Categorial Grammars, Combinatory Logic and the Korean Language Processing

Juyeon Kang, Jean-Pierre Desclés

LaLIC Laboratory, Paris-Sorbonne University
28, Rue Serpente, 75006, Paris, France
kjuyeon79@yahoo.fr, jean-pierre.descles@pars4.sorbonne.fr

Abstract

In this paper we propose a new approach to Categorial Grammars based on Combinatory Logic to solve some syntactic and semantic problems in the Korean language processing. We handle particularly the problems of cases, free word order structure and coordination by developing a formalism of the extended Categorial Grammar that was originally introduced by J.-P. Desclés, and I. Biskri. We call this extended Categorial Grammar **"Applicative and Combinatory Categorial Grammar (ACCG)"**. The **ACCG** formalism allows us to analyze syntactically and semantically the cases in Korean, in particular the linguistic phenomenon of double cases. In spite of the importance of the cases in the processing of the Korean language, this topic has not been well studied in view to its automatic processing and there are still many difficulties in parsing Korean texts. This article shows some robust solutions for the analysis of the free word order structure and of the coordination by introducing combinators, such as **B, C*, Φ**, of Combinatory Logic developed by H.-B. Curry and R. Feys.

1. Introduction to Categorical Grammars

Categorial Grammars and their extensions are deeply linked to logic, linguistic, computer science and mathematics.

We can find the origins of the Categorial Grammars (Oehrle, Bach and Wheeler, 1988) in the works of philosopher E. Husserl, the polish logicians S. Lesniewski and K. Ajdukiewicz. They proposed the first categorical systems by developing the theory of meaning category by Husserl and the grammar of semantic categories by Lesniewski. Then, the mathematician Y. Bar-Hillel established the model of Categorial Grammars and applied it to the analysis of natural languages. In this model, the relation between functors and arguments was viewed as an important element for the sentence analysis, so that an operator (functor) can be applied to an operand (argument) to form a new operator (new functor) by using applicative languages. The purpose of Categorial Grammars is to calculate and verify correctly a syntactic connection between basic syntactic categories such as S (sentence) and N (nominal) using algebraic notations: $X/Y * Y \rightarrow X$ or $Y * X\backslash Y \rightarrow X$ (Steedman's notation[1]).

1.1 Extensions of Categorial Grammars

Since the introduction of simple Categorial Grammars, different propositions were made to improve them by adopting applicative languages such as the calculus of syntactic types proposed by J. Lambek (1961), the lambda-calculus proposed by A. Church, the combinatory logic created by the mathematician H.-B. Curry (1958), and some attempts by the logician W. V. O. Quine, etc. These works are based on the mechanism of the application of an operator to an operand. Combinatory logic and lambda-calculus were applied to the analysis of grammatical and lexical meanings of natural languages by S. K. Shaumyan (1987) with his model of the Universal Applicational Grammar using Curry's combinatory logic, that extends simple Categorial Grammars: this model is easily implementable on computational tools using functional programming languages such as CAML, HASKELL and SCHEME. In a close way, Z. Harris (1982) in his grammar uses analysis by operators with different types. In the 80's, important extensions were given by R. Montague, M. Moortgat (1988), J. Lambek (1961) and M. Steedman (1989). Combinatory Categorial Grammar developed by Steedman (1989, 2001) was most often quoted and studied for the analysis of Korean sentences (Cho and Park, 2000, Cha, 2001, Cha and Lee, 2002, Lee and Park, 2003).

1.2 Applicative and Combinatory Categorial Grammar

In this paper[2] we propose a new approach of Categorial Grammars "**Applicative and Combinatory Categorial Grammar (ACCG)**" that is an extension of the

[1] Compared to Bar-Hillel's notation ($X/Y * Y \rightarrow X$ or $Y * Y\backslash X \rightarrow X$), Steedman's notation allows to fix the positions of operators and operands. For this reason we use this notation in this study.

[2] This study is included in a Franco-Korean project on "Categorial Grammars and Anti-anti relativism" with K.S. Cheong, J-P. Desclés, B. Djioua, J.Y Kang, H.G. Son.

Combinatory Categorial Grammar elaborated by Steedman. This **ACCG** was originally developed by J.-P. Desclés and I. Biskri (1995, 1996) with the tools of the combinatory logic by introducing a canonical association between some rules and the combinators.
The purpose of this work is the automatic analysis of Korean sentences in which there exist the problems of case, free word order structure and coordination structure. Firstly, the **ACCG** provides the possibility to go beyond the well-known limits (such as the processing of a coordination, etc.) of simple Categorial Grammars. Secondly, this formalism allows the construction of logico-grammtical representations that open a way to building semantico-cognitive representations in the general model of Applicative and Cognitive Grammar developed by J.-P. Desclés (1990, 2003) with the three following levels; 1) morpho-syntactic configurations, 2) logico-grammatical representations, 3) semantico-cognitive representations. The ACCG builds applicative representations on the second level from concatenated expressions given on the first level.

We present here the rules[3] of the **ACCG,** to analyze Korean sentences.

Application rules

[X/Y : u1]-[Y : u2]	[Y : u1]-[X\Y : u2]
---------->	----------<
[X : (u1 u2)]	[X : (u2 u1)]

Type raising rules

[X : u]	[X : u]
----------**>T**	----------**<T**
[Y/(Y\X) : (**C*** u)]	[Y\(Y/X) : (**C*** u)]

Functional composition rules

[X/Y : u1]-[Y/Z : u2]	[Y\Z : u1]-[X\Y : u2]
----------**>B**	----------**<B**
[X/Z : (**B** u1 u2)]	[X\Z : (**B** u2 u1)]

Consider the following analysis of a Korean sentence in the **ACCG**.

Sumi-ga *Minju-lil* *man-ass-da.*
Sumi-NOM Minju-ACC meet-PS-DC.
(Sumi met Minju.)

1. [N*[4]: Sumi-ga]-[5][(N*: Minju-lil]-[(S\N*)\N*: man-ass-da]
2. [S/(S\N*):(**C***Sumi-ga)]-[(N*:Minju-lil]-[(S\N*)\N*:man-ass-da] (>**T**)
3. [S/(S\N*):(**C***Sumi-ga)]-[(S\N*)/((S\N*)\N*):(**C***Minju-lil)]- [(S\N*)\N*:man-ass-da] (>**T**)
4. [S/((S\N*)\N*):(**B**(**C***Sumi-ga)(**C***Minju-lil))]-[(S\N*)\N*:man-ass-da] (>**B**)
5. **[S: ((B(C*Sumi-ga)(C*Minju-lil))(man-ass-da))]** **(>)**
6. **[S: ((C*Sumi-ga)((C*Minju-lil)(man-ass-da)))]** **(B)**
7. **[S: ((C*Minju-lil)(man-ass-da))Sumi-ga]** **(C*)**
8. **[S: (((man-ass-da)Minju-lil)Sumi-ga)]** **(C*)**

[3] **B** is a composition combinator. Its β-reduction is: **Bfgx→f(gx)**.

[4] The N*s are qualified Nouns such as N*nom, N*acc, etc.

[5] The hyphen (-) means a concatenation at the syntactic level.

We start from the concatenated sentence with assigned syntactic types. Then, we apply consecutively the type raising rules to *"Sumi-ga"* and *"Minju-lil"* which are operands, by introducing the combinator **C***. This operation allows us to transform an operand into an operator. Then, we apply the functional composition rule to form a new operator "(**B**(**C***Sumi-ga)(**C***Minju-lil))" that will be applied to the operand *"man-ass-da"* at step 5. We reduce (in Combinatory Logic formalism) consecutively the combinators **B** and **C***s to build a well formed applicative expression at step 8. This expression gives a formal interpretation in terms of predicates, arguments and cases.

2. Applications of ACCG to the Korean language processing

2.1 Case's problems

The Korean is an agglutinative language in which the words are formed by the linking of affixes to a radical such as the cases (or postpositions). In the syntactic and semantic analysis of the Korean sentence, the cases determine the grammatical roles of nominal syntagms (Sung, 1999, Hong, 1999, Nam, 2001). In this paper we study five essential cases[6]: *-ga* as nominative marker, *-lil/-ul* as accusative marker, *-eke* as dative marker, *-uy* as genitive marker and *-eso* as locative marker.
We use predefined notations to facilitate our categorial analysis.
X^o=S
X^1=(S\N*)
X^2=(S\N*)\N*
X^3=((S\N*)\N*)\N*
To the two classical basic types N(nominal) and S(sentence), we add a new basic type N* for the complete nominal syntagms.

Let us analyze the following sentence:

Gyosil-eso, Sumi-ga Minju-eke na-uy chaek-ul ju-aes'-da.
Class-LOC Sumi-NOM Minju-DAT me-GEN book-ACC give –PS-DC
In the class, Sumi gave my book to Minju.)

1. [N:Gyosil]-[(S/S)\N):-eso]-[N:Sumi]-[N*\N:-ga]-[N:Minju]-[N*\N:-eke]-[N:na]-[(N*/N*)\N:-uy]-[N:chaek]-[N*\N:-ul]-[X³: ju-aes'-da]
2. [S/S:-eso Gyosil]-[N:Sumi]-[N*\N:-ga]-[N:Minju]-[N*\N:-eke]-[N:na]-[(N*/N*)\N:-uy]-[N:chaek]-[N*\N:-ul]-[X³:ju-aes'-da] (<)
3. [S/S:-eso Gyosil]- [N*:-ga Sumi]- [N:Minju]-[N*\N:-eke]-[N:na]-[(N*/N*)\N:-uy]-[N:chaek]-[N*\N:-ul]- [X³: ju-aes'-da] (<)
4. [S/S:-eso Gyosil]- [N*:-ga Sumi]- [N*:-eke Minju]- [N:na]-[(N*/N*)\N:-uy]-[N:chaek]- [N*\N:-ul]- [X³: ju-aes'-da] (<)
5. [S/S:-eso Gyosil]- [N*:-ga Sumi]- [N*:-eke Minju]- [N*/N*: -uy na]- [N:chaek]-[N*\N:-ul]-[X³:ju-aes'-da] (<)
6. [S/S:-eso Gyosil]- [N*:-ga Sumi]- [N*:-eke Minju]- [N*/N*: -uy na]- [N*: -ul chaek]- [X³:ju-aes'-da] (<)

[6] A description of Korean cases and their categorial analyses in the **ACCG** are presented in detail and with more examples in the Master's thesis of KANG (2005).

7. [S/S:-eso Gyosil]- [N*:-ga Sumi]- [N*: -eke Minju]- [N*: ((-uy na)-ul chaek)]- [X³ ju-aes'-da] (>)
8. [S/S:-eso Gyosil]- [S/X¹: **C***-ga Sumi]- [N*: -eke Minju]- [N*: ((-uy na)-ul chaek)]- [X³:ju-aes'-da] (>**T**)
9. [S/S:-eso Gyosil]- [S/X¹: **C***-ga Sumi]- [X¹/X²: **C***-eke Minju]- [N*: ((-uy na)-ul chaek)]- [X³: ju-aes'-da] (>**T**)
10. [S/S:-eso Gyosil]- [S/X¹: **C***-ga Sumi]- [X¹/X²: **C***-eke Minju]- [X²/X³: **C***((-uy na)-ul chaek)]-[X³:ju-aes'-da] (>**T**)
11. [S/S:-eso Gyosil]- [S/X²: **B**((**C***-ga Sumi)(**C***-eke Minju))]- [X²/X³: **C***((-uy na)-ul chaek)]- [X³:ju-aes'-da] (>**B**)
12. [S/S:-eso Gyosil]- [S/X³: (**B**(**B**((**C***-ga Sumi)(**C***-eke Minju)))(**C***((-uy na)-ul chaek))]- [X³:ju-aes'-da] (>**B**)
13. [S/S:-eso Gyosil]- [S: ((**B**(**B**((**C***-ga Sumi)(**C***-eke Minju)))(**C***((-uy na)-ul chaek))ju-aes'- da)] (>)

---

14. **[S:(-eso Gyosil ((B(B((C*-ga Sumi)(C*-eke Minju)))(C*((-uy na)-ul chaek))ju-aes'-da))]** (>)
15. **[S:(-eso Gyosil ((B(C*-ga Sumi)(C*-eke Minju))((C*((-uy na)-ul chaek)))ju-aes'-da))]** (**B***)
16. **[S:(-eso Gyosil ((C*-ga Sumi)((C*-eke Minju))((C*((-uy na)-ul chaek))ju-aes'-da)))]** (**B***)
17. **[S:(-eso Gyosil ((C*-eke Minju)((((C*((-uy na)-ul chaek))ju-aes'-da))-ga Sumi)]** (**C***)
18. **[S:(-eso Gyosil ((((C*((-uy na)-ul chaek))ju-aes'-da)-eke Minju)-ga Sumi)]** (**C***)
19. **[S:(-eso Gyosil((((ju-aes'-da)((-uy na)-ul chaek))-eke Minju)-ga Sumi))]** (**C***)

We show in the above analysis that the categorical calculus of the given sentence allows us, on one hand, to verify the correct syntactic structure of the sentence by obtaining the result "S" at step 10, and on the other hand, to obtain an applicative expression that underlies this sentence structure. Furthermore, this analysis allows us to deduct syntactic types of the used cases as follows:

Nominative marker	(S/X1)\N
Dative marker	(X1/X2)\N or (X2/X3)\N
Accusative marker	(X2/X3)\N or (X1/X2)\N
Genitive marker	((X2/X3)/(X2/X3))\N or ((X1/X2)\(X1/X2))\N or ((S/X1)/(S/X1))\N
Locative marker1	(X^1/X^1)\N
Locative marker2	(S/S)\N

This means that the cases in Korean function as operators that are applied to operands such as nouns including proper nouns, common nouns, collective nouns, materials nouns, etc. The types of cases are given here as a first approximation, namely we intend to go deeper into the assignation of types to cases. Our purpose is to find some invariants of these types in order to reduce the ambiguity and the complexity in the choice of one of the assigned types to each case during their application. This paper presents the first step in this direction.

In the next sections, we handle the problem of the double cases, the free word order structure and the coordination considering the proposed types of cases.

2.1.1 Double nominative

The problem of a double subject corresponding to a double nominative and of a double object corresponding to a double accusative is actually very important in the syntactic and semantic study of Korean (Sung, 1999, Hong, 1999, Lee, 2002, Chung, 2003). This subject has not been well studied in the aspect of the natural language processing, so it is worth proposing a new analysis in the computational linguistic point of view.

In this paper we give a solution to the problem of the double nominative marker such as *-ga*.

Sumi-ga maumsi-ga jo-ta.
Sumi-NOM heart-NOM be good-DC
(Sumi has a good heart.)
(Sumi's heart is good.)

The above sentence having double nominative (-ga) of *Sumi-ga* and (-ga) of *maumsi-ga* can be interpreted in the other form *"Sumi-uy maumsi-ga jo-ta"*, namely the first *-ga* can be replaced by the genitive marker *-uy* without changing the meaning of the original sentence. This analysis leads us to calculate the syntactic types of the two occurrences of *-ga*.

We consider now the above sentence with a double nominative and the following processing.

Sumi-ga maumsi-ga jo-ta.

1.[N :Sumi]-[(N*/N*)\N :-ga]-[N :maumsi]-[N*\N :-ga]-[X^1 :jo-ta]
2.[N*/N* :-ga Sumi]- [N :maumsi]-[N*\N :-ga]-[X^1 :jo-ta] (<)
3.[N*/N* :-ga Sumi]- [N*:-ga maumsi]-[X^1 :jo-ta] (<)
4.[N* : (-ga Sumi)-ga maumsi]- [X^1 :jo-ta] (>)
5.[S/X^1 : **C***((-ga Sumi)-ga maumsi)]- [X^1 :jo-ta] (>**T**)

6.[S : C*((-ga Sumi)-ga maumsi)jo-ta] (**>**)
7.[S : (jo-ta((-ga Sumi)-ga maumsi))] (**C***)

This analysis shows that the second *"-ga"* is a real nominative marker that forms the subject of the sentence *"maumsi-ga"* and the first *"-ga"* is used as an element that forms the determinant *"Sumi-ga"* of the subject. Thus, we assign the syntactic type ((S/X1)/(S/X1))\N to the first *"-ga"* and the syntactic type (S/X1)\N to the second *"-ga"*.

The nominative case becomes more complex when we must process double nominative. The new assignation of type is: (S/X1)\N or ((S/X1)/(S/X1))\N.

The above analysis is new and different of the proposed analyses for the formal double cases in Korean (Cha, 2001, Kang, 2001) by the Categorial Grammars, it corresponds to the intuitive interpretation.

2.2 Free word order structure

The free word order is a widespread phenomenon in Korean, so the processing of the Korean language becomes difficult. The position of the linguistic elements in the Korean sentence does not play an essential grammatical role because it is generally possible to permute these linguistic elements, namely subjects, direct/indirect objects, etc. but verbs take always a position at the end of the sentence.

Consider the following examples having the free word order structure.

a. *Sumi-ga Minju-eke jilmun-ul hae-ss-da.*

Sumi-NOM Minju-DAT question-ACC do-PS-DC
(Sumi asked to Minju a question.)

b. *Sumi-ga jilmun-ul Minju-eke hae-ss-da.*
c. *Minju-eke Sumi-ga jilmun-ul hae-ss-da.*
d. *Minju-eke jilmun-ul Sumi-ga hae-ss-da.*
e. *Jilmun-ul Minju-eke Sumi-ga hae-ss-da.*
f. *Jilmun-ul Sumi-ga Minju-eke hae-ss-da.*

In the above examples the sentences have the same predicative interpretation (but the topic/comment interpretation is not invariant). So, the position of the elements is not very important in the Korean analysis. We propose now to analyze these sentences in the ACCG.

a) *Sumi-ga Minju-eke jilmun-ul hae-ss-da.*
Sumi-NOM Minju-DAT question-ACC do-PS-DC
« Sumi asked to Minju a question. »

1.[N: Sumi]-[N*\N:-ga]-[N:Minju]-[N*\N:-eke]-[N:jilmun]-[N*\N:-ul]-[X³: hae-ss-da]
2.[N*:-ga Sumi]-[N*:-eke Minju]-[N*:-ul jilmun]-[X³: hae-ss-da] (<)
3.[S/X¹:(**C***-ga Sumi)]-[N*:-eke Minju]-[N*:-ul jilmun]-[X³:hae-ss-da] (>**T**)
4.[S/X¹:(**C***-ga Sumi)]-[X¹/X²:(**C***-eke Minju)]-[N*:-ul jilmun]-[X³:hae-ss-da] (>**T**)
5.[S/X¹:(**C***-ga Sumi)]-[X¹/X²:(**C***-eke Minju)]-[X²/X³: (**C***-ul jilmun)]-[X³:hae-ss-da] (>**T**)
6.[S/X²:**B**(**C***-ga Sumi)(**C***-eke Minju)]-[X²/X³: (**C***-ul jilmun)]-[X³:hae-ss-da] (>**B**)
7.[S/X³ :**B**(**B**(**C***-ga Sumi)(**C***-eke Minju)) (**C***-ul jilmun)]-[X³:hae-ss-da] (>**B**)

8.[S: B(B(C*-ga Sumi)(C*-eke Minju)) (C*-ul jilmun) hae-ss-da] **(>)**
9.[S: (B(C*-ga Sumi)(C*-eke Minju)) ((C*-ul jilmun) hae-ss-da)] **(>)**
10.[S: (C*-ga Sumi)((C*-eke Minju) ((C*-ul jilmun) hae-ss-da))] **(B)**
11.[S: ((C*-eke Minju) ((C*-ul jilmun) hae-ss-da)) -ga Sumi] **(C*)**
12.[S: (((C*-ul jilmun) hae-ss-da) -eke Minju) -ga Sumi] **(C*)**
13.[S: ((((hae-ss-da) -ul jilmun) -eke Minju) -ga Sumi)] **(C*)**

b) *Sumi-ga jilmun-ul Minju-eke hae-ss-da.*
Sumi-NOM question-ACC Minju-DAT do-PS-DC
« Sumi asked a question to Minju. »

1.[N:Sumi]-[N*\N:-ga]-[N:jilmun]-[N*\N:-ul]-[N:Minju]-[N*\N:-eke]-[X³: hae-ss-da]
2.[N*:-ga Sumi]-[N*:-ul jilmun]-[N*:-eke Minju]-[X³: hae-ss-da] (<)
3.[S/X¹:(**C***-ga Sumi)]-[N*:-ul jilmun]-[N*:-eke Minju]-[X³:hae-ss-da] (>**T**)
4.[S/X¹:(**C***-ga Sumi)]-[X²/X³:(**C***-ul jilmun)]-[N*:-eke Minju]-[X³:hae-ss-da] (>**T**)
5.[S/X¹:(**C***-ga Sumi)]-[X²/X³:(**C***-ul jilmun)]-[X¹/X²:-eke Minju]-[X³:hae-ss-da] (>**T**)
6.[S/X¹ :(**C***-ga Sumi)]-[X¹/X³:(**B**(**C***-eke Minju)(**C***-ul jilmun))]-[X³:hae-ss-da] (<**B**)
7.[S/X³:**B**((**C***-ga Sumi-ga)(**B**(**C***-eke Minju-eke)(**C***-ul jilmun)))]-[X³:hae-ss-da] (<**B**)

8.[S:B((C*-ga Sumi)(B(C*-eke Minju)(C*-ul jilmun)))hae-ss-da)] **(>)**
9.[S:(C*-ga Sumi)((B(C*-eke Minju)(C*-ul jilmun))hae-ss-da)] **(B)**
10.[S:((B(C*-eke Minju)(C*-ul jilmun)) hae-ss-da) -ga Sumi] **(C*)**
11.[S:((C*-eke Minju)((C*-ul jilmun) hae-ss-da)) -ga Sumi] **(B)**
12.[S: (((C*-ul jilmun) hae-ss-da) -eke Minju) -ga Sumi] **(C*)**
13.[S: ((((hae-ss-da) –ul jilmun) -eke Minju) -ga Sumi)] **(C*)**

c) *Minju-eke Sumi-ga jilmun-ul hae-ss-da.*
Minju-DAT Sumi-NOM question-ACC do-PS-DC
« It's to Minju that Sumi asked a question. »

1. [N:Minju]-[N*\N:-eke]-[N:Sumi]-[N*\N:-ga]-[N:jilmun]-[N*\N:-ul]-[X³: hae-ss-da]
2. [N* : -eke Minju] – [N* : -ga Sumi] – [N*:-ul jilmun] - [X³: hae-ss-da] (<)
3. [X¹/X²:(**C***-eke Minju)]-[N*:-ga Sumi]-[N*:-ul jilmun]-[X³:hae-ss-da] (>**T**)
4. [X¹/X²:(**C***-eke Minju)]-[S/X¹:(**C***-ga Sumi)]-[N*:-ul jilmun]-[X³:hae-ss-da] (>**T**)
5. [X¹/X²:(**C***-eke Minju)]-[S/X¹:(**C*** -ga Sumi)]-[X²/X³:(**C***-ul jilmun)]-[X³:hae-ss-da] (>**T**)
6. [S/X²:(**B**(**C***-ga Sumi)(**C***-eke Minju))]-[X²/X³:(**C***-ul jilmun)]–[X³:hae-ss-da] (>**B**)
7. [S/X³:(**B**(**C***-ga Sumi)(**C***-eke Minju))(**C***-ul jilmun)]- [X³:hae-ss-da] (>)

8. [S:(B(C*-ga Sumi)(C*-eke Minju))(C*-ul jilmun)hae-ss-da] **(>)**
9. [S:(C*-ga Sumi)((C*-eke Minju)(C*-ul jilmun))hae-ss-da)] **(B)**
10. [S:(((C*-eke Minju)(C*-ul jilmun))hae-ss-da) -ga Sumi] **(C*)**
11. [S: :(((C*-ul jilmun)hae-ss-da) -eke Minju) -ga Sumi] **(C*)**
12. [S: ((((hae-ss-da) -ul jilmun) -eke Minju) -ga Sumi)] **(C*)**

d) *Minju-eke jilmun-ul Sumi-ga hae-ss-da.*
Minju-DAT question-ACC Sumi-Nom do-PS-DC
« It's to Minju that Sumi asked a question. »

1. [N:Minju]-[N*\N:-eke]-[N:jilmun]-[N*\N:-ul]-[N:Sumi]-[N*\N:-ga]-[X³: hae-ss-da]
2. [N* : -eke Minju] – [N* : -ul jilmun] – [N*:-ga Sumi] - [X³: hae-ss-da] (<)
3. [X¹/X²:(**C***-eke Minju)]-[N*:-ul jilmun]-[N*:-ga Sumi]-[X³: hae-ss-da] (>**T**)
4. [X¹/X²:(**C***-eke Minju)]-[X²/X³:(**C***-ul jilmun)]-[N*:-ga Sumi]-[X³: hae-ss-da] (>**T**)
5. [X¹/X²:(**C***-eke Minju)]-[X²/X³:(**C***-ul jilmun)]-[S/X¹:(**C***-ga Sumi)]-[X³: hae-ss-da] (>**T**)
6. [X¹/X³:**B**(**C***-eke Minju)(**C***-ul jilmun)]-[S/X¹:(**C***-ga Sumi)]–[X³: hae-ss-da] (>**B**)
7. [S/X³: **B**(**B**(**C***-ga Sumi)(**C***-eke Minju))(**C***-ul jilmun)]-[X³: hae-ss-da] (>**B**)

8. [S: B(B(C*-ga Sumi)(C*-eke Minju))(C*-ul jilmun) hae-ss-da] **(>)**
9. [S: B(C*-ga Sumi)(C*-eke Minju)((C*-ul jilmun)hae-ss-da)] **(B)**
10. [S: (C*-ga Sumi)((C*-eke Minju)(C*-ul jilmun)hae-ss-da)] **(B)**
11. [S: ((C*-eke Minju)(C*-ul jilmun)hae-ss-da) -ga Sumi] **(C*)**
12. [S: (((C*-ul jilmun)hae-ss-da) -eke Minju)-ga Sumi] **(C*)**
13. [S: ((((hae-ss-da) -ul jilmun) -eke Minju) -ga Sumi)] **(C*)**

e) *Jilmun-ul Sumi-ga Minju-eke hae-ss-da*
question-ACC Sumi-NOM Minju-DAT do-PS-DC
« It's a question that Sumi asked to Minju. »

1. [N:jilmun]-[N*\N:-ul]-[N:Sumi]-[N*\N:-ga]-[N:Minju]-[N*\N:-eke]-[X³: hae-ss-da]
2. [N*:-ul jilmun] – [N*: -ga Sumi] – [N*:-eke Minju] - [X³: hae-ss-da] (<)
3. [X²/X³:(**C***-ul jilmun)]-[N*:-ga Sumi]-[N*:-eke Minju]-[X³: hae-ss-da] (>**T**)
4. [X²/X³:(**C***-ul jilmun)]-[S/X¹:(**C***-ga Sumi)]-[N*:-eke Minju]-[X³: hae-ss-da] (>**T**)
5. [X²/X³:(**C***-ul jilmun)]-[S/X¹:(**C***-ga Sumi)]-[X¹/X²: -eke Minju]-[X³:hae-ss-da] (>**T**)
6. [X²/X³:(**C***-ul jilmun)]-[S/X²:**B**(**C***-ga Sumi)(**C***-eke Minju)]-[X³: hae-ss-da] (>**B**)
7. [S/X³:(**B**(**B**(**C***-ga Sumi)(**C***-eke Minju))(**C***-ul jilmun))]-[X³: hae-ss-da] (>**B**)

8. [S: B((B(C*-ga Sumi)(C*-eke Minju))(C*-ul jilmun))hae-ss-da] **(>)**
9. [S: (B(C*-ga Sumi)(C*-eke Minju))((C*-ul jilmun)hae-ss-da))] **(B)**
10. [S: (C*-ga Sumi)((C*-eke Minju)((C*-ul jilmun)hae-ss-da)))] **(B)**
11. [S:((C*-eke Minju)((C*-ul jilmun)hae-ss-da))) -ga Sumi] **(C*)**
12. [S:((C*-ul jilmun)hae-ss-da)) -eke Minju) -ga Sumi] **(C*)**
13. [S: ((((hae-ss-da)-ul jilmun) -eke Minju) -ga Sumi)] **(C*)**

f) *Jilmun-ul Minju-eke Sumi-ga hae-ss-da*
question-ACC Minju-DAT Sumi-NOM do-PS-DC
« It's a question that Sumi asked to Minju. »

1. [N:jilmun]-[N*\N:-ul]-[N:Minju]-[N*\N:-eke]-[N:Sumi]-[N*\N:-ga]-[X³: hae-ss-da]
2. [N*:-ul jilmun] – [N*: -eke Minju] – [N*:-ga Sumi] - [X³: hae-ss-da] (<)
3. [X²/X³:(**C***-ul jilmun)]-[N*:-eke Minju]-[N*:-ga Sumi]-[X³: hae-ss-da] (>**T**)
4. [X²/X³:(**C***-ul jilmun)]-[X¹/X²:(**C***-eke Minju)]-[N*:-ga Sumi]-[X³: hae-ss-da] (>**T**)
5. [X²/X³:(**C***-ul jilmun)]-[X¹/X²:(**C***-eke Minju)]-[S/X¹:(**C***-ga Sumi)]-[X³: hae-ss-da] (>**T**)
6. [X²/X³:(**C***-ul jilmun)]-[S/X²: **B**(**C***-ga Sumi)(**C***-eke Minju)]-[X³: hae-ss-da] (>**B**)
7. [S/X³: (**B**(**B**(**C***-ga Sumi)(**C***-eke Minju))(**C***-ul jilmun))]-[X³: hae-ss-da] (>**B**)

8. [S: (B(B(C*-ga Sumi)(C*-eke Minju))(C*-ul jilmun))hae-ss-da] **(>)**
9. [S: (B(C*-ga Sumi)((C*-eke Minju)(C*-ul jilmun)))hae-ss-da] **(B)**
10. [S: (C*-ga Sumi)((((C*-eke Minju)(C*-ul jilmun))hae-ss-da)] **(B)**
11. [S: (C*-ga Sumi)((C*-eke Minju)(hae-ss-da(-ul jilmun)))] **(C*)**
12. [S: ((C*-ga Sumi)((hae-ss-da(-ul jilmun)) -eke Minju))] **(C*)**
13. [S: ((((hae-ss-da)-ul jilmun) -eke Minju) -ga Sumi)] **(C*)**

For the six sentences, the six different analyses lead us to one unique and identical applicative expression. Thus, we observe that the **ACCG** formalism gives a new and more efficient approach to the processing of the free word order structure. Moreover, these results can be automatically obtained by computer.

2.3 Coordination

During the automatic processing of the Korean language, the coordination structure is one of the most difficult characteristics of natural languages (in French: Biskri, 2005 and 2007, Desclés, 2005, in English: Steedman, 2001, in Korean: Cho and Park, 2000) because of the ellipse of predicates, etc.

We are particularly interested in the distributive coordination and the non-distributive coordination. We handle precisely here the propositional coordination with the conjunction "**,**" between the propositions and the predicative coordination with the conjunction *"-myeonseo"* between the predicates. The combinators introduced for the coordination analysis are **B**, **C*** and notably **Φ**[7]. This combinator can be explained in the following coordination rules proposed for Korean.

Coordination rules[8] for Korean

Distiributive coordination rule

[X :u1]-[CONJD : **,**]-[X :u2]

----------<CONJD>

[X: (**Φ ,** u1u2)]

Non-distributive coordination rule

[X :u1]-[CONJN: **-** *myeonseo*]-[X :u2]

----------<CONJN>

[X: (**-** *myeonseo* u1u2)]

Consider the following example[9] of the Korean distributive coordination.

Sumi-ga gongbu-lil , Minju-ga yori-lil han-da.

Sumi-NOM study-LOC CONJD Minju-NOM cook-LOC do-DC

(Sumi studies, Minju cooks.)

1. [N:Sumi]-[N*\N:-ga]-[N:gonbu]-[N*\N:-lil]-[CONJD:,]-[N:Minju]-[N*\N:-ga]-[N:yori]-[N*\N:-lil]-[X²:han-da]
2. [N*:-ga Sumi]-[N*:-lil gongbu]-[CONJD:,]-[N*:-ga Minju]-[N*:-lil yori]-[X²:han-da] (<)
3. [S/X¹:(**C***-ga Sumi)]-[X¹/X²:(**C***-lil gongbu)]-[CONJD:,]-[S/X¹:(**C***-ga Minju)]-[X¹/X²:(**C***-lil yori)]-[X²:han-da] (>**T**)
4. [S/X²:(**B**(**C***-ga Sumi)(**C***-lil gongbu))]-[CONJD:,]-[S/X²:(**B**(**C***-ga Minju)(**C***-lil yori))]-[X²:han-da] **(>B)**
5. [S/X²:**Φ**,(**B**(**C***-ga Sumi)(**C***-lil gongbu))(**B**(**C***-ga Minju)(**C***-lil yori))]-[X²:han-da] **(Φ)**
6. **[S:(Φ,(B(C*-ga Sumi)(C*-lil gongbu))(B(C*-ga Minju)(C*-lil yori)))han-da]** **(>)**
7. **[S:,((B(C*-ga Sumi)(C*-lil gongbu))han-da)((B(C*-ga Minju)(C*-lil yori)))han-da)]** **(<CONJD>)**
8. **[S:,((C*-ga Sumi)((han-da)-lil gonbu))((B(C*-ga Minju)(C*-lil yori)))han-da]** **(B)**
9. **[S:,(((han-da)-lil gongbu)-ga Sumi)((B(C*-ga Minju)(C*-lil yori)))han-da]** **(C*)**
10. **[S:,(((han-da)-lil gongbu)-ga Sumi)((C*-ga Minju)((han-da)-lil yori)]** **(C*)**
11. **[S:,(((han-da)-lil gongbu)-ga Sumi)(((han-da)-lil yori)-ga Minju)]** **(C*)**

The conjunction "," coordinates two expressions having the same type: **(((han-da)-lil gongbu)-ga Sumi)** and **(((han-da)-lil yori)-ga Minju)**. These expressions are applicative expressions of propositions: "***Sumi studies"*** and "***Minju cooks"*** that were presented here in the form of a logico-grammatical predicative relation. The intorduction of the combinator **Φ**, at step 4, builds a new complex operator "**Φ**,(**B**(**C***-ga Sumi)(**C***-lil gongbu))(**B**(**C***-ga Minju)(**C***-lil yori))". Then, at step 6, the reduction of the combinator allows us to apply the operators "**B**(**C***-ga Sumi)(**C***-lil gongbu)" and "**B**(**C***-ga Minju)(**C***-lil yori)" to the operand "han-da".

Then, consider the example of the korean non-distributive coordination.

Haebaragi-nun nora-myeonseo parass-da.

sunflower-NOM be yellow-CONJN be green-DC

(A sunflower is yellow and green.)

1. [N:haebaragi]-[N*\N:-nun]-[(S\N*):nora]-[CONJN:-myeonseo]-[(S\N*): parass-da]
2. [N*:-nun haebaragi]-[(S\N*):nora]-[CONJN:-myeonseo]-[(S\N*):parass-da] (<)
3. [S/(S\N*):**C***-nun haebaragi]-[(S\N*):nora]-[CONJN:-myeonseo]-[(S\N*):parass-da] (>T)
4. [S/(S\N*):(**C*** -nun haebaragi)]-[(S\N*):(-myeonseo nora parass-da)] **(<CONJN>)**
5. **[S:(C*-nun haebaragi)(-myeonseo nora parass-da)]** **(>)**
6. **[S: ((-myeonseo nora parass-da)-nun haebaragi)]** **(C*)**

This analysis allows us to handle the conjunction *-myeonseo* that is one of the conjunctions generally used in the predicative coordination. In this analysis of the non-distributive coordination, we do not use the combinator **Φ** that allows the distribution of two coordinated propositions. The simple application of the non-distributive coordination rule contributes to analyze correctly this sentence that is very difficult to handle automatically by simple Categorial Grammars and by most of the formal models.

We manage thus to reorganize an utterance having a coordination structure in the form of the applicative expression.

3. Conclusion and Perspectives

In this paper we give a solution to the problems of cases, free word order structure and coordination for Korean by using the **Applicative and Combinatory Categorial Grammar**. Our analyses, especially the analyses of free word order structure, show that different calculations lead to a unique result, namely a logico-grammatical representation. Compared to other related works for the Korean language processing using exclusively Steedman's

[7] This is a combinator of coordination in the sense of Desclés and Biskri. The β-reduction rule is defined as follows: **Φfgh→f(gx)(hx). f,g,h,** are operators that will form a new complex operator **Φfgh** and **x** is its operand.

[8] These coordination rules were originally presented for the analysis of a French coordination by Biskri and Desclés in the paper "*Analyse de la coordination et de la subordination au moyen de la Grammaire Catégorielle Combinatoire Applicative"* (written in French, 2005).

[9] In Korean, this sentence will be more acceptable with the marker *–nun* than with the marker *–ga*. But because of the complex linguistic characteristics of *–nun,* which denotes generally a comparison/contrast between two sentences in the coordination structure, we use here *–ga*.

Combinatory Categorial Grammar, our attempt is a considerable challenge and a new approach resulting in a calculus of the Korean sentence that improves the above mentioned problems.

For the moment, it will be possible to assign the several syntactic types to one case in this formalism. To resolve this complexity and ambiguity, we try to find some invariants for each case by calculating more complex sentences. The proposed results from our analyses can be obtained by automatic computation by using the Caml language.

References

Biskri, I. 1995. *La Grammaire Catégorielle Combinatoire Applicative dans le cadre de la Grammaire Applicative et Cognitive*, Ph.D. thesis, EHESS, Paris.

Biskri, I. and Desclés, J.-P. 2005. Analyse de la coordination et de la subordination au moyen de la Grammaire Catégorielle Combinatoire Applicative, *Colloque Typologie et modélisation de la coordination et de la subordination*, LACITO - ParisIII (UMR 7107, CNRS), Paris.

Biskri, I. and Rochette, M. A. 2007. The operational annotation and the analysis of the correlative coordination in French, *In Proceedings of the 2007 FLorida Artificial Intelligence Society*, AAAI Press.

Cha, J. W. 2001. *Statistical Parsing for Korean Categorial Grammar*, Ph.D. thesis, Pohang University of Science & Technology, Pohang, Korea.

Cha, J. W. and Lee, G. B. 2002. Structural disambiguation of morpho-syntactic categorial parsing for Korean, *International Conference On Computational Linguistics*, Proceedings of the 18th conference on Computational linguistics – Vol. 2, 1002~1006, Saarbrücken, Germany.

Cho, H. J. and Park, J. C. 2000. Combinatory Categorial Grammar for the Syntactic, Semantic and Discourse Analyses of Coordination in Korean, *Journal of KISS: Software and Applications,* 27(4), 448-462, Korea.

Chung, G. 2003. *Analyse des constructions à double nominatif/accusatif par l'opération de restructuration en Coréen. Classification syntaxique des constructions à adjectifs sôngsang,* Ph.D. thesis, Marne-La-Vallée University, Paris.

Curry, H. and Feys, R. 1958. *Combinatory Logic*, Vol.I, North-Holland.

Desclés, J.-P. 1990. *Langages applicatifs, langues naturelles et cognition*, Hermès, Paris.

Desclés, J.-P. and Biskri, I. 1996. Logique Combinatoire et linguistique : La Grammaire Catégorielle Combinatoire Applicative, *Mathématiques, Informatiques et Sciences Humaines*, n. 132, 39-68, Paris.

Desclés, J.-P. 2003. *La Grammaire Applicative et Cognitive construit-elle des représentations universelles ?*, LINX 48, Université Paris X, 139-160, Paris.

Harris, Z. 1982. *A Grammar of English. Based on Mathematical Principles,* John Wiley, New York.

Hong, J. S. 1999. Generative Grammar and Korean Cases, in *Korean Cases and Particles* by Institute of Korean Language, Wolin, 83-112, Korea.

Kang, B. M. 2001. *Categorial Grammar* (written in Korean), Institute of Korean Culture, Korea University, Seoul, Korea.

Kang, J. Y. 2005. *Grammaires Catégorielles appliquées au coréen en comparaison avec le français*, Master's thesis, Paris-Sorbonne University, Paris.

Lambek, J. 1961. On the calculus syntactic types, *Proceeding of symposia in Applied Mathematics,* vol. XII, America Mathematical Society, Providence, Rhode Island, 166-178.

Lee, H. J. and Park, J. C. 2003. Morphological Analysis of Irregular Conjugation in Korean with Micro Combinatory Categorial Grammar, *Proceedings of the Korea Information Science Society (KISS) Spring Conference*, volume 30, number 1(B), 531-533, Korea.

Lee, S. H. 2002. Korean Though Constructions and Double Nominative Constructions, *In Proceedings of the 9th International Conference On HPSG*, Seoul, Korea.

Moortgat, M. 1988. *Categorial Investigation, Logical and Linguistic Aspects of the Lambek calculus*, Foris Publications.

Nam, G. S. 2001. *The Syntax* (written in Korean), Taehaksa, Seoul, Korea.

Oehrle, R. T., Bach, E. and Wheeler, D. 1988. *Categorial Grammars and Natural Languages Structures*, D. Reidel.

Shaumyan, S. K. 1987. *A Semiotic Theory of Natural Language*, Bloomington, Indiana Univ. Press.

Steedman, M. 1989. *Work in Progress: Combinators and grammars in natural language understanding*, Summer institute of linguistic, Tucson University.

Steedman, M. 2001. *The Syntactic Process*, MIT Press/Bradford Books.

Sung, K. S. 1999. *Expression of Cases and Meaning of Particles* (written in Korean), Wolin, Seoul, Korea.