

# On the Syntax and Semantics of the Bound Noun Constructions: With a Computational Implementation\*

Jong-Bok Kim and Jaehyung Yang

Kyung Hee University, School of English, jongbok@khu.ac.kr  
and  
Kangnam University, School of Computer Engineering, jhyang@kangnam.ac.kr

**Abstract.** The so-called Korean BNC (bound noun construction) displays complex syntactic, semantic, and constructional properties. This paper, couched upon a constraint-based approach, two different syntactic structures for the construction with articulated lexical properties for the BNs and relevant predicates. The paper reports an implementation of this analysis in the LKB (Linguistic Knowledge Building) system and shows us that this direction is robust enough to pare relevant sentences.

**Keywords:** bound noun construction, complex predicate, implementation

## 1. Introduction

Bound nouns (BN) exhibit various peculiar properties, not found in common nouns in the language. For example, unlike canonical nouns, bound nouns cannot occur independently: they obligatory select a complement (determiner or sentence). This is rather unusual when considering the language allows most of the arguments to be freely omitted with proper context:

- (1) a. \*(i) kes  
          this thing  
      b. \*(wuli-ka motu nollass-ten) kes  
          we-NOM all surprise-MOD BN  
          ‘the thing that we all surprised’

Bound nouns also place restrictions on the types of their complements. There are at least two different types of BNs with respect to their complements: BNs selecting only a dependent clause (Type I) and those selecting either a dependent clause or a determiner phrase (Type II) (cf. Cha 2001):<sup>1</sup>

- (2) a. Type I: cheyk (‘pretense’), cwul (‘method’), li (‘reason’), cek (‘experience’), ppen (‘being close to doing something’), ba (‘way’), etc.  
      b. Type II: swu (‘possibility’), hwu (‘after’), cen (‘before’), etc

For example, unlike Type II BN *hwu*, Type I BN *li* selects only a dependent sentence, as observed in the following contrast:

- (3) a. [John-i cam-ul ca-n]/ku hwu-ka mwusep-ta  
          John-TOP sleep-ACC sleep-MOD/that BN-NOM not.exist-PAST-DECL  
          ‘the time after John was sleeping/after the time’

---

\*Copyright 2007 by Jong-Bok Kim and Jaehyung Yang

<sup>1</sup>The dependent clause, different from an independent sentence, has the head verb in a dependent verb form. The clause thus cannot be used as an independent sentence. See Kim (2004).

- b. [John-i cam-ul ca-l]/\*ku li-ka eps-ess-ta  
 John-NOM sleep-ACC sleep-MOD/that BN-NOM not.exist-PAST-DECL  
 ‘It is not possible that John was sleeping’

Bound nouns also place tight restrictions on the verb forms of their sentential complement. In the noun complement construction (NCC), the dependent clause places no strict constraints on the head verb’s VFORM value:

- (4) [John-i cam-ul ca-n/ca-ss-ta-nun] sasil  
 John-TOP sleep-ACC sleep-PNE/sleep-PAST-DECL-PNE fact  
 ‘the fact that John slept

As in (4), the head verb of the dependent clause, functioning as the complement of the factive noun *sasil* ‘fact’, can be either in a short form *ca-n* or in a full form *ca-n-ss-ta-nun* with the declarative ending. Meanwhile, in the BN construction, the head verb of the dependent clause cannot be in a full verb form: only a short form with restricted tense is allowed:

- (5) a. John-un cam-ul \*ca-n-ta-nun/ca-l/\*ca-n  
 John-TOP sleep-ACC \*sleep-PRES-DECL-PNE/sleep-FUT/\*sleep-PRES  
 swu-ka eps-ess-ta  
 BN-NOM not.exist-PAST-DECL  
 ‘John couldn’t sleep.’
- b. John-un cam-ul \*ca-n-ta-nun/\*ca-l/ca-nun  
 John-TOP sleep-ACC \*sleep-PNE-DECL-PNE/s\*leep-FUT/sleep-PNE  
 chey hayess-ta  
 BN did-DECL  
 ‘John pretended to sleep.’

In both cases, the full dependent form (with the declarative marking) is not possible. In addition, the BN *swu* requires the head verb of its complement clause be in the future form, whereas *chey* restricts the dependent verb to be marked with the present tense.

Bound nouns are also peculiar in that they impose restrictions on the types of the predicates following them. For example, the BN *swu* can combine only with the predicate *iss-* ‘exist’ or *eps-* ‘not exist’ whereas the BN *li* requires only the latter *eps-*. Also BNs like *tes*, *ppen*, and *ccek* occur only with *ha-* ‘do’, whereas BNs like *kes* can be followed only by the auxiliary verb *kath-ta* ‘seem’:

- (6) a. [John-i o-l li-ka] eps-ta/\*iss-ta/\*kath-ta  
 John-i come-MOD BN-NOM not.exist-DECL/exist-DECL/seem-DECL  
 ‘It is unlikely that John will come.’
- b. [John-i o-l tus] ha-ta/\*eps-ta/\*kath-ta  
 John-NOM come-MOD BN do-DECL/not.exist/seem-DECL  
 ‘It seems that John will come.’
- c. [John-i o-l kes] kath-ta/\*issta  
 John-NOM come-MOD BN seem-DECL/exist.  
 ‘It seems that John will come.’

Bound nouns have an additional restriction on the occurrence with case markers, which may be related to the function of the dependent clause:

- (7) a. Either NOM or ACC can be attached to the BN: *tey* ('place'), *pa* ('way'), *cco hk* ('side'), etc  
 b. Only NOM: *nawi* ('degree'), *li* ('reason'), *swu* ('possibility'), *ci* ('whether'), etc  
 c. Only ACC: *tung* ('so forth'), *yang* ('pretense'), *cwul* ('way'), *chey* ('pretense'), etc  
 d. Only DEL: *tus* ('seem'), *man* ('possible'), *sang* ('seem'), *kes* ('possible'), etc

The fact that the BNs in (7b) can occur only with NOM and those in (7c) only with ACC can be expected when considering the possible predicate they can be followed. Though all the BNs can occur with a delimiter, those in (7d) allow no case markers at all:

- (8) a. *pi-ka o-nun tus-\*ul/\*i ha-ta*  
 rain-NOM come-PNE BN-ACC/NOM do-DECL  
 'It seems to rain.'  
 b. *John-i o-l \*kes-i kath-ta*  
 John-NOM come-MOD BN-\*NOM seem-DECL/exist.  
 'It seems that John will come.'

As observed here and in the literature, BNs display complex combinatory possibilities with their complements, case/delimiter markers, and predicates following them. In addition to the constructional properties that each BNC shares, each BN also has its own idiosyncratic lexical properties. This implies that to process these BN constructions with intriguing properties, we need to develop an explicit syntactic and semantic analysis.

## 2. Syntax and Semantics of the Bound Noun Construction

### 2.1. Two Different Types

Based on the observations we have seen earlier, one could argue that a BN forms a morphological unit with the following predicate or undergoes a lexical process. However, this fails empirically in several respects: the BN's occurrence with a case or delimiter marker evidences both its syntactic and phonological independence from the following predicate:

- (9) a. *sensayngnim-i o-si-l swu-(ka/cocha/man) eps-ta*  
 teacher-NOM come-HON-PNE BN-NOM/even/only not exist  
 'It is not possible that the teacher comes.'  
 b. *sensayngnim-i o-si-l kes-(to/man) kathta*  
 teacher-NOM come-HON-PNE KES-also/only seem  
 'It seems that the teacher also will come.'

If one takes the BN with the following predicate as an inseparable lexical unit (e.g., *swu-iss-ta*), we would ignore the traditional wisdom of wordhood and not account for such productive processes.

One basic syntactic property in the BNC is that the BN forms a tight syntactic unit with its complement: no element can intervene between the two:

- (10) a. [wuli-ka kwanye ha-l] (\*cincca) pa-ka ani-ta  
 we-NOM intervention do-MOD really BN-NOM not-DECL  
 ‘This is not the case where we can intervene.’
- b. [wuli-nun ku-ka ka-l] (\*cal) cwul-un moll-ass-ta  
 we-TOP he-NOM go-MOD BN-TOP not.know-PAST-DECL  
 ‘We didn’t know that we would leave.’

Further note that there are at least two different types of the BNC. In the *kes* BNC, for example, there is strong syntactic cohesion between the BN and the following predicate. However, such a strong syntactic unity is not found in the *cwul* BNC:

- (11) a. pi-ka o-l kes (\*cengmal) kah-ta  
 rain-NOM come-PNE BN really seem-DECL  
 ‘It seems that it will rain.’
- b. wuli-nun ku-ka ka-l cwul-ul/un (cengmal) moll-ass-ta  
 we-TOP he-NOM go-MOD BN-ACC/TOP really not.know-PAST-DECL  
 ‘We really didn’t know that we would leave.’

These two types are also different with respect to the projection of a full NP. That is, when BNs combine with its complement (dependent clause or determiner), not all can function as a full NP. Observe the following coordination data:

- (12) a. \*[[sensayngnim-i ka-l swu] kuliko  
 teacher-NOM go-PNE BN and  
 [haksayngtul-i o-l swu]] iss-ta  
 student-NOM come-PNE BN exist  
 ‘The teacher can go and students can come.’
- b. [[sensayngnim-i ka-ko] [hasayngtul-i o-l] swu] issta  
 teacher-NOM go-or student come-PNE BN exist

As noted in (12a), the *swu* BNC cannot be coordinated with another BNC. This implies that this type of BNC cannot project a full NP even with its sentential complement. This can be further evidenced by the fact that the BNC cannot be used as an NP fragment, either:

- (13) a. \*[[sensayngnim-i ka-l swu] ‘(lit.) the possibility that the teacher goes’  
 b. [sensayngnim-i ka-l hakkyo] ‘the school where the teacher will go’

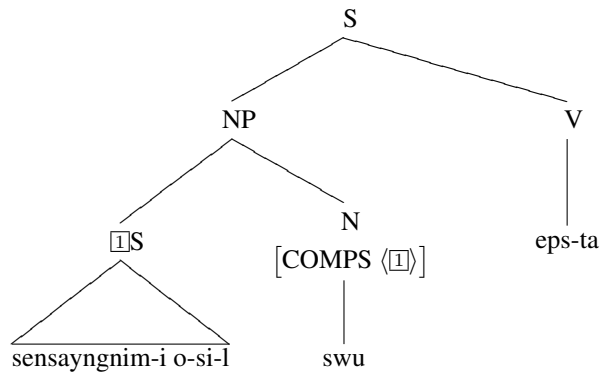
Unlike this, there are examples where the BN with its sentential complement behaves like a constituent:<sup>2</sup>

- (14) a. [[aitul-un ca-nun chek], [eleu-un cwuk-un chek]] hayessta  
 children-TOP sleep-PNE BN adult die-PNE pne did  
 ‘Children pretended to be sleeping whereas adults pretended to be dead.’
- b. [[aitul-un ka-nun cwul], [elun-un o-nun cwul]] alassta  
 children-TOP go-sleep-PNE BN adult-TOP come-PNE BN know  
 ‘(We) thought that children go while adults come.’

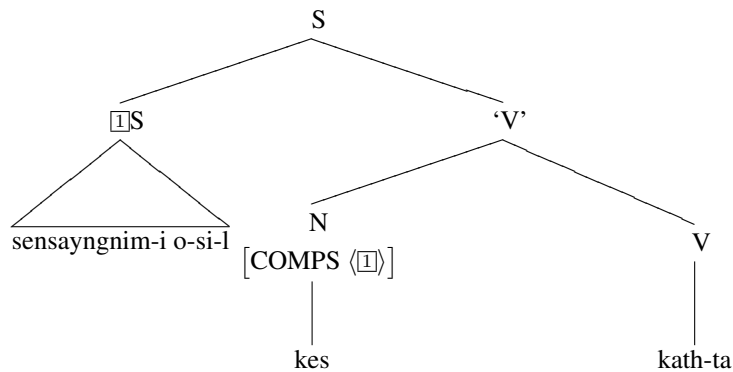
<sup>2</sup>Google search provides even a few examples where the BN *swu* projects an independent NP with its sentential complement. However, such examples are hardly found in canonical usages.

These syntactic differences imply that we need both (9a) and (9b) for possible syntactic structures for the Korean BNCs:

(15) Head-Complement Structure



(16) Complex Predicate Structure:



The structure in (15) assumes that the matrix predicate *eps-ta* ‘not.exist’ selects one argument projected from the BN *swu*. This approach takes the BN, combined with its sentential complement, combined projects an independent NP. Meanwhile, the structure in (16) assumes that the matrix predicate *kath-* first combines with the BN *kes*, forming a complex-predicate like unit.

## 2.2. Head-Complement Type BNC

The structure in (15) is a canonical head-complement phrase licensed by the Head-Complement Rule in (17b). This rule, along with the other grammar rules given here, licenses well-formed phrasal combinations in the language:

(17) a. Head-Subject Rule:

$$XP[hd-subj-ph] \rightarrow \boxed{1}, \mathbf{H}[\text{SUBJ } \langle \boxed{1} \rangle]$$

b. Head-Complement Rule:

$$XP[hd-comp-ph] \rightarrow \boxed{1}, \mathbf{H}[\text{COMPS } \langle \dots, \boxed{1}, \dots \rangle]$$

c. Head-Modifier Rule:

$$XP[hd-mod-ph] \rightarrow [\text{MOD } \langle \boxed{1} \rangle], \boxed{1}\mathbf{H}$$

The Head-Subject Rule, generating a *hd-subj-ph*, allows a VP to combine with its subject. The Head-Complement Rule ensures a head to combine with one of its COMPS elements, forming a *hd-comp-ph*. The Head-Modifier Rule allows a head to form a well-formed phrase with an adverbial element that modifies the head, resulting in *hd-mod-ph*.<sup>3</sup>

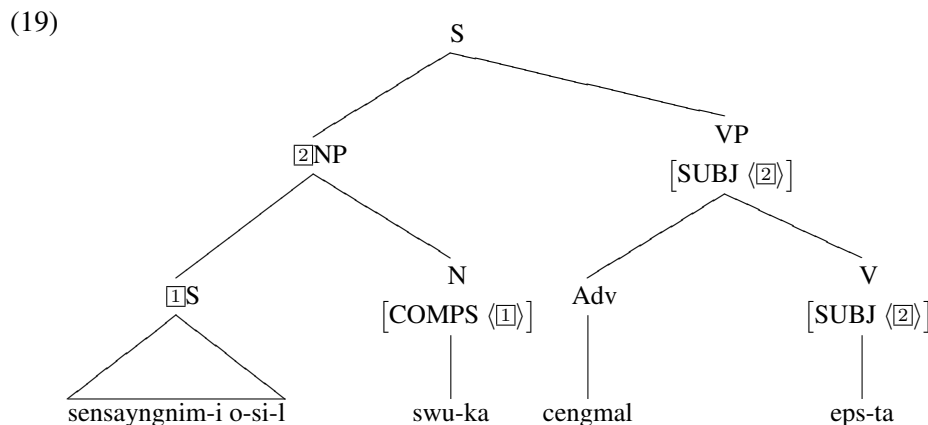
We also posit the following lexical information for the BN *swu* ‘possibility’ and *eps-ta* ‘not.exist’:

$$(18) \left[ \begin{array}{l} \text{bn-nonlex} \\ \text{HEAD} \left[ \begin{array}{l} \text{POS } \textit{noun} \\ \text{NFORM } \textit{swu} \end{array} \right] \\ \text{a. ARG-ST} \left\langle \text{S} \left[ \begin{array}{l} \text{MOD } \langle \text{N} \rangle \\ \text{IND } \textit{s0} \end{array} \right] \right\rangle \\ \text{SEM} \left[ \begin{array}{l} \text{IND } \textit{i} \\ \text{RELS} \left\langle \left[ \begin{array}{l} \text{PRED } \textit{possibility} \\ \text{ARG0 } \textit{s0} \end{array} \right] \right\rangle \end{array} \right] \end{array} \right]$$

$$\left[ \begin{array}{l} \text{int-v-bn} \\ \langle \textit{eps-} \rangle \\ \text{HEAD} \mid \text{POS } \textit{verb} \\ \text{b. ARG-ST} \left\langle \text{NP} \left[ \begin{array}{l} \text{IND } \textit{i} \\ \text{FORM } \textit{swu} \end{array} \right] \right\rangle \\ \text{SEM} \left[ \text{RELS} \left\langle \left[ \begin{array}{l} \text{PRED } \textit{not.exist} \\ \text{ARG0 } \textit{i} \end{array} \right] \right\rangle \right] \end{array} \right]$$

As the lexical information tells us, *swu* is a bound noun selecting a dependent sentence denoting a situation ‘s0’ which functions as its semantic argument. Meanwhile *eps-ta* selects an NP whose FORM value is *swu*.

Together with the grammar system, these lexical entries will project the following structure for a sentence like (9a):



The verb *eps-ta* selects one argument realized as the SUBJ in syntax. In the structure, this verb first combines with the adverb, forming a *hd-mod-ph*. Meanwhile, the BN *swu* combines with its sole complement, the dependent clause marked with a prenominal ending. The resulting NP then will serve as the subject of the verb *eps-ta*. The structure thus involves a canonical head-modifier, head-complement, and head-subject phrase, respectively.

### 2.3. Complex-Predicate Type BNC

Notice that the language, unlike English, also employs a grammar rule forming a complex predicate like the auxiliary verb construction (AVC). As noted in the literature (cf. Kim (2004)), in the AVC, the main verb and the following auxiliary show a tight syntactic cohesion and form a complex predicate:

<sup>3</sup>Note that the grammar rules here place no restriction on the SUBJ value: this allows the head to combine with the subject before combining with a complement. One great advantage of this is to allow sentential internal scrambling with no further operation or mechanism. See Kim (2004), Kim and Yang (2004) for details.

- (20) John-i sakwa-ka/lul mek-ko (\*cengmal) siph-ess-ta  
 John-NOM apple-NOM/ACC eat-COMP really would.like  
 ‘John would really like to eat apples.’

As argued and shown by Kim and Yang (2004), one effective way of capturing such complex predicate-like properties of the AVC is to introduce the Head-Lexical Rule given in (21):

- (21) Head-Lexical Rule:

$$\left[ \begin{array}{l} hd\text{-}lex\text{-}ex \\ COMPS \boxed{A} \end{array} \right] \rightarrow \boxed{I} \left[ \begin{array}{l} LEX + \\ COMPS \boxed{A} \end{array} \right], H \left[ \begin{array}{l} AUX + \\ COMPS \langle \boxed{I} \rangle \end{array} \right]$$

The rule specifies that the auxiliary head combines with a lexical (LEX) complement ( $\boxed{I}$ ), and that to the resulting combination the COMPS value ( $\boxed{A}$ ) of this lexical complement is composed.<sup>4</sup> This system, interacting with appropriate lexical entries for auxiliary verbs, will allow the auxiliary verb to combine with the preceding main verb, forming a complex predicate.

We also take BNs like *kes*, *tus*, *man* to form a complex predicate with the following predicate. These BNs also display a tight syntactic unit with the following predicates: no elements can intervene between the two.

- (22) a. \*sensayngnim-i o-si-l kes cengmal kathta  
 teacher-NOM come-HON-MOD BN really seem  
 ‘It seems that the teacher will really come.’  
 b. \*sensayngnim-i o-si-l tus cengmal hata  
 teacher-NOM come-HON-MOD BN really do  
 ‘It seems that the teacher will really come.’

The complex predicate analysis can also reflect the mono clausal property with respect to NPI.

- (23) a. \*John-un [amuwto ossta-ko] mit-ci anh-ass-ta  
 John-TOP anybody came-COMP] believe-COMP neg-PAST-DECL  
 ‘John didn’t believe anybody came’  
 b. amwuto o-n kes kath-ci anh-ta  
 anybody come-MOD BN seem-COMP NEG-DECL  
 ‘It seems that no one has come.’

Given that the expression *kes kath-ci anh-ta* forms a complex predicate, we could expect the NPI *amwuto* in the same clause.

Just like the Head-Complement type of BNs, the complex-predicate BNs also select a dependent sentence realized as the complement:

(24) 
$$\left[ \begin{array}{l} bn\text{-}lex \\ \langle kes \rangle \\ HEAD \left[ \begin{array}{l} POS \textit{noun} \\ NFORM \textit{kes} \end{array} \right] \\ LEX + \\ ARG\text{-}ST \left\langle S \left[ \begin{array}{l} MOD \langle N \rangle \\ SEM \boxed{2} \end{array} \right] \right\rangle \\ SEM \boxed{2} \end{array} \right]$$

<sup>4</sup>This kind of argument composition is different from the previous analyses, mainly in that the composition happens in syntax rather than in the lexicon.

Notice that there is one difference from BNs like *swu*. That is, complex-predicate BNs are marked as carrying the feature LEX to reflect that it will form a *head-lex-ph*. In addition, its semantics is identified with the complement, reflecting the fact that it behaves like a sentential complementizer, even though it is categorically a noun.

But how about the predicate *kath-ta* ‘seem’? Does this select only one argument? Unlike nominal elements, all verbal elements have a subject. What is the subject of this verb? Can it be identical with the subject of the dependent clause? The subject of the dependent clause cannot be identified with that of *kath-ta*. Observe the following between the AVC and the BNC:

- (25) a. *sensayng-nim-i o-si-ko siph-(usi)-ta*  
 teacher-HON-NOM come-HON-COMP would.like-HON-DECL  
 ‘The teacher would like to come.’
- b. *sensayng-nim-i o-si-l kes kath-(\*usi)-ta*  
 teacher-HON-NOM come-HON-PNE BN seem-HON-DECL  
 ‘The teacher would like to come.’

If the subject of *kath-* ‘seem’ is *sensayngnim* ‘teacher-HON-NOM, there is no reason why we cannot have the form *kath-usi-ta* as in *siph-usi-ta*. This appears that the BNC does not have any overt subject, but a covert subject not realized in syntax. As a way of reflecting this idea, we posit the following lexical entry for *kath-ta* ‘seem’:

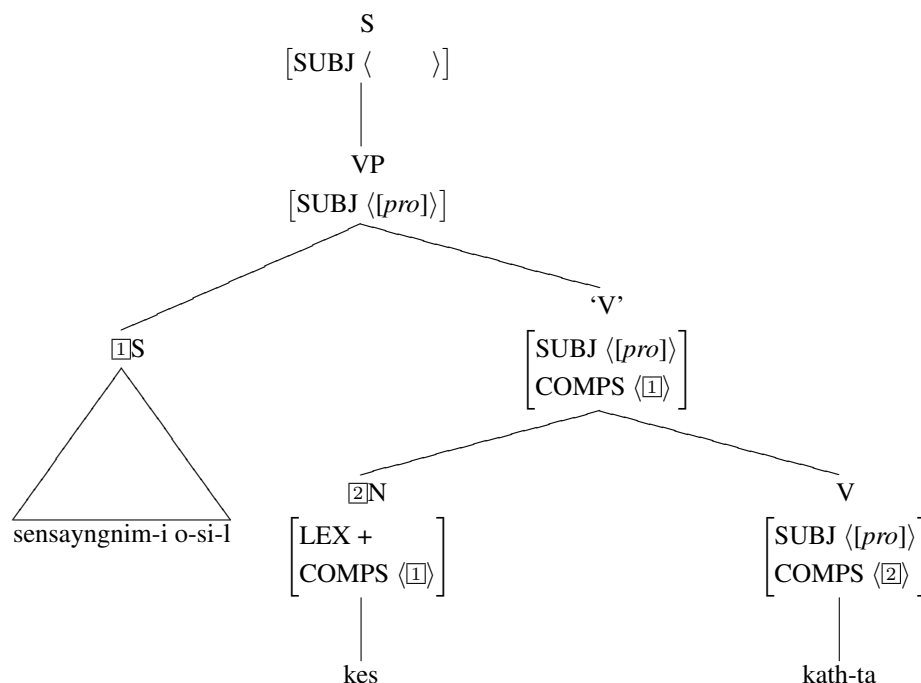
- (26) 
$$\left[ \begin{array}{l} \textit{aux-v-bn} \\ \langle \textit{kath-} \rangle \\ \text{HEAD} \left[ \begin{array}{l} \text{POS } \textit{verb} \\ \text{AUX } + \end{array} \right] \\ \text{ARG-ST} \left\langle \text{NP}[\textit{pro}], \left[ \begin{array}{l} \text{IND } \boxed{3} \\ \text{NFORM } \textit{kes} \end{array} \right] \right\rangle \\ \text{SEM} \left[ \text{RELS} \left\langle \left[ \begin{array}{l} \text{PRED } \textit{seem} \\ \text{ARG0 } \boxed{3} \end{array} \right] \right\rangle \right] \end{array} \right]$$

This lexical entry specifies that the auxiliary verb *kath-ta* selects two arguments: the first one is realized as the *pro* subject whereas the second one is realized as the COMPS whose NFORM value is *kes*.

Given these, we will then have a more elaborated structure like the following for a complex-predicate BNC given in (9):



(27)



The auxiliary-like verb *kath-ta* takes two arguments. The first argument is the *pro* subject whereas the second argument is the BN *kes*. The verb, combining with this BN complement, forms the complex predicate that inherits the BN's COMPS value. This complex predicate then in turn combines with the dependent sentence, forming a VP with the *pro* subject unsaturated. The language specific rule then allows this VP to project into a complete S.

### 3. An Implementation and Concluding Remarks

The analysis we have presented so far has been incorporated into the typed-feature structure grammar HPSG for Korean (Korean Resource Grammar) aiming at working with real-world data. To test the performance and feasibility of the analysis, we have implemented it into the LKB (Linguistic Knowledge Building).<sup>5</sup>

In representing the semantics, we employ Minimal Recursion Semantics (MRS) developed by Copestake et al. (2005). The MRS is a framework for computational semantics designed to enable semantic composition using only the unification of type feature structures (Bender, Flickinger, and Oepen, 2002; Flickinger and Bender, 2003). For example, Figure 1 and Figure 2 are the parsed results for the sentence (9a) and (9b) in our system.

We can see here that the MRS that the grammar generates provides us with enriched semantic information as well. The value of LTOP is the local top handle, the handle of the relation with the widest scope within the sentence. The INDEX value here is identified with the ARG0 value of the *prpstn\_m\_rel* (propositional message). The attribute RELS is basically a bag of elementary predications (EP) each of whose values is a *relation*.<sup>6</sup> Each of the types *relation* has at least three features LBL, PRED (represented here as a type), and ARG0. For example, we can see that in Fig. 1 the semantic relation *not\_exist\_rel* selects *possibility* as its ARG1 value. The relation *possibility* also takes the *come\_rel* as its argument, inducing the correct semantics. In Fig 2, we can observe

<sup>5</sup>The current Korean Resource Grammar has 394 type definitions, 36 grammar rules, 77 inflectional rules, 1,500 lexical entries, and 2100 test-suite sentences, and aims to expand its coverage on real-life data.

<sup>6</sup>The attribute HCONS is to represent quantificational information. The *undef\_q\_rel* means an undefined quantificational meaning assigned to the unexpressed determiner. See Bender et al. (2002).

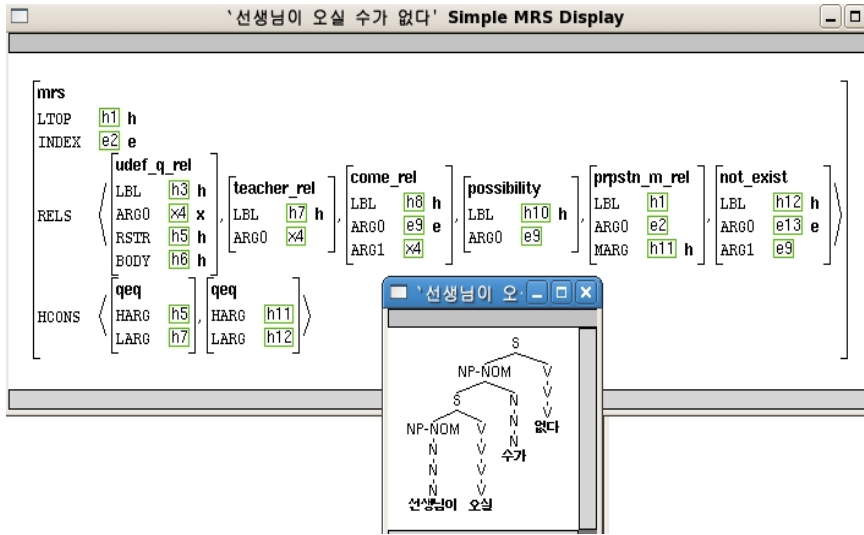


Figure 1: Parsed Tree and MRS for 'There is no possibility that the teacher will come.'

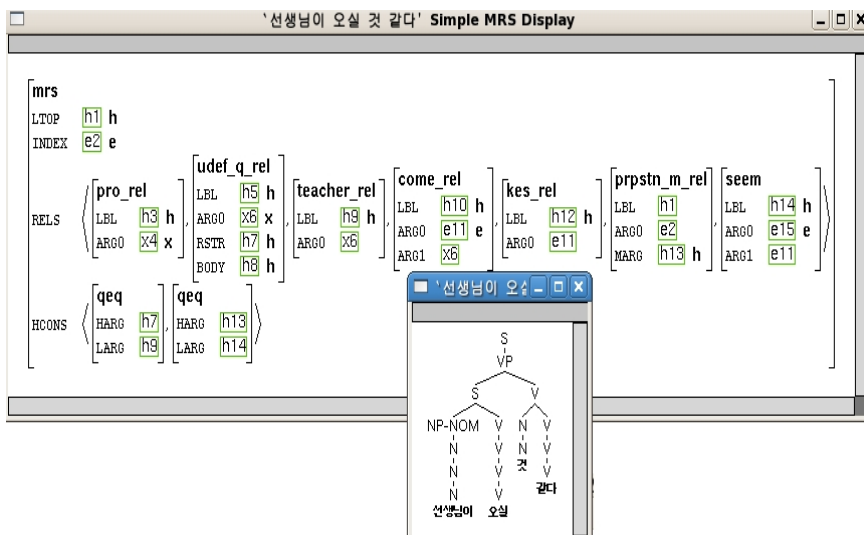


Figure 2: Parsed Tree and MRS for 'It seems that the teacher will come.'

that the ARG1 value of *seem* is ‘e9’ which is also the ARG0 value of the BN’s meaning *kes\_rel*. This ‘e9’ is in fact the event that *come\_rel* denotes.

#### 4. Conclusion

The Korean BNC (bound noun construction) display complex syntactic, semantic, and constructional properties. In particular, their combinatorial possibilities with respect to the complement and predicate types call for a much finer-grained syntax. The BNCs can even be classified into two types, depending on the syntactic coherence with the following predicate.

This paper has developed a constraint-based approach that can dissolve such issues. In terms of syntax, we postulated two different syntactic structures: head-complement and complex-predicate structures and then specify articulated lexical properties for the BNs and relevant predicates. This system has been implemented in the LKB system, which gave us robust parsing results for the given sentences.

#### References

- Bender, Emily M., Dan Flickinger, and Stephan Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on computational Linguistics*, pages 8–14, Taiwan. Taipei.
- Cha, Jong-Yul. 2001. *Constraints on Clausal Complex Noun Phrases in Korean with Focus on the Gapless Relative Clause Construction*. Ph.D. thesis, UIUC.
- Copestake, Ann. 2002. *Implementing Typed Feature Structures*. CSLI Publications, Stanford.
- Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal recursion semantics: an introduction. *Research on Language and Computation*, 3(4):281–332.
- Flickinger, Dan and Emily M. Bender. 2003. Compositional semantics in a multilingual grammar resource. In *Proceedings of the ESSLLI 2003 Workshop “Ideas and Strategies for Multilingual Grammar Development”*, Vienna, Austria.
- Kim, Jong-Bok. 2004. *Korean Phrase Structure Grammar*. Hankook Publishing. In Korean.
- Kim, Jong-Bok and Jaehyung Yang. 2004. Projections from morphology to syntax in the Korean resource grammar: Implementing typed feature structures. *Lecture Notes in Computer Science*, Vol.2945, pp.13–24, Springer-Verlag, 2004.2.