Floating numeral classifiers in Korean: A thematic-structure perspective

Jong-Bok Kim
Kyung Hee University

Proceedings of the 18th International Conference on
Head-Driven Phrase Structure Grammar

University of Washington

Stefan Müller (Editor)

2011

Stanford, CA: CSLI Publications

pages 302–313

Kim, Jong-Bok. 2011. Floating numeral classifiers in Korean: A thematic-structure perspective. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 302–313. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2011.17.

Abstract

The syntactic and semantic complexity of the so-called numeral classifier (NUM-CL) constructions in languages like Korean (Japanese and Chinese as well) has much challenged theoretical as well as computational approaches. Among several types of the NUM-CL constructions, the most complicated type includes the so-called FQ (floated numeral classifier/quantifier) construction where the NUM-CL 'floats' away from its antecedent. This paper, couched upon the non-derivational VP-modifier view, shows that in addition to the grammatical function of the host NP and types of the main predicate, properties of the intervening expression between the FQ and its host NP also play an important role in licensing the FQ's distribution. In particular, we show that the FQ introduces new information in discourse and as default sets off rheme in the thematic structure. This functional analysis can provide an answer to several puzzling contrasts we observe in the distribution of the FQ.

1 The Issues

There exist at least three different environments where numeral classifiers (NUM-CL) in Korean can appear:

(1) a. Genitive-Case (GC) Type:

sey myeng-uy pemin-i iss-ta
three CL-GEN criminal-NOM exist-DECL
'There are three criminals.'

b. Noun Initial (NI) Type:

pemin sey myeng-i iss-ta
criminals three CL-NOM exist-DECL

c. Floated Numeral Classifier (FNC) Type:

pemin-i sey myeng iss-ta
criminals-NOM three CL exist-DECL

Though these three types of NUM-CL constructions behave similarly with respect to the propositional meaning, they are different in many syntactic and semantic respects. In the GC type, the NUM-CL appears with the genitive case marking, preceding the head noun *pemin* 'criminal' whereas in the NI, the NUM-CL sequence

†This work was supported by the National Research Foundation of Korea (NRF-2011-32A-A00066).

follows the head noun. Meanwhile, in the FNC type, the head noun is case-marked, followed by the NUM-CL. In this case, the NUM-CL can further 'float' away from the associated NP:

(2) **pemin-i** cengmal **sey** **myeng** te iss-ta
criminal-NOM really three CL more exist-DECL
'There are really three more criminals.'

In this example, the NUM-CL *sey myeng* and its antecedent NP *pemin-i* are not adjacent, but are separated by an intervening adverb, *cengmal* 'really'.

In the FNC type there are several constraints where the NUM-CL can be floated and with which argument the floated NUM-CL can be associated. For example, the NUM-CL just like adverbial elements, canonically has a free distribution, but cannot precede its host NP (cf. Kim 1984, Choi 1988, Lee 1989, Shi 2000, Kang 2002):

(3) *__sey myeng__ cengmal **pemin-i** te iss-ta
three CL really criminal-NOM more exist-DECL

Matters become complicated when an argument intervenes between the two. Literature has often noted that there is an asymmetry between subject and object (see Park and Sohn 1993, Kang 2002, Ko 2007 for Korean and Saito 1985 and Miyagawa 1989 for the same paradigm in Japanese):

(4) a. ??/*__haksayng-tul-i__ chayk-ul **sey** **myeng** ilkessta
student-PL-NOM book-ACC three CL read
'(int) Three students read books.'

b. **chayk-ul** haksayng-tul-i **sey** **kwen** ilkessta
book-ACC student-PL-NOM three CL read
'Students read three books.'

As seen from (4a), the object cannot intervene between the subject and its NUM-CL whereas as illustrated in (4b) such an effect disappears when the subject intervenes between the scrambled object and its NUM-CL.

Numerous attempts have been made to understand the grammatical properties of numeral classifier constructions, mainly focusing on how to generate the three types of NUM-CL and figure out the syntactic relations among these three if there are any. The generation of the GC and NI construction has been rather simple, but that of the FNC has been controversial. In the traditional 'stranding' view, the FNC construction is derived from the NI by moving the NP antecedent out of the VP, leaving the FNC and its trace behind (e.g., Sportiche 1988, Koopman and Sportiche 1991, Bošković 2004, Miyagawa 1989, Miyagawa and Arikawa 2007, for Japanese, Park and Sohn 1993, Ko 2005, 2007 for Korean). However, there are many facts arguing against this kind of movement assumption, but support for

the base-generated VP-modifier view (e.g., Bobaljik 2003 for English, Fukushima 1991, Gunji and Hasida 1998, Kuno and Takami 2003 for Japanese, Kang 2002, Shi 2000, Kim and Yang 2007 for Korean). In the VP-modifier view we support here too, the floated-away NUM-CL is simply taken to modify a verbal predicate in situ and quantifies over the event that the VP denotes.

In addition to supporting this VP-modifier view, this paper also suggests that the main function of the floated NUM-CL is to start off rheme in the thematic structure. This functional approach, accompanied by the VP-modifier view, can provide the subtle contrasts previous literature have tried to capture. In so doing, in what follows, we will first review some formal properties of the three types and then discuss the pros and cons of the stranding and VP-modifier view. We then discuss how the functionally-motivated thematic structure can account for the phenomena in question.

2 Some Main Properties of the Three Types

As indicated by the name of the three NUM-CL constructions, the possible case value on the NUM-CL in each is different. In particular, even though the NI type can host almost any semantic case marker, the FNC type allows only nominative or accusative on the NUM-CL (cf. Choi 2001). In terms of syntactic structures, the three types also display clear differences. For example, coordination shows us a main difference with respect to constituenthood: the GC and the NI type can participate in coordinate constructions, but the FNC type cannot:

(5) a. Kim-un [[sey kwen-uy kongchayk]-kwa [twu calwu-uy
Kim-TOP three CL-GEN notebook-CONJ two CL-GEN
yenphil]]-ul sassta.
pencil-ACC bought
'Kim bought three notes and two pencils.'

b. Kim-un [[kongchayk sey kwen]-kwa [yenphil twu calwu]]-lul
Kim-TOP notebook three CL-CONJ pencil two CL-ACC
sassta
bought

c. ??*Kim-un [[kongchayk-ul sey kwen]-kwa [yenpil-ul twu
Kim-TOP notebook-ACC three CL-CONJ pencil-ACC two
calwu]] sassta.
CL-ACC bought

The syntactic differences among the three types also lead to subtle semantic and pragmatic differences. Unlike the GC and NI type, the FNC construction licenses a partitive reading. Consider the following set of data:

(6) a. Seoul-lo tomangka-n tases myeng-uy haksayng-i tolawassta
Seoul-to run-away-PNE five CL-GEN student-NOM returned
'The five students who ran away for Seoul returned.'

b. Seoul-lo tomangka-n haksayng tases myeng-i tolawassta.
Seoul-to run-away-PNE student five CL-NOM returned
'The five students who ran away for Seoul returned.'

c. Seoul-lo tomangka-n haksayng-i tases myeng-(i) tolawassta.
Seoul-to run-away-PNE student-NOM five CL-NOM returned
'Of those who ran away for Seoul, just five returned.'

The examples (6a) and (6b) are true in the situation where there are five students who left for Seoul, and they all came back. Meanwhile, the preferred reading of (6c) is such that there are more than five students who left for Seoul and of them just five returned, thus licensing a partitive reading here.

We can also observe a difference in the specific and nonspecific reading. The NI allows either a specific or nonspecific reading whereas the FNC allows only a nonspecific reading (cf. Lee 1989, Kim 2005):

(7) a. pemin twu myeng-i ecey tomangkassta
criminal two CL-NOM yesterday ran.away
'Two (specific or nonspecific) criminals ran away yesterday.'

b. pemin-i ecey twu myeng-i tomangkassta
criminal-NOM yesterday two CL-NOM ran.away
'Of the criminals, two (nonspecific) ran away.'

As given in the English glosses here, in the NI type, the two criminals can be either specific or nonspecific whereas in the FNC, they can be only nonspecific. With respect to this reading, (7b) can be interpreted as having a partitive and nonspecific reading such that there are a set of criminals and of the members in this set, two unspecific criminals ran away. No such reading is available in the NI type (or the GC) type.

3 Two Different Approaches for the FNC Construction

Stranding Approaches: The traditional wisdom of dealing with a FNC example has been the stranding approach, trying to link the NI or GC type to the FNC type. For instance, the FNC type is derived from the following source with movement processes (cf. Miyagawa 1989, Miyagawa and Arikawa 2007, Lee 1989, etc).

(8) Mia-nun [$_{VP}$chayk$_i$-ul [$_{VP}$ ceketo [$_{VP}$t$_i$ sey kwen-(ul) ilkessta]]]

The NP *chayk* 'book', being in the same local domain (e.g., mutual c-commanded) with the NUM-CL is moved out of the VP, stranding behind the NUM-CL in the original position. The claimed argument for this stranding view follows from the strict locality condition between a NUM-CL and its associated NP. That is, if they two are not adjacent to each other, the NUM-CL has been 'stranded' by the NP. This locality requirement has been motivated from the contrast between subject and object, which we have seen earlier. However, as even the proponents of the stranding approach acknowledge (cf. Miyagawa and Arikawa 2007), there is a question if such an example is really unacceptable. There are many examples where a similar ordering is acceptable. In particular, a case marking or a delimiter marker on the NUM-CL, makes the following acceptable:

(9) **haksayng-tul-i** [maykcwu-lul [**sey myeng-i/ina/man**] masiessta]
students-NOM beer-ACC three CL-NOM/even/only drank
'Even/Only three of the students drank beer.'

VP-modifier Analyses: Unlike the standing analysis, the VP modifier analysis assumes that there is no transformation relation between the NI or GC and FNC version (Fukushima 1991, Gunji and Hasida 1998 for Japanese, Kang 2002 and Kim and Yang 2007 for Korean). Contrary to the stranding view, the VP-modifier view assumes the NUM-CL directly combines with a verbal predicate in syntax and semantically modifies an event structure the predicate denotes:

(10) pemin-i cengmal [$_{VP}$ sey myeng [$_{VP}$ te iss-ta]]
criminal-NOM really three CL more exist-DECL
'There are really at least three more criminals.'

As given in the structure, there is no movement: the NUM-CL just modifies the VP.

Several welcome predictions follow from this view. First of all, the VP-modifier view will predict the distributional possibilities of the NUM-CL as an adverbial element. As we have seen, the NUM-CL can appear where an adverbial element can occur otherwise constraint such as it cannot precede its associate NP. In addition, since there are no direct links between the FNC type with the other two types, we expect each will behave differently in many syntactic and semantic aspects. This has been true as we have observed so far. Additional support can also find from semantic aspects: the NI or GC type induces either specific or non-specific whereas the FNC has only nonspecific. This also has to do with the fact that the FNC allows only a narrow scope reading when interacting with another scope operator such as negation:

(11) a. NI Type: $\exists 3 >$ NOT or NOT $> \exists 3$

haksayng sey myeng-i acik ttenaci anh-ass-ta
students three CL-NOM still leave-COMP not-PAST-DECL
'Three students haven't left yet or
these three students still didn't left yet.'

b. FNC Type: *∃3 > NOT or NOT > ∃3

haksayng-i acik **sey** **myeng-i** ttenaci anh-ass-ta
student-NOM still three CL-NOM leave-COMP not-PAST-DECL
'Three students haven't left yet,
(even though more had already left)'

The VP modifier approach allows us to specify that the floated NUM-CL has a narrower scope reading than the VP it modifies, which may not be an easy task within a stranding approach.

4 A Functional Account

Even if we adopt the VP-modifier approach, puzzles still remain: why certain FNC examples are bad or at least unnatural. As we have seen so far, when the FNC is marked with a case marker or a delimiter, the intervening effects or contrasts (subject and object asymmetry, unaccusative/unergative contrast) disappear. In this paper, we suggest that the floated NUM-CL and the modified verbal predicate serve as rheme in the thematic structure. In particular, we claim that the floated NUM-CL starts off the rheme in a given clause.

Based on the interactions between information and intonation structure in partitioning theme and rheme, we assume that the FNC is subject to the functional constraint that the number of referents it denotes conveys new information, part of the rheme in the thematic structure. Together with this notion of thematic structure, we suggest that the floated NUM-CL sets apart theme and rheme whose constraint can be paraphrased as in (12):

(12) Thematic Constraint in Korean:
A floated NUM-CL in Korean sets off rheme in the thematic structure.

The constraint states that the floated NUM-CL marks the beginning of rheme which may contain both old and new information, but tells about the information about the theme.

According to this, the subject-object asymmetry follows immediately. What sets off the rheme in a given clause determines the degree of acceptability: In (4a), it is not the NUM-CL but the intervening object that marks the beginning of rheme, which violates the constraint in (12). Note that unlike (4a), we have seen that in (4b) the subject can intervene between the object and its NUM-CL, whose example we repeat here:

(13) chayk-ul haksayngtul-i ‖ sey kwen ilkessta
book-ACC student-NOM three CL read
'Students read three books.'

In this example, both the fronted object and the subject are theme elements, and thus the NUM-CL starts the rheme of the sentence (marked with the symbol ||), observing the thematic constraint.

Further welcoming effects of this constraint can be observed from the following set of examples:

(14) a. haksayngtul-i || **sey myeng(-i)** tosekwan-eyese chayk-ul
students-NOM three CL-NOM libary-at book-ACC
ilkessta
read
'As for the students, there were three who read the book at the library.'

b. haksayngtul-i tosekwan-eyese || **sey myeng(-i)** chayk-ul ilkessta
'As for the students at the libary, there were three who read the book.'

c. haksayngtul-i tosekwan-eyese chayk-ul || **sey myeng(-i)** ilkessta
'As for the students at the libary and as for the book, there were three who read it.'

As pointed out earlier and illustrated here, the adverbial NUM-CL can appear in various places, but induces subtle differences in the thematic structure. Given our thematic constraint, the expressions preceding the NUM-CL are themes in the clause whereas those following it are the members of rheme. This results in the subtle meaning differences here as indicated in the English glosses (see Kim 2005 for similar meaning differences among these), which would be hard to capture otherwise.

Note that the thematic constraint can also account for the difference between high and low adverb with respect to the distribution of a floated NUM-CL:

(15) a. ai-tul-i ecey sey myeng kyosil-eyse wusessta
child-PL-NOM yesterday three CL classroom-at laughed
'Three children laughed at the classroom yesterday.'

b. *ai-tul-i khu-key sey myeng wusessta
child-PL-NOM loudly three CL laughed
'Three children laughed loudly at the classroom.'

Within the traditional stranding view, the contrast follows from the following derivations:

(16) a. $[_{TP}$ ai-tul-i $[_{VP}$ ecey $[_{VP}[t_i$ sey myeng] kyosil-eyse wusessta]]]

b. *$[_{TP}$ ai-tul-i $[_{VP}$ t_i $[_{VP}$[khu-key sey myeng wuessta]]]]

In (16a) with the high adverb 'yesterday', the NUM-CL is in the same local domain with its associated subject here, observing the strict locality requirement. However, in (16b) with the low adverb 'loudly', the subject and its NUM-CL are not in the same local domain. Once again, note that when the NUM-CL has a focus marker, indicating the starting point of the rheme, the grammaticality improves a lot:

(17) a. ai-tul-i || sey myeng-i khu-key wusessta
child-PL-NOM three CL-NOM loudly laughed
'Three children laughed loudly.'

b. ai-tul-i khu-key || sey myeng-ina wusessta

c. ai-tul-i ecey khu-key || twu myeng-i/ina wuessta

In our thematic constraint, given the assumption that either a manner adverb or a floated NUM-CL can set off the rheme, the accceptablity of all these examples then follows straightforwardly.

Our account, resorting to the thematic structure, can also get support from the claimed contrast between unergative and unaccusative (cf. Ko 2007):

(18) a. koyangi-ka pyeng-ulo sey mali cwukessta
cat-NOM illness-of three CL died
'Three cats died of illness.'

b. ?*haksayng-tul-i caki-uy ton-ulo twu myeng
student-PL-NOM self-GEN money-with two CL
cenhwahayessta
phoned
'Two students made a phone call with their own money.'

In the stranding view, (18b) with the unergative verb 'phoned' violates the strict locality condiiton betwen the subject and its NUM-CL. However, note that the grammaticality of (18b) improves greatly with supporting elements:

(19) haksayng-tul-i caki ton-ulo cikcep Seoul-ey || twu
student-PL-NOM self money-with without.help Seoul-at || two
myeng cenhwahayessta
CL phoned
'Two students made a phone call to Seoul with their own money without any help.'

In the context where it is important to see how many students made a phone call to Seoul by themselves, such a sentence is more than acceptable, supporting our analysis. It is also not difficult to construct acceptable unergative examples with the same configuration:

(20) a. haksayng-tul-i himtulkey caki ton-ulo || twu myeng
student-PL-NOM hard self money-with two CL
mikwuk-ey kassta
America-to went
'Two students went to America with difficulties with their own money.'

b. haksayng-tul-i pwumonim towum epsi caki ton-ulo ||
student-PL-NOM parents help without self money-with
twu myeng mikwuk-ey kassta
two CL America-to went
'Two students went to America for themselves
with their own money and without their parents' help.'

What this means is that being the subject of an unergative verb does not block its NUM-CL from being floated or being in a nonlocal position. Our conjecture is that the unacceptability of (18b) is rather related to the thematic constriant: the phrase, *caki ton-ulo*, is rheme, so that the NUM-CL cannot starts the rheme component, violating our thematic constraint.

5 Conclusion

The syntactic and semantic complexity of the so-called numeral classifier (NUM-CL) constructions in languages like Korean (Japanese and Chinese as well) has much challenged theoretical as well as computational approaches. Among several types of the NUM-CL constructions, the most complicated type includes the one where the NUM-CL 'floats' away from its antecedent.

This paper supports a non-movement approach for the NUM-CL constructions, in particular, a VP-modifier approach for the floated NUM-CL (FNC) construction. In the paper, we claim that the main function of the FNC is to set off rheme in the thematic structure, cued by both information and intonation tunes. Further supported by a pilot prosodic test, this functional-based approach can provide us with a streamlined analysis for various distributional possibilities of the FNC without resorting to movement operations.

References

Bobaljik, Jonathan. 2003. Floating Quantifiers: Handle with Care. In L. Cheng and R. Sybesma (eds.), *The Second Glot International State-of-the-Article Book: The Latest in Linguistics*, 107-148. Berlin: Mouton de Gruyter.

Bošković, Aeljko. 2004. Be Careful Where You Float Your Quantifiers. *Natural Language and Linguistic Theory* 22: 681 – 742.

Choi, Kiyong. 2001. The Structure and Interpretation of Non-genitive Numeral Classifier Constructions in Korean. *Language Research* 37.3: 445–480.

Choi, Young-Seok. 1988. A Reanalysis of Quantifier Floating in Korean. *Korean Journal of Linguistics* 13: 109–147.

Downing, Pamela. 1993. Pragmatic and Semantic Constraints on Numeral Quantifier Position in Japanese. *Linguistics* 29: 65-93.

Fitzpatrick, Justin. 2006. Two Types of Floating Quantifiers and Their A/A-bar Properties. In the *Proceedings of the 36th Conference of the North East Linguistic Society (NELS 36)*, 253–256.

Fukushima, Kazuhiko. 1991. Phrase Structure Grammar, Montague Semantics, and Floating Quantifiers in Japanese. *Linguistics and Philosophy* 14, 581-628.

Gunji, Takao and Koiti Hasida. 1998. Measurement and Quantification. In T. Gunji and K. Hasida (eds.) *Topics in Constraint-Based Grammar of Japanese,* 39–79. Dordrecht: Kluwer Academic Publishers.

Kang, Beom-mo. 2002. Categories and Meanings of Korean Floating Quantifiers-with Some Reference to Japanese. *Journal of East Asian Linguistics* 11: 375-398.

Kim, Alan Hyun-Oak. 1988. Why Quantifiers Float. Paper Presented at the LSA Meeting, San Diego, CA.

Kim, Christine. 2005. Order and Meaning: Numeral Classifiers and Specificity in Korean. In Alderete, John (ed.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 218-226. Somerville, MA: Cascadilla Proceedings Project.

Kim, Jong-Bok and Jaehyung Yang. 2007. Syntax and Semantics of Korean Numeral Classifier Constructions. Müller, Stefan (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 163–172. Stanford: CSLI On-line Publications.

Kim, Young-Hee. 1984. *A Syntax of Quantification in Korean*. Tower Press, Seoul. [written in Korean]

Ko, Heejeong. 2005. *Syntactic Edges and Linearization.* Doctoral dissertation, MIT, Cambridge, MA.

Ko, Heejeong. 2007. Asymmetries in Scrambling and Cyclic Linearization. *Linguistic Inquiry* 38.1: 49-83.

Koopman, Hilda, and Dominique Sportiche. 1991. The Position of Subjects. *Lingua* 85: 211-258.

Kuno, Susumu, and Ken-Ichi Takami. 2003. Remarks on Unaccusativity and Unergativity in Japanese and Korean. In William McClure (ed.), *Japanese/Korean linguistics 12*, 280–294. Stanford, Calif.: CSLI Publications.

Lee, Chungmin. 1989. (In)definites, Case markers, Classifiers and Quantifiers in Korean. In Susumu Kuno et al. (eds.), *Harvard Studies in Korean Linguistics*

3, 469–487. Cambridge, Mass.: Harvard University, Department of Linguistics.

Miyagawa, Shigeru. 1989. *Structure and Case Marking in Japanese*. Academic Press, New York.

Miyagawa, Shigeru, and Koji Arikawa. 2007. Locality in Syntax and Floated Numeral Quantifiers. *Linguistic Inquiry* 38.4: 645–670.

Park, Myung-Kwan, and Keun-Won Sohn. 1993. Floating Quantifiers, Scrambling, and the ECP. In Choi, Soonja (ed.), *Japanese/Korean linguistics* 3, 187–203. Stanford, Calif.: CSLI Publications.

Saito, Mamoru. 1985. *Some Asymmetries in Japanese and their Theoretical Implications.* Ph.D. thesis, MIT.

Shi, Chung-Kon. 2000. The Syntactic structure of Quantifier Phrase in Korean. *Korean Journal of Linguistics*, 25.1: 73–101.

Sportiche, Dominique. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19.3: 425–449.