

Parsing Two Types of Multiple Nominative Constructions: A Constructional Approach

Jong-Bok Kim*, Peter Sells[†], and Jaehyung Yang[‡]

Kyung Hee University, Stanford University, and Kangnam University

Jong-Bok Kim, Peter Sells, and Jaehyung Yang. 2007. Multiple Nominative Constructions¹. *Language and Information 11.1*, 1–##. So-called multiple nominative constructions (MNCs) in Korean are quite theoretically as well as computationally puzzling. This paper shows that a grammar allowing the interaction of declarative constraints on types of signs – in particular, constructions (phrases and clauses) – can provide a robust and efficient way of encoding generalizations for two different MNCs. The feasibility of the grammar developed has been checked with its implementation into the LKB (Linguistic Knowledge Building) system. (Kyung Hee University, Stanford University, and Kangnam University)

Key words: multiple nominative, poessessive, adjunct, HPSG, LKB

1. Recognizing Two Types: Similarities and Differences

So-called ‘multiple’ nominative constructions (henceforth MNCs) exemplified in (1)a and (1)b are some of the more puzzling phenomena in topic-prominent languages like Korean, Japanese, and Chinese ((Yoon, 2004), (O’Grady, 1991)).²

* School of English, Kyung Hee University, Seoul, 130-701, Korea

[†] Department of Linguistics, Stanford University, CA 94305

[‡] School of Computer Engineering, Kangnam University, Kyunggi, 446-702, Korea

¹ An earlier version of this paper was presented at the 5th Korean-Japanese Workshop on Linguistics and Language Processing held at Kyung Hee University on Dec 9, 2006. We thank Jae-Woong Choe, Francis Bond, Yasunari Harada, and Kazuhiko Fukushima for helpful comments. We also thank three anonymous reviewers for questions and criticisms. This research was supported by the Korea Research Foundation (KRF-2005-A00577)

² The abbreviations for the glosses and attributes used in this paper are ACC (accusative), ARG (argument), C-CONT (constructional content), DAT (dative), DECL (declarative), LBL (label), LOC (locative), LTOP (local top), NOM (nominative), PL (plural), PRE (predicate), PST (past), IND (index), RELS (relations), TOP (topic).

- (1) a. John-i/-uy son-i khu-ta
 John-NOM/GEN hand-NOM big-DECL
 ‘John’s hand is big.’
- b. yelum-i/-ey/*-uy maykcwu-ka choyko-i-ta
 summer-NOM/-LOC/-GEN beer-NOM best-COP-DECL
 ‘Summer is the best time to have beer.’

In both examples, it is not the first but the second nominative (NOM) phrase that is the argument of the intransitive matrix predicate: it is the hand that is big, and it is the beer that tastes good in summer. *John* and *summer* are not direct arguments of the matrix predicate. Considering that a clause usually contains at most one subject, expressed as a NOM phrase, the function of the first NOM is then a puzzle.

In terms of pragmatic conditions, the first NOM phrase in both cases characterizes the remaining part (which is often called ‘sentential predicate’). For example, in (1)a having a big hand is a characterizing property of John whereas in (1)b, tasty beer is a characteristic of summer. If there is no such relation, the first phrase cannot be NOM, though it can be a genitive modifier:

- (2) a. John-uy/*-i [swuep-i ttapwunha-ta]
 John-GEN/-NOM class-NOM boring-DECL
 ‘John’s class is boring.’
- b. yelum-ey/*-i [John-ka congcong mikwuk-ul ka-n-ta]
 summer-LOC John-NOM often America-ACC go-PRES-DECL
 ‘In summer, John often goes to America.’

The boring class isn’t a characterizing property of John, though it could be a temporary property.

However, the first NOM in these examples also behaves differently. In examples like (1)a (which we call the possessive nominative construction (PNC)), the two consecutive NOM phrases are in a possessive relation, as attested by the alternation with the possessive marker on the first NOM. Meanwhile, in examples like (1)b (which we call the adjunct nominative construction (ANC)), there is no such a relation. The first phrase functions more like an adjunct, as indicated by the locative marker in (1)b.

There are also other differences between the first NOM phrase in the PNC and the ANC. For example, only the former can function as a raised object:

- (3) a. Mary-nun [John-ul] son-i khu-ta-lako mitessta
 Mary-TOP John-ACC hand-NOM big-DECL-COMP believed
 ‘Mary believed John’s hand is big.’
- b. ??/*Na-nun ecey-lul nalssi-ka acwu tewu-ess-ta-ko
 I-TOP yesterday-ACC weather-NOM very hot-PAST-DECL-COMP
 sayngkakha-n-ta
 think-PRES-DECL
 ‘I think yesterday the weather was really hot.’

The first NOM in the PNC can also serve as the antecedent of a floating quantifier, whereas this is not possible in the ANC:

- (4) a. haksayng-tul-i khi-ka [sey myeng-i] khu-ta
 students-NOM height-NOM three CL-NOM tall
 ‘Three students are tall.’
- b. *tosi-ka nalssi-ka [sey kos-i] cwup-ta
 city-NOM weather-NOM three CL-NOM cold
 ‘In three cities, the weather is cold.’

The first NOM expression *haksayng-tul-i* in (4)a can function as the antecedent of the classifier *three-CL-NOM*, but the NOM expression *tosi-ka* in (4)b cannot antecede the classifier *sey kos-i*.

Another clear difference between the two constructions comes from the possibility of invoking an idiomatic reading. Only the PNC allows an idiomatic reading, as illustrated in (5):

- (5) a. John-i pal-i nelp-ta
 John-NOM feet-NOM wide-DECL
 ‘John’s feet are wide.’ or ‘John has more contacts (idiomatic).’
- b. John-i kapang kkun-i ceyil kil-ta
 John-NOM bag strap-NOM most long-DECL
 ‘The straps of John’s bag are longest.’ or ‘John is the most highly educated (idiomatic).’

These two PNC sentences induce idiomatic meanings as well, but to our knowledge there are no ANC sentences which have an idiomatic meaning.

These differences indicate that the language has at least two different MNCs. However, this does not mean that the two do not share some properties. As noted earlier, the first NOM in both the PNC and ANC is in a characterizing relation with the remaining parts ('sentential predicate'). In addition, we can show that the first NOM in both constructions is the realization of information focus. The evidence that the first NOM marks focus can be drawn from several phenomena. For example, only the first NOM-marked phrase can be *wh*-questioned, as shown in (6)a. It is not possible to *wh*-question the second one as shown in (6)b.³

- (6) a. Nwu-ka apeci-ka kyoswu-i-si-ni?
 who-NOM father-NOM professor-COP-HON-Q
 '(lit.) Who is it whose father is a professor?'
- b. *John-i nwu-ka kyoswu-i-si-ni?
 John-NOM who-NOM professor-COP-HON-Q
 '(lit.) John's 'who' is a a professor?'

In addition, an MNC example like (8) can be an appropriate answer only for a MNC-style question with an initial nominative *wh*-phrase like (7)a, but not one like (7)b:

- (7) a. Nwu-ka apeci-ka kyoswu-i-si-ni?
 who-NOM father-NOM professor-COP-HON-Q
- b. Nwukwu-uy apeci-ka kyoswu-i-si-ni?
 who-GEN father-NOM professor-COP-HON-Q
- (8) John-i apeci-ka kyoswu-i-si-ta.
 John-NOM father-NOM professor-COP-HON-Q
 'It is John whose father is a professor.'

³ As a reviewer pointed out, in sentences like the following, we can question the second nominative:

- (i) a. Nwu-ka ton-i manh-ni
 who-NOM money-NOM many-Q
 'Who has a lot of money?'
- b. John-i mwues-i manh-ni?
 John-NOM what-NOM many-Q
 'What does John have a lot?'

We take that such an example is not a double nominative construction but a transitive sentence with two independent arguments.

Elliptical answers, often tested as focus information, also indicate the focus status of the first nominative phrase. The elliptical expression with no case marker as given in (9) can be an answer to the question (7a), but not to the question (7)b:⁴

(9) John-(i)/-*uy. ‘John-NOM/-GEN’

These differences imply that the initial *i/ka* marked NP serves as an independent focus phrase. Further, the first nominative (unlike a genitive NP) receives an exhaustive reading, a canonical property of focus. The impossibility of having the exclamatory expression *ceki* ‘there’ in (10)a, which is generally not used for exhaustive listing, but rather for neutral description, could be attributed to the exhaustive list reading of *John-i*.

- (10) a. *Ceki John-i apeci-ka o-si-nta!
 over.there John-NOM father-NOM come-HON-DECL
- b. Ceki John-uy apeci-ka o-si-nta!
 over.there John-GEN father-NOM come-HON-DECL

A similar point can be made with the definiteness of the initial phrase. It is not possible to have an indefinite pronoun as the initial nominative:⁵

- (11) *Etten salam-i apeci-ka pwuca-i-si-ta.
 some person-NOM father-NOM rich-COP-HON-DECL
 ‘(lit.) It is some person whose father is rich.’

Observing the similarities and differences between the two constructions we have discussed so far, the questions that arise with respect to parsing such constructions are (a) how to license the first NOM phrase which is not an argument of the main predicate, (b) how to process its semantic and pragmatic contributions to the sentence as a whole, and (c) how to recognize and reflect the different properties of these two constructions.

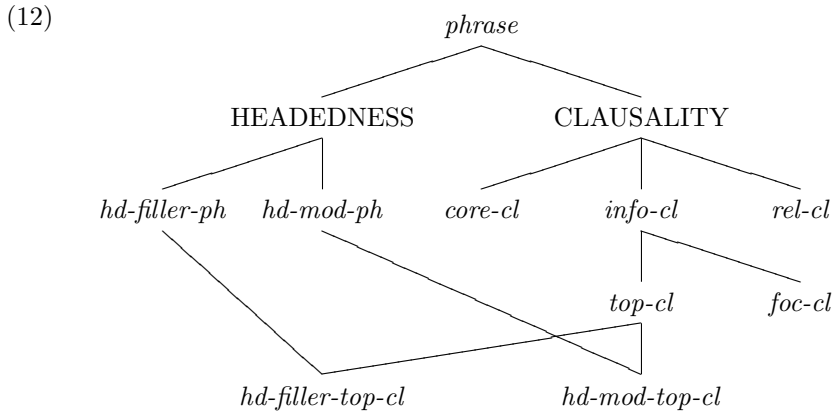
⁴ In a spoken context, *John-uy* could serve as an answer to (7)b.

⁵ The example is grammatical with a ‘specific’ reading. In a similar manner, a sentence like the following given by a reviewer is possible since the sentence refers to a specific person in the class.

- (i) (I pan-uy) nwukwunka-ka apeci-ka pwuca-i-ta.
 this class-GEN someone-NOM father-NOM rich-COP-DECL
 ‘(lit.) The father of someone in this classroom is rich.’

2. A Construction-Based Analysis

As a way of capturing generalizations about the shared properties of diverse construction types (including the MNCs here), our grammar adopts the notion of constructions from (Ginzburg and Sag, 2001) and classifies phrases in terms of HEADEDNESS and CLAUSALITY, as represented in (12):



As shown in the hierarchy here, each type of phrase is cross-classified, inheriting both from a CLAUSALITY type and from a HEADEDNESS type.⁶ The constraints on the subtypes of the HEADEDNESS will license well-formed phrases in the language (see (Kim, 2004)):

- (13) a. $XP[hd-spr-ph] \rightarrow \boxed{1}, \mathbf{H}[\text{SPR } \langle \boxed{1} \rangle]$
- b. $XP[hd-mod-ph] \rightarrow [\text{MOD } \langle \boxed{1} \rangle], \boxed{1}\mathbf{H}$
- c. $S[hd-filler-ph] \rightarrow \boxed{1}XP, S[\text{GAP } \langle \boxed{1} \rangle]$

These constraints on well-formed phrases, playing the same role as the structure licensed by X' -theory in transformational approaches, allow the combination of a head and its specifier, a head and its modifier, and a head and its filler, respectively. These constraints inherit to their subtypes like *hd-filler-top-cl* and *hd-mod-top-cl*, which also function as the subtypes of CLAUSALITY.

⁶ In addition to these well-formed phrases, the language has *hd-subj-ph*, *hd-comp-ph*, and *hd-lex-ex* for the combination of head with its subject, head with its complement, and head with another lexical element to form a complex predicate, respectively.

The subtypes of CLASUALITY include *core-cl*, *rel(ative)-cl*, and *info-cl*. The *core-cl* type includes canonical clauses like declarative and imperative. The constraints on *info-cl* are the locus of our treatment of the PNC and ANC. The type *info-cl* has at least two subtypes: *top-cl* and *foc-cl*, which have either a positive TOP(IC) or FOC(US) value. Each has its own constraints that will be inherited to its subtypes. For example, *top-cl* and *foc-cl* are declared to have the following constraints which will be inherited to their subtypes:⁷

(14) a. *top-cl*:

$$\left[\text{C-CONT} \mid \text{RELS} \left\langle \left[\begin{array}{l} \text{PRED } \textit{about} \\ \text{ARG1 } h3 \\ \text{ARG2 } h4 \end{array} \right] \right\rangle \right] \rightarrow \left[\begin{array}{l} \text{LBL } h3 \\ \text{TOP } + \end{array} \right], \text{S} \left[\begin{array}{l} \text{MOOD } \textit{decl} \\ \text{LBL } h4 \\ \text{IC } + \\ \text{SUBJ } \langle \quad \rangle \end{array} \right]$$

b. *foc-cl*:

$$\left[\text{C-CONT} \mid \text{RELS} \left\langle \left[\begin{array}{l} \text{PRED } \textit{characterized-by} \\ \text{ARG1 } h3 \\ \text{ARG2 } h4 \end{array} \right] \right\rangle \right] \rightarrow \left[\begin{array}{l} \text{GCASE } \textit{nom} \\ \text{FOC } + \\ \text{LBL } h3 \end{array} \right], \text{S} \left[\text{LBL } h4 \right]$$

The topic clause (*top-cl*) has as its constructional content (C-CONT) an *about-relation*: the topic phrase tells us what the main clause is about. The value of LBL is a handle, which is a token to its elementary predicate (EP) in the MRS system. We can see that the ARG values of *about* are the value of the topic phrase's LBL (h3) and that of the head S (h4). Meanwhile, the focus phrase (*foc-cl*) also has a constructional constraint indicated by the relation *characterizing*. That is, in a *foc-cl*, the focused initial phrase (having a grammatical case (GCASE) such as nominative and also being marked as a FOC phrase) is characterized by the following S. This in turn can mean that the focus value is given to the individual whose characterization is represented by the following S.

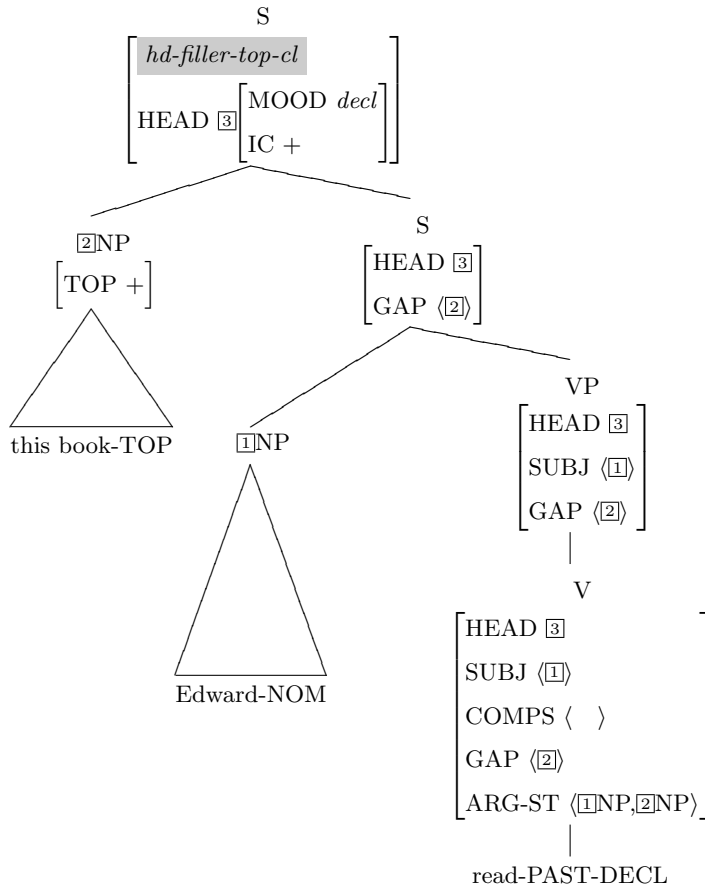
Now notice that the *top-cl* has two subtypes: *hd-filler-top-cl* and *hd-mod-top-cl*. The existence of two types of topic clause has been well attested:

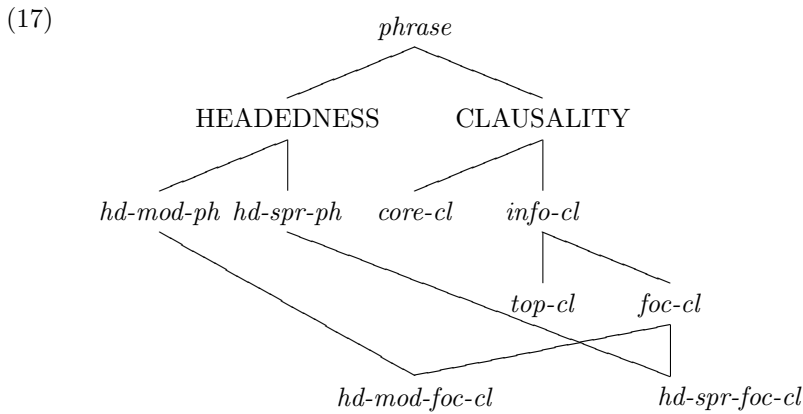
⁷ The meaning representation that we adopt here is MRS (Minimal Recursion Semantics) developed by (Copestake et al., 2003). See section 3.

- (15) a. ku chayk-un [Edward-ka ___ ilk-ess-ta] (hd-filler-top-cl)
 the book-TOP Edward-NOM read-PAST-DECL
 ‘The book, Edward read ___.’
- b. [ecey-nun [nalssi-ka chwu-ess-ta]] (hd-mod-top-cl)
 yesterday-TOP weather-NOM cold-PAST-DECL
 ‘As for yesterday, it was cold.’

In (15)a, the topic phrase *ku chayk-un* is an argument of the main predicate *ilk-ess-ta* and enters into a head-filler relation, whereas in (15)b, the topic *ecey-nun* is just an adjunct. The present grammar then assigns the following structures to these two examples:

- (16) a.





As defined here, the PNC is an instance of *hd-spr-foc-cl* whereas the ANC is an instance of *hd-mod-foc-cl*. This classification is motivated by the fact that in the PNC the first NOM functions as the specifier of the second NOM NP, whereas in the ANC it is just an adjunct, as exemplified in (18):

- (18) a. [John-i/uy [son-i khu-ta]] (*hd-spr-foc-cl*)
 John-NOM hand-NOM big
 ‘It is John’s hand which is big.’
- b. [yelum-i/-ey [maykcwu-ka mas-iss-ta]] (*hd-mod-foc-cl*)
 summer-NOM beer-NOM delicious-COP-DECL
 ‘It is summer when beer tastes delicious.’

This kind of multiple inheritance system for clausal types allows us to capture the generalizations among constructions. The constructional constraints on *foc-cl* are inherited to its subtypes, *hd-spr-foc-cl* and *hd-mod-foc-cl*. One thing to notice here is that in the *hd-mod-foc-cl* (ANC), the first NOM can be freely introduced if it has a positive MOD value. Meanwhile, in the *hd-spr-foc-cl* (PNC), the first NOM phrase is introduced as a specifier. The introduction of a new specifier, being in a *subordinate* relation with the subject, is introduced in accordance with the following lexical rule:⁸

⁸ As a reviewer points out, the current analysis does not address examples with multiple specifiers as in the following:

- (i) a. John-i chinkwu-ka apeci-ka pwuca-i-ta.
 John-NOM friend-NOM father-NOM rich-COP-DECL
 ‘It is the father of John’s friend who is rich.’

(19) SPR Lexical Rule:

$$v\text{-stative} \rightarrow \left[\begin{array}{l} v\text{-spr} \\ \text{VAL} \left[\begin{array}{l} \text{SPR } \langle \underline{2}_i \rangle \\ \text{SUBJ} \langle \left[\begin{array}{l} \text{SPR } \langle \underline{2} \rangle \\ \text{LBL } h6 \end{array} \right]_j \rangle \end{array} \right] \\ \text{SEM | RELS} \left\langle \dots, \left[\begin{array}{l} \text{PRED } \textit{subordinate} \\ \text{ARG1 } j \\ \text{ARG2 } i \end{array} \right], \dots \right\rangle \end{array} \right]$$

The effects of this lexical rule are as follows. It allows a stative verb taking one argument to be turned into a verb that selects an additional specifier which is in a *subordinate* relation with the subject.⁹ The two consecutive NOM phrases need to be in a certain semantic relation (e.g., the subordinate relation) in the PNC, as can be seen from the evidence in (20):

- (20) a. Pyeng-uy/*-i akhwak-ka i kyolkwa-lul cholayhayessta.
 illness-GEN/NOM worsening this result caused
 ‘The worsening of the illness caused this condition.’
- b. John-uy/*-i iphak-i wuli-lul nolla-key hayessta
 John-GEN/NOM admission-NOM we-ACC surprise-COMP did
 ‘John’s admission surprised us.’

An intransitive predicate like ‘big’ will be turned into a *v-spr* word by the lexical rule above:

-
- b. John-i elkwul-i oynccok-i alay-ka aphuta
 John-NOM face-NOM left-NOM bottom-NOM sick
 ‘John is sick in the face on the left side in the bottom.’

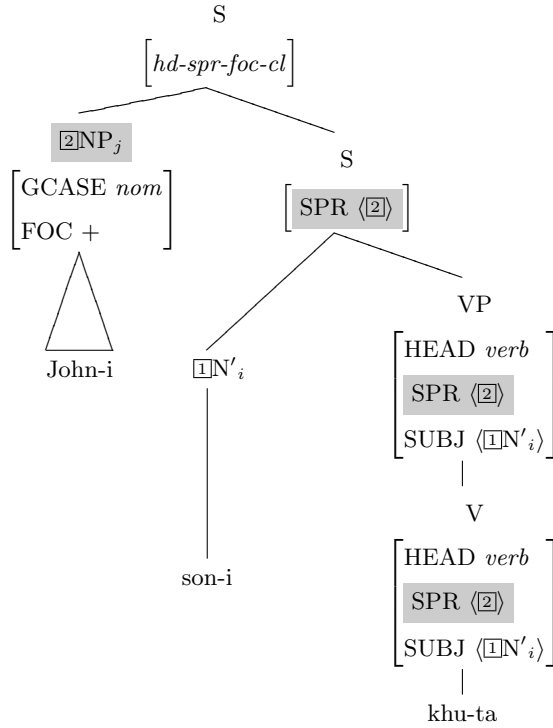
If we want to allow such examples (though hardly found in corpus examples), we need to modify the SPR Lexical Rule. See (Kim, 2001) for a concrete analysis for this direction.

⁹ The term *subordination* is borrowed from (Na and Huck, 1993). X is *thematically subordinate* to an entity Y iff Y’s having the properties that it does entails that X has the properties that it does. (Na and Huck, 1993) classify these thematic subordination relations into five types: *part-whole* (cover vs. book, voice vs. man, tail vs. dog), *quality-to-entity* (use vs. tool, color vs. eyes, taste vs. food), *conventional* ((hat vs. boy, nest vs. bird), *hierarchical* (parent vs. child, doctor vs. patient), and *taxonomic* (apple vs. fruit, chair vs. furniture, shirt vs. clothes, soccer vs. game).

- (21) a.
$$\left[\begin{array}{l} \text{PHON} \quad \langle \text{khu-} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{HEAD} \quad \textit{verb} \\ \text{VAL} \mid \text{SUBJ} \quad \langle \boxed{1} \text{NP}_i \rangle \end{array} \right] \\ \text{ARG-ST} \quad \langle \boxed{1} \rangle \\ \text{SEM} \mid \text{RELS} \left\langle \left[\begin{array}{l} \text{PRED} \quad \textit{big} \\ \text{ARG0} \quad \textit{s1} \\ \text{ARG1} \quad \textit{i} \end{array} \right] \right\rangle \end{array} \right]$$
- b.
$$\left[\begin{array}{l} \textit{v-spr} \\ \text{PHON} \quad \langle \text{khu} \rangle \\ \text{SYN} \left[\begin{array}{l} \text{HEAD} \quad \textit{verb} \\ \text{VAL} \left[\begin{array}{l} \text{SPR} \langle \boxed{3} \text{NP}_i \rangle \\ \text{SUBJ} \langle \text{N}'_j [\text{SPR} \langle \boxed{3} \rangle] \rangle \end{array} \right] \end{array} \right] \\ \text{ARG-ST} \quad \langle \boxed{1} \rangle \\ \text{SEM} \left[\begin{array}{l} \text{INDEX} \quad \textit{s1} \\ \text{RELS} \left\langle \left[\begin{array}{l} \text{PRED} \quad \textit{big} \\ \text{ARG0} \quad \textit{s1} \\ \text{ARG1} \quad \textit{i} \end{array} \right], \left[\begin{array}{l} \text{PRED} \quad \textit{subordinate} \\ \text{ARG1} \quad \textit{j} \\ \text{ARG2} \quad \textit{i} \end{array} \right] \right\rangle \end{array} \right] \end{array} \right]$$

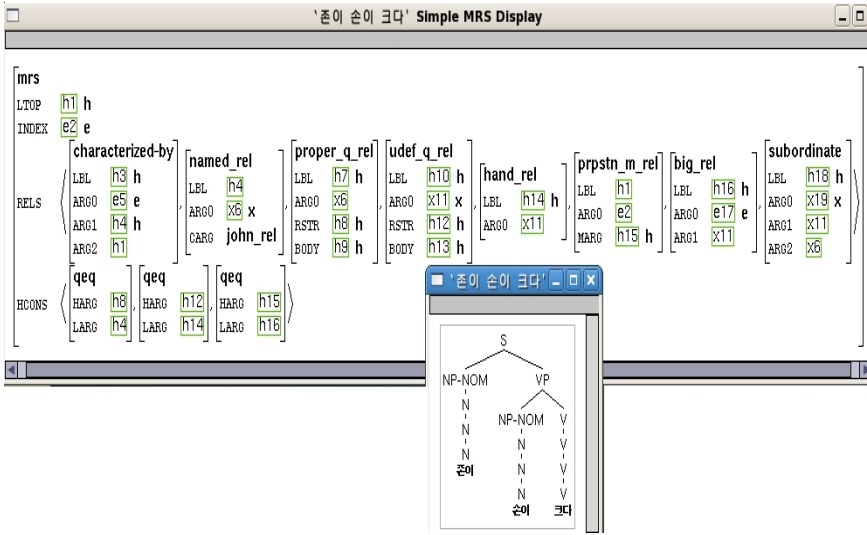
As sketched here, the generation of the PNC and the ANC is dependent upon interactions among different grammatical components, assigning the following structures to the examples (18):

(22)

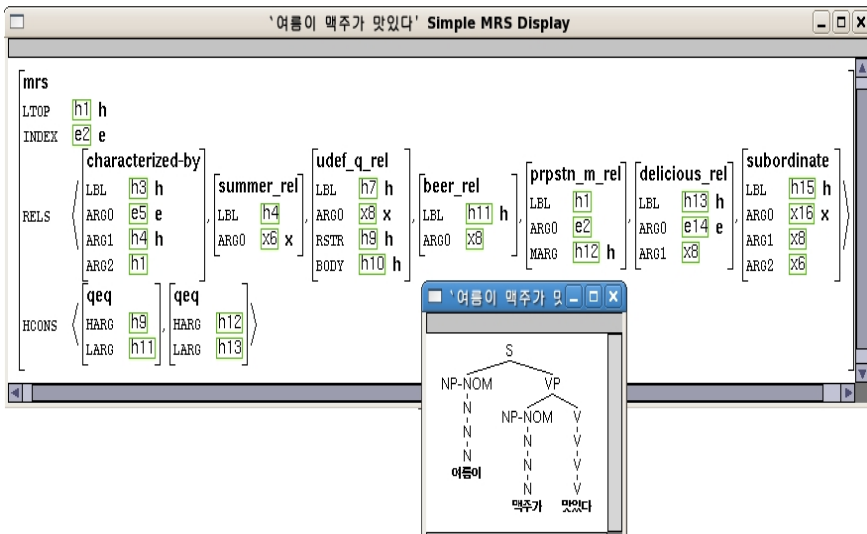


The SPR value of the lexical head *kuh-ta* is introduced by the SPR Lexical Rule. The SPR *John-i* is also in a *subordinate* relation with the subject *son-i*, all of which constitutes a well-formed *hd-spr-foc-cl*.

Now consider the structure of an ANC:



[Figure 1] Parsed Tree and MRS for ‘It is John whose hand is big.’



[Figure 2] Parsed Tree and MRS for ‘It is summer when beer tastes delicious.’

semantics designed to enable semantic composition using only the unification of type feature structures. (See (Copestake et al., 2003) and (Bender, Flickinger, and Oepen, 2002)) For example, Figure 1 and Figure 2 are the parsed results for the sentence (1a) and (1b) in our system.

We can see here that the MRS that 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 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*.¹¹ Each of the types *relation* has at least three features LBL, PRED (represented here as a type), and ARG0. We can see that the LBL value of *named_rel* and that of the *prpstn_m_rel* are both the arguments of the PRED relation *characterized-by*, capturing the pragmatic relations in the MNC. The two NOM phrases are also linked by the relation *subordinate*.

4. Conclusion

The so-called multiple nominative constructions present challenges to theoretical as well as computational linguists. In particular, the functions of the first NOM phrase in MNCs are not straightforward. The first NOM can be either a specifier or an adjunct, and it has a specific semantic relation with regard to the remaining sentence – it is ‘characterized’ by the rest of the sentence.

This paper shows that a grammar allowing interactions of declarative constraints on types of signs – in particular, constructions (phrases and clauses) – can provide an robust and efficient way of parsing these two different types of MNC. Though there is need for extending the current grammar to a wider range of authentic corpus data, displaying more complex properties of the language, the parsing results indicate that the current grammatical system is feasible enough.

<References>

- Bender, Emily M., Daniel P. 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 John Carroll, Nelleke Oostdijk, and Richard Sutcliffe (eds.), *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, pp. 8–14, Taipei, Taiwan.
- Copestake, Ann, Daniel Flickinger, Ivan Sag, and Carl Pollard. 2003. Minimal Recursion Semantics: An introduction. Manuscript.
- Ginzburg, Jonathan and Ivan A. Sag. 2001. *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. CSLI Publications, Stanford, California.

¹¹ The attribute HCONS is to represent quantificational information. See (Bender, Flickinger, and Oepen, 2002).

- Kim, Jong-Bok. 2001. A Constraint-Based and Head-driven Approach to Multiple Nominative Constructions. In Dan Flickinger and Andreas Kathol (eds.), *Proceedings of the HPSG-2000 Conference, University of California, Berkeley*, pp. 166–181, Stanford. CSLI Publications.
- Kim, Jong-Bok. 2004. *Korean Phrase Structure Grammar*. Hankwuk Publishing, Seoul. In Korean.
- Kim, Jong-Bok and Jaehyung Yang. 2004. Projections from Morphology to Syntax in the Korean Resource Grammar: Implementing Typed Feature Structures. In *Lecture Notes in Computer Science*, 2945. Springer-Verlag, pp. 13–24.
- Na, Younghee and G. J. Huck. 1993. On the Status of Certain Island Violations in Korean. *Linguistics and Philosophy* 16, 181–229.
- O’Grady, William. 1991. *Categories and Case*. John Benjamins Publishing, Amsterdam.
- Yoon, James. 2004. Non-nominative (Major) Subjects and Case Stacking in Korean. In P. Bhaskararao and K. V. Subbarao (eds.), *Non-nominative Subjects*, pp. 265–314, Amsterdam. John Benjamins.