1994

Parsing Korean morphological expressions of syntactic functions

Inrae Kim Shin
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

---

Recommended Citation

Shin, Inrae Kim, "Parsing Korean morphological expressions of syntactic functions" (1994). *Master's Theses*. 804.

DOI: <https://doi.org/10.31979/etd.du86-gu3g>
https://scholarworks.sjsu.edu/etd_theses/804

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 1358228

Parsing Korean morphological expressions of syntactic functions

Shin, Inrae Kim, M.A.

San Jose State University, 1994

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

PARSING KOREAN MORPHOLOGICAL
EXPRESSIONS OF SYNTACTIC FUNCTIONS

A Thesis

Presented to

The Faculty of the Department of Linguistics
and Language Development
San Jose State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

by

Inrae Kim Shin

May, 1994

©

1994

Inrae Kim Shin

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF LINGUISTICS
AND LANGUAGE DEVELOPMENT

Carol J. Justus

Dr. Carol Justus

Peter Lowenberg

Dr. Peter Lowenberg

Estrella Calimag

Dr. Estrella Calimag

APPROVED FOR THE UNIVERSITY

M. Lou Lewandowski

ABSTRACT

**PARSING KOREAN MORPHOLOGICAL
EXPRESSIONS OF SYNTACTIC FUNCTIONS**

BY

INRAE KIM SHIN

This thesis shows how a unification-based formalism for parsing natural language as exemplified by English can also parse Korean language structures. In an agglutinative language such as Korean, where each morpheme has an independent grammatical meaning, it also has semantic scope over structures that are syntactically independent of its position in the sentence. Functional composition, a property of unification-based grammar, allows syntactically separate entities to fuse semantically.

Unification-based grammar has the property of leaving open a large number of different approaches to the parsing of free word order languages. It is possible within this formalism to separate a grammar into a number of components. This formalism can unify semantic components of the grammar which may be expressed in the morphology or syntax of a language.

ACKNOWLEDGEMENTS

My special thanks go to Prof. Carol Justus, my advisor, for her advice and guidance. She has given her time to discuss problems and to suggest improvements in style, organization, and other matters.

My sincere gratitude also goes to Prof. Peter Lowenberg and Prof. Estrella Calimag for their helpful criticism and advice in improving this study.

Last, but not least, I am grateful to my parents, my husband and my son, without whose help this work could never have been done.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION.....	1
CHAPTER 2. COMBINATORY CATEGORIAL GRAMMAR IN A UNIFICATION-BASED FORMALISM.....	3
2.1. Combinatory Categorical Grammar.....	3
2.1.1. Grammatical Categories.....	4
2.1.2. Directionality.....	6
2.1.3. Functional Composition.....	8
2.2. A Unification-Based Formalism.....	11
2.2.1. The Information Domain.....	11
2.2.2. A Graph Unification-Based Formalism	13
2.2.3. CCC in a Graph Unification-Based Formalism..	14
CHAPTER 3. KOREAN LANGUAGE STRUCTURES.....	17
3.1. Case Structure.....	18
3.1.1. Case Markers and Postpositions.....	18
3.1.2. The Case System.....	24
3.2. Verbal Suffixes.....	34
3.2.1. Passive.....	34
3.2.2. Tense.....	40
3.2.3. Aspect.....	46
3.2.4. Mood.....	51

CHAPTER 4. UNIFICATION-BASED KOREAN GRAMMAR.....	53
4.1. Parametric Rules for Korean.....	53
4.1.1. Constituent Structures and Functional Structures.....	53
4.1.2. Combinatory Rules.....	57
4.1.3. Bottom-up Parsing	60
4.2. Parsing Templates.....	61
4.3. Lexicon.....	62
 CHAPTER 5. CONCLUSION.....	 63
 APPENDIX.....	 66
FOOTNOTES.....	78
REFERENCES.....	78

Chapter 1. INTRODUCTION

Over the last decade or so, unification-based grammar formalisms have attracted a lot of attention and have been used productively in both computational linguistics and theoretical linguistics as alternatives to previous grammar formalisms. Concurrent studies in both theoretical and computational linguistics have led to various grammar frameworks based on those formalisms and to computational efforts to use the formalisms for natural language processing.

This thesis aims to show how efficiently natural language parsing of Korean language structures could be achieved with a unification-based formalism. Korean is a free word order, verb-final language with agglutinative morphology, in which most grammatical functions are expressed by single morphemes. Korean uses morphological forms to express functions that are expressed syntactically in English. For a long time, Korean structure has been discussed in the context of Transformational Generative Grammar(TG). TG is excellent for handling issues of formal syntax, but a problem with TG is that it is unable to handle morphological expression of grammatical functions adequately.

In Chapter 2 I examine a Combinatory Category Grammar (CCG) which has the advantage that it handles both formal (syntactic) constituent structures and underlying functional categories. I discuss the use of this syntactic model with a unification-based formalism for natural language processing purposes.

In Chapter 3 I introduce Korean structures which cannot be handled without a framework such as that of Categorical Grammar. These include the distribution and function of Korean case markers and postpositions and their propositional relations to the verb. Chapter 3 then describes each category of verbal suffix. Verbal suffixes fall into categories on the basis of the positional slots that they fill. Verbal suffix categories include passive, aspect, tense and mood.

In Chapter 4 I analyze Korean structures using a unification-based grammar formalism. This analysis lays the basis for a future system for parsing Korean syntax. It includes the composition of module files which are given in greater detail in the Appendix. Here I examine ways in which a parser can use a well-coded lexicon to access functional grammatical information encoded in Korean as morphology or in English as syntax.

CHAPTER 2. COMBINATORY CATEGORIAL GRAMMAR IN A UNIFICATION-BASED FORMALISM

In this chapter, I describe a Combinatory Categorical Grammar (CCG) and the use of this syntactic model with a unification-based formalism for natural language processing purposes.

2.1. Combinatory Categorical Grammar

Combinatory Categorical Grammar (CCG), first proposed by Ades and Steedman (1982) within the general framework of Categorical Grammar (Adjukiewicz 1935, Bar-Hillel 1953, Montague 1974, Bach 1984), has the advantage that it has (i) the devices to represent unbounded extraction, i.e., constructions such as questions, relative clauses, and topicalized sentences in which a grammatical dependency exists between what are called fillers and their associated gaps, (ii) the ability to capture certain types of conjunction which have been, until now, beyond the capabilities of phrase-structure grammar, and the methods by which one represents word order (Wittenburg 1986). For example, in the sentence 'who(m) do you think he loves?', there is a gap after 'loves', where one expects the declarative sentence to have a direct object. The interrogative 'who(m)' is the filler for that gap.

2.1.1. Grammatical Categories

The difference between CCG and phrase structure grammars is that the constituency information in phrase structure grammars is realized in the formulation of lexically assigned grammatical categories. CCG instead consists of a lexicon and a set of combinatorial operations. Categories that are defined in terms of the type of element with which they combine ('argument') and the type of element that results are called 'functors'. Functor categories are of the form X/Y , which is viewed as a function from categories of type Y to categories of type X . Thus, for instance, in English, a determiner such as 'the' might be assigned the category NP/CN , an indication that it is a function from common nouns to noun phrases. Thus, in the mathematical sense, a grammatical derivation for a phrase consisting of a determiner followed by a common noun would apply the function signified by NP/CN to the argument signified by CN , yielding as a value the category NP as the category assigned to the string 'the man'.

English tensed transitive verbs might carry the category $(S\backslash NP)/NP$. This can be viewed as a function from object noun phrases to another function, namely $S\backslash NP$, which is a function from subject noun phrases to sentences. Derivations of sentences with categories like these and only

a simple 'cancellation' operation reflect a structuring of functional applications. A derivation of the sentence 'the man watches the game' is shown in (1).

In (1), the functor category assignments at the lexical level already have in place the category of some higher phrase. There are no atomic labels corresponding to lexical and phrasal categories that are typically considered the heads of a higher phrase. The verb 'watches' at the lexical level is assigned a category that is an incomplete S, which must find two noun phrases in order to complete itself. This notational difference is the formal difference between basic categorial grammars with only functional application defined over adjacent elements and context-free phrase structure grammars.

2.1.2. Directionality

Directionality in categories has been disputed among linguists. Those who have taken the position that the categories are directional include Bar-Hillel (1953), Lyons (1971), Bach (1984), and Steedman (1985b). Those who do not assume that categories are directional take the view that word order is a function of the combination rules and that there is a separate ordering component of the grammar that generalizes across individual categories and rules (Montague 1974, Flynn 1983). Assuming that categories are inherently directional amounts to a partial description of word order. Directionality indicates whether each argument in a functor category is found to the right or to the left of the functor. Steedman (1985b) uses the following notation: a right-slanting slash indicates that the argument must appear to the right of the functor as in X/Y ; a left-slanting slash indicates that the argument must appear to the left of the functor, as in $X\backslash Y$ (X is the functor and Y is the argument).

Imposing directionality on categories entails that we have two versions of a basic functional application rule. Forward functional application ('fa>') is shown in (2a), backward functional application ('fa<') in (2b).

(2) a. Forward Functional Application (fa>)

$$X/Y \quad Y \quad \Rightarrow \quad X$$

b. Backward Functional Application (fa<)

$$Y \quad X \backslash Y \quad \Rightarrow \quad X$$

Therefore, the following sentences (3) may be derived by the forward functional application or the backward functional application.

(3) a. John   went

NP   S \ NP

-----fa<

S

b. Mary   ate   an   apple

NP   (S \ NP) / NP   NP / CN   CN

-----fa>

NP

-----fa>

S \ NP

-----fa<

S

The pattern derivation shown by the underlines has the same structure as a conventional phrase-structure tree for these sentences, but no syntactic structure is produced. The string of categories is progressively reduced towards S, with the semantic interpretation developed at each step, but

there is no level of syntactic analysis. In Categorical Grammar, every rule of syntax is, inherently, a rule of semantics.

2.1.3. Functional Composition

The difference between CCG and other forms of Categorical Grammar is the use of functional composition as a primitive syntactic operation in the grammar. Functional composition is a combinatory operation in which a functor category is composed out of two functor categories. Steedman (1985b) assumes the four basic versions of forward and backward functional composition as in (4).

(4) Forward Functional Composition (fc>)

- a. $X/Y \quad Y/Z \Rightarrow X/Z$
- b. $X/Y \quad Y \backslash Z \Rightarrow X \backslash Z$

Backward Functional Composition (fc<)

- c. $Y/Z \quad X \backslash Y \Rightarrow X/Z$
- D. $Y \backslash Z \quad X/Y \Rightarrow X \backslash Z$

A classical CG offered no obvious way to deal with 'discontinuous dependency', i.e., constructions like the fronting of objects and the separation of verbs from their particles. The solution adopted in TG was to posit movement rules which applied to deep structures like 'you did see

who?', 'he looked up it' and produced the desired surface output 'who did you see?', 'he looked it up'. CCG accounts for topicalization, wh-questions, relative clauses, right-node raising and non-constituent conjunction in a different way. Let's look at the examples of how the operation can be exploited to account for unbounded extraction in English¹.

Ades and Steedman (1982) use the recursive power of generalized composition to build onwards from the main clause until a derivational constituent is formed which can apply (backwards) to the fronted object and give the correct semantic interpretation. Steedman's (1985a) basic account of extraction in English develops his earlier proposal slightly, using type-raising as an alternative to backward combination rules. This gives derivations like (6), where the fronted topic NP has been raised to the (non-directional) category of a function into a topicalized

sentence from a function-into-a sentence from an NP
 (S/(S/NP)), that is, a subject and transitive verb together
 looking for their object. The \$ notation is a variable in
 rules that can match any sequence of categories and slashes.

(6) Those cakes I can believe that she will eat

2.2. A Unification-Based Formalism

The representation language for encoding linguistic information is based on a unification-based formalism developed at SRI (Stanford Research Institute) in the PATR project. Some of PATR's mathematical properties are further detailed in Pereira and Shieber (1984), where the PATR formalism is used as a basis for comparing and contrasting the formal properties of the feature-based parts of Generalized Phrase Structure Grammar (Gazdar and Pullum 1985) and Lexical Functional Grammar (Kaplan and Bresnan 1982). Shieber (1986) gives an overview of unification-based formalism. He describes the PATR grammar formalism as the least common denominator of the unification-based formalisms to be found in Lexical Functional Grammar (Kaplan and Bresnan 1982), Generalized Phrase Structure Grammar (Gazdar et al. 1985), Functional Unification Grammar (Kay 1985), Definite Clause Grammars (Pereira and Warren 1980), and Head Grammar (Pollard 1985).

2.2.1. The Information Domain

Unification-based formalisms use as their informational domain a system based on features and their values. The elements of this domain are referred to as feature structures. A feature structure is a partial

function from features to their values. For instance, we might have a function mapping the feature 'number' onto the value 'singular' and mapping 'person' onto 'third'. The common notation for such a feature structure is D3sg (Shieber 1986:13):

```
[number: singular
  person: third]
```

One of the distinguishing characteristics of unification-based formalisms is that the feature values may themselves be structured. For example, D3sg might be just one component of a larger structure associated with a third-person singular noun phrase.

```
[cat: NP
  agreement : [number: singular
               person: third]]      (Dnp3sg)
```

This notion of combining the information from two feature structures to obtain a feature structure that includes all the information of both is central to unification-based formalisms.

2.2.2. A Graph Unification-based Formalism

Feature structures can be viewed as rooted, directed graph structures whose arcs are labeled with feature names. The important characteristics of a graph unification-based formalism is in the string rewriting portion of the formalism. For example, the feature structure 'Dnp3sgsubj' would be expressed in a graph-structural notation as the following:

Sets of labeled arcs point to nodes whose content may be either atomic or else an additional set of labeled arcs. The graph represented in (7) is fully equivalent to the graph represented by the matrix notation in which reentrant structures are indicated by a kind of anaphoric indexing device in (8).

```

(8) [cat: NP
      agreement: 1[number: singular
                    person: third]
      subject: <agreement 1>]

```

The graph in (8) begins with an entry point onto a node consisting of a set of attributes (also referred to as features), in this case, 'cat' 'agreement' and 'subject', whose values may be either atomic, as with 'NP', or complex, as with 'number: singular, person: third'. Reentrancy is indicated by an integer within brackets '<1>', which is to be read much like a pronoun whose antecedent is the value following the same integer without the brackets.

2.2.3. CCC in a Graph Unification-Based Formalism

All grammatical categories of the grammar begin with two arcs at the entry point in their graphs: an arc labeled 'struct' and an arc labeled 'subcat'. This is the basic distinction between a basic non-terminal categorial structure and a list of subcategorized-for arguments of that structure.

The maximal set of arcs that can be present as value of the 'struct' node also happens to be the same set that can appear at any subconstituent node in the structural representation. There are three such arcs in the set:

'cat' which takes as value some atomic grammatical category, such as 'S', or 'NP'; 'feats'; and 'constits'.

As for the sets of arcs that may appear as values of 'constits', there is a complementary distribution between two sets. When the constituent is terminal, the attribute 'lex' is the only arc that may appear after 'constits', followed by the root form of the lexical item. On the other hand, if the constituent is nonterminal, underneath 'constits' there may appear arc labels such as 'head', 'obj', 'mod', etc., each pointing to subconstituents. Each grammatical category has a particular set of these functional labels that can appear as paths to its subconstituents.

At the 'subcat' node value, the maximal set of arc labels is '{first rest}' which is used to form an ordered list. As values of each of these attributes, either the atomic symbol '#' may appear, or else the maximal arc label set '{struct subcat}'. 'Struct' and 'subcat' are used to access those same paths in the argument categories themselves that are unified during rule applications. For example, the transitive verb 'eat' might be described as the following:

(9) eat

```
[struct: [cat: VP
          feats: [vform: []
                  mood: []
                  tns: []
                  constits: [v head: [constits: [lex: eat]]]
                  compls: [first: <1>
                           rest: #]]]
  subcat: [first: [struct: 1[cat: NP
                           feats: [case: acc]
                           subcat: #]
          rest: #]]]
```

What happens during a derivation is that the rules will effect a unification at the value of '<subcat first struct>', and since this value is shared by the '<compls first>' of the VP, the properties of the NP argument are passed up into the resulting structural description. The argument to NPs is constrained with accusative case by including these features in this shared structure.

CHAPTER 3. KOREAN LANGUAGE STRUCTURES

Various theoretical aspects of Korean grammar have been studied for a long time. Many Korean linguists, I.S. Yang (1972), D.W. Yang (1975), H.B. Lee (1975), among others, have studied various aspects of Korean syntax and semantics. All the recent analyses (in the '70's and '80's), however, have been studied in the context of Transformational Generative Grammar (TG), which has been predominant in formal syntax in the past thirty years. But a problem with TG is that it is unable to handle grammatical function adequately. In this chapter, I am going to discuss Korean structure in the framework of Categorical Grammar that represents language form and function.

Korean is a free word order, verb-final language with agglutinative morphology, in which most grammatical functions can be mapped onto single morphemes. For example, the verb form used in a sentence such as 'The man has been caught' has the structure in (1):

- (1) ku   pun-un   cap -hi -epeli -si -ess -ta
the man -topic 1 -2 -3   -4 -5 -6
1: catch   2: passive   3: aspect
4: honorific   5: past   6: declarative marker

Although Korean is a 'head-final' language that favors SOV sentence patterns, the grammar allows relatively free ordering of a verb's arguments and adverbial modifiers. Hence the following two sentences are both acceptable:

- (2) a. Haksang-i chaek-ul ilk-ess-ta.
student-N book-Acc read-Past-Decl
'The student read the book.'
- b. Chaek-ul haksang-i ilk-ess-ta.
book-Acc student-N read-Past-Decl
'The student read the book.'

This 'free word order' option is made possible by the existence of the case suffixes that represent the sentence function.

3.1. Case Structure

In this section, I am going to discuss the distribution and function of Korean case markers and postpositions and their propositional relations to the verb.

3.1.1. Case Markers and Postpositions

Korean case markers are the nominative suffix *-ka* (after stems ending in a vowel, *-i* elsewhere), the accusative *-lul*

(after stems ending in a vowel, -ul elsewhere), the genitive -uy, and the topic suffix -nun (after stems ending in a vowel, -un elsewhere). The following sentences exemplify the use of Korean case suffixes with NPs bearing a variety of thematic roles (cf also O'Grady 1991).

(3) a. Nominative suffix marking an actor

Kay-ka John-ul mwul-ess -ta.

dog-N -Acc bite-tense

'The dog bit John.'

b. Nominative suffix marking a patient

John-i mwul -li-ess-ta.

-N bite -passivization

'John was bitten.'

c. Nominative suffix marking a possessor

Mary-ka dari-ka aphu-ta.

-N foot-N hurt

'Mary's foot hurts.'

(4) a. Accusative suffix marking a patient

Ai-ka pap-ul mek-ess-ta.

child-N rice-Acc eat-tense

'The child ate the food.'

b. Accusative suffix marking an actor

Na-ka ai-lul hakkyo-ey ka-keyhay-da.

I-N child-Acc school-to go-make

'I let the child go to school.'

- c. Accusative suffix marking a possessor

Kay-ka John-ul bal-ul mwul-ess-ta.

dog-N -Acc foot-Acc bite

'The dog bit John's foot.'

- (5) a. Genitive suffix marking a possessor

Na-uy chaek

I-Gen book

'my book'

- b. Genitive suffix marking an actor and a patient

John-uy casin-uy manjok

-Gen self-Gen satisfaction

'John's satisfaction of himself'

The major difference between case markers and postpositions² that include -eyse 'in, at, on', -eyuyhayse 'by', (u)lo 'by means of', -hako 'with', is categorial: whereas the case suffixes mark NPs, the postpositions head PPs. As the following sentences show, case suffixes differ from postpositions in not being able to precede another postnominal particle.

- (6) Case markers cannot precede other particles

a. * Ai-ka-man pap-ul mek-ess-ta.

child-N-only food-Acc eat

'Only the child ate the food.'

b. *Kay-ka John-ul-un ilk-ess-ta.

dog-N -Acc-Topic read-past

'The dog bit John, not someone else.'

c. *Na-uy-man chaek

'my only book'

(7) Postpositions can precede other particles

a. Hakkyo-eyse-nun haksang-i kongpwu hann-da.

school-at -Topic student-N study do

'The students are studying at school.'

b. John-i gicha-lo-man yohangha-ss-ta.

-N train-by-only travel

'John travelled only by train.'

A second characteristic distinguishing case suffixes from postpositions is that the occurrence of case markers in sentences is optional in that a marker may be freely dropped when the case of its host can be understood from context.

(8) Ai(-ka) pap(-ul) mek-ess-e?

child-N food-Acc eat-tense

'Did the child eat the food?'

In contrast, deletion of postpositions is generally not permitted.

- (9) Haksang-i Hakkyo-(*eyse) kongpwu han-ta.
 student-N school-(at) study
 'The student is studying at school.'

Dative particle -eykey, which marks the NP as goal and recipient, is sometimes classified as a postposition and sometimes as a case marker (Hong 1985). The following sentences (10a through 10b) in which -eykey may co-occur with a topic marker and a postposition show that -eykey patterns like a postposition rather than a case marker.

- (10) a. -eykey with a topic marker
 John-eykey-nun nay-ka chayk-ul cwu-ess-ta.
 -Dat -Topic I-N book-Acc give-tense
 'As for John, I gave a book to (him).'
- b. -eykey with a postposition
 John-eykey-man nay-ka chayk-ul cwu-ess-ta.
 Dative-only I -N book-Acc give-tense
 'I gave a book to only John.'

As the following sentence shows, -eykey cannot be deleted optionally:

(11) Na-ka John-(*eykey) malhay-ta.

I-N   -Dat   speak

'I speak to John.'

Topic suffix *-n(un)* which marks discourse notions such as 'topic' or 'contrast', is attached to an NP and does not distinguish among grammatical relations such as subject and direct object, as in (12):

(12) (a) Topic marker associated with a subject

Na-nun   hakkyo-ey   kan-da.

I-Topic   school-to   go

'I go to school.'

(b) Topic marker associated with a direct object

Saka-nun   John-i   cohahan-da.

apple-Topic   -N   like

'John likes apples (not oranges).'

As these sentences show, the prototypical position of a topic-marked NP is at the beginning of the sentence, preceding even the subject. At first glance, topics seem to be a type of sentential adjunct, occurring as sisters of S. I will return to this matter in the next section.

3.1.2. The Case System

Korean case suffixes and certain postnominal particles combine with the NP as an argument of the functor category (S\NP or VP\NP). The following structure illustrates how the case suffixes work:

In this structure, the NP bearing the nominative suffix is an argument of the functor S while the accusative-marked NP is an argument of the functor VP. In other words, the nominative case suffix *-ka* marks an NP that combines with VP category while the accusative case suffix *-lul* marks an NP that combines with TV category. As mentioned before, since Korean is a free word order language, word order is not helpful in determining the combinatory category. Instead, case suffixes have a crucial role in identifying the category.

The nominative suffix *-ka* and the accusative suffix *-lul* might be assigned the category NP\N in the NP structure. However, the NP bearing the nominative suffix and the NP bearing the accusative suffix are totally different in determining category. Sentence (13) can be analyzed as follows in (14) using functional composition:

As shown in the structure, functional composition is not enough to differentiate case markers *ka* and *lul*. A mechanism is needed to add information about the NP. A unification-based formalism can serve as this mechanism to add information about the features of NP. I will discuss this matter in Chapter 4.

The Genitive suffix

Now consider the genitive suffix *-uy*, which marks an NP bearing the possessor role. It is generally assumed in categorial grammar (Dowty 1982), that genitives are 'modifiers' that combine with an N to give a larger nominal phrase. The genitive is a dependant category of the type N/N (i.e. a function from noun to noun). Let's look at the following structure.

In (15), the N 'John' is converted into the NP that can combine with the adjacent nominal *cha* 'car'. The genitive case marks an NP that combines with an N category.

The Topic Suffix

Consider the topic marker *-(n)un*, in the following structures:

(16) John-nun   pap-ul   mek-ess-ta.

-Topic   food-Acc   eat

'As for John, he ate the food.'

(17) Mary-nun   John-i   cohahan-da.

-Topic   -N   like

'As for Mary, John likes her.'

As these sentences show, the prototypical position of a topic-marked NP is at the beginning of the sentence, preceding even the subject if this latter element is not itself topicalized. It seems clear that these sentences must have the structures (18) and (19).

(18)

(19)

It is possible to propose a parallel structure for topic constructions that contain an apparent gap which corresponds to a null pronoun. The empty pronoun not only fills the subject position (18) or the object position (19) but also, by virtue of its coreference with the topic, helps to establish the relation between a topic and the S. The NP bearing the topic suffix is a sentence adjunct (a function from S to S).

The Dative Postposition

The postposition *-eykey* (or its free variant *-hanthey*) is attached to a recipient or goal nominal in sentence (20).

(20) John-i   Mary-eykey   chaek-ul   cwu-ess-ta.
         -N   -Dat   book-Acc   give

'John gave a book to Mary.'

be atomic predicates when they serve as complements, as in the ditransitive case in the following sentence:

The Korean ditransitive verb *cwu-ess-ta* 'gave' seems to take PP as a predicate rather than a modifier. Thus, sentence (20) will have the structure (23).

In (23), *-ul* marks the NP combining with a TV while *-i* marks the NP combining with the VP category. The dative

postposition -eykey heads a PP. The complement PP combines with the TV *cwu-ess-ta*. Another approach is to introduce feature-marking on NPs or TV. This is the central power of unification-based grammar, which will be discussed. The sentence (20) can be formulated as (24) by functional composition:

(24)	John-i	Mary-eykey	chaek -ul	cwu-ess-ta.
	-N	-Dat	book -Acc	give
	-----	-----	-----	-----
	NP	PP	NP	((S\NP)\PP)\NP

			(S\NP)\PP	

		S\NP		

	S			

A major advantage of treating the dative suffix as a PP-creating particle is represented in passive sentences such as (25), in which -eykey can be used to mark the demoted actor. (The preposition -eyuyhayse 'by' can also be used to mark the actor in the passive.)

(25)	Mary-ka	kay-eykey	mwul-li-ess-ta.
	-N	dog-Dat	bite-Pass

'Mary was bitten by the dog.'

In this sentence, *kay-eykey* 'by the dog' combines with the passive IV *mwul-li-ess-ta* 'was bitten'. In this sentence PP *kay-eykey* is different from the PP in (20). In (25), PP functions as an adverb ((S\NP)\(S\NP)) while in (20), PP functions as a complement of a verb. Therefore, sentence (25) has a syntactic structure (26), and can be formulated as (27) by functional composition:

Although the dative -eykey belongs to the postposition category, it has a special status in Korean structure. The dative -eykey seems to mark NPs as subcategorized complements and demoted actors in passive constructions.

3.2. Verbal Suffixes

A Korean verb stem consists of a root together with zero or more verbal suffixes. The Korean finite verb complex consists of the following positional categories, in the order given, as illustrated in example (1).

- (1) Cap   -hi   -epeli   -si   -ess   -ta
     catch   -Passive   -Aspect   -Honorific -Tense -Mood

The rest of this chapter describes each category of verbal suffix by positional categories--passive, aspect, honorific, tense and mood.

Several of the verb suffixes of various categories have vowel alternations -a or -e in initial position conditioned by the last phoneme of the preceding stem (e.g. -epeli, -apeli for perfective aspect, -ass, -ess for past tense).

3.2.1. Passive

The passive verb form of Korean is not so uniform as that of English. Korean exhibits two types of passive structure: a lexically restricted 'morphological' passive formed with the suffix -i (-li after lateral [l]; -ki after nasals, [kk] and [s]; and -hi after other fricatives and stops); and a more productive 'auxiliary passive' formed

with the help of the inchoative morpheme *-ci* 'get/become' (*-ci* after stems ending *-y*, otherwise as *-eci*)³. The demoted agent, when expressed, is typically marked by either *-eykey* or *-eyuyhayse*:

Morphological passive

- (2) Bumin-i kyungchal-eykey cap-hi-ess-ta.
criminal-N policeman-Dat arrest-pass
'The criminal was arrested by the policeman.'

Auxiliary passive formed with *-ci*

- (3) Ai-ka yomma-eyuyhayse katwu-e- ci-ess-ta.
child-N mother-by lock up-Inf-get
'The child was locked up by mother.'

Still another type of passive structure is used when the main predicate is a noun. Here the auxiliary verb *toy-ta* 'become' is used to signal a passive structure.

- (4) Bumin-i kyungchal-eykey cheypho(-ka) toy-ess-ta.
criminal-N police -Dat arrest -N become
'The criminal was arrested by the policeman.'

The dominant view in the literature on Korean seems to be that only structures formed with the suffix *-i* or the morphemes *-ci* and *toy* are true passives (Shin 1982, Hwang

1975). Nonetheless, the *-tangha* or *-pat* construction, which contains a patient-like subject, might be analyzed as a compound passive (Lee and Kim 1988).

- (5) Mary-ka John-eykey kwangkan tanghay-ess-ta.

N   Dat   rape   suffer

'Mary was raped by John.'

- (6) Mary-ka John-eykey conkyeng pat-ass-ta.

N	Dat	respect	receive

'Mary received respect from John.'

Passivization converts a TV category that takes an actor and a theme as an argument into a IV category that takes a theme as its argument. A typical Korean passive sentence can be shown in the following structure (7):

Passive morpheme *-li* combines with TV *mwul-ta* 'bite' converting TV into an IV that takes a nominative NP as an argument (a function from TV to IV). The NP bearing an actor role becomes the PP marked by the postposition *-eykey*. In passive sentences, the PP marked by *-eykey* functions as a modifier of a verb. Sentence(7) could be analyzed by functional composition as the following:

(8)	Na -ka	bam-eykey	mwul	-li	-ess	-ta.
	I -N	sanke-Dat	bite	pass		
	NP	(S\NP)\(S\NP)	(S\NP)\NP	(S\NP)\((S\NP)\NP)		

				(S\NP)		

				S\NP		

				S		

Now, let's look at the passive construction of ditransitive verbs. In English, there are two types of passive constructions: a direct passive (9b) and an indirect passive (9c). The indirect passive, in which a dative object of an active sentence becomes a passive subject, is more common than the direct passive in modern English (Quirk, et al. 1985). Compare:

- (9) a. John gave Mary a book.
b. A book was given to Mary by John.
c. Mary was given a book by John.
-

In Korean, on the other hand, the dative object cannot normally be passivized (Kim 1985), because transitives have no passive form for the indirect passive. Instead, a notional passive using a contrastive verb with passive meaning (10d) is used, as shown in the following examples.

- (10) a. John -i   Mary -ekey   chaek -ul   cwu-ess-ta.
              -N   -Dat   -Acc   give
              'John gave a book to Mary.'
- b. Chaek-i   Mary-ekey   John-eyuyhayse   cwu-a-ci-ess-ta.
              -N   -Dat   -by   give-Pass
              'A book was given to Mary by John.'
- c. *Mary-ka   chaek-ul   John-eyuyhayse   cwu-a-ci-ess-ta.
              -N   -Acc   -by   give-Pass
              'Mary was given a book by John.'
- d. Mary-ka   chaek-ul   John-eykeyse   pat-ass-ta.
              -N   -Acc   -from   receive
              'Mary received a book from John.'

The ditransitive verb *cwu-ta* 'give' needs the inchoative morpheme *-ci* to be passivized in (10b). The Korean ditransitive verb has a function $((S \backslash NP) \backslash PP) \backslash NP$ ⁴. The passive morphemes convert a TV category into an IV category. Therefore, *cwue-ci-ess-ta* has the following analysis:

(11)	<i>cwue</i>	<i>-ci</i>	<i>-ess</i>	<i>-ta.</i>
	give		pass	
	$((S \backslash NP) \backslash PP) \backslash NP$	$(S \backslash NP) \backslash ((S \backslash NP) \backslash NP)$		

	$(S \backslash NP) \backslash PP$			

Sentence (10b) can be analyzed by functional composition as follows:

(12)	<i>Chak-i</i>	<i>Mary-eykey</i>	<i>John-eyuyhayse</i>	<i>cwue-ci-ess-ta.</i>
	book-N	-Dat	-by	give -Pass
	-----	-----	-----	-----
	NP	PP	$(S \backslash NP) \backslash (S \backslash NP)$	$(S \backslash NP) \backslash PP$

	$(S \backslash NP) \backslash PP$			

	$S \backslash NP$			

	S			

3.2.2. Tense

Korean tenses are indicated by the morphemes -n 'present', -ess 'simple past', and -keyss 'future'. Since these morphemes occur paradigmatically in a verb form, a general statement can be made as follows: tense suffixes come after passive, aspect, honorific suffixes and before the mood suffix.

The predicate with the morpheme -n (as -nun after consonants) is said to have the meaning of 'present' time. The meaning of present tense in Korean, however, is not limited to the time of the situation in a single place. A more characteristic use of the Korean present tense -n is referring to situations which occupy a much longer period of time, including future situations. Compare:

(13) a. 'Universal time'

Han kang-un   Seoul-lo   hulu -n -da.

river-Topic   -to   flow -pre T

'The Han river flows into Seoul.'

b. 'Habitual'

John-un   mayil   hakkyo-ey   ka -n -da.

-Topic   everyday   school-to   go -Pre T

'John goes to school every day.'

c. 'Future'

Na-nun   nayil   New york-ey   ka -n -da.

I-Topic   tomorrow   -to   go -Pre T

'I go to New York tomorrow.'

The morpheme -ess has alternations depending upon the final element of a verb stem. It can be realized as -ass (preceded by -o) or -yess (preceded by -ha, or -ss). The predicate with the morpheme -ess designates any action or description which has been finished before speech time. Compare:

(14) a. Na-nun   ecey   kongwon-ey   ka -ess-ta.

I-Topic   yesterday   park -to   go -Past T

'I went to park yesterday.'

b. Mary-un   eryuswutae   yepp -ess-ta.

-Topic   childhood   pretty -Past T

'Mary was pretty when she was young.'

The function of the future morpheme -keyss has been disputed. The morpheme -keyss can be interpreted either modally or as a future tense depending upon the subject and the absence/presence of tense or aspect marker in the sentence. Consider the following examples:

- (15) a. Na-nun   nayil   John-ul   manna-keyss-da.
          I-Topic tomorrow   -Acc   meet -Fut T
          'I will meet John tomorrow.'
- b. Ne-nun   nayil   edi -ey   ka-keyss-ni?
          you-Topic   tomorrow   where-to   go-Fut T
          'Where are you going to go tomorrow?'

If the marker *-keyss* is used with the first person pronoun *na*, it indicates the speaker's intention for the future. Also, if *-keyss* is used in an interrogative form with second person *ne*, the speaker is asking the addressee about his/her intention or opinion of the future. In the absence of other markers of tense and aspect, *-keyss* is interpreted as a future time. On the other hand, if the subject or the topic of the sentence is other than the speaker or the addressee, the declarative sentence which includes *-keyss* denotes an opinion or presumption about the subject or the topic regardless of any specific tense (Lee and Im 1983).

Compare:

- (16) John-i   Seoul-lo   ka-epelie-keyss-ta.
          -N   -to   go-Aspect-Modal
          'John has gone to Seoul.'

In sentence (16), -keyss does not represent any future meaning. Rather than representing future meaning, -keyss gives the meaning of 'prediction' or 'intention'. That is, in the presence of markers of tense or aspect, -keyss is interpreted as modal. We can say that the dispute⁵ on the morpheme -keyss (i.e. future tense/modal) comes from the confusion surrounding its same syntactic position in both meanings (after the honorific and before the mood suffix).

Now, let's think about the grammatical and semantic function of tense morphemes. In English, Chomsky (1957) sets up the rule of Aux: Aux--->Tense (Modal) (have-en) (be-ing). In deep structure, Aux is a node of S, not a node of VP.

(17) Deep structure according to Chomsky (1957) analysis

In subsequent X-bar theory, generative transformationalists (Radford 1988) have assumed that Modals are generated under a finite I', and that a finite I' carries Tense' (Past/Present) and Agreement (Person, Number) features. Koopman (1984) proposes the rule of V-to-I' movement by

which the head of V of VP moves into the empty I' position as in (18) below:

McCawley (1988:230) has a different view. He posits tense above the propositional content of S. Therefore, tense dominates the whole sentence S. He analyzes the sentence 'Has Ann read the report?' as follows:

As shown in the above structure, the domain of tense has been disputed for a long time. In Korean structure, many Korean verbal morphemes make up typical cases of 'bracketing paradoxes' (Kang 1987). For example, although they attach to a verb, the grammatical and semantic functions have scope over VP or S (whole clauses or parts of clauses).

Korean passive morphemes which we have discussed in 3.2.1. have scope over VP (they are functions from TV to IV). However, the grammatical semantic functions of tense morphemes have scope over S. Compare the English sentence 'Mary slept' which has the phonological structure (20a) and the semantic interpretation (20b):

- (20) a. (phonology)
Mary-ka	[ja	-ess-ta]
-N	sleep	-Past T
- b. (semantics)
[Mary-ka	ja]	-ess -ta

The tense morpheme -ess does not belong to the propositional S node in the syntactic structure. However, the morpheme -ess gives the PAST feature to S. It functions from S to S[tns]. The verb ja-ess 'slept' can be analyzed as the following by functional composition:

(21) Ja   -ess
       S\NP   S[tns]\S

       S[tns]\NP

Here the argument NP of the verb is inherited through the morphological inflection, so that *ja-ess* has become a one place function which takes an argument to give a proposition with the PAST operator.

3.2.3. Aspect

Korean aspectual meanings are expressed by the simple marker *-peli* 'perfective' and the complex markers *-ko iss/cwung i* 'imperfective'. *Peli-* as an independent verb means 'to throw away' or 'to discard'. As a verbal suffix on the stem of an action verb, it denotes 'finality' or 'completion' of the action of the verb. Compare:

- (22) a. John-i   ku   chaek- ul   peli   -ess-ta.
               -N   that book -Acc   throw -Past T
               'John threw away that book.'
- b. Mary-nun   ka   -peli   -ess   -ta
               -N   go   -Perf   -Past T
               'Mary has gone away.'

c. Na-ka ku conwi -lul peli -epeli -ess-ta.
 I-N that paper-Acc throw -Perf -Past T
 'I have thrown away that paper.'

The forms *-peli* and *-epeli* in (22b) and (22c) are different realizations of same morpheme, *-peli* (*-peli* after *-a*, *-apeli* after *-o*; otherwise, it is realized as *-epeli*). The morpheme *-peli* in the sentence (22b) gives the meaning of 'completeness'. Example (22c) demonstrates the dual function of *peli*. The first *peli* in (22c) functions as the lexical main verb, and the second *-epeli* in (22c) functions as an aspectual marker.

The progressive aspect is indicated by a combination of the two morphemes *-ko* and *iss*, with an intervening word boundary. If the verb *iss* 'to exist' immediately follows *ko* 'and', it denotes that the action of the verb in the *ko* form is in the process of occurring. Since *iss* is morphologically a verb stem, tense and mood suffixes occur after it. Compare:

(23) (a) Na-nun chaek-ul ilk -ko iss -ess-ta.
 I-Topic book -Acc read -Prog -Past T
 'I was reading a book.'

- (b) John-un   pap -ul   mek -ko iss -ta.
                       -Topic   food -Acc   eat -Prog
                       'John is eating the food.'

The construction consisting of *cwung* 'middle' followed by the copular expression *-i* means that something is in progress or is in the midst of a certain action. Compare:

- (24) (a) John-i   cenwha   -hanun   cwung-i-ta.
                       -N   telephone   -do   Prog
                       'John is calling.'
- (b) Mary-ka   kongpwu   -hanun   cwung-i-ess-ta.
                       -N   study   -do   Prog   -Past T
                       'Mary was studying.'

Now, consider the function of Aspect. In English, Aspect belongs to the Aux category (Aux--> Tense (Modal) (have-en) (be-ing)). For a long time the constituency Aux has been disputed (Chomsky 1957, Radford 1988, McCawley 1988). McCawley (1988) posits a verb structure by using the configuration [v' V V']. The English sentence 'Fred may have been singing songs' has the following syntactic structure (McCawley 1988:213):

As shown in the structure, Aspect combines with a V' to make another V'. It functions from V' to V' $((S \backslash NP) / (S \backslash NP))$. In Korean verb structure, the aspect morpheme *-epli* combines with a verb stem in active sentence (26a) while it combines with a passive morpheme in passive sentence (26b). Compare:

- (26) a. Mek   -epeli   -ess   -ta
          eat   -Aspect   -Past T
          'have eaten'
- b. Jap   -hi   -epeli   -ess   -ta
          arrest   -Pass   -Aspect   -Past T
          'have been arrested'

Aspect morpheme *-epli* functions from V' to V'. However, after combination, the verb takes on the feature [Aspect]. Applying functional composition to example (26) one derives (27):

- (27) a.   Mek   -epli   -ess   -ta
           eat   -Aspect   -Past T
           (S\NP)\NP   (S[asp]\NP)\(S\NP)   S[tns,asp]\S[asp]

                           (S[asp]\NP)\NP

                                   (S[tns,asp]\NP)\NP
           'have eaten (something)'
- b.   Jap   -hi   -epli   -ess
           arrest   -Pass   -Aspect   -Past T
           (S\NP)\NP   (S\NP)\((S\NP)\NP)

                           S\NP   (S[asp]\NP)\(S\NP)

                                   S[asp]\NP   S[tns,asp]\S[asp]

                                                   S[tns,asp]\NP
           'have been arrested'

We see here that the verbal suffixes for passive, aspect, and tense have different kinds of scope. Passive has a function from TV to IV. Aspect functions from V' to V'. Tense behaves as a function from S to S.

3.2.4. Mood

The mood suffix indicates not only the kind of sentence (declarative *-ta*, interrogative *-ni*, imperative *-la*, exclamatory *-nunkuna*, e.g.) but also the level of speech (plain, polite *-yo*, *-kunyo*, e.g.). Values for the level of speech and for mood are simultaneously realized by the sentence-final suffix. These morphemes follow the tense markers. Although these morphemes are formally verbal endings, they behave as functions from S[tense] to S[mood]. In the sentence *Na-ka ja-ass-ta* 'I slept', we have suffixes for mood and tense. Functional composition unifies verbal suffixes with underlying categories which have scope over the entire sentence. Compare:

(28)	Na	-ka	ja	-ass	-ta
	I	-N	sleep	-Past T	-Mood

	NP	S\NP	S[tns]\S	S[mood,tns]\S[tns]	

		S[tns]\NP			

		S[mood, tns]\NP			

	S[mood, tns]				

Using unification-based formalisms, we see the difference between categories expressed by suffixes for tense and those for mood. In transformational grammar, we cannot make these distinctions.

CHAPTER 4. UNIFICATION-BASED KOREAN GRAMMAR

In this chapter, I use a unification-based grammar formalism to suggest how features for each category of Korean verbal suffix might be attached to syntactic constituents.

In a Unification-based Korean Grammar (Kim 1989), we can assume the following set of module files.

- (1) Parametric rules for Korean
- (2) Parsing templates
- (3) Lexicon

4.1. Parametric Rules for Korean

In this section, I discuss constituent structures, functional structures and combinatory rules which are basis of a Unification-based Korean Grammar (UKG).

4.1.1. Constituent Structures and Functional Structures

A Unification-based Korean Grammar would assign two levels of syntactic description to every sentence. Phrase structure configurations are represented in a constituent structure. A constituent structure (or 'c-structure') is a conventional phrase structure tree. Surface grammatical functions are represented explicitly at the other level of

description, called functional structure (or 'f-structure'). The f-structure is the sole input to the semantic component, which may translate the f-structure into the appropriate formulas in a logical language.

Constituent structures are formally quite different from functional structures. C-structures are defined in terms of syntactic categories, terminal strings, and their dominance and precedence relationships, whereas f-structures are composed of grammatical function names, semantic forms and feature symbols such as those needed for categories of case suffixes and postpositions or passive, tense, aspect and mood expressed by verbal suffixes. Compare the c-structure derivation in (1-3):

- (1) John-i Mary-ekey chaek -ul cwu-ess-ta.
       -N   -Dat   book -Acc give
       'John gave a book to Mary.'

- (2) S -----> NP VP
     NP -----> N P
     VP -----> PP NP V

The c-structure (3) is assigned to sentence (1) by the rule in (2). A c-structure is determined by a grammar that characterizes all possible surface structures.

The functional structure for a sentence encodes its meaningful grammatical relations and provides sufficient information for the semantic component to determine the appropriate predicate-argument formulas. The f-structure for (1) would indicate that the NP *John-i* is the grammatical subject, *cwu-ess-ta* conveys the predicate, the NP *chaek-ul* is the grammatical object, the PP *Mary-ekey* serves as the complement. The f-structure represents this information as a set of ordered pairs each of which consists of an attribute and a specification of that attribute's value for this sentence. An attribute is the name of a grammatical function (SUBJ, PRED, OBJ, COMP, etc.). The following is a f-structure for sentence (1).

```
(4)  SUBJ  [CASE:  NOM
           THEME:  AGENT
           HONORIFIC:  PLAIN
           LEX:  JOHN-I]
      PRED  [TNS:  PAST
           VOICE:  ACT
           MOOD:  DECL
           LEX:  CWU-ESS-TA(((S\NP)\PP)\NP)]
```

OBJ   [CASE: ACC
           THEME: PAT
           LEX: CHA EK-UL]
 COMP [FEATS: DAT
           THEME: GOAL
           LEX: MARY-EKEY]

The values of the PRED functions are semantic forms. Semantic forms usually arise in the lexicon and are carried along by the syntactic component as unanalyzable atomic elements. When the f-structure is semantically interpreted, these forms are treated as patterns for composing the logical formulas encoding the meaning of the sentence. Thus, the semantic interpretation for this sentence is obtained from the value of its PRED function. The semantic-form information can be thought of as a compact way of representing subcategorization and semantic information. The predicate *cwu-ess-ta* ((S\NP)\PP)\NP) expresses that the predicate is preceded by the arguments NP,PP. The argument list specification defines a mapping between the logical or thematic arguments of the three-place predicate *cwu-ess-ta* (agent, patient, goal) and the grammatical functions of the f-structure.

4.1.2. Combinatory Rules

There are two combination rules in UKG, namely, backward-functional-application, and backward-functional-composition. In backward functional application, a functor category which is looking for some category to its left combines with that category. After the appropriate unifications have taken place, the subcategories of the category resulting from this rule application will take on the subcategories of the functor category minus the first, which has just been found. Let's look at the following:

(5) Backward-functional-application

```
X ----> Y   Z
" [X:  [struct:  1[]
      subcat:  2[]]
  Y:  [struct:  <3>
      subcat:  <4>]
  Z:  [struct:  <1>
      subcat:  [first:  [struct:  3[]
                        dir:  left
                        subcat:  4[]]
rest:  <2>]]]"
```

The directionality constraint for backward functional application on Z's first argument is enforced by including '[dir: left]' in the rule graph within Z, a feature which must be unifiable with the same feature in the functor category's graph. Korean case suffixes and postpositions can be analyzed by backward functional application. See the examples in the Appendix.

In backward functional composition, a functor category combines with some argument even though that argument is not yet complete. In other words, the argument category is itself a functor. It still has subcategories that need to be filled. What happens is that the result of this rule application becomes a functor which is looking for any subcats that the argument has not filled. Thus, a new functor is composed out of two original functors. This rule is designed to work only for simple functors, i.e., those whose arguments are of the form $X \backslash Y$. As discussed in Section 3.2., Korean verbal suffixes such as passive, tense, aspect and mood can be analyzed by backward functional composition.

(6) Backward-functional-composition

X -----> Y Z

```
" [X:  [struct:  1 [cat:  {vp s}]
      subcat:  <2>]
  Y:  [struct:  <3>
      subcat:  2[first:  [struct:  [cat:  {np pp}]]
  Z:  [struct:  <1>
      subcat:  [first:  [struct:  3[cat:  {vp s}]
                    subcat:  #
                    dir:  left]
      rest:  #]]]
```

Unlike functional application, the main functor Z in (6) is restricted to categories with one and only one argument of a simple type by the use of the '#' symbol. The argument Y, on the other hand, must be a functor which has some remaining arguments. This condition is enforced by including the value '[first: [...]]' at its subcat node.

Backward functional application applies with case suffixes and postpositions and backward functional composition with verbal suffixes. We see thus that backward functional application applied to case suffixes and backward functional composition applied to verbal suffixes reflect the complementary distribution obtaining between the case and verbal suffixes in the Korean language.

4.1.3. Bottom-up Parsing

In a unification-based formalism, we can think of proceeding bottom up, combining subphrases by concatenation and their informational elements by unification thereby yielding the minimal feature structures that conform to all the identities.

One method of bottom-up parsing employs a shift-reduce parser. As it proceeds through the sentence, the parser either shifts the next word to the stack, displacing downward what is already there, or reduces by substituting other categories using rules of the grammar, or combines the top most items. It places the result of each reduction on the stack, and when the only item left on the stack is an S the parse has succeeded. Bottom-up parsing allows the incremental pairing of syntactic with semantic rules and hence allows for the timely application of semantic constraints. Let's look at the Korean verb structure.

(7) Mek -epeli -ess -ta
eat -Aspect -Past Tns -Mood
'have eaten (something)'

According to the c-rules based on X-bar theory, the Korean verb structure (7) would be (8):

By the bottom-up interpretation, V(1) combines with V(2), and proceeds following the numeral order through a functional composition yielding the minimal feature structures that conform to all the identities.

4.2. Parsing Templates

Following PATR, a template file is included as one of three basic grammar files. The templates file imposes the form of the graphs used in the grammatical descriptions.

The general form of the graphs of the unification-based grammar is overviewed in 2.2.3. They are intended to be structural characterizations of the sentences, i.e., the syntactic constituent structure together with the functional roles that each of the syntactic constituents plays.

The basic nodes in the structural graphs are constituent nodes, which are the only nodes that have a 'cat' arc leading from them. Besides a category label, these nodes have a collection of atomic features ('feats') and an arc 'constits' that lead to any subconstituents. The subconstituents will hang from arcs that are named with functional terms such as 'subj', 'obj', etc. See the examples in the Appendix.

4.3. Lexicon

The third and last of the basic grammar files is the lexicon. When the word is encountered in the input, the graph corresponding to each template name is retrieved and unified with any other template definitions in the word's listing. Besides template names, the word may include information to search for a given attribute in the graph and fill in a value. When the value in such a form is '*', it is an indication to use the form of the term being defined as the value of that attribute. Thus, since the 'lex' feature in this grammar usually corresponds to the root form of these words, most morphologically unmarked words here have 'lex*' in their entries whereas the inflectionally marked forms typically use the root form of the word as 'lex' value. See the examples in the Appendix.

Chapter 5. CONCLUSION

The goal of this thesis has been to explore the extent to which a unification-based formalism could parse Korean language structures. Korean, as a free word order, verb-final language with agglutinative morphology, has posed problems for more traditional transformational grammar models. The unification-based devices explored here solve those problems because they account, not just for syntactic constituent structures, but also for the more abstract functional structures expressed by Korean morphology.

Unification-based grammar has the property - which some will consider an advantage and some a liability - of leaving open a large number of different approaches to the parsing of free word order languages. As we have seen, it is convenient within this formalism to separate a grammar into a number of components. The categorial component can be separated from the logical component, and both of these can be distinct from a functional component. Compositional semantics can be described within the formalism, and this can constitute a separate component of the grammar. In this last case, the advantages of keeping the components separate are particularly clear because we should expect the semantic component to be largely universal, allowing itself to be

combined through unification with categorial and other components from the grammars of particular languages.

In an agglutinative language such as Korean, where each morpheme has its independent grammatical meaning, we can understand why semantic scope can be partially independent from its bearer's syntactic appearance by allowing functional composition. Korean cannot be described in a revealing way within a framework that contains only ordered rules or that does not allow separate statements to be made about logical and functional aspects of the language. In English, constituent order expresses the syntactic and semantic functions of the sentence. But, in Korean, constituent order alone does not have the same function that it does in English. In Korean, morphemes instead express the syntactic and semantic functions of the sentence. As it has been shown, in a unification-based grammar, not only the constituent structure but also the functional information can be accounted for at least for languages in English and German, where they are expressed syntactically. The question remained as to whether morphological structures like those of Korean could be accounted for in a formal syntactic framework. As discussed in Section 3.2., Korean verbal morphemes attach to a verb; however, the grammatical semantic functions of each morpheme have scope over VP or S.

This study has shown that Korean morphological structures can also be accounted for in unification-based formalisms.

We conclude that unification-based formalisms justify two of the principal claims made for them, namely that they enable a perspicuous account to be given of the so-called functional aspects of language and that they provide a firm basis for performance models and computational procedures.

Appendix

Parsing Templates

-Sentence:

```
[struct: [cat: S
          feats: [mood: {decl, imp, interr, excl}
          constits: [subj: [cat: NP]
                      pred: []]]
subcat: []]
```

-Subject:

```
[struct: 1[cat: NP
           feats: [case: {nom, topic}
                   theme: {agent, pat}
                   honorific: {polite, plain}
           constits: [subj: <1>
                      pred: []]]
subcat: []]
```

-Predicate:

```
[struct: [cat: {vp, ap}
          feats: [tns: {pres past}
                  voice: {act pas}
                  aspect: {perf imperf}
                  honorific: {+ -}
                  mood: {decl interr imp excl}}
          constits: [head: [finite]
                     compls:[]]]
subcat: []]
```

-Korean inflectional class for predicates include:

-present tense:

```
[struct: [feats: [tns: pres]]]
```

-past tense:

```
[struct: [feats: [tns: past]]]
```

-passive:

```
[struct: [feats: [voice: [pas: {morp aux compo}]]]]]
```

-morphological passive:

```
[struct: [feats: [voice: [pas: morp]]]]]
```

-auxiliary passive:

```
[struct: [feats: [voice: [pas: aux]]]]]
```

-compound passive:

```
[struct: [feats: [voice: [pas: compo]]]]]
```


-Transitive verb:

```
[struct: [constits: [head: [constits: [lex: []]]
                                compls: [first: <1>
                                         rest: #]]]
subcat: [first: [dir: left
                struct: 1[cat: NP
                        feats: [case: acc]]
                subcat: #]
rest: #]]
```

-Noun Phrases:

```
[struct: [cat: np
          feats: [case: 1{nom acc gen top}
                  type: {pro name}
                  theme: {agent pat poss}
                  honorific: {plain polite}}
          constits: [head: n
                     case suffix: <1>]]
subcat: #]
```

-Common Nouns:

```
[struct: [cat: cn
          constits: [lex: []]]
subcat: #]
```

-Case Suffixs:

```
[struct: [cat: p
          feats: 1{nom acc gen top}
          constits: [head: 2[cat: n
                     feats:<1>
                     constits: [lex: []]]]
subcat: [first: [struct: <2>
                dir: left
                subcat:#]
rest: #]]
```

-rules for case suffixes in Korean apply to CN's as well as
PRO's and include Nominative, Accusative, Genitive and
Topic:

```
-Nom: [struct: [feats: [sub-case: nom]]]
-Acc: [struct: [feats: [obj-case: acc]]]
-Gen: [struct: [feats: [case: gen]]]
-Top: [struct: [feats: [sub-case: top]]]
```

-Postpositions

```
[struct: [cat: p
          feats: 1{ dat loc ins causa comita temp}
          constits: [head: 2[cat: n
                      feats: <1>
                      constits: [lex: []]]]
          subcat: [first: [struct: <2>
                          dir: left
                          subcat:#]
                  rest: #]]
```

-rules for postpositions apply to CN's and include Dative,
Locative, Instrumental, Causative, Comitative, Temporal:

```
-Dat: [struct: [feats: dat]]
-Loc: [struct: [feats: loc]]
-Ins: [struct: [feats: ins]]
-Causa: [struct: [feats: causa]]
-Comita: [struct: [feats: comita]]
-Temp: [struct: [feats: temp]]
```

Lexicon

VERB PHRASES

Ja-ta (S\NP) : V intransitive verb decl-mood

'sleep' (lex *)

Ja-n-ta: V intransitive verb pres-tense decl-mood

(lex 'ja-ta')

Ja-ess-ta: V intransitive verb past-tense decl-mood

(lex 'ja-ta')

Ja-peli-ess-ta: V intransitive verb perf-aspect past-tense
decl-mood

(lex 'ja-ta')

Ja-koiss-yess-ta: V intransitive verb imperf-aspect past-
tense decl-mood

(lex 'ja-ta')

Ja-koiss-si-ess-ta: V intransitive verb imperf-aspect
honorific past-tense decl-mood

(lex 'ja-ta')

Ja-ess-yo: V intransitive verb past-tense interr-mood

(lex 'ja-ta')

Ja-ra: V intransitive verb imp-mood

(lex 'ja-ta')

Ja-kuna: V intransitive verb excl-mood

(lex 'ja-ta')

Ilk-ta ((S\NP)\NP): V transitive verb decl-mood
 'read' (lex *)
 Ilk-nun-ta: V transitive verb pres-tense decl-mood
 (lex 'ilk-ta')
 Ilk-ess-ta: V transitive verb past-tense decl-mood
 (lex 'ilk-ta')
 Ilk-epeli-ess-ta: V transitive verb perf-aspect past-
 tense decl-mood
 (lex 'ilk-ta')
 Ilk-ysi-ess-ta: V transitive verb honorific past-tense
 decl-mood
 (lex 'ilk-ta')
 Ilk-koiss-si-ess-ta: V transitive verb imperf-aspect
 honorific past-tense decl-mood
 (lex 'ilk-ta')

Jap-ta ((S\NP)\NP): V transitive verb decl-mood
 'catch' (lex *)
 Jap-nun-ta: V transitive verb pres-tense decl-mood
 (lex 'jap-ta')
 Jap-ess-ta: V transitive verb past-tense decl-mood
 (lex 'jap-ta')

Jap-epeli-ess-ta: V transitive verb perf-aspect past-
tense decl-mood

(lex 'jap-ta')

Jap-hi-ess-ta: V transitive verb morph-passive past-
tense decl-mood

(lex 'jap-ta')

Jap-hi-si-ess-ta: V transitive verb morph-passive
honorific past-tense decl-mood

(lex 'jap-ta')

Jap-hi-epeli-si-ess-ta: V transitive verb morph-passive
perf-aspect honorific past-tense
decl-mood

(lex 'jap-ta')

Mul-ta ((S\NP)\NP): V transitive verb decl-mood

'bite' (lex *)

Mul-li-ess-ta: V transitive verb morph-passive past-
tense decl-mood

(lex 'mul-ta')

cwu-ta (((S\NP)\PP)\NP): V transitive verb decl-mood

'give' (lex *)

cwue-ci-ess-ta: V transitive verb aux-passive past-tense
decl-mood

(lex 'cwu-ta')

Ha-ta: V auxiliary verb decl-mood
 'do' (lex *)
 Ha-n-ta: V auxiliary verb pres-tense decl-mood
 (lex 'ha-da')
 Ha-ess-ta: V auxiliary verb past-tense decl-mood
 (lex 'ha-da')
 Ha-peli-ess-yo: V auxiliary verb perf-aspect past-
 tense interr-mood
 (lex 'ha-da')
 Ha-koiss-yess-yo: V auxiliary verb imperf-aspect past-
 tense interr-mood
 (lex 'ha-da')
 Ha-nun-cwung i-ta: V auxiliary verb imperf-aspect pres-
 tense decl-mood
 (lex 'ha-da')

 Chepo-ha-da ((S\NP)\NP): V transitive verb decl-mood
 'arrest' (lex *)
 Chepo-ha-n-da: V transitive verb pres-tense decl-mood
 (lex 'chepo-ha-da')
 Chepo-ha-ess-da: V transitive verb past-tense decl-mood
 (lex 'chepo-ha-da')

Chepo-toy-ess-da: V transitive verb aux-passive past-
tense decl-mood
(lex 'chepo-ha-da')

Sarang-ha-da ((S\NP)\NP): V transitive verb decl-mood
'love' (lex *)

Sarang-ha-n-da: V transitive verb pres-tense decl-mood
(lex 'sarang-ha-da')

Sarang-ha-ess-da: V transitive verb past-tense decl-mood
(lex 'sarang-ha-da')

Sarang-pat-ess-da: V transitive verb compo-passive past-
tense decl-mood
(lex 'sarang-ha-da')

NOUN PHRASES

Haksang 'student' : Common-noun sing (lex *)

Haksang-dul: Common-noun plur (lex 'haksang')

Haksang-i: Common-noun sing nom-case (lex 'haksang')

Haksang-ul: Common-noun sing acc-case (lex 'haksang')

Haksang-un: Common-noun sing top-case (lex 'haksang')

Haksang-uy: Common-noun sing gen-case (lex 'haksang')

Haksang-ekey: Common-noun sing dat-postposition
(lex 'haksang')

Haksang-hanthey: Common-noun sing dat-postposition

(lex 'haksang')

Ai 'child' : Common-noun sing (lex *)

Ai-dul : Common-noun plur (lex 'ai')

Ai-ka: Common-noun sing nom-case (lex 'ai')

Ai-lul: Common-noun sing acc-case (lex 'ai')

Ai-nun: Common-noun sing top-case (lex 'ai')

Ai-uy: Common-noun sing gen-case (lex 'ai')

Ai-ekey: Common-noun sing dat-postposition (lex 'ai')

Ai-hanthey: Common-noun sing dat-postposition (lex 'ai')

Ai-wa: Common-noun sing comita-postposition (lex 'ai')

Hakkyo 'school' : Common-noun sing (lex *)

Hakkyo-ka: Common-noun sing nom-case (lex 'hakkyo')

Hakkyo-lul: Common-noun sing acc-case (lex 'hakkyo')

Hakkyo-nun: Common-noun sing top-case (lex 'hakkyo')

Hakkyo-eyse: Common-noun sing loc-postposition (lex
'hakkyo')

Achim 'morning' : Common-noun sing (lex *)

Achim-ey: Common-noun sing temp-postposition (lex 'achim')

Kicha 'train' : Common-noun sing (lex *)

Kicha-lo: Common-noun sing ins-postposition (lex 'kicha')

NOTES

1. The verb 'think' takes S as an argument while the verb 'love' takes NP as an argument in this structure. The subcategory for verbs which have direct objects may be either NP or S.
2. Postpositions also include 'delimiters' (I.S.Yang 1972): '-man'(only), '-to' (also), '-pakkey' (nothing but), '-kkaci'(even).
3. Lee and Kim (1988) classify '-ci' as an auxiliary passive morpheme, while O'Grady (1991) classifies it as a compound passive morpheme.
4. See 3.1.2. 'dative postposition'.
5. Lee and Im (1983) exclude the suffix 'keyss' from Korean tense while Nam and Ko (1985) include it in tense category.

REFERENCES

- Ades, Anthony and Mark Steedman. 1982. "On the Order of Words." *Linguistics and Philosophy* 4. 517-559.
- Bach, Emmon. 1984. "Some Generalizations of Categorical Grammars." *Varieties of Formal Semantics, Proceedings of the Fourth Amsterdam Colloquium*, ed. by Landman, F. and Veltman, F., 1-24. Dordrecht:Foris.
- Bar-Hillel, Yehoshua. 1953. "A Quasi-Arithmetical Notation for Syntactic Description." *Language* 29. 47-58.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague:Mouton.

- Chung, Hee-sung. 1989. "A Phonological Theory Using Unification-Based Grammar Formalism." *Language Research* 25. 7-17.
- Dowty, David. 1982. "Grammatical relations and Montague Grammar." *The Nature of Syntactic Representation*, ed. by Pauline Jacobson and Geoffrey Pullum, 79-130. Dordrecht: Reidel.
- Dowty, David. 1988. "Type raising, Functional Composition, and Non-Constituent conjunction." *Categorial Grammars and Natural Language Structures*, ed. by Oehrle, R., Bach, E. and Wheeler, D., 153-197. Dordrecht:Reidel.
- Dowty, David. 1989. "Notes on Categorical Grammar and X-Bar Syntax: Some Fundamental Differences and Similarities." *Ohio State Univ. working Papers in Linguistics* 37. 1-10.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. London: Blackwell.
- Flynn, Michael. 1983. "A Categorical Theory of Structure Building." *Order, Concord and Constituency*, ed. by Gazdar, G. et al., 139-174. Dordrecht:Foris.
- Hong, Yun-sook. 1985. "A quantitative analysis of KA/I and (N)UN as subject markers in spoken and written Korean." *Harvard studies in Korean Linguistics II*, ed. by Kuno, S. et al., 145-153. Seoul:Hanshin.
- Hwang, Shin-ja Joo. 1975. *Korean clause structure*. University of Oklahoma:Summer Institute of Linguistics.
- Kang, Beom-mo. 1987. "Constraining Category Lifting in Categorical Morphology." *Eastern States Conference on Linguistics* 4. 164-175.
- Kaplan, Ronald and Joan Bresnan. 1982. "Lexical-Functional Grammar: A Formal System for Grammatical Representation." *The Mental Representation of Grammatical Relations*, ed. by Bresnan, J., 173-281. MIT press.
- Kay, Martin. 1985. "Parsing in Functional Unification Grammar." *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives*. Cambridge

- University Press, 252-278.
- Kim, Han-kon. 1989. "Some Consideration in Parsing Korean in Terms of Computational Unification." *Language Research* 25. 211-224.
- Kim, Il-kon. 1985. "The Need for Grammatical Relations and Syntactic Levels." *Linguistic Journal of Korea* 10.
- Koopman, H. 1984. *The Syntax of Verbs*. Dordrecht:Foris.
- Lee, Chang-in. 1989. *Dynamic Model of Comprehension with Applications to Korean and English*. Rice University:Ph.D. Dissertation.
- Lee, Hong-bae. 1975. "Kwuke-uy Kwankyeyhwa-ey Tayhaye (On Korean Relativization)." *Language Research* 11:2. 189-306.
- Lee, Hong-bae and Mi-kyung Kim. 1988. "Double-nominative constructions in Korean." *Linguistics in the morning calm* Vol. 2, ed. by the Linguistic Society of Korea, 489-506. Seoul:Hanshin.
- Lee, Ik-hwan and Hong-bin Im. 1983. *Kwukemwumpeplon (A Korean Grammar)*. Seoul:Hankwuksa.
- Lee, Jong-hyeok and Chang-gil Kim. 1988. "Voice Generation from Conceptual Representation: Syntactic, Semantic, and Pragmatic Aspects." *Literary & Linguistic Computing*. 250-254.
- Lee, Kang-hyuk. 1992. *Table-driven Parsing of Natural Language with Constraint-Based Grammar*. Urbana, Illinois:Ph.D. Dissertation.
- Lyons, John. 1971. *Theoretical Linguistics*. Cambridge: Cambridge University Press.
- McCawley, James D. 1988. *The Syntactic Phenomena of English Vol. I*. The University of Chicago Press.
- Montague, Richard. 1973. "The Proper treatment of quantification in ordinary English." *Formal Philosophy: Selected Papers of Richard Montague*, ed. by Richmond Thomason, 247-270. Yale University Press, 1974.
- Moortgat, Michael. 1988. "Categorial Investigations:

- logical and linguistic aspects of the Lambek calculus." Dordrecht:Foris.
- Nam, Ki-shim and Y.K. Ko. 1985. *Khyocwun Kwukemwunpeplon (A Modern Korean Grammar)*. Seoul:Top Press.
- O'Grady, William D. 1991. *Categories and Case: The Sentence Structure of Korean*. Philadelphia:John Benjamins Publishing Co.
- Oh, Jonathan C. and Steven Graham. 1989. "Problems with Parsing Free Ordered, Elliptic Languages." *Language Research* 25. 793-801.
- Pereira, Fernando and David Warren. 1980. "Definite clause Grammars for Language Analysis- A Survey of the Formalism and a Comparison with Augmented Transition Networks." *Artificial Intelligence* 13. 231- 278.
- Pereira, Fernando and Stuart Shieber. 1984. "The Semantics of Grammar Formalism Seen as Computer Languages." in *Proceedings of the Tenth International Conference on Computational Linguistics*, 123-129. Stanford University.
- Pollard, Carl. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and Natural Languages*. Stanford University:Ph.D. dissertation.
- Pollard, Carl. 1985a. *Lecture notes on Head-driven Phrase-structure Grammar*. Stanford University:Center for the Study of Language and Information.
- Quirk, R. et al. 1985. *A Comprehensive Grammar of the English Language*. NewYork:Longman.
- Radford, Andrew. 1988. *Transformational Grammar*. Cambridge University Press.
- Shieber, Stuart M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. CLSI Lecture Notes No. 4.
- Shin, Kyung-koo. 1982. "Passive construction in Korean." *Linguistic Journal of Korea* 7. 199-240.
- Smith, George W. 1991. *Computers and Human Language*. Oxford University Press.

- Steedman, Mark. 1985a. "Dependancy and coordination in the grammar of Dutch and English." *Language* 61. 523-68.
- Steedman, Mark. 1985b. "Combinators, Categorical Grammars, and Parasitic Gaps." Paper presented at the Tuscon Conference on Categorical Grammar.
- Yang, Dong-whee. 1975. *Topicalization and Relativization in Korean*. Seoul:Pan Korean Book Corp.
- Yang, In-seok. 1972. *Korean syntax: Case Markers, Delimiters, Complementation and Relativization*. The University of Hawaii:Ph.D Dissertation.
- Yang, In-seok. 1976. "The Semantics of Korean Causation." *Foundations of Language* 14. 55-87.
- Yang, In-seok. 1977. "Progressive and Perfective Aspects in Korean." *Linguistic Journal of Korea* 2.1. 25-40.
- Winograd, Terry. 1983. *Language as a Cognitive Process*. Addison-Wesley Publishing Company.
- Wittenburg, Kent B. 1986. *Natural Language Parsing with Combinatory Categorical Grammar in a Graph-Unification-Based Formalism*. The University of Texas at Austin:Ph.D. Dissertation.
- Wood, Mary McGee. 1993. *Categorical Grammar*. NewYork: Routledge.