Processing Three Types of Korean Cleft Constructions in a Typed Feature Structure Grammar[*]

Jong-Bok Kim

Kyung Hee University

Jaehyung Yang

Kangnam University

The expression KES, one of the most commonly used words in the Korean language, has various usages. This expression is also used to express English-like cleft constructions. Korean seems to employ at least three different types of cleft constructions: predicational, identificational, and eventual. The paper tries to provide a constraint-based analysis of these three types of Korean cleft constructions and implement the analysis in the LKB (Linguistic Knowledge Building) system to check the feasibility of the analysis. In particular, the paper shows how a typed feature structure grammar, couched upon HPSG, can provide a robust basis for parsing Korean cleft constructions.

Keywords : KES, cleft, computational implementation, predicational, identificational

---

† 교신저자: 김종복. 경희대학교 영어학부 & 언어연구소. 연구분야: 통사, 의미론, 전산언어학. E-mail: jongbok@khu.ac.kr

Introduction

Cleft constructions are employed to mark a certain constituent as a discourse prominent element.[1] In this respect, Korean has at least three main types of clefts:[2]

(1) a. Predicational:

    [John-i   I   ilk-un   kes-un]   [kacca]$_i$-i-ta

    John-NOM   read-MOD   KES-TPC   fake-COP-DECL

    'What John read is a fake.'

    b. Identificational:

    [i chayk]$_i$-i   palo [John-i $_i$ ilk-un   kes-i-ta]

    this book-NOM   very John-NOM   read-MOD   KES-COP-DECL

    'This book is what John read.'

    c. Eventual:

    kuttay   [John-i   cip-ey   o-n]   kes-i-ess-ta

    the moment John-NOM home-LOC   come-MOD   KES-COP-PST-DECL

    'It is at the very moment that John came home.'

These three types of cleft mainly consist of a cleft clause, a pivot XP, and the copula verb. The predicational cleft in (1a) consists of a cleft clause with the missing object coindexed with the precopula expression *kacca* 'fake' whereas the identificational cleft in (1b) has the nominative phrase *i chayk* 'this book' as the pivot XP coindexed with the missing object in the following cleft clause. In these two clefts, the pivot XP is linked to the content of the cleft clause introduced by KES, though the exact semantic function is different. For example, in the predicational cleft (1a), the XP is

---

1) Much of the theoretical discussion follows from Kim 2008.
2) See Kim 2008 and Kim and Sells 2007 for the distinction between predicational and identificational cleft.

predicated of the individual that the cleft clause refers to, whereas in the identificational one, the XP and the individual are in a identity relation. Meanwhile, in the event cleft, the whole clause preceding the KES expression is clefted, functioning as the pivot phrase. In this sense, the whole clause is focused. The structure of these three types of clefts can be schematized as follows:

(2) a. Predicational: [[$_S$ ___$_i$]-KES]-TOP XP$_i$-COP-DECL

b. Identificational: XP$_i$-TOP [$_S$ ___$_i$]-KES-COP-DECL

c. Eventual: [[adverbial], [$_S$ 'saturated clause']-KES]-COP-DECL

As represented here, both of the predicational and identificational have a pivot or highlighted expression like the English cleft constructions whereas in the event cleft it is the whole clause that seems to be clefted. The event clause is thus different from the other two in that the clause has no missing element.[3]

This paper aims to review the basic properties of these three different types of Korean cleft constructions and provide a constraint-based analysis. We also show a brief summary of the results of implementing this analysis in the LKB (Linguistic Knowledge Building) system (Copestake 2002).

Formal Properties of the Cleft Constructions

Predicational and Identificational Cleft

As observed, the cleft constructions mainly consist of a cleft clause, a focus element, and the copula. The canonical cleft clause, usually representing given or discourse-old

---

3) Some transformational analyses assume that the predicational and identificational clefts are derived from the eventual clefts. See Jhang 1995 and Sohn 2004.

information, has a syntactically missing element.[4] In this respect, Korean clefts behave like relative clauses, but differently from topic constructions which can be either gapped or gapless. Consider these three constructions:

(3) a. ku   chayk-un [John-i ____ ilk-ess-ta]
        the   book-TOP   John-NOM   read-PST-DECL
        'This book, John read.'

    b. kkoch-un   [cangmi-ka   yepputa]
        flowers-TOP rose-NOM   pretty
        'As for flowers, roses are pretty.'

(4) a. [John-i ____ ilk-un]   ku chayk
        John-NOM   read-MOD the book
        'the book that John read'

    b. *[John-i   ku sosel-ul   ilk-un]   ku chayk
        John-NOM   the novel-ACC read-MOD the book

(5) a. [John-i ____ ilk-un   kes]-un   palo i   chayk-i-ta
        John-NOM   read-MOD KES-TOP   very   this book-COP-DECL
        'What John read is this very book.'

    b. *[John-i   ku sosel-ul   ilk-un   kes]-un   palo i   chayk-i-ta
        John-NOM the novel-ACC read-MOD KES-TOP very this book-COP-DECL

As illustrated here, the topic construction can have either a gap or a non-empty gap whereas the relative clause and cleft construction requires its clause to have a missing gap element. In this sense, clefts are like relative clauses, rather than topic clauses.

As for the possible type of the pivot or focus phrase, in the predicational copula, the focused XP can be either an argument or an adjunct. The postposition or

_____

4) Much of the discussion in this section follows from Kim 2008 and Kim and Sells 2007.

semantic marker of the focused expression is optional:

(6) a. [John-i   Mary-lul   manna-n   kes]-un   {kongwen-(eyse)}-i-ta
       [John-NOM   Mary-ACC   meet-MOD   KES]-TPC park-at-COP-DECL
       'It was at the park that John met Mary.'

    b. John-i   Mary-lul   manna-n   kes-un   [tosekwan-(eyse)}-i-ta
       John-NOM   Mary-ACC   meet-MOD   KES-TPC library-at-COP-DECL
       'Where John met Mary is (at) the library.'

    c. swum-i   taptapha-n   kes-un   [sanso-ka   pwucokhayse]-i-ta
       breath-NOM   choking-MOD   KES-TOP oxygen-NOM   short.do-COP-DECL
       'Because of lack of oxygen, it is hard to breathe.'

    d. John-i   Mary-eykey senmwul-ul cwu-n   kes-un   [wupyen-(ulo)}-i-ta
       John-NOM   Mary-DAT   present-ACC give-MOD KES-TOP mail-by-COP-DECL
       'The way John gave Mary a present is by mail.'

In the predicational cleft, an adverbial element also can be focused as long as it is categorically nominal:[5]

(7) a. John-i   Mary-lul   manna-n   kes-un   [ecey]-i-ta
       John-NOM   Mary-ACC   meet-MOD   KES-TPC yesterday-COP-DECL
       'It is yesterday when John met Mary.'

---

5) However, true adverbs cannot be focused:

(i) a. *John-i   talli-n   kes-un   [chenchenhi]-i-ta
        John-NOM   run-MOD   KES-TOP   slowly-COP-DECL
        '(lit.) The way John ran was slowly.'

    b. *[chenchenhi]-ka John-i   talli-n   kes-i-ta
        slowly-NOM   John-NOM run-MOD   KES-COP-DECL

As noted here, neither the predicational nor identificational cleft allows a true adverb to be focused.

b. John-i   Mary-lul   manna-n   kes-un   [siksa-lul   ha-ko   nase]-i-ta

John-NOM   Mary-ACC   met-MOD   KES-TPC   meal-ACC   do-COMP   after-COP-DECL

'It is after having a meal when John met Mary.'

Meanwhile, the identificational cleft does not allow the PP adjunct to be focused, regardless of the presence of the postposition:

(8)   a. *[kongwen-(eyse)]PP-ka [John-i   Mary-lul   manna-n   kes]-i-ta

park-at-NOM   [John-NOM   Mary-ACC meet-MOD   KES]-COP-DECL

b. [sanso-ka   pwucokhayse]-ka swum-i   taptapha-n   kes-i-ta

oxygen-NOM   short.do-NOM   breath-NOM   choking-MOD   KES-COP-DECL

c. [tosekwan-(eyse)]-ka John-i   Mary-lul   manna-n   kes-i-ta

library-at-NOM   John-NOM   Mary-ACC   meet-MOD   KES-COP-DECL

d. *[ecey]-ka   John-i   Mary-lum   manna-n   kes-i-ta

yesterday-NOM   John-NOM   Mary-ACC   meet-MOD KES-COP-DECL

This indicates that unlike the predicational cleft, the identificational cleft allows only an NP argument to serve as its XP focus.

The gapped element in the cleft clause can be in the embedded clause, allowing a long dependency relationship between the gap and the linked XP:

(9)   a. [John-i   [Mary-ka ____   cohahanta-ko] sayngkakha-nun   kes]-un

John-NOM   Mary-NOM   like-COMP   think-MOD   KES-TOP

i   kulim-i-ta

this picture-COP-DECL

'What John thought Mary likes is this picture.'

b. i   kulim-i   [John-i   [Mary-ka ____   cohaha-n-ta-ko]

this picture-NOM John-NOM Mary-NOM   like-PRES-DECL-COMP

sayngkakha-nun   kes]-i-ta

think-MOD   KES

'This picture is what John thought Mary likes.'

In both predicational and identificational clefts, the pivot phrase *i kulim* 'this picture' is linked to the object of the embedded clause. This pivot XP, however, cannot be an adjunct in the embedded clause. This is once again similar to relative clauses:

(10) a. [John-i   [Mary-ka   ku chak-ul   ilkessta-ko]] sayngkakha-n   ecey

       John-NOM Mary-NOM the book-ACC   read-COMP   think-MOD   yesterday

       'the time when John thought Mary read the book'

    b. [John-i   [Mary-ka   ku chak-ul   ilkessta-ko] sayngkakha-nun   kes]-un

       John-NOM Mary-NOM the book-ACC read-COMP   think-MOD   KES-TOP

       ecey-i-ta

       yesterday-COP-DECL

       'The time when John thought Mary read the book was yesterday'

In both relative and cleft examples here, the relativized and cleft adjunct is linked to the higher main clause, not to the embedded clause.

We can further observe that just like relative clauses, the cleft observes the CNPC (complex noun phrase constraint):

(11) a. [John-i   ___   piphanha-n   kes-un]   ku   nonmwun-i-ta

       John-NOM   criticize-MOD KES-TOP   the article-COP-DECL

       'What John criticized is the article.'

    b. *[John-i   [[___ ssu-n]   salam-ul]] piphanha-n]   kes-un

       John-NOM   write-MOD person-ACC criticize-MOD KES-MOD

       ku nonmwun-i-ta

       the article-COP-DECL

'(lit.) What John criticized the person who wrote ___ was the article.'

This indicates that the cleft clause introduced by KES behaves like a nominal clause that forms an island even though internally it is a clause.

Eventual Cleft Constructions

Unlike the predicational and identificational constructions, KES can nominalize a whole preceding S, highlighting an event, as in (12).[6]

(12) a. [ku   ttay]   [sako-ka   na-n]   kes-i-ya
         that moment accident-NOM   happen-MODpast KES-COP-DECL
         'It is at that moment that an accident happened.'
     b. [ku yeca-ka   John-ul   manna-n]   kes-i-ya
         that woman-NOM   John-ACC   meet-MODpast   KES-COP-DECL
         'The fact is that [that woman met John].'

Such an event cleft cannot be used discourse initially:

(13) cal   iss-ess-e?   nay-ka   tola   o-ass-e/*o-n ke-ya
     well   exist-PST-QUES i-NOM   return come-PST-DECL
     'How have you been? I came back

This kind of cleft construction conveys the meaning of 'cause, reason, explanation,

6) It is possible to present the new information as a canonical VP:

(i) chelwsu-nun   [onul   hyuil-i-nci   moll-ass-ten   kes]-i-ess-ta
    Chelsoo-TPC [today   holiday-COP-COMP not.know-PAST-MOD KES]-COP-PAST-DECL
    'The fact is that Chelsoo did not know that today is a holiday.'

Intuitively, this VP focus example presents a noteworthy fact about a given individual.

or consequence', focusing the information in the cleft clause.[7] Notice that there is no syntactic gap in the event clause. The clause also is all presented as new information, as can be attested by the fact that these examples can be an appropriate answer to a question like (14):[8]

(14) mwusun   il-i-ni?

　　　what   thing-COP-Q

　　　'What happened?'

Unlike the predicational and identificational ones, the KES in the event cleft cannot be replaced by a common noun:

(15) a. ku   ttay   sako-ka   na-n   kes/*iyu-i-ta

　　　　　then   accident   occur-MOD   KES/reason-COP-DECL

　　　　　'And then the accident occurred.'

　　　b. kuliko nase hyung-i   os-ul   twici-nun

　　　　　and   then brother-NOM   clothes-ACC search-MOD

　　　　　kes/*swunkan-i-ess-ta

　　　　　KES/moment-COP-PST-DECL

　　　　　'And then, brother was searching the clothes.'

---

7) We leave out the exact semantic and pragmatic functions of this construction.

8) Any phrase within the event cleft can have a narrow focus interpretation with a phonological prominence on it.

　　(i) a. kuliko nase   HYUNG-i   os-ul   twici-nun   kes-i-ess-ta

　　　　　　and   then   brother-NOM clothes-ACC search-MOD   KES-COP-PST-DECL

　　　　　　'And then my brother was searching the clothes.'

　　　　b. kuliko nase hyung-i OS-ul twici-nun kes-i-ess-ta

　　　　c. kuliko nase hyung-i os-ul TWICI-NUN kes-i-ess-ta

In addition, the cleft clause induces a freezing effect in that no element can be extracted out of the clause. For example, relativization is disallowed from the event cleft clause:

(16) a. os-ul   twici-n   hyung
        clothes-ACC   search-MOD   brother
        'brother who is searching the clothes'
     b. akha-n   haksayng-i-n   hyung
        honest-MOD   student-COP-MOD   brother
        'the bother who is honest'
     c. *kuliko nase os-ul   twici-nun   kes-i-n   hyung
         and   then clothes-ACC search-MOD   KES-COP-MOD   brother

Syntax and Semantics of the Cleft Constructions

Predicational and Identificational Cleft

The observations we have seen in the previous section have shown us that the cleft clause exhibits nominal properties externally though it displays verbal properties internally. With the aim of implementing the analysis for computational purposes, the challenges are thus how we capture these mixed properties with less stipulations.

The first issue in processing cleft constructions concerns the categorical status of KES. Consider the main uses of KES:

(17) a. nay kes-i   ne kes-pota   khu-ta
        my   thing-NOM your thing-more big-DECL
        '(Lit.) My thing is bigger than your thing.'

b. [[John-i _____ mek-un} kes]-ul mek-ess-ta

    John-NOM eat-MOD KES-ACC eat-PST-DECL

'(We) ate the thing that John ate.'

c. [[John-i talli-nun} kes]-ul moll-ass-ta

    John-NOM run-MOD KES-ACC not.know-PAST-DECL

'(We) didn't know that John was running.'

As noted here, KES in (17a) combines with a specifier whereas the one in (17b) combines with the relative clause with one missing argument. In both of these examples, KES refers to a 'thing'. Meanwhile, KES in (17c) combines with a complete sentence, referring to the event denoted by the clause.[9] In terms of meaning, we can assume KES to have at least the following two lexical entries:

(18) a.
$$\begin{bmatrix} bn \\ \langle \text{kes} \rangle \\ \text{SYN} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{POS } noun \\ \text{NFORM } kes \end{bmatrix} \\ \text{VAL}|\text{COMPS} \left\langle \text{S} \begin{bmatrix} \text{GAP } \langle \ \rangle \\ \text{IND } e1 \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{SEM}|\text{INDEX } i \end{bmatrix}$$
b.
$$\begin{bmatrix} cn \\ \langle \text{kes} \rangle \\ \text{SYN} \begin{bmatrix} \text{HEAD}|\text{POS } noun \\ \text{VAL}|\text{SPR} \langle (\text{Det P}) \rangle \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{INDEX } i \\ \text{RELS} \left\langle \begin{bmatrix} \text{PRED } one\_rel \\ \text{ARG0 } i \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The lexical entry (18a) means that KES refers to an individual equivalent to the meaning of *one* and combines with an optional DetP.[10] This kind of treatment, in

9) In the literature, KES in (17c) has been treated as a complementizer. (cf. Jhang 1995 and Sohn 2004)

10) Though it may be premature to link KES with the substitute pronoun *one*, there exist many cases where the two behave similar (cf. Quirk et al. 1985):

(i) a. The one with chocolate frosting has cream filling.

which KES is taken to be a type of common noun diverges from the traditional view treating KES as only a bound noun. Corpus data reveal that KES can be used like a common noun in various contexts:

(19) saylowun siswul-ul   hanta-ko, yatan-i-ess-nuntey, kes-to   chwisotoy-ess-e.
     new   operation-ACC do-COMP lowsy-COP-PST-but, KES-also cancel-PST-DECL
     'People were talking about the new way of operation, but it was also canceled.'
     (Sejong Corpus)

This in turn means that we predict examples like the following:

(20) [John-i   ___   ilk-un   kes/ku   kes]-ul   **ilk-ess-ta**
     John-NOM   read-MOD   KES/the   KES-ACC read-PAST-DECL
     '(We) read the thing that John ate.'

Like other common nouns, the expression KES, regardless of combining with its specifier or not, can thus be modified by the relative clause.

Unlike such a common noun KES, the lexical entry for KES in (18b) specifies that KES is a bound noun, selecting a saturated sentence. In this case, its INDEX value is identified with that of the sentential complement, insuring that KES denotes an even t.[11] One clear example where KES is linked to an event can be found from an example like the following:

---

b. Is this the one you want to meet?

c. Do you want these ones?

d. These donuts look delicious; I think I will choose this one.

The expression *one* here can refer to an entity as well as a human; it can be pluralized; it can be a member of the set in the given context.

11) The bound noun KES also has the NFORM value. See the discussion of event cleft for its use.

(21) [John-i   sakwa-lul   mek-un kes-ul]   moll-ass-ta

    John-NOM   apple-ACC eat-MOD KES-ACC not.know-PAST-DECL

   '(He) didn't know that John ate an apple.'

Here the argument of predicate 'not.know' is the sentence introduced by KES, similar to the English complementizer *that*. In this sense we can assume that KES denotes an event identified with its sentential complement.

Now let's consider the following relative clause and predicative cleft example, respectively:

(22) a. [John-i   ____   mek-un   kes/sakwa]-ul   **mek-ess-ta**

      John-NOM   eat-MOD   KES/apple-ACC eat-PST-DECL

      '(We) ate the thing that John ate.'

   b. [John-i   ____   mek-un   kes/kwail]-un   **sakwa-i-ta**

      John-NOM   eat-MOD   KES/fruit-TOP   apple-COP-DECL

      'What John ate is an apple.'

The only difference we can observe here is the predicate. The sentence with a relative clause has a transitive verb whereas the one with a cleft clause has a copula. As we have seen earlier, both the relative and cleft clause are identical in the sense that each clause has a syntactic gap here. In both cases, KES can be replaced by a common noun, indicating that there is no semantic difference. Let's consider the structure of (22a) with a relative clause first:

(23)
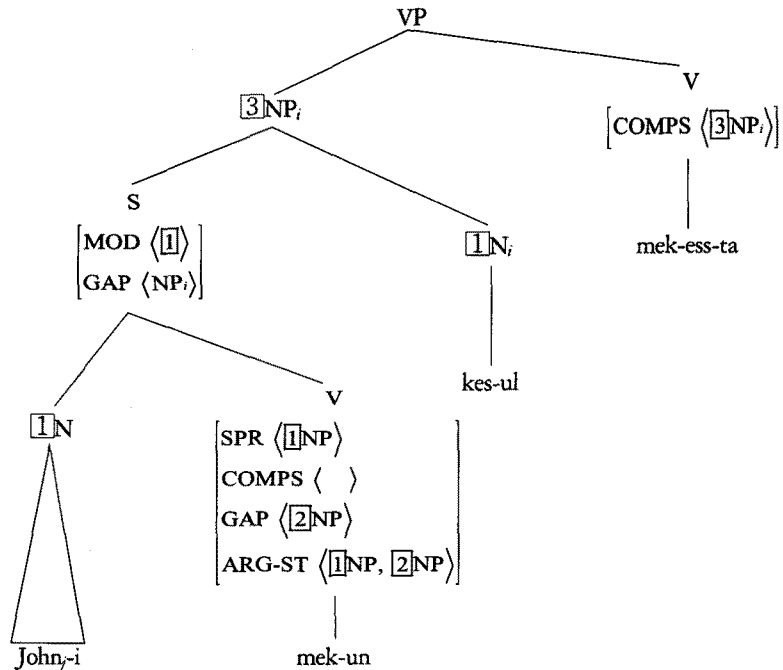
As we noted here, the noun KES in relative clauses is a common noun referring to an individual: this information is passed up to the NP projection (the index value 'i'). Since the verb *mek-un* 'ate' also requires its object to be a referential individual, there is no mismatch between these two requirements. To observe how we obtain the semantics correctly, let us consider the shorthand MRS (minimal recursion semantics) representation of this NP.[12]

---

12) Minimal Recursion Semantics, developed by Copestake 2003, is a framework of computational semantics designed to enable semantic composition using only the unification of type feature structures. See Copestake 2003 and Bender et al. 2002. The value of the attribute SEM(ANTICS) in our system represents a simplified MRS. Also see Kim 2004 for the analysis of Korean relative clauses. ARG0 canonically refers to the index value of the EP (elementary predicate) itself whereas ARG1 or ARG2 refers to the predicate's semantic arguments. CARG refers to constant arguments whose value can be a name.

(24)

$$\begin{bmatrix} \text{INDEX } i \\ \text{RELS } \left\langle \begin{bmatrix} \text{PRED } one\_rel \\ \text{ARG1 } i \end{bmatrix}, \begin{bmatrix} \text{PRED } eat\_rel \\ \text{ARG0 } e1 \\ \text{ARG1 } j \\ \text{ARG2 } i \end{bmatrix}, \begin{bmatrix} \text{PRED } name\_rel \\ \text{ARG1 } j \\ \text{CARG } john \end{bmatrix} \right\rangle \end{bmatrix}$$

This semantic representation simply means that there is an individual 'i' which the person named John eats. This index value is linked to the ARG1 value of 'one' coming from KES. This index value is the semantic head information visible at the NP level, functioning as the object of the matrix clause.

How about the predicative cleft sentence? Before we provide its structure, consider the lexical entry for the copula *i-ta*. We have seen that there are two different cleft constructions. This leads us to assume that there are two different copula uses: predicational and identificational (also see Kim and Sells 2007). The difference of these two different copulas is represented in the following lexical entries:
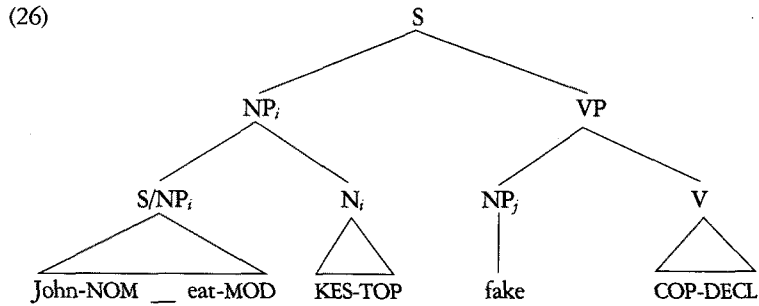
(25) a. Predicational Copula:

$$\begin{bmatrix} \langle \text{i-ta} \rangle \\ \text{ARG-ST} \langle \text{NP}_i, \text{XP}_j[\text{PRD }+] \rangle \\ \text{SEM}|\text{RELS} \left\langle \begin{bmatrix} \text{PRED } predicative\_rel \\ \text{ARG1 } i \\ \text{ARG2 } j \end{bmatrix} \right\rangle \end{bmatrix}$$

b. Identificational Copula:

$$\begin{bmatrix} \langle \text{i-ta} \rangle \\ \text{ARG-ST} \langle \text{NP}_i, \text{NP}_j \rangle \\ \text{SEM}|\text{RELS} \left\langle \begin{bmatrix} \text{PRED } identity\_rel \\ \text{ARG1 } i \\ \text{ARG2 } j \end{bmatrix} \right\rangle \end{bmatrix}$$

The predicational copula in (25a) requires its second argument to carry the positive PRD feature, ensuring that this expression predicates of the first argument (subject). The semantics also reflects this. Meanwhile, the identificational copula in (25b) requires the INDEX value of the first argument is in the *identity_rel* with that of the second argument. This lexical specification implies that the two expressions here have identical referential types.

Given these, we then can generate a structure like (26) for the predicational cleft (1a):

(26)

```
                             S
                 ┌───────────┴───────────┐
               NP_i                      VP
          ┌─────┴─────┐            ┌──────┴──────┐
        S/NP_i        N_i        NP_j            V
      ┌───┴───┐        △          │              △
John-NOM __ eat-MOD  KES-TOP    fake          COP-DECL
```

This structure, including the cleft clause as the subject and the predicative expression, will then induce the meaning similar to (25a). The predicative expression 'fake' will predicate of this nominal element, inducing a semantic representation like the following:[13]

$$(27) \quad \left[\text{RELS} \left\langle \begin{bmatrix} \text{PRED } one\_rel \\ \text{ARG0 } i \end{bmatrix}, \begin{bmatrix} \text{PRED } fake\_rel \\ \text{ARG1 } j \end{bmatrix}, \begin{bmatrix} \text{PRED } predicative\_rel \\ \text{ARG1 } i \\ \text{ARG2 } j \end{bmatrix} \right\rangle \right]$$

The meaning of the copula *i-ta* is relevant to the variable missing in the cleft. This

---

13) The index value of a predicative expression is identified with that of the gapped element in the cleft clause.
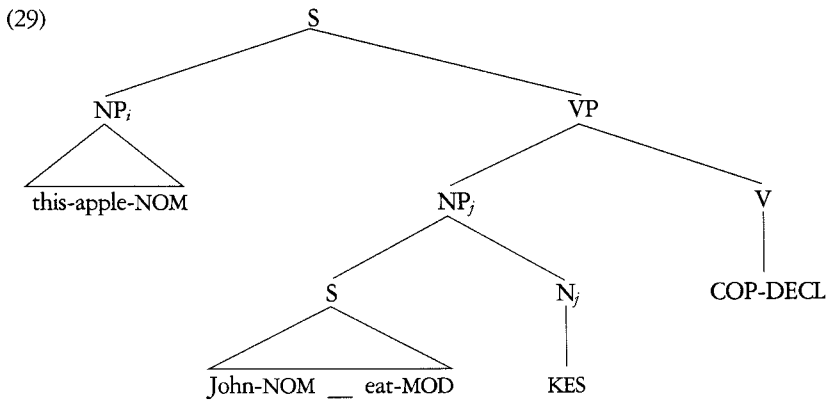
index value is 'i', which is in the *predicative* semantic relation with the index value of 'fake'. This in turn means that as long as the precopular expression can be predicated of the cleft-clause subject, there is no categorial restriction on the type of the precopular expression. This is why we allow other than an NP in this position as we have seen before.
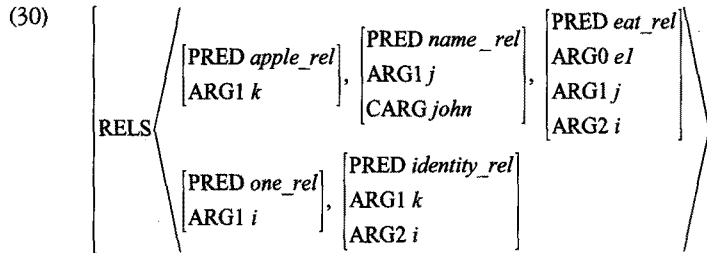
Now consider the structure of an identificational cleft sentence:

(28) I   sakwa-ka   John-i   mek-un   kes-i-ta

 this apple-NOM   John-NOM   eat-MOD KES-COP-DECL

 'This apple is what John ate.'

A simple tree representation will be something like the following:

(29)

```
                              S
                  _____
                 /                        \
              NP_i                          VP
              /\                    _____
             /  \                  /                    \
    this-apple-NOM              NP_j                      V
                          _____                 |
                         /               \             COP-DECL
                        S                 N_j
                      _____               |
                     /     \             KES
              John-NOM __ eat-MOD
```

The lexical constraints of the identificational copula insure that the index value of the subject is identified with that of KES as represented in the expected semantics of this sentence:

(30)

$$
\text{RELS} \left\langle
\begin{bmatrix} \text{PRED } apple\_rel \\ \text{ARG1 } k \end{bmatrix},
\begin{bmatrix} \text{PRED } name\_rel \\ \text{ARG1 } j \\ \text{CARG } john \end{bmatrix},
\begin{bmatrix} \text{PRED } eat\_rel \\ \text{ARG0 } e1 \\ \text{ARG1 } j \\ \text{ARG2 } i \end{bmatrix},
\begin{bmatrix} \text{PRED } one\_rel \\ \text{ARG1 } i \end{bmatrix},
\begin{bmatrix} \text{PRED } identity\_rel \\ \text{ARG1 } k \\ \text{ARG2 } i \end{bmatrix}
\right\rangle
$$

Unlike the predicational one, the identificational one thus requires the identity of two index values. This is why neither the PP or an adverbial element can occur as the pivot XP in the identificational cleft, whose data we repeat here:

(31) a. [[John-i   Mary-lul   manna-n]   kes-nun]   kongwon(eyse)-i-ta

      John-NOM   Mary-ACC   meet-MOD   KES-TPC   park.at-COP-DECL

      '(lit.) When John met Mary was at the park.'

    b. *[kongwen-(eyse)]PP-ka [John-i   Mary-lul   manna-n   kes]-i-ta

      park-at-NOM   [John-NOM   Mary-ACC   meet-MOD   KES]-COP-DECL

Since the predicational cleft only requires the precopular element to be predicative, regardless of its referential property, we can have a locative element in the pivot clause as well as an adverbial element as given in the following:

(32) [[John-i   Mary-lul   manna-n]   kes-nun]   ecey-i-ta

      John-NOM   Mary-ACC   meet-MOD   KES-TPC   yesterday-COP-DECL

      '(lit.) When John met Mary was yesterday.'

This event denoting KES clause here is in a predicative relation with the adjunct *ecey* 'yesterday'. That is, the semantics the analysis generates is something like the following:

(33)
$$\left[\text{RELS}\left\langle \begin{bmatrix} \text{PRED } meet\_rel \\ \text{ARG0 } e1 \\ \text{ARG1 } j \\ \text{ARG2 } m \end{bmatrix}, \begin{bmatrix} \text{PRED } yesterday\_rel \\ \text{ARG0 } t1 \\ \text{ARGF1 } e1 \end{bmatrix}, \begin{bmatrix} \text{PRED } predicative\_rel \\ \text{ARG1 } e1 \\ \text{ARG2 } t1 \end{bmatrix} \right\rangle \right]$$

We have shown that clefts are like relative clauses. It is then natural to expect that clefts are also involved in long distance dependency just like relative clauses since the analysis takes the internal structure of the cleft as that of a relative clause. In particular, the gap in the cleft clause is an argument (cf. See Kim 2004). This means in the present analysis, the argument-gapped cleft and adjunct gapped cleft are thus different: only the former is treated as a kind of unbounded dependency, as we have seen earlier. Let's consider similar data again:

(34) a. John-i   [Mary-ka   ___   ilkessta-ko]   sayngkakha-n   kes-un

John-NOM   Mary-NOM   read-COMP   think-MOD   KES-TOP

i   chayk-i-ta

this   book-COP-DECL

'The one that John thinks Mary read is this book.'

b. John-i   [Mary-ka   i   chaky-ul   ilkessta-ko]   malha-n   kes-un

John-NOM   Mary-NOM this book   read-COMP   said-MOD   KES-TOP

i   kos-(eyse)-i-ta

this   place-COP-DECL

'The place that John said Mary read this book is in this place.'

Though the precopular in (34a) expression is linked to the argument gap in the embedded clause, the adjunct precopular one in (35b) modifies only the matrix predicate 'said'.
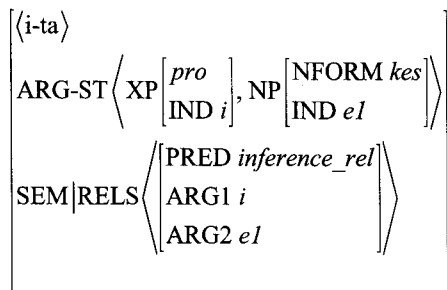
Eventual Cleft

As noted earlier, the event cleft construction is a construction where the entire matrix clause is headed by the nominalizer KES followed by the copula. Even though the KES here nominalizes the whole sentence, corpus search shows us that the event cleft commonly is preceded by a temporal or reason adverb that introduces this event clause.
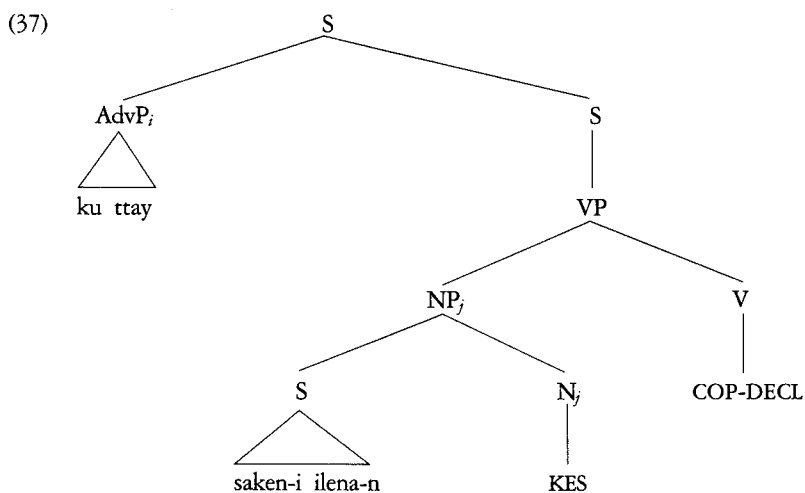
(35) a. ku ttay, saken-i ilena-n   kes/*iyu-i-ta
    then   accident occur-MOD KES/reason-COP-DECL
    'And then the accident occurred.'
  b. kuliko nase, hyung-i   os-ul   twici-nun kes-i-ess-ta
    and after brother-NOM clothes-ACC search-MOD KES-COP-PST-DECL
    'After that, brother searched the clothes.'

Such an event cleft clause does not occur in the beginning of context as in English examples like *It is then that Tom ate the big apple*. It appears that the event cleft sentence is interpreted based on the speaker's prior knowledge. This semantic relation can be 'cause, conclusion, reason, or explanation'. Following Declerck 1992, we assume that such relations are subtypes of the semantic relation 'inference'. That is, the interpretation of this construction is inferred from the speaker's prior knowledge. To reflect these properties, we assume that the copula here is minimally different from the copula in the predicational or identificational cleft in that the first argument is an unrealized *pro* element linked to the speaker's prior knowledge.

(36) Eventual:

$$\begin{bmatrix} \langle \text{i-ta} \rangle \\ \text{ARG-ST} \left\langle \text{XP} \begin{bmatrix} pro \\ \text{IND } i \end{bmatrix}, \text{NP} \begin{bmatrix} \text{NFORM } kes \\ \text{IND } e1 \end{bmatrix} \right\rangle \\ \text{SEM}|\text{RELS} \left\langle \begin{bmatrix} \text{PRED } inference\_rel \\ \text{ARG1 } i \\ \text{ARG2 } e1 \end{bmatrix} \right\rangle \end{bmatrix}$$

This lexical entry tells us that the two arguments selected by this peculiar copula are eventually related by the semantic relation *inference_rel*. Even though the first argument functions as *pro* and thus will not be realized in syntax, the argument is linked to the cleft clause by this semantic/pragmatic relation. This lexical entry will then allow a structure like the following:

(37)

```
                              S
                 ┌────────────┴──────────┐
             AdvP_i                       S
              ╱╲                          │
           ku ttay                        VP
                              ┌───────────┴──────┐
                            NP_j                  V
                     ┌────────┴────┐              │
                     S            N_j          COP-DECL
                    ╱╲             │
             saken-i ilena-n      KES
```

The lexical constraints of the eventual copula insures that we have the following MRS for this sentence:

(38)

$$\left[\text{RELS} \left\langle \begin{array}{l} \begin{bmatrix} \text{PRED } \textit{the moment\_rel} \\ \text{ARG0 } m \\ \text{ARG1 } el \end{bmatrix}, \begin{bmatrix} \text{PRED } \textit{pro\_rel} \\ \text{ARG1 } i \end{bmatrix}, \begin{bmatrix} \text{PRED } \textit{inference\_rel} \\ \text{ARG1 } i \\ \text{ARG2 } el \end{bmatrix} \\ \begin{bmatrix} \text{PRED } \textit{accident\_rel} \\ \text{ARG1 } k \end{bmatrix}, \begin{bmatrix} \text{PRED } \textit{occur\_rel} \\ \text{ARG1 } el \\ \text{ARG2 } k \end{bmatrix} \end{array} \right\rangle \right]$$

As given in the MRS, the sentence means that there occurred an accident at a given moment. This event is linked to the speaker's prior knowledge by the 'inference' relation.

Results of the Implementation

The analysis we have presented so far has been incorporated in the typed-feature structure grammar HPSG for Korean (Korean Resource Grammar) aiming at working with real-world data (cf. Kim and Yang 2004 and Kim 2004). To test its performance and feasibility, it has been implemented into the LKB (Linguistic Knowledge Building).[14] The test results give the proper syntactic as well as semantic structures for the three types of Korean cleft constructions.

Figure 1 and 2 are the parsing results of the two main types of cleft constructions we have seen so far. The small boxes in the figures indicate parsed tree structures whereas the big boxes denote the semantic representations.

---

14) The current Korean Resource Grammar has 394 type definitions, 36 grammar rules, 77 inflectional rules, 1100 lexical entries, and 2100 test-suite sentences, and aims to expand its coverage on real-life data.
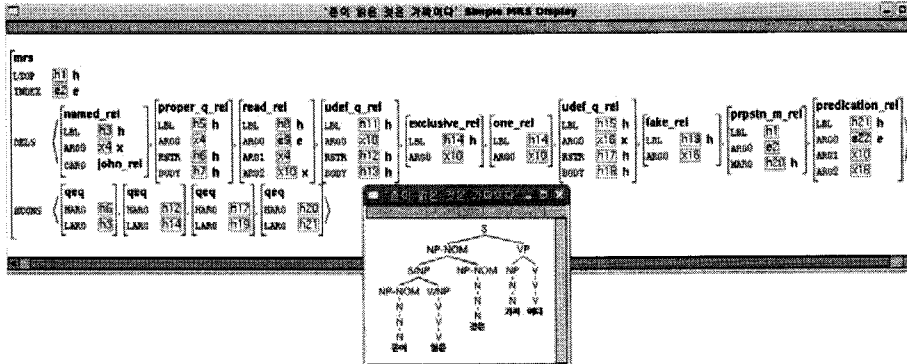
Fig 1. Parsed Trees and MRS for Predicational cleft (1a): 'What John read is a fake'

Consider Figure 1 first. In terms of the syntactic structure, the grammar generates the structure in which KES combines with the relative clause 'John read ____'. This functions as the subject. The copula verb selects this cleft-clause like element as its subject and the noun 'fake' as its predicative element. We can notice here that the MRS the grammar generates provides enriched information of the phrase. The value of LTOP is the local top handle, the handle of the relation with the widest scope within the constituent. The attribute RELS is basically a bag of elementary predications (EP) each of whose value is a *relation*.[15] Each of the types *relation* has at least three features LBL, PRED (represented here as a type), and ARG0. For the proper noun John, it has two related EPs: *named_rel* and *proper_q_rel*. The relation *udef_q_rel* is related to the projection of the common noun KES not combining with any specifier, whereas *exclusive_rel* concerns the topic marker *-nun*.[16] The meaning of KES is represented as *one_q_rel*. We can observe that the EP *read_rel* has two ARGs: one is

---

15) The attribute HCONS is to represent quantificational information. See Bender et al. 2002.

16) Korean common nouns do not require a determiner to project an NP. Even though a determiner is not available, we need to express an underspecified quantification on the noun in order to make the semantics compatible with the semantic output of other languages, and to make scope restrictions work. Such a move is essential in deep processing aimed at multilingual applications.

linked to the argument of *named_q_rel* and the other linked to *one_rel* (X10). This indicates that we have the meaning 'John reads one'. The relation *predication_rel* selects two arguments: X10 and X16. The second value X16 is linked to 'fake'. The parsing system thus gives us an enriched semantics for the predicational cleft sentence.
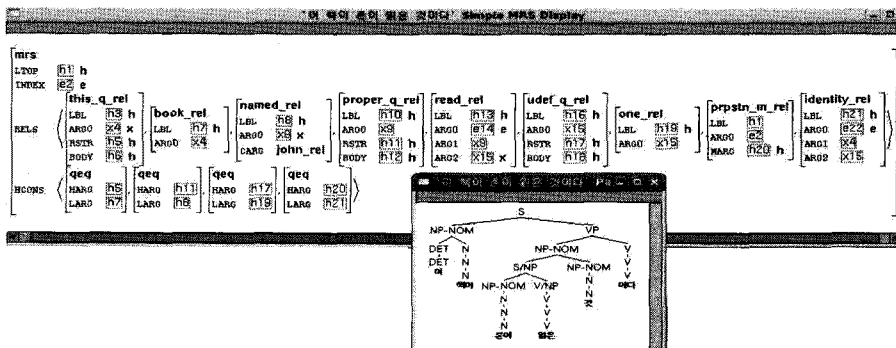

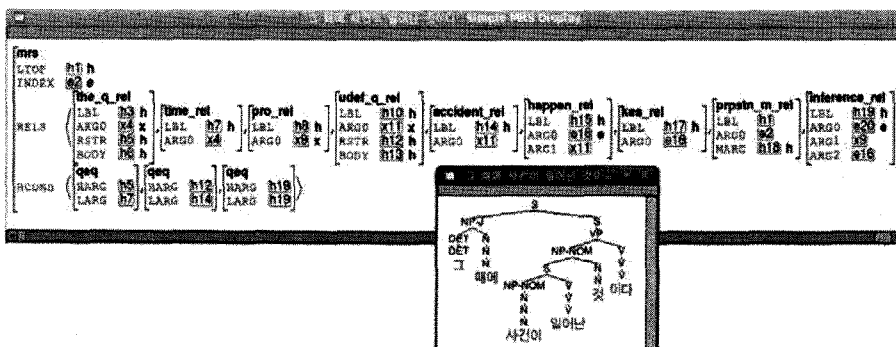Fig 2. Parsed Trees and MRS for Identificational cleft (1b): 'This book is what John read'

Fig 3. Parsed Trees and MRS for Eventual cleft (1c): 'At the moment, the accident occurred'

Figure 2 for the identificational cleft is similar to the predicational one. The syntactic structure in the small box shows us that 'this book' functions' as the subject whereas as the cleft clause as the precopular complement. In the MRS representation,

we can observe that *identity_q_rel* selects two arguments, ARG1 and ARG2. The first argument ARG1 is linked to the argument of the book (*book_rel* whereas the second argument ARG2 is identical with the argument of KES (*one_rel*).

Now let us consider the results of parsing an event cleft. Figure 3 is the results of parsing the sentence. The syntactic structure gives us the information that *ku ttay-ey* is modifying the matrix sentence which is missing the subject. We assumed that the subject is lexically a pronoun not realized in syntax. In terms of semantics, this unrealized pronoun (x9) is in an 'inference' semantic relation with the proposition (e16) 'the accident occurred'. The inference relation is context-determined.

Conclusion

We have seen that there are three types of Korean clefts constituted of a cleft clause, a focused expression, and a copula. These predicational, identificational, and eventual cleft are closely related to corresponding copula constructions.

The proper syntactic and semantic treatment of cleft constructions has been challenges even to theoretical aspects. Based on the simple assumption that KES is treated as a nominal element as a morphosyntactic category but refers to either an individual or an event. The pronoun KES in the cleft clause refers to an individual entity as in the relative clause. Given these basic assumptions, we have built a constraint-based grammar couched upon HPSG. The grammar we have built within a typed-feature structure system and well-defined constraints, eventually aiming at working with real-world data, has been implemented in the LKB (Linguistic Knowledge Building) system. We have shown that the grammar can parse the appropriate syntactic and semantic aspects of the three types of cleft constructions. Even though the test data set we used in checking the feasibility of the system is limited, the test results shows us that the grammar, built upon the typed feature

structure system, is efficient enough to build semantic representations for the complex cleft constructions.

References

Bender, Emily M., and Dan Flickenger, 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*, 8-14, Taipei, Taiwan.

Chung, Chan and Jong-Bok Kim. 2003. Differences between externally and internally headed relative clause constructions. In J.B. Kim and Stephen Wechsler (eds.) *Proceedings of HPSG 2002*, 3-25.

Collins, Peter C. 1991. *Cleft and Pseudo-Cleft Constructions in English*. London and New York: Routledge.

Copestake, Ann. 2002. *Implementing Typed Feature Structures*. Stanford: CSLI Publications.

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.

Davidse, Kristin. 2000. A Constructional Approach to Clefts. *Linguistics* 38.6: 1101-1131.

Heycock, Caroline and Anthony Kroch. 1999. Pseudo-cleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30, 365-397.

Jo, Jung-Min. 2004. *Grammatical Effects of Topic and Focus Information*. Ph.D. Dissertation. University of Illinois, Urbana-Champaign.

Kang, Bosook. 2006. Some pecularities of Korean 'kes' cleft constructions. *Studia Linguistica* 60.3: 251-281.

Kim, Jong-Bok. 2004. *Korean Phrase Structure Grammar* (In Korean). Hankook Publishing.

Kim, Jong-Bok. 2006. Minimal Recursion Semantics: An Application into Korean. *Journal of*

the Linguistic Association of Korea, 14.2: 59-85.

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.

Kim, Jong-Bok, Peter Sells and Michael T. Wescoat. 2004. Korean Copular Constructions: A Lexical Sharing Approach. To appear in M. Hudson, S.-A. Jun and P. Sells (eds.) *Proceedings of the 13th Japanese/Korean Linguistics Conference*. Stanford, CSLI Publications.

Kim, Jong-Bok and Peter Sells. 2007. Nominalizer *kes* and Information Structure. Paper presented in the 12th ISOKL, Harvard University, August 3-5, 2007.

Jhang, Sea-Eun. 1994. *Headed nominalizations in Korean: Relative clauses, clefts, and comparatives*. Ph.D. Dissertation. Simon Fraser University.

Quirk, Randolph, S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Sohn, Keun-Won. 2004. Kes-clefts, Connectedness effects, and the implications thereof. *Studies in Generative Grammar*, 14: 561-571.

요 약

유형화된 자질문법에서의
한국어 분열구문의 전산학적 처리

김 종 복   양 재 형
경희대학교   강남대학교

한국어 표현 '것'은 실제 언어사용에서도 가장 빈번히 사용되며, 다양한 문법적 기능을 가지고 있다. 영어의 분열문에 상응하는 기능으로도 사용되는 '것'은, 서술(predicational), 동일(identificational), 그리고 사건(eventual) 분열문을 도입할 수 있다. 본 논문은 이러한 '것'을 포함한 세 가지 분열구문에 대한 제약기반 분석을 제공하고, 이를 전산언어학적으로 구현한 결과를 제공한다. 유형화된(typed) 자질구조를 중요한 매개로 사용하는 제약기반 분석 방법은 분열구문에 대한 적절한 통사적, 의미적 처리(parsing) 결과를 제공한다.

주제어 : 분열구문, 서술분열문, 동일분열문, 사건분열문, 유형화된 자질구조문법, '것'