure 1. Three embedded levels of the topological structure in Korean⁸

⁶ There is also a nominal cohesive constituent corresponding to the verb cluster, i.e. a noun cluster. The noun cluster is a topological unit with a strong cohesion among nouns.

⁷ Korean has two negations: short negation such as *an* appearing in front of verbs, and long negation such as an auxiliary verb *anh-*.

⁸ There is no such field proposed for the complementizer. This is related to the fact that Korean is an agglutinative language in which it is the morphemes that play a role of complementizer.

3.2 Topological Dependency Grammar

In this section, we develop a topological dependency grammar for Korean, based on Gerdes & Kahane (2001) in which the parameters of topological dependency grammar are defined as follows:

- ▶ Six components of a grammar
 - the vocabulary V
 - the set of (lexical) categories C
 - the set of syntactic relations R
 - the set of box names B
 - the set of field names F
 - the field of initialization i
- ▶ Order of permeability of the boxes (which is a partial ordering on B used for emancipation⁹)
- ▶ Four sets of rules
 - box description rules
 - field description rules
 - correspondence rules
 - box creation rules

Note that in a topological approach, non-projectivity phenomena are related to the notion of “emancipation”, which means that “the dependents of a verb do not have to be placed in their governor’s domain” (Gerdes & Kahane 2007). We will give the order rules for linear placement of nominal dependents in terms of emancipation.

We now present the six components of the Korean grammar, and two of the four sets of rules, i.e. the box description rules and the field description rules, in a formalized manner. For clarity, the correspondence rules and the box creation rules are going to be described in natural language, and at the same time we will show the steps of the derivation of a declarative sentence.

▶ Six components of the Korean grammar

- V = the Korean words
 C = {V, V-*eo*, V-*ji*, V-*ko*, V-*myeonseo*, *neun*-phrase...Y}
 R = {subj, obj, obji, attr, mod, cv¹⁰}
 B = {macro-domain, micro-domain, verb cluster}

⁹ The definition of the notion of emancipation is going to be followed after this presentation of parameters.

¹⁰ This means “verbal complement”, for verbal dependents such as infinitive, completive.

- F = {pre-kernel field, kernel field, post-kernel field, principal field, head field, dep-V field, lexical field}
 i is the field of initialization

▶ Permeability order

micro-domain > verb cluster

This formula means that it is easier for the nominal dependent to be emancipated from the micro-domain than from the verb cluster.

▶ Box description rules

This rule indicates that such a box is composed of the list of fields.

- macro-domain → pre-kernel field, kernel field, post-kernel field
 micro-domain → principal field, head field
 verb cluster → dep-V field, ADV field, NEG field, V field

▶ Field description rules

Still following Gerdes & Kahane (2001), we present the field description in the form of pair (f,ε) in which f is a field and ε is a symbol among {!,?,+,*}. The pair means that the field f has to contain exactly one element (!), at most one element (?), at least one element (+) or any number of elements (*):

- (i, !), (lexical field, !), (pre-kernel field, *), (post-kernel field, *), (kernel field, !), (head field, !), (principal field, *), (dep-V field, ?)

▶ Correspondence rules and box creation rules

We have established the following correspondence rules and box creation rules for the linearization of verbs and their dependents.

- 1) The root of the dependency tree opens a macro-domain containing three fields, i.e. the pre-kernel field, the kernel field, and the post-kernel field. Then, the principal verb opens a micro-domain in the kernel field having two fields, the principal field and the head field. It finally opens a verb cluster in the head field:

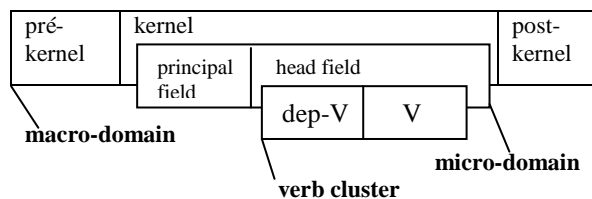


Figure 2. Illustration of the topological structure of Korean

2) The principal verb opens a field for its dependent verb, after being placed in the verb cluster. The latter can occupy this dep-V field in the verb cluster, where it opens an embedded verb cluster:

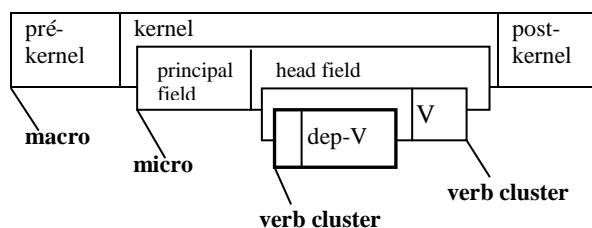


Figure 3. Dependent verb in the verb cluster

If the dependent verb of the root has its dependent verb, the latter proposes anew a place for its possible dependent. This process is *recursive*.

3) The dependent verb is not obliged to stay with its governor in the verb cluster:

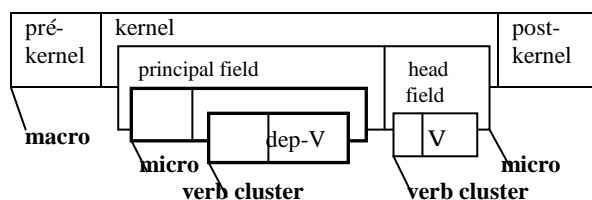


Figure 4. Micro-domain of the dependent verb in the principal field

The place of the dependent verb largely depends on the type of verbal morphemes with which it is combined, and on the communicative structure:

- The *V-eo/ji* obligatorily goes into the verb cluster;
- The *V-ko* has two possibilities: to stay in the verb cluster or to create a micro-domain in terms of the communicative structure;

- The *V-myeonseo*¹¹ is obliged to create a micro-domain in one of the three major fields (the pre-kernel field, the principal field and the post-kernel field).

4) Other non-verbal predicative dependents such as the predicative noun can join the dep-V field in the verb cluster. In this case, it is preferable that the predicative noun does not bear any markers (cf. example 7)

5) All dependents of a verb can create a subdomain that should be placed in one of the three major fields in terms of the communicative constraints:

- Any dependent can go into the principal field without emancipation;
- Any dependent can go into the pre-kernel field with possibly emancipation from a verb cluster;
- Any dependent can go into the post-kernel field with possibly emancipation from a micro-domain.

6) The *neun*-phrase interpreted as a topic has to be emancipated from the micro-domain; the *neun*-phrase interpreted as a contrastive focus should remain in the principal field without emancipation;

7) In the case of extraction, the verbs belonging to the verbal nucleus (cf. section 4 for its definition) governing the extracted element must form a verb cluster.

We now try to solve our problem presented in the examples (1) and (2). In the following section, we will show how the correspondence rules and box creation rules are applied from a given dependency tree. In particular, we will see that the dependent verb should go into the verb cluster, created by its governor, in the case of extraction of its nominal dependent.

4 Non-projectivity and Verb Cluster

In this section, based on our topological dependency grammar, we will insist on the utility of the verb cluster in the syntax-topology interface for solving constraint on the relation between extraction and the placement of verbs in sentences with relatives. We will also show that in our analysis, unlike Ross (1967), it is

¹¹ This is considered as a morpheme which marks an adverbial clause.

not necessary to consider the concept of movement or “island constraint” phenomena. Throughout this section, we refer to the rule 7 which allows for describing non-projectivity phenomena in a simple way.

We have shown the relation between extraction and restriction on the placement of verbs contrary to the case of declarative sentences (The examples (2a) and (2b) are reproduced here as (9) for convenience of the reader):

(9) a. 영이가 시골에 타고 간
 Yeongi-ka sigol-e ta-ko ka-n
 Yeongi-NOM country-LOC take-VM go-REL
 엄마 차는 검은색이다
 eomm cha-neun keomeunsaek-i-da
 mother car-TOP black-C-DEC
 ‘the mother’s car that Yeongi took
 for going to the countryside is black’

b. ?* 영이가 타고 시골에 간
 Yeongi-ka ta-ko sigol-e ka-n
 Yeongi-NOM take-VM country-LOC go-REL
 엄마 차는 검은색이다
 eomm cha-neun keomeunsaek-i-da
 mother car-TOP black-C-DEC

The problem is whether two verbs form a constituent or not. How can we account for this phenomenon? We postulate the notion of “verbal nucleus”, a syntactic position of a single verb, which can also receive a sequence of verbs, a notion introduced by Kahane (1997), for modeling of non-projectivity phenomena. This means that in the dependency tree, we postulate that the syntactic position of verbs or complex units containing a sequence of verbs corresponds to one verb. For example, verbal nuclei in English are auxiliary-participles (*be eating, have eaten*), verb-infinitives (*want to eat*), verb-conjunction-verbs (*think that...eat*), and verb-prepositions (*look for*) (cf. Kahane 1997). Our hypothesis is that in the case of extraction, the verbal nucleus tends to form a topological constituent, i.e. the verb cluster. The following figure shows a dependency tree with a relative, and a topological constituent containing the two verbs is superimposed on this tree. The dotted oval represents the verbal nucleus:

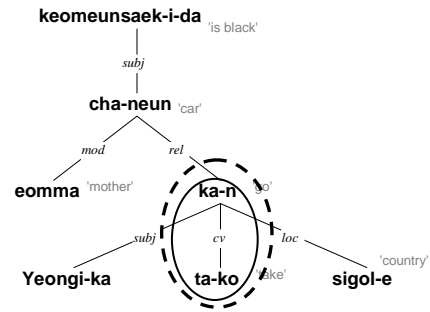
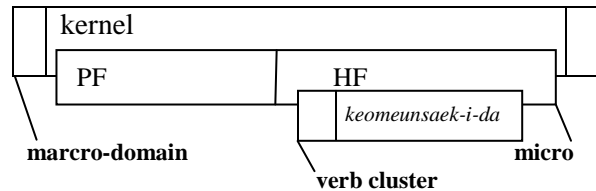


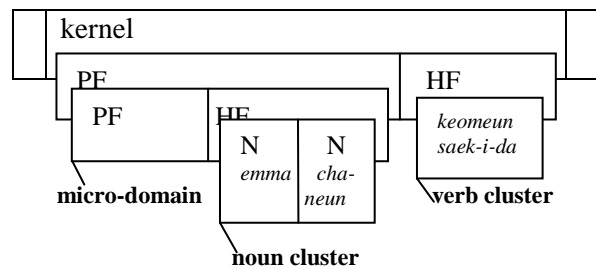
Figure 5. Dependency tree of the examples (9a)

Let us show how the correspondence is established from this dependency tree of the example (9a), referring to our grammar. Recall our correspondence rules and box creation rules.

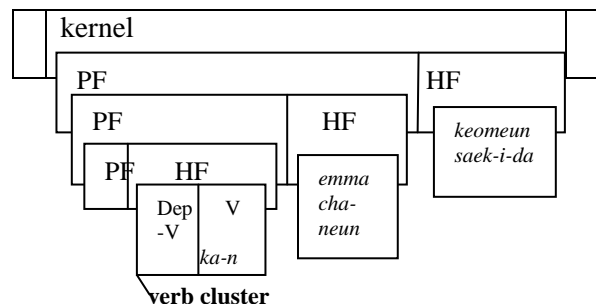
- 1) The root of this tree opens a verb cluster after creating a macro-domain and a micro-domain. And then it is placed in the head field:



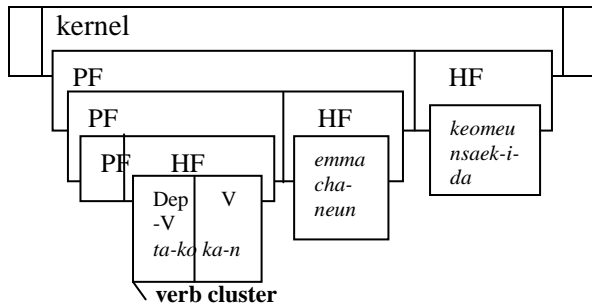
- 2) The nominal head of the relative opens a micro-domain in the principal field as a nominal dependent of the root. It is placed in the field proposed for nouns in the noun cluster. Its dependent *eomma* ‘mother’ rejoins the noun cluster:



- 3) *ka-n* ‘go’ opens a verb cluster in the head field:



- 4) *ta-ko* ‘take’ **has to** be placed in the dep-V field of the verb cluster opened by its governor *ka-n* ‘go’, instead of creating an independent constituent:



- 5) Finally, the nominal dependents of *ka-n* ‘go’ and those of *ta-ko* ‘take’ rejoin the principal field:

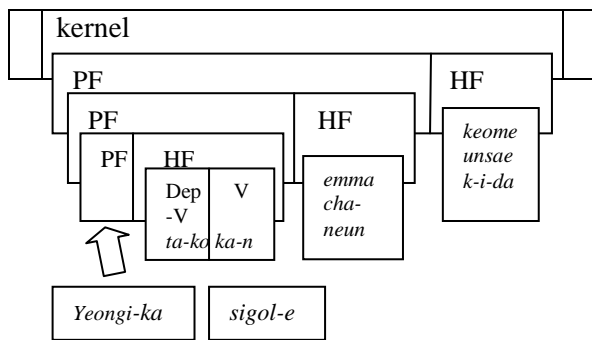


Figure6. Topological structure of the example (9a)

Let us consider other possible linearization if the two verbs in a verbal nucleus do not form a topological constituent, i.e. a verb cluster.

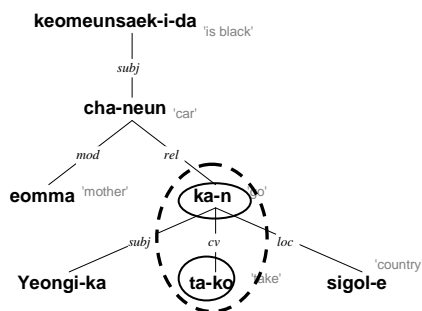


Figure7. Dependency tree of the examples (9b)

In this case, we could not have the example (9a). One of the word orders from the Figure 7 corresponds to example (9b), which is not natural, rather ungrammatical. This point enhances our hypothesis above. In corpora, we can find data where the relation between extraction and the placement of verbs is attested:

- (10) a. 평소에 그거 입고 다니는
 pyeongso-e keukeo ip-ko dani-neun
 usually-LOC this wear-VM go-REL
 아저씨 많이 봤어
 ajeossi manhi bo-ass-eo
 man much see-P-DEC
 ‘I saw a man who was walking wearing this’

- b. 사회자가 이끌어 가는
 sahoija-ka ikkeul-eo ka-neun
 announcer-NOM lead-VM go-REL
 대화가 큰 비중을
 daewha-ka keu-n bijung-eul
 conversation-NOM be big-REL importance-ACC
 차지한다
 chajiha-nda
 occupy-PRES.DEC
 ‘the conversation that the announcer lead has a great importance’

We thereby believe that the notion of verb cluster is useful to describe non-projectivity phenomena.

5 Conclusion

We have discussed and proposed a simple solution of the description of extraction in terms of the verb cluster, in the framework of the topological dependency grammar, a simple modeling of Korean word order.

On the level of modeling word order variation, we have shown that our model allows us to determine the order of nominal and verbal dependents, with a small number of correspondence rules or box creation rules. Moreover, we have shown that we can describe the word order variation, postulating only three types of boxes: the macro-domain, the micro-domain, and the verb cluster. The internal structure of these boxes is much simpler than those of German, considered as a language with similar word order properties as Korean (Choi 1999, Gerdes 2002). For example, we have shown that in Korean, it is sufficient for the micro-domain to have only two fields (principal field and head field) for the relative, the completive and the nominal groups.

This paper is the early stage in a study of the topology for Korean. We should investigate further in various directions. Especially, we are aware of the importance to understand the communicative (or information) structure of sentences, which plays a crucial role in linearization. The word order rules should be further developed to include constraints on the communicative structure.

We have simply mentioned the topological behaviors of the *neun* marker, referring to Chun (2013). Korean is an agglutinative language. This means that in addition to this morpheme *neun*, further work will have to be done to study the topological behaviors of other morphemes such as *eul*, traditionally considered as an accusative marker, but more recently as a marker of communicative values (Han 1999) which, of course, is related to its linear position. In other words, understanding its syntactic behaviors and communicative values could allow for characterizing its topological position as being in the macro- or micro-domain.

References

- Claire Blanche-Benveniste. 1990. *Le français parlé: études grammaticales*, CNRS Editions, Paris.
- Hye-Won Choi. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*, CSLI Publications, Stanford.
- Jihye Chun. 2013. *Interface syntaxe-topologie et amas verbal en coréen et en français*, Ph.D. Dissertation, University of Paris Ouest Nanterre La Défense, Paris.
- Chan Chung. 1998. Argument Composition and Long-distance Scrambling in Korean, in Erhard Hinrichs(ed.), *Complex predicates in Nonderivational Syntax : Syntax and Semantics*, 30:158-220, Academic Press, New York.
- Kim Gerdes. 2002. *Topologie et grammaires formelles de l'allemand*, Ph.D. Dissertation, University of Denis Diderot, Paris.
- Kim Gerdes and Sylvain Kahane. 2001. Word order in German: a Formal Dependency Grammar Using a Topological Hierarchy, *Proceedings Association for Computational Linguistic*, Toulouse.
- Kim Gerdes and Sylvain Kahane. 2007. Phrasing it differently, in Leo Wanner (ed). *Selected lexical and grammatical issues in the Meaning-Text Theory*, 297-335.
- Kim Gerdes, Sylvain Kahane and Hiyon Yoo. 2005. On the descriptive adequacy of topology, *Proceedings of MTT'05*, Moscou.
- Joseph Greenberg. 1963. *Universals of Language* (2nd edition), MIT Press, Cambridge.
- Chung-Hye Han. 1998. Asymmetry in the Interpretation of -(n)un in Korean, *Japanese and Korean Linguistics*, 7:1-15.
- Jeonghan Han. 1999. *Morphosyntactic Coding of Information Structure in Korean*, Ph.D. Dissertation, State University of New York at Buffalo.
- Son-Moon Hwang. 2002. Hankukeo hwaje kumun-eui unyul-jeok kochal (Study on Prosody of the Korean topic construction), *Eumseong Kwahak*, 9(2):59-68.
- Sylvain Kahane. 1997. Bubble trees and syntactic representations, in Tilman Becker and Hans-Ulrich Krieger(eds), *Proceedings of 5th Meeting the Mathematics of Language (MOL5)*, Saarbrücken.
- Jong-Bok Kim and Minghaeng Lee. 2001. Realizations of Information Structure and Its Projection in Korean, *Harvard Studies in Korean Linguistics*, IX:463-494, Hanshin Publishing Company, Seoul.
- Mi-Young Kim. 2003. *An Optimality Approach to the Referential Interpretation of Zero Anaphora in Korean*, Ph.D. Dissertation, Seoul National University, Seoul.
- Charles Li and Sandra Thompson. 1976. Subject and Topics: A New Typology, in Charles Li (ed), *Subject and Topic*, Academic Press, New York.
- Igor Mel'čuk. 1988. *Dependency Syntax: Theory and Practice*. The SUNY Press, Albany, N.Y.
- John Ross. 1967. Constraints on Variables in Syntax, Ph.D. Dissertation, MIT.
- Cheol-Jae Seong and YoonKyoung Song. 1997. Jueojali josa-eui unyoulpaeteon-e kwanhan siheomeumseonghak-jeok yeonku (Experimental study on particles in the subject position), *Malsoli*, 33/34:23-42.
- Lucien Tesnière. 1959. *Éléments de syntaxe structurale*. Klincksieck, Paris.